

# TRANSFORMATION OF INDEPENDENT VARIABLES IN POLYNOMIAL REGRESSION MODELS

---

A. OKOLO

(Received 24 July 2008; Revision Accepted 5 December 2008)

## ABSTRACT

In representing a relationship between a response and a number of independent variables, it is preferable when possible to work with a simple functional form in transformed variables rather than with a more complicated form in the original variables. In this paper, it is shown that linear transformations applied to independent variables in polynomial regression models affect the t ratio and hence the statistical significance for certain parameters of the polynomial regression models.

**KEY WORDS:** Transformation, Polynomial regression, Independent variables.

## INTRODUCTION

Transformations in experimental, mathematical and statistical work have found use in two major areas:

- i. that of providing theoretical approximations, and
- ii. that of bending the data to conform to assumptions underlying conventional analysis.

The usual techniques, according to Draper and Smith (1998), for the analysis of linear models as exemplified by regression analysis are usually justified by assuming:

1. Simplicity of structure for average values for the dependent variable,  $y$ ,
2. constancy of error variance about these average values,
3. normality of distributions, and
4. independence of observations.

Bates and Watts (2007) reported that if assumptions 1 and 3 are not satisfied in terms of the original observations,  $y$ , a linear or non-linear transformation of variables might be desirable to produce the simplest possible regression model in the transformed variables.

In regression problems, assumption 1 might be that  $E(y)$  is adequately represented by a rather simple empirical function of the independent variables  $x_1, x_2, \dots, x_n$  and we would want to transform so that this assumption together with assumptions 2 and 3, is approximately satisfied. Each of the assumptions 1 and 3 can and has been used separately to select a suitable candidate from a parametric family of transformations (see Neter, Wasserman and Kutner, 1989). The majority of the literature on transformation is concerned with transforming the  $y$ 's to achieve simplicity when the necessary assumptions could not otherwise be realistically made (Hocking, 1983; Bates and Watts, 2007). Most emphasis has been placed therefore on transformations of  $E(y)$  which may be expected to stabilize the variance or reduce the regression function to linearity in the parameters.

In regression problems, one can transform both the dependent and the independent variables (Bernhardt and Jung, 1979). It has been pointed out in Weisberg (2005) that replacing either the dependent, the independent, or both by nonlinear transformations of them is an important tool that the analyst can use to extend the number of problems for which linear regression methodology is appropriate. However, when the independent variables are transformed, it is shown in this paper that additivity is a concern in tests of statistical significance for individual parameters.

## TRANSFORMATIONS ON THE X'S

Suppose we have a model under consideration, which can be written in the form

$$Y = x\beta + \varepsilon, \quad (1)$$

where  $Y$  is the  $n \times 1$  vector of dependent variables,  $x$  is the  $n \times k$  matrix of independent variables,  $\beta$  is the  $k \times 1$  vector of parameters to be estimated and  $\varepsilon$  is an  $n \times 1$  vector of errors. In the present investigation we suppose that the errors in the  $Y$ 's are at least approximately normally and independently distributed with constant variance  $\sigma^2$ , where  $E(\varepsilon) = 0$  and  $\text{Var}(\varepsilon) = I\sigma^2$ .

The least squares estimators of  $\beta$  and  $\sigma^2$  are

$$\hat{\beta} = (x'x)^{-1}x'Y \quad (2)$$

and

$$\hat{\sigma}^2 = (y - x\hat{\beta})'(y - x\hat{\beta})/(n - k) \quad (3)$$

respectively (Draper and Smith, 1998).

Consider a transformation  $Z = xT$  of the matrix  $x$ . The transformed model becomes

$$Y = Z\gamma + \varepsilon = xT\gamma + \varepsilon = x(T\gamma) + \varepsilon = x\beta + \varepsilon,$$

so  $\gamma$  is a  $k \times 1$  vector of coefficient parameters, equal to  $T^{-1}\beta$ . It then follows that the least squares estimator of  $\gamma$  is  $\hat{\gamma} = T^{-1}\hat{\beta}$ . Since  $Z\hat{\gamma} = x\hat{\beta}$ , the two models have the same predicted

values and residuals; values of  $R^2$  and  $\hat{\sigma}^2$  also are identical. Since  $\hat{\sigma}^2$  is the same, for simplicity sake, let it be unity. The respective variance – covariance matrices of the estimators  $\hat{\beta}$  and  $\hat{\gamma}$  are  $(x'x)^{-1}$  and  $T^{-1}(x'x)^{-1}(T^{-1})'$  respectively.

## TRANSFORMATIONS IN POLYNOMIAL REGRESSION MODELS ILLUSTRATIVE EXAMPLE

Consider quadratic models in (4) and (5) with the transformation  $Z_i = a + bx_i$ :

$$Y_i = \beta_0 + \beta_1x_i + \beta_2x_i^2 + \varepsilon_i \quad (4)$$

$$Y_i = \gamma_0 + \gamma_1z_i + \gamma_2z_i^2 + \varepsilon_i \quad (5)$$

$$= \gamma_0 + \gamma_1(a + bx_i) + \gamma_2(a + bx_i)^2 + \varepsilon_i$$

$$= (\gamma_0 + a\gamma_1 + a^2\gamma_2) + (b\gamma_1 + 2ab\gamma_2)x_i + b^2\gamma_2x_i^2 + \varepsilon_i$$

From above,  $\beta = T\gamma$  when

$$T = \begin{bmatrix} 1 & a & a^2 \\ 0 & b & 2ab \\ 0 & 0 & b^2 \end{bmatrix},$$

so that

$$T^{-1} = \begin{bmatrix} 1 & -a/b & a^2/b^2 \\ 0 & 1/b & -2a/b^2 \\ 0 & 0 & 1/b^2 \end{bmatrix}.$$

Using the previous result that  $\hat{\gamma} = T^{-1}\hat{\beta}$ , it follows that

$$\hat{\gamma} = \begin{bmatrix} \hat{\gamma}_0 \\ \hat{\gamma}_1 \\ \hat{\gamma}_2 \end{bmatrix} = T^{-1}\hat{\beta} = \begin{bmatrix} 1 & -a/b & a^2/b^2 \\ 0 & 1/b & -2a/b^2 \\ 0 & 0 & 1/b^2 \end{bmatrix} \begin{bmatrix} \hat{\beta}_0 \\ \hat{\beta}_1 \\ \hat{\beta}_2 \end{bmatrix}$$

$$= \begin{pmatrix} \hat{\beta}_0 - \frac{a}{b}\hat{\beta}_1 + \frac{a^2}{b^2}\hat{\beta}_2 \\ \frac{1}{b}\hat{\beta}_1 - \frac{a^2}{b^2}\hat{\beta}_2 \\ \frac{1}{b^2}\hat{\beta}_2 \end{pmatrix} \quad (6)$$

Denote the elements of  $(x'x)^{-1}$  by  $(C_{ij})$ ,  $i, j = 0, 1, 2$ .

Then

$$(x'x)^{-1} = \begin{pmatrix} C_{00} & C_{01} & C_{02} \\ & C_{11} & C_{12} \\ & & C_{22} \end{pmatrix} \quad (7)$$

where  $(C_{ij}) = (C_{ji})$ , so that

$$\text{Var}(\hat{\gamma}) = T^{-1} (x'x)^{-1} (T^{-1})^1$$

$$= \begin{pmatrix} 1 & -a/b & a^2/b^2 \\ 0 & 1/b & -2a/b^2 \\ 0 & 0 & 1/b^2 \end{pmatrix} \begin{pmatrix} C_{00} & C_{01} & C_{02} \\ C_{10} & C_{11} & C_{12} \\ C_{20} & C_{21} & C_{22} \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ -a/b & 1/b & 0 \\ a^2/b^2 & -2a/b^2 & 1/b^2 \end{pmatrix}$$

$$= \begin{pmatrix} C_{00} - \frac{a}{b}C_{10} + \frac{a^2}{b^2}C_{20} & C_{01} - \frac{a}{b}C_{11} + \frac{a^2}{b^2}C_{21} & C_{02} - \frac{a}{b}C_{12} + \frac{a^2}{b^2}C_{22} \\ \frac{1}{b}C_{10} - \frac{2a}{b^2}C_{20} & \frac{1}{b}C_{11} - \frac{2a}{b^2}C_{21} & \frac{1}{b}C_{12} - \frac{2a}{b^2}C_{22} \\ \frac{1}{b^2}C_{20} & \frac{1}{b^2}C_{21} & \frac{1}{b^2}C_{22} \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ -a/b & 1/b & 0 \\ a^2/b^2 & -2a/b^2 & 1/b^2 \end{pmatrix}$$

$$\begin{pmatrix} C_{00} - \frac{2a}{b}C_{01} + \frac{2a^2}{b^2}C_{02} + \frac{a^2}{b^2}C_{11} - \frac{2a^3}{b^3}C_{12} + \frac{a^4}{b^4}C_{22} & \frac{1}{b}C_{01} - \frac{a}{b^2}C_{11} + \frac{3a^2}{b^3}C_{12} - \frac{2a}{b^2}C_{02} - \frac{2a^3}{b^4}C_{22} & \frac{1}{b^2}C_{02} - \frac{a}{b^3}C_{12} + \frac{a^2}{b^4}C_{22} \\ & \frac{1}{b^2}C_{11} - \frac{4a}{b^3}C_{12} + \frac{4a^2}{b^4}C_{22} & \frac{1}{b^3}C_{12} - \frac{2a}{b^4}C_{22} \\ & & \frac{1}{b^4}C_{22} \end{pmatrix} \cdot 8$$

Let  $t(\beta_i)$  and  $t(\gamma_i)$  denote the  $t$  ratio associated with  $\beta_i$  and  $\gamma_i$  respectively. Using (7),

$$t(\beta_1) = \frac{\hat{\beta}_1}{\sqrt{C_{11}}} \text{ and } t(\beta_2) = \frac{\hat{\beta}_2}{\sqrt{C_{22}}}. \quad (9)$$

Furthermore, using (6) and (8), we obtain

$$t(\gamma_1) = \frac{\frac{1}{b} \left( \hat{\beta}_1 - \frac{2a}{b} \hat{\beta}_2 \right)}{\frac{1}{b} \sqrt{C_{11} - \frac{4a}{b} C_{12} + \frac{4a^2}{b^2} C_{22}}} \text{ and } t(\gamma_2) = \frac{\frac{1}{b^2} \hat{\beta}_2}{\frac{1}{b^2} \sqrt{C_{22}}}. \quad (10)$$

We observe that  $t(\beta_1) \neq t(\gamma_1)$ , while  $t(\beta_2) = t(\gamma_2)$ . Notice that  $t(\gamma_1)$  will equal  $t(\beta_1)$  when  $a$  is zero. That is, the additive quantity  $a$ , in the transformation  $Z = a + bx$  is responsible for the non-equality of the  $t$  ratio for the lower order coefficient.

## CONCLUSION

The most frequent purpose of transformations is to achieve a mean function that is linear in the transformed scale. Over the years, emphasis for transformation has tended to be on obtaining a constant error variance. In all cases, we are concerned not merely to find a transformation which will justify assumptions but rather to find, where possible, a metric in terms of which the findings may be succinctly expressed. Simplicity and ease commend transformations because it is much simpler to transform the variables, estimate the coefficients, and merely inspect the  $t$  ratios for possible differences.

It has been shown that additivity is a problem when using linear transformation in polynomial regression models.

## REFERENCES

- Bates, D. M. and Watts, D. G., 2007. *Nonlinear Regression Analysis and Its Applications*, John Wiley and Sons, Inc., New York, 32pp.
- Bernhardt, I. and Jung, B. S., 1979. The Interpretation of Least Squares Regression with Interaction or Polynomial Terms. *The Review of Economics and Statistics*, 61: 481 – 483.
- Draper, N. R. and Smith, H., 1998. *Applied Regression Analysis*. 3<sup>rd</sup> Edition, John Wiley and Sons, Inc., New York, 253pp.
- Hocking, R. R., 1983. Developments in Linear Regression Methodology (with discussion). *Technometrics*, 25 (3): 219 – 249.
- Neter, J., Wasserman, W. and Kutner, M. H., 1989. *Applied Linear Regression Models*. 2<sup>nd</sup> Edition, Homewood, NJ, Irwin, 506pp.
- Weisberg, S., 2005. *Applied Linear Regression*. 3<sup>rd</sup> Edition, John Wiley and Sons, Inc., New York, 146pp.