Haybach, Kuechenhoff:

Testing for a Breakpoint in Two-Phase Linear and Logistic Regression Models

Sonderforschungsbereich 386, Paper 77 (1997)

Online unter: http://epub.ub.uni-muenchen.de/

Testing for a Breakpoint in Two–Phase Linear and Logistic Regression Models

Institute of Statistics
University of Munich

by

Gerhard Haybach
and
Helmut Küchenhoff

Contents

Abstract

In many practical problems, it is of interest to check whether a functional relationship between an explanatory and a response variable remains unchanged over the whole domain of the explanatory variable or whether the functional form changes at certain unknown points, the so–called breakpoints. Thus, testing for the existence of a breakpoint is often an essential task.

In this paper, we consider likelihood–ratio tests for different regression models such as broken line and threshold models. The problem related to the use of likelihood–ratio tests in this context concerns the determination of the null distribution of the likelihood–ratio statistic which has not been solved yet analytically. It is shown by means of Monte–Carlo experiments that the proposals of a limiting distribution discussed in the literature often yield unreliable results. It is therefore recommended to determine appropriate critical values by simulating the null distribution according to the data situation under investigation.

1 Introduction

Frequently in regression problems a model is used which assumes that the regression function is of a single parametric form. However, in many practical situations it is necessary to consider regression functions f which have different parameters in different regions of the domain of the independent variable. The functional relationship between the response Y and the explanatory variable X changes at certain points of the domain of X, the so–called *breakpoints*, also named *change-points*. In the literature, we find a wide application of different types of these segmented regression models, see for example Seber and Wild (1989, Chapter 9).

A problem which is strongly related to segmented regression models or models with a structural change concerns the derivation of an appropriate statistical test for checking if there is indeed a change in the parametric form of the regression function. Apart from the likelihood–ratio test being of interest in this paper there are other possibilities for testing a breakpoint as for instance a suggestion by Davies (1977, 1987) about hypothesis testing when a nuisance parameter is present only under the alternative. In this context, Davies also discusses the case of broken line regression. A further test for shifting slope coefficient in a linear model can be found in Farley and Hinich (1970). Krämer, Ploberger and Schlüter (1991) use CUSUM–tests for structural change. The fluctuation test is discussed by Ploberger, Krämer and Kontrus (1989). For a review concerning the different methods we refer to Müller (1995, Chapter 4).

In this paper we focus on linear and logistic regression models with *two* segments and with *continuous* regression function f in the breakpoint τ. Important special cases and the motivation for our work are *threshold models* typically used in epidemiology. A threshold effect indicates an association between a risk factor X (e.g. a chemical agent) and a defined outcome Y (e.g. a certain disease) above the threshold value but none below it. The threshold value is represented by τ, the function f is described by a constant below the threshold value where no risk should be related to the agent and by a usual logistic or linear regression function for values of X exceeding the threshold value. For a discussion of the logistic threshold model see Ulm (1991).

In occupational medicine the assessment of a threshold value for certain chemical compounds is of special interest. An important task is checking the question whether a threshold value really does exist. From a statistical point of view that means to find out whether a unique or two segmented regression lines with horizontal first segment are appropriate in the given situation. In the more general case of broken line regression in which each segment is a different line we test whether the two slope parameters of the regression function are identical implying a simple regression model. In any case, the null hypothesis is that of no breakpoint, which means that the model is a unique regression over the whole domain of the independent variable. The alternative hypothesis then can be both a general broken line regression model or a threshold model.

For deriving an appropriate statistical test we resort to the likelihood–ratio statistic λ which compares the unrestricted maximum of the log–likelihood function (model fit assuming a breakpoint) with the maximum obtained for the restricted ML–estimator computed under the restriction of the null hypothesis (lack of a breakpoint). Large values of λ indicate the existence of a breakpoint. The distribution of λ under the null hypothesis is, however, unknown up to now.

This problem is addressed in this paper by means of Monte–Carlo experiments where especially different regression models and different assumptions concerning the explanatory variable X are investigated. We first focus on broken linear regressions with normally distributed response for getting to know the asymptotic behaviour of the likelihood–ratio statistic (Section 2). In case a breakpoint exists we have a different slope, both unknown in each segment. Although the problem concerning the existence of a breakpoint related to this model is quite similar to that of a threshold model also having two different slopes but with the slope of the first segment being known as zero, the limiting distributions of λ are substantially different. Models of the latter type are investigated in Section 3. To ensure a comparison between the resulting limiting distributions, we consider a linear threshold model in Section 3.1, although the more interesting case is that of a logistic threshold model. Such models being of major concern in epidemiological studies are therefore separately treated in Section 3.2. In Section 4, we give an example for the application of these methods in forest science. We analyse the relationship between the proportion of deciduous trees in a forest and the concentration of certain minerals in the ground. In Section 5, we summarize our results.

2 The Broken Line Model

We consider the general *broken line regression model* for normally distributed response Y with $Y|X \sim N(\mu, \sigma^2)$:

$$E(Y|X = x) = \beta_0 + \beta_1(x - \tau)_- + \beta_2(x - \tau)_+, \tag{1}$$

where $(x - \tau)_- = \min(x - \tau, 0)$ and $(x - \tau)_+ = \max(x - \tau, 0)$.

We are interested in the question whether the two segment functions might be identical. The testing problem with a simple linear model as null hypothesis, i.e. no breakpoint exists, can be stated as

$$H_0 : \beta_1 = \beta_2 \quad \text{against} \quad H_1 : \beta_1 \neq \beta_2. \tag{2}$$

As already mentioned, the test is to be based on the likelihood–ratio statistic

$$\lambda = -2 \cdot \{l(H_0) - l(H_1)\} \tag{3}$$

with l denoting the logarithm of the likelihood function.

In case of a linear model maximizing the log–likelihood function is equivalent to minimizing the residual sum of squares which is given by $SSR(\beta_0, \beta_1, \beta_2, \tau) = \sum_{i=1}^{n}(y_i - f_i)^2$ by $(x_i, y_i)_{i=1,\ldots,n}$ denoting independent, identically distributed observations of X and Y and $f_i = E(Y|X = x_i)$ denoting the regression function. The likelihood–ratio for independent and identically normally distributed errors is of the following form for equal error variances in the two segments

$$\mathrm{LR} = \left(\frac{SSR\{H_1\}}{SSR\{H_0\}}\right)^{\frac{n}{2}}$$

with the residual sum of squares based on the full model in the numerator and the residual sum of squares based on the restricted model, i.e. under H_0, in the denominator. Under H_0 the regression function is given as $f_i = \breve{\beta}_0 + \breve{\beta}_1 x_i$ and $SSR\{H_0\}$ has to be minimized regarding $\breve{\beta}_0$ and $\breve{\beta}_1$. The according test statistic then is given as

$$\lambda = -2 \log \mathrm{LR}. \tag{4}$$

In order to calculate λ we have to estimate both a model without breakpoint and model (1) with breakpoint τ and insert the two residual sums of squares in equation (4). The null hypothesis is rejected, if the value of λ exceeds the $(1-\alpha)$–quantile of the null distribution of λ. However, this distribution is unknown.

This problem is also discussed in detail in the literature. Feder (1975) remarks that the distribution of λ is not asymptotically χ^2 with an appropriate number of degrees of freedom. Such an asymptotic χ^2–distribution has only be derived under the null hypothesis of a special type of a broken linear regression with two different slope parameters. Under the null hypothesis of no breakpoint the asymptotic behaviour is, however, more complicated and depends on the configuration of the datapoints of the independent variable. Concerning this aspect Feder gives an example of a two–phase linear model with the first segment function being equal to zero, which means the case of a threshold model. But in contrast to our test situation in Section 3 he tests the null hypothesis that also the second segment function is equal to zero. He obtains as null distribution of the likelihood–ratio statistic that of the maximum of correlated χ^2_1– and χ^2_2–distributed random variables. Nevertheless, he states that this behaviour is typical for the general case of broken line regression. Despite the problem of determining the limiting distribution of λ, Hawkins (1980) shows its existence. Hinkley (1969) reports from empirical studies which indicate that the distribution of λ is very close to a χ^2–distribution with three degrees of freedom, although the dimension of reduction of the parameter space by the null hypothesis which usually is decisive for the degrees of freedom here takes the value of two. In model (1) namely the breakpoint and one slope parameter loose identifiability under H_0.

To get a deeper insight into this problem, we simulate the distribution of λ. For each situation, we create $R = 1000$ datasets with sample size n each from a linear model of the form $E(Y|X = x) = \breve{\beta}_0 + \breve{\beta}_1 x$ and $V(Y|X = x) = \sigma^2_{Y|X}$. In each replication, a one–phase model and a broken line model of the form (1) are estimated and the statistic λ is calculated. The estimation of a broken line model frequently is performed by a grid–search–type algorithm. For the purpose of simulating the distribution of the test statistic, the grid–search procedure does not lead to reliable results. For that reason we use an exact algorithm for estimating breakpoints in segmented generalized linear models which is proposed by Küchenhoff (1997).

The parameters are chosen as $\breve{\beta}_0 = 0$, $\breve{\beta}_1 = 1$, $V(Y|X = x) = 0.3^2$, sample sizes vary with $n = 100$, $n = 300$, $n = 500$ and $n = 1000$. In addition, two different distributions for the independent variable are considered: a standard normal one and a uniform distribution on the interval $[-2.5, 2.5]$. We also investigate the behaviour of the likelihood–ratio statistics in case the values of X are fixed at certain points.

In case of a standard normal X we find a quite good agreement of the distribution of λ with the χ^2_3–distribution which confirms the results of Hinkley. For illustration, we refer to the quantile–quantile–plot (qq–plot) for $n = 300$ (Figure 1). Looking, however, at a uniformly distributed X the results are not convincing

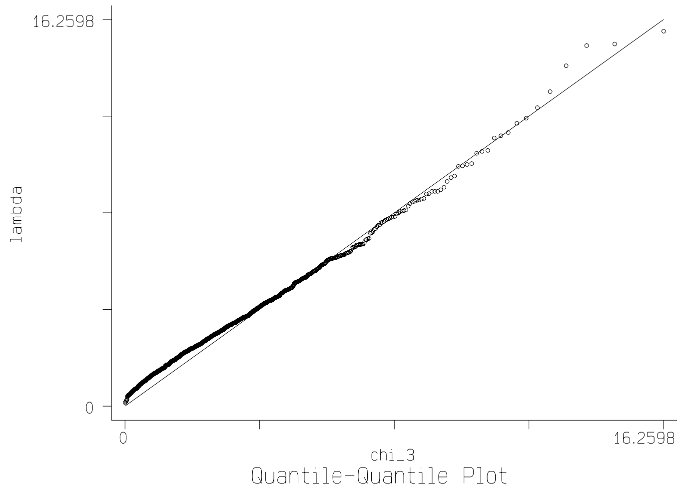

Figure 1: *QQ–plot for λ (lambda) against the theoretical quantiles of a χ^2_3–distribution (chi 3) under the null hypothesis of a linear model without breakpoint against the alternative of a broken line model. Sample size $n = 300$, $\breve{\beta}_0 = 0$, $\breve{\beta}_1 = 1$, $\sigma_{Y|X} = 0.3$, $X \sim N(0,1)$, 1000 replications each.*

at all. From the qq–plot for $n = 1000$ in Figure 2, we notice a systematic difference from the supposed χ^2_3–distribution. For a fixed X–design with 500 observations at each of the points $x_i = -2.375 + i \cdot 0.25$, $i = 0, \ldots, 19$, $(n = 10000)$ and remaining parameters as above, we also find essential deviations from the χ^2_3–quantiles, but in the opposite direction. The qq–plot here is below the diagonal.

Table 1 additionally illustrates the quality of the approximation by the χ^2_3–distribution. Here, the actual rejection probabilities α_{sim} of tests with theoretical significance levels $\alpha = 0.01, 0.025, 0.05, 0.1, 0.2$ based on this distribution are

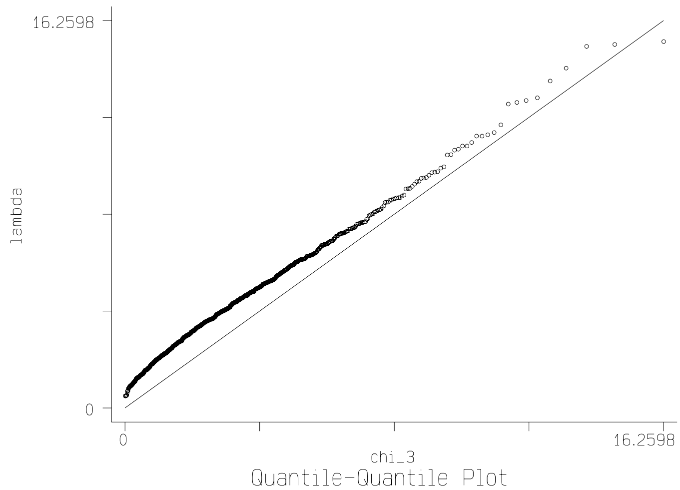

Figure 2: *QQ–plot for* λ *(lambda) against the theoretical quantiles of a* χ^2_3*–distribution (chi 3) under the null hypothesis of a linear model without breakpoint against the alternative of a broken line model. Sample size* $n = 1000$*,* $\breve{\beta}_0 = 0$*,* $\breve{\beta}_1 = 1$*,* $\sigma_{Y|X} = 0.3$*,* X *uniformly distributed on* $[-2.5, 2.5]$*, 1000 replications each.*

shown for varying sample sizes ($n = 100, 300, 1000$). In the case of a standard normal X and for predetermined values of $\alpha \leq 0.1$ the suggested distribution of Hinkley can be used as approximation. Since the actual values are even smaller, the relevant tests are conservative. For greater values of α, this statement is no longer valid and thus the approximation quality is no longer satisfactory. For uniformly distributed X we observe large deviations from the theoretical quantiles. The corresponding tests often exceed the fixed significance level and are therefore not reliable in practical situations.

Because of the above results it cannot be generally recommended to work with the quantiles of a χ^2_3–distribution. We propose to simulate the quantiles by a Monte–Carlo algorithm for each X–design and actual parameter configuration. Based on these results tables of quantiles can be constructed, which then can be used for testing.

Significance level α	0.01	0.025	0.05	0.1	0.2
$X \sim N(0,1)$					
level α_{sim} for $n = 100$	0.009	0.02	0.039	0.084	0.166
level α_{sim} for $n = 300$	0.01	0.021	0.049	0.095	0.198
level α_{sim} for $n = 1000$	0.006	0.018	0.044	0.093	0.225
X uniformly distributed on $[-2.5, 2.5]$					
level α_{sim} for $n = 100$	0.002	0.022	0.05	0.106	0.237
level α_{sim} for $n = 300$	0.009	0.027	0.057	0.119	0.271
level α_{sim} for $n = 1000$	0.014	0.033	0.064	0.144	0.311

Table 1: *Percentage α_{sim} of the simulated values of λ exceeding the $(1-\alpha)$–quantile of the χ^2_3–distribution in the case of the broken line model under the null hypothesis of a linear model without breakpoint, $\breve{\beta}_0 = 0$, $\breve{\beta}_1 = 1$, $\sigma_{Y|X} = 0.3$, 1000 replications each.*

3 The Linear and Logistic Threshold Model

In the following, we consider the linear threshold model with a normally distributed response Y, i.e. $Y|X \sim N(\mu, \sigma^2)$:

$$E(Y|X = x) = \beta_0 + \beta_2(x - \tau)_+ \qquad (5)$$

and a logistic threshold model with a binary response Y, coded by 0 and 1, i.e. $Y|X \sim B(1, \pi)$:

$$E(Y|X = x) = P(Y = 1|X = x). \qquad (6)$$

Here, we are interested in getting to know whether there is in fact a threshold value below that the regression function equals a constant. Because we can estimate a potential threshold value only within the observed domain of the independent variable, the corresponding testing problem can be formulated as one–sided with a threshold value below the smallest X–observation as null hypothesis. Then, under H_0 there is an effect of X on Y, but no threshold. Thus, we get (see also Ulm, 1991):

$$H_0 : \tau \leq \min_{i=1,\ldots,n} x_i \quad \text{against} \quad H_1 : \tau > \min_{i=1,\ldots,n} x_i. \qquad (7)$$

For applying the likelihood–ratio test the null distribution of λ has again to be specified but now under the model assumptions (5) and (6). The dimension of reduction of the parameter space here takes the value $q = 1$, and we could suspect the χ^2–distribution with one degree of freedom as limiting distribution. But as additional problem it has to be taken into account that the question concerning the existence of a breakpoint is now formulated as a one–sided testing problem. Ulm points out that a level–α test for a certain parameter ϑ has to reject the null hypothesis $H_0 : \vartheta = 0$ in favour of the one–sided alternative $H_1 : \vartheta > 0$ if the likelihood–ratio statistic exceeds the $1-2\alpha$–quantile of the χ^2_1–distribution where the significance level α is limited to values between 0 and 0.5. This result cannot directly be extended to our testing problem because it is derived for parameters without any restrictions. The breakpoint, however, being of interest here has to lie between the smallest and the greatest observation. Thus, additional effort is necessary to check whether the so–called *quasi one–sided* χ^2_1*–distribution* is also the correct limiting distribution in our case. The quasi one–sided χ^2_1–distribution is a mixed discrete and continuous distribution with probability 0.5 for the value of zero. The remaining 50%–probability is taken by a χ^2_1–distribution.

In a response to Ulm's (1989) proposal using the quasi one–sided χ^2_1–distribution as limiting distribution Cox (1989) comments that it is difficult to be sure of the claimed justification of that proposal. Silvapulle (1991), however, confirms the statement of Ulm because in his opinion it could be easily shown that $P(\lambda \geq c|H_0) \approx 0.5 \cdot P(\chi^2_1 \geq c)$ for large samples. Thus, although giving no formal proof, Silvapulle also suggests to use the $1 - 2\alpha$–quantile of the χ^2_1–distribution for a level–α test.

In the next two sections, we present our simulation results for the linear (Section 3.1) and the logistic (Section 3.2) threshold model, where Ulm restricts his statements to the latter one. The obtained results are discussed in the light of the above statements.

3.1 Simulation Results for the Linear Threshold Model

In fact, our simulations have confirmed the above speculations that in about 50% of the estimated models of one simulation study the likelihood–ratio statistic takes the value of zero. In these cases, the same models are fitted under H_0 and H_1. The breakpoint then is the smallest X–observation.

Looking at the simulation procedure it can be observed, that in most cases the one y–value belonging to the smallest X–realization is decisive for the realization of the test statistic. If this y-value lies below the regression line fitted under H_0, then the regression lines under H_0 and H_1 are mostly the same and the smallest

x is the estimated threshold value. The minimal residual sum of squares then coincide and $\lambda = 0$. Vice versa, the residual sum of squares of the H_1 fit is smaller than that of the H_0 fit, if the named y–value lies above the regression line. The probability that such a situation occurs is 0.5, because of the assumed distribution for the error variable as normal with expectation 0. But of course, the position of the smallest x–observation is not always decisive and therefore we cannot derive any theoretical conclusions from this.

Figures 3 (X standard normal and $n = 300$) and 4 (X uniformly distributed and $n = 1000$) show qq–plots with the sorted values of λ on the vertical axis and the (v/R)–quantiles, $v = 1, \ldots, R$, of a quasi one–sided χ^2_1–distribution on the horizontal axis. These quantiles equal zero for $v = 1, \ldots, R/2$. For $v = (R/2)+1, \ldots, R$ and for R even they are the $(v-(R/2))/(R/2)$–quantiles of a χ^2_1–distribution. The parameter configuration under H_0 is the same as in Section 2. We notice a very good agreement with the quasi one–sided χ^2_1–distribution.

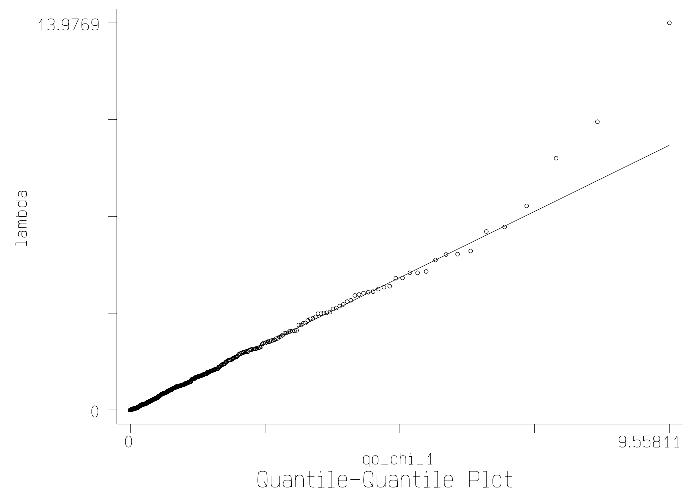

Figure 3: *QQ–plot for λ (lambda) against the theoretical quantiles of a quasi one–sided χ^2_1–distribution (qo chi 1) under the null hypothesis of a linear model without breakpoint against the alternative of a linear threshold model. Sample size $n = 300$, $\breve{\beta}_0 = 0$, $\breve{\beta}_1 = 1$, $\sigma_{Y|X} = 0.3$, $X \sim N(0, 1)$, 1000 replications each.*

In contrast to the results for the broken line model (1), here there is nearly no dependence observable of the results on the configuration of the independent variable.

Table 2 again gives an overview over the actual significance levels from tests

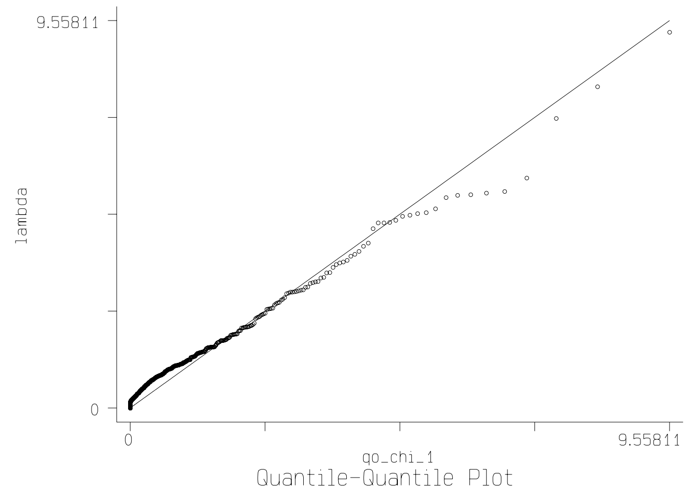

Figure 4: *QQ–plot for λ (lambda) against the theoretical quantiles of a quasi one–sided χ^2_1–distribution (qo chi 1) under the null hypothesis of a linear model without breakpoint against the alternative of a linear threshold model. Sample size $n = 1000$, $\breve{\beta}_0 = 0$, $\breve{\beta}_1 = 1$, $\sigma_{Y|X} = 0.3$, X uniformly distributed on $[-2.5, 2.5]$, 1000 replications each.*

based on the $1 - 2\alpha$–quantiles of the χ^2_1–distribution. Again, for a fixed level of $\alpha = 0.2$ the actual values are too high. But the differences between the two designs are not such large as before. For smaller values of α the tests can be carried out using the χ^2_1–distribution. These accordingly conducted tests also have good statistical properties.

In addition to the considered X–configurations we also carry out simulations for different fixed X–values and different parameter configurations. All modifications have shown similar results and agree very well with the reproduced qq–plots.

In a further simulation for $n = 500$ with $\breve{\beta}_0 = 0$ and $\breve{\beta}_1 = 1$, the error variance is set nearly to zero, namely $\sigma_{Y|X} = 0.0001$. The result shows an almost perfect agreement of the distribution of the likelihood–ratio statistic with the quasi one–sided χ^2_1–distribution. Although this way of looking at the limiting distribution should be handled carefully, because increasing the sample size and decreasing the error variance is not the same thing, but they yield similar effects, the quasi one–sided χ^2_1–distribution seems to be the correct asymptotic distribution. But also with finite sample sizes, the test decision to reject H_0 can be carried out at a $\alpha\%$–level for values of λ when exceeding the $1 - 2\alpha$–quantile of the χ^2_1–

Significance level α	0.0125	0.025	0.05	0.1	0.2
$X \sim N(0,1)$					
level α_{sim} for $n = 100$	0.024	0.044	0.068	0.129	0.240
level α_{sim} for $n = 300$	0.01	0.025	0.049	0.1	0.202
level α_{sim} for $n = 1000$	0.017	0.033	0.058	0.113	0.226
X uniformly distributed on $[-2.5, 2.5]$					
level α_{sim} for $n = 100$	0.018	0.027	0.060	0.128	0.250
level α_{sim} for $n = 300$	0.011	0.023	0.052	0.106	0.243
level α_{sim} for $n = 1000$	0.010	0.021	0.049	0.102	0.234

Table 2: *Percentage α_{sim} of the simulated values of λ exceeding the $(1 - 2\alpha)$–quantile of the χ^2_1–distribution in the case of the linear threshold model under the null hypothesis of a linear model without breakpoint, $\breve{\beta}_0 = 0$, $\breve{\beta}_1 = 1$, $\sigma_{Y|X} = 0.3$, 1000 replications each.*

distribution. Based on the results of our simulations we thus recommend the use of the quasi one–sided χ^2_1–distribution for testing a threshold value in a linear model.

3.2 Simulation Results for the Logistic Threshold Model

It should be mentioned first that the results by Feder (1975) can be confirmed for a logistic model with two different segment functions which can be seen from simulation results obtained by Haybach (1996). Thus, in this case the likelihood–ratio statistic behaves as being asymptotically χ^2_q–distributed. But again it may be a quite different problem for a logistic threshold model which we discuss now. In order to find out whether λ is distributed according to a quasi one–sided χ^2_1–distribution, Ulm (1991) (cf. Ulm, 1989) conducts a simulation study. For three different X–designs (uniformly distributed, normally distributed and log–normally distributed, X–values between 3 and 12 in each situation) and three different $\breve{\beta}_1$–values (0.1, 0.3, 0.5) 1000 replications are run in each situation with $\breve{\beta}_0 = 0$ and $n = 1000$ and the empirical distribution functions of λ under H_0 are plotted. The results show only a slight difference between the curves for the different X–designs and $\breve{\beta}_1$–values. Further, the model fit under H_1 is better than the H_0 fit in about 50% of all replications which means $\lambda > 0$ in such cases. In the other 50% the value of λ equals zero and $\hat{\tau} = \min x$. The empirical distribution functions are finally compared with the theoretical distribution functions of the χ_1^2–distribution and of the quasi one–sided χ_1^2–distribution. The course of the χ_1^2–distribution is below the empirical curves, its use is therefore too conservative. In contrast, the quasi one–sided χ_1^2–distribution shows a much better agreement with the empirical curves. For that reason Ulm recommends in general to reject H_0 if λ exceeds the $1-2\alpha$–quantile of the χ_1^2–distribution.

Our results derived from numerous simulations for the case considered here are not at all such reliable as in the case of the linear threshold model. The distribution of the likelihood–ratio statistic heavily depends on the sample size and on the parameter configuration. Thus, general use of the quasi one–sided χ_1^2–distribution for the testing problem (7) cannot be advised.

For our simulations we use a sample of $n = 1000$ or more, because in a logistic model results for smaller n fluctuate too much and are not useful. The first aspect of our investigations concerns the percentage of simulated λ–values with result zero which means identical fit under H_0 and H_1. All simulated data sets are generated from a logistic model of the form $\log(\pi/(1-\pi)) = \breve{\beta}_0 + \breve{\beta}_1 x$ with $E(Y|X = x) = P(Y = 1|X = x)$. The number of replications is 1000 in each situation. With a parameter choice of $\breve{\beta}_0 = -2$ and $\breve{\beta}_1 = 2$, where the probability for $Y = 1$ is greater than 0.5 only for X–values greater than 1, we find for $n = 1000$ and standard normal X that in about 78% of all replications the estimated breakpoint is the smallest X–realization and $\lambda = 0$. Also for a uniformly distributed X on the interval $[-2.5, 2.5]$ more than 50% of the λ–values are zero. We get a similar result with a percentage of λ being equal to zero of about 72% for fixed X–values, which are simulated for the case of 111 observations at each of the points $x_i = -2 + i \cdot 0.5$, $i = 0, \ldots, 8$. A percentage near 50% is only reached for a much larger sample size and a fixed X–design.

For constellations of $\breve{\beta}_0$ and $\breve{\beta}_1$ where the realizations $y = 1$ and $y = 0$ occur more balanced, we get results comparable to those reported by Ulm, i.e. about 50% zeros. This is also the case for $\breve{\beta}_0 = 0$ and $\breve{\beta}_1 \in \{0.5, 0.75\}$ as well as for $\breve{\beta}_0 = 1$ and $\breve{\beta}_1 = 1$ with sample sizes of $n = 1000, 5000, 8500$. Here, the results do not depend on the design of X.

In Figures 5 and 6 the qq–plots for $\breve{\beta}_0 = -2$, $\breve{\beta}_1 = 2$ and standard normal X (Figure 5) resp. fixed X (Figure 6) are depicted. While Figure 5 shows a very bad agreement between theoretical and empirical quantiles, Figure 6 seems to confirm Ulm's proposal. Figure 7 shows the qq–plot for $\breve{\beta}_0 = 1$ and $\breve{\beta}_1 = 1$ and fixed X–design. We find a very good agreement with the quasi one–sided χ_1^2–distribution. Finally, we look again at the actual significance levels of tests

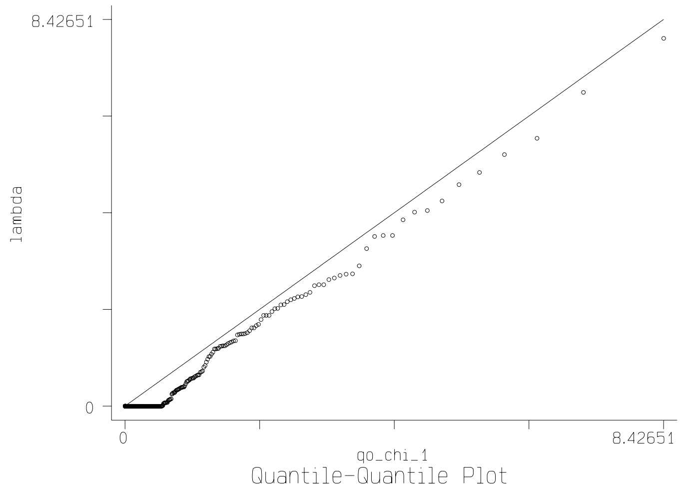

Figure 5: *QQ–Plot for λ (lambda) against the theoretical quantiles of a quasi one–sided χ^2_1–distribution (qo chi 1) under the null hypothesis of a logistic model without breakpoint against the alternative of a logistic threshold model, sample size $n = 1000$, $\breve{\beta}_0 = -2$, $\breve{\beta}_1 = 2$, $X \sim N(0,1)$, 1000 replications each.*

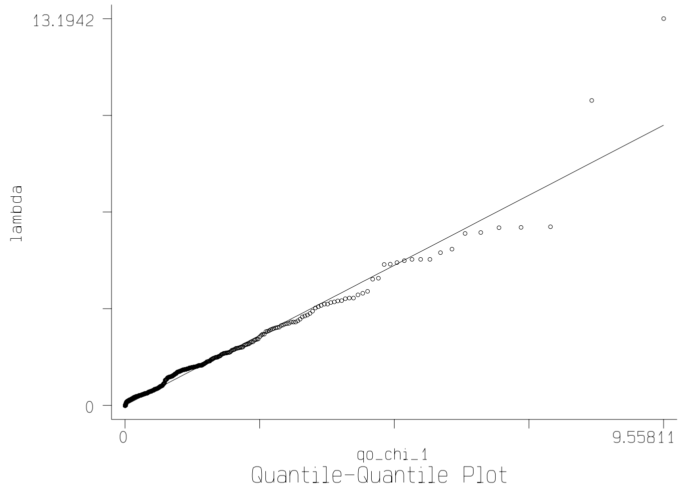

Figure 6: *QQ–Plot for λ (lambda) against the theoretical quantiles of a quasi one–sided χ^2_1–distribution (qo chi 1) under the null hypothesis of a logistic model without breakpoint against the alternative of a logistic threshold model, sample size $n = 3774$, $\breve{\beta}_0 = -2$, $\breve{\beta}_1 = 2$, X fixed with 222 observations at each of the points, $x_i = -2 + i \cdot 0.25$, $i = 0, \ldots, 16$, 1000 replications each.*

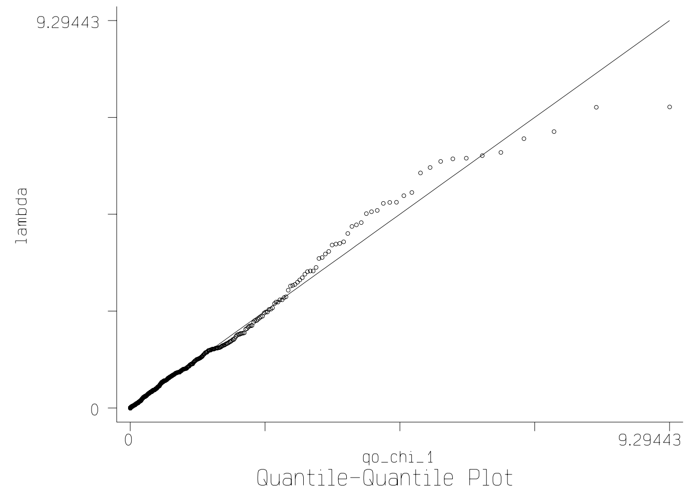

Figure 7: *QQ–Plot for* λ *(lambda) against the theoretical quantiles of a quasi one–sided* χ_1^2*–distribution (qo chi 1) under the null hypothesis of a logistic model without breakpoint against the alternative of a logistic threshold model, sample size* $n = 8500$*,* $\breve{\beta}_0 = 1$*,* $\breve{\beta}_1 = 1$*, X fixed with 500 observations at each of the points,* $x_i = -2 + i \cdot 0.25$*,* $i = 0, \ldots, 16$*, 1000 replications each.*

which are based on this distribution. Table 3 presents the percentage α_{sim} of the simulated λ–values exceeding the $1 - 2\alpha$–quantile of the χ_1^2–distribution for the three situations of the qq–plots 5, 6 and 7. We notice that in many situations the tests are quite acceptable. In the first case (situation of Figure 5), however, the actual significance levels are generally smaller than the theoretical ones.

Summarizing the above results the quasi–one–sided–χ_1^2–distribution should not be used in all situations. Therefore, we again recommend to run a small simulation to find out the null distribution of the likelihood–ratio statistic.

4 Application to the Forest Data

In a study conducted by Rothe (1997) the influence of tree species composition on rooting, hydrology, elemental turnover, and growth in a mixed spruce–beech stand in southern Germany (Hoeglwald) is analysed. Data are available from 71 points of a sampling grid in a mixed spruce-beech stand with an area of 1 hectar. The proportion of tree species is calculated using growing space in a radius of 10 m around the sampling points. Also many ecological properties are measured.

significance level α	0.0125	0.025	0.05	0.1	0.2
$\breve{\beta}_0 = -2$, $\breve{\beta}_1 = 2$, $X \sim N$, $n = 1000$	0.009	0.019	0.037	0.080	0.139
$\breve{\beta}_0 = -2$, $\breve{\beta}_1 = 2$, X fixed, $n = 3774$	0.011	0.023	0.048	0.094	0.180
$\breve{\beta}_0 = 1$, $\breve{\beta}_1 = 1$, X fixed, $n = 8500$	0.015	0.031	0.049	0.090	0.209

Table 3: *Percentage α_{sim} of the simulated λ–values exceeding the $(1-2\alpha)$–quantile of the χ_1^2–distribution in the case of the logistic threshold model under the null hypothesis of a logistic model without breakpoint, sample sizes, parameters and designs as in Figures 5 with 7.*

The main question is whether the quantified properties of the mixed stand equal the sum of the quantified properties of the single–species stands, i.e. whether the linear model

$$E(\text{ECO}|\text{PROP}) = \beta_0 + \beta_1(\text{PROP}) \tag{8}$$

is valid for describing the data. Here, ECO is the quantified ecological characteristic and PROP is the proportion of spruce. Based on this model β_0 is the expected value of ECO in case of a pure beech forest and $\beta_0 + 100 * \beta_1$ the expected value of ECO for spruce forest.

A sensible alternative to the simple linear model is the broken line regression. Thus, we consider the model (1) with the ecological characteristic as response and the proportion of spruce as explanatory variable, i.e.

$$E(\text{ECO}|\text{PROP}) = \beta_0 + \beta_1(\text{PROP} - \tau)_- + \beta_2(\text{PROP} - \tau)_+ \tag{9}$$

with a possible breakpoint τ between 0 and 100.

Beside other ecological characteristics the humus morphology and soil acidity is measured. The characteristics are the thickness of the surface humus, the pH (measured in $CaCl_2$) in the surface humus (Of–layer) and in the top soil (Ah–horizon). Since pH is a logarithmic measure it is reconverted to the proton concentration, named by Prot Ah and Prot Of.

The statistical analyses are conducted for each of these three ecological characteristics where we especially test for the existence of a breakpoint. As null distribution of the likelihood–ratio statistic we use the χ_3^2–distribution and also a distribution derived from a simulation reflecting the special setting of our data. In all cases the null hypothesis of no breakpoint is rejected using the p–values from the χ_3^2–distribution at the 5%–level. As can be seen from Table 4 the p–values

ECO	$\hat{\tau}$	p-value (χ_3^2)	p-value (simulation)
humus	31	0.04	0.0556
Prot Ah	28	0.0006	0.001
Prot Of	31	0.0000	0.001

Table 4: *P–values from the χ_3^2–distribution and from a simulation study for testing the null hypothesis of no breakpoint ($\alpha = 0.05$).*

based on the simulation are quite similar, but for humus the null hypothesis can no longer be rejected. More detailed results are thus reported for humus, where a breakpoint at about 31 % is estimated, see Table 5 and Figure 8 below.

	broken linear		linear	
parameter	estimate	$\hat{\sigma}$	estimate	$\hat{\sigma}$
τ	31.19	11.50	–	–
α	6.41	0.62	3.97	0.39
β_1	0.11	0.05	0.03	0.01
β_2	-0.004	0.02	–	–
R^2	0.30	–	0.21	–

Table 5: *Parameter estimates for the relationship between the proportion of spruce and the concentration of humus in the ground.*

5 Discussion

Summarizing our simulation results and the results of our analysis we can conclude that especially in the broken line model and in the logistic threshold model the appropriateness of the limiting distribution of the likelihood–ratio statistic heavily depends on the design of the explanatory variable and the parameter configuration. Only in case of a linear threshold model the distribution of the likelihood–ratio statistic is not such sensitive to changes in the design of X or different parameter configurations. Nevertheless, the unreflected use of a χ^2–distribution or a quasi one–sided χ_1^2–distribution, respectively, as asymptotical null distribution should be avoided. If possible we recommend to simulate the

Figure 8: *Estimated broken line regression for the relationship between the proportion of spruce and the concentration of humus in the ground.*

distribution of the likelihood ratio statistic for the specific data situation under investigation as presented in Section 4.

References

Cox, Ch. (1989): *On the estimation of threshold values (reply)*, Biometrics **45:** 1327–1328.

Davies, R.B. (1977): *Hypothesis testing when a nuisance parameter is present only under the alternative*, Biometrika **64:** 247–254.

Davies, R.B. (1987): *Hypothesis testing when a nuisance parameter is present only under the alternative*, Biometrika **74:** 33–43.

Farley, J.U. and Hinich, M.J. (1970): *A test for shifting slope coefficient in a linear model*, Journal of the American Statistical Association **65:** 1320–1329.

Feder, P. (1975): *The log–likelihood–ratio in segmented regression*, Annals of Statistics **3:** 84–97.

Hawkins, D.M. (1980): *A note on continuous and discontinuous segmented regressions*, Technometrics **22:** 443–444.

Haybach, G. (1996): *Konfidenzintervalle und Tests in Regressionsmodellen mit Bruchpunkten*, Diplomarbeit, Institute of Statistics, Ludwig–Maximilians–University of Munich.

Hinkley, D.V. (1969): *Inference about the change point in a sequence of random variables*, Biometrika **57:** 1–17.

Krämer, W., Ploberger, W. and Schlüter, I. (1991): *Recursive vs. OLS residuals in the CUSUM test*, in P. Hackl and A. H. Westlund (eds), Economic Structural Change. Analysis and Forecasting, Springer–Verlag, pp. 36–47.

Küchenhoff, H. (1997): *An exact algorithm for estimating breakpoints in segmented Generalized Linear Models*, Computational Statistics 77 **12:** 235–247.

Müller, R. (1995): *Mehrphasige Regression: Das Change Point Problem*, Diplomarbeit, Institute of Statistics, Ludwig–Maximilians–University of Munich.

Ploberger, W., Krämer, W. and Kontrus, K. (1989): *A new test for structural stability in the linear regression model*, Journal of Econometrics **40**: 307–318.

Rothe (1997): *Einfluss des Baumartenanteils auf Durchwurzelung, Wasserhaushalt, Stoffhaushalt und Zuwachsleistung eines Fichten–Buchen–Mischbestandes am Standort Höglwald* Dissertation, Forstwissenschaftliche Fakultät, University of Munich.

Seber, G.A.F. and Wild, C.J. (1989): *Nonlinear Regression*, John Wiley & Sons, New York.

Silvapulle, J.M. (1991): *On testing for threshold values (correspondence)* Biometrics **47:** 1628–1629.

Ulm, K. (1989): *On the estimation of threshold values (correspondence)*, Biometrics **45:** 1324–1326.

Ulm, K. (1991): *A statistical method for assessing a threshold in epidemiological studies*, Statistics in Medicine **10:** 341–349.