UPPSALA
UNIVERSITET

Digital Comprehensive Summaries of Uppsala Dissertations from the Faculty of Social Sciences 18

Estimation and Inference for Quantile Regression of Longitudinal Data

With Applications in Biostatistics

ANDREAS KARLSSON

ACTA
UNIVERSITATIS
UPSALIENSIS
UPPSALA
2006

ISSN 1652-9030
ISBN 91-554-6678-8
urn:nbn:se:uu:diva-7186

Dissertation presented at Uppsala University to be publicly examined in Hörsal 2, Ekonomikum, Kyrkogårdsgatan 10, Uppsala, Friday, November 10, 2006 at 13:15 for the degree of Doctor of Philosophy. The examination will be conducted in Swedish.

Abstract
Karlsson, A. 2006. Estimation and Inference for Quantile Regression of Longitudinal Data. With Applications in Biostatistics. Acta Universitatis Upsaliensis. *Digital Comprehensive Summaries of Uppsala Dissertations from the Faculty of Social Sciences* 18. 36 pp. Uppsala. ISBN 91-554-6678-8.

This thesis consists of four papers dealing with estimation and inference for quantile regression of longitudinal data, with an emphasis on nonlinear models.

The first paper extends the idea of quantile regression estimation from the case of cross-sectional data with independent errors to the case of linear or nonlinear longitudinal data with dependent errors, using a weighted estimator. The performance of different weights is evaluated, and a comparison is also made with the corresponding mean regression estimator using the same weights.

The second paper examines the use of bootstrapping for bias correction and calculations of confidence intervals for parameters of the quantile regression estimator when longitudinal data are used. Different weights, bootstrap methods, and confidence interval methods are used.

The third paper is devoted to evaluating bootstrap methods for constructing hypothesis tests for parameters of the quantile regression estimator using longitudinal data. The focus is on testing the equality between two groups of one or all of the parameters in a regression model for some quantile using single or joint restrictions. The tests are evaluated regarding both their significance level and their power.

The fourth paper analyzes seven longitudinal data sets from different parts of the biostatistics area by quantile regression methods in order to demonstrate how new insights can emerge on the properties of longitudinal data from using quantile regression methods. The quantile regression estimates are also compared and contrasted with the least squares mean regression estimates for the same data set. In addition to looking at the estimates, confidence intervals and hypothesis testing procedures are examined.

Keywords: Bias correction, Bootstrap, Dependent errors, Hypothesis testing, Nonlinear model, Simulation study

Andreas Karlsson, Department of Information Science, Box 513, Uppsala University, SE-75120 Uppsala, Sweden

ISSN 1652-9030
ISBN 91-554-6678-8
urn:nbn:se:uu:diva-7186 (http://urn.kb.se/resolve?urn=urn:nbn:se:uu:diva-7186)

To Marika, with all my love

Contents

Acknowledgements

I have to admit that five years ago, when I first entered my office at the Department of Information Science after being admitted to the Ph.D. program in statistics, it was hard to imagine the day when I had finished my thesis and would write my acknowledgements. But now that day has finally arrived, and although I have very much enjoyed my years as a postgraduate student, it feels really great to at last write the final words of acknowledgement.

My first and foremost thanks go to my supervisor Dr. Johan Lyhagen for many useful advises, fruitful suggestions and rewarding discussions. A second thanks goes to Professor Rolf Larsson and my assistant supervisor Dr. Ulf Olsson for reading and commenting earlier versions of the papers in this thesis, as well as the other members of the supervising board at the Division of Statistics at the Department of Information Science, Professor Anders Christoffersson, Dr. Dag Sörbom, Dr. Bo Wallentin, and Dr. Fan Yang Wallentin, for reading and commenting the almost finished versions. Thanks also to Professor Roger Koenker for pointing me to a paper on the consistency of nonlinear quantile regression estimators for data with dependent errors. Moreover, I would like to thank Dr. Jonas Andersson for interesting discussions during my defence of the licentiate thesis consisting of the first two papers in this thesis, and to Professor emeritus Anders Ågren for supervising my bachelor and master theses, which inspired me to apply to the Ph.D. program in statistics, and for many interesting discussions during my years as a postgraduate student.

For making the life at the department pleasant and enjoyable I also thank all current and former colleagues at the department, and especially those at the Division of Statistics: Bertil Andersson, Peder Blom, Joakim Ekström, Dr. Anders Eriksson, Dr. Lars Forsberg, Dr. Anna Gunsjö, Dr. Lisbeth Hansson, Anna Hermansson, Professor emeritus K. G. Jöreskog, Nicklas Korsell, Katrin Kraus, Mats Nyfjäll, Dr. Inger Persson, Dr. Roland Pettersson, Tomas Pettersson, Daniel Preve, Professor emeritus Adam Taube, and Yu Wang.

An academic institution would not work properly without the invaluable help of the administrative staff. A great thanks to all these: Davoud Emamjomeh, Eva Enefjord, Ann Gunnarsson, Pierre Hjälm, Gunilla Klaar, Anna-Lena Kåberg, Ingrid Lukell, and Lars-Göran Svensk.

Although my supervisor, my colleagues, and my friends at the Department of Information Science have provided much help, support and inspiration that have made the completion of this thesis possible, it would not have been completed without the love and support of my family. I want to express my great gratitude to my father Rune and my mother Miriam, my sister Maria, my brother David, and my nieces Salome and Tilda.

I have saved my greatest thanks to the last: Thanks Marika, my love, for making me happy and joyful, thanks for your kindness and support, thanks for all your love. I love you with all my heart.

Andreas Karlsson

List of papers included in this thesis

I. Karlsson, A. (2006). Nonlinear Quantile Regression Estimation of Longitudinal Data. Unpublished manuscript, Division of Statistics, Department of Information Science, Uppsala University.

II. Karlsson, A. (2006). Bootstrap Methods for Bias Correction and Confidence Interval Estimation for Nonlinear Quantile Regression of Longitudinal Data. Unpublished manuscript, Division of Statistics, Department of Information Science, Uppsala University.

III. Karlsson, A. (2006). Bootstrap-Based Hypothesis Tests for Nonlinear Quantile Regression of Longitudinal Data. Unpublished manuscript, Division of Statistics, Department of Information Science, Uppsala University.

IV. Karlsson, A. (2006). New Insights on Longitudinal Biostatistics Data from Using Quantile Regressions. Unpublished manuscript, Division of Statistics, Department of Information Science, Uppsala University.

1 Introduction

Repeated measurements is a method of collecting data in which the response and explanatory variables for each experimental unit are observed on multiple occasions. The collected data are called longitudinal data, repeated measures or, in some disciplines, panel data. It can be seen as a collection of cross-sectional data, with each cross-section taken at different time points. In contrast to ordinary time series, which usually consists of a single long series, longitudinal data usually consist of many short time series.

Longitudinal data often occurs in applied sciences, such as medicine, pharmacology and agricultural science. The focus of interest is often in how some variable (e.g., the pattern of growth for the subjects) changes over time. The main advantages of longitudinal data are that they make it possible to follow the individual patterns of change for each subject and that in the inference process one can borrow strength across subjects (for details about the merits of longitudinal data, see, e.g., Zeger and Liang, 1992, Davis, 2002, and Diggle *et al.*, 2002).

1.1 Advantages of using quantile regression to analyze longitudinal data

In the statistical literature there are a number of different methods available for analyzing longitudinal data. Examples of these are MANOVA, repeated measures ANOVA, mixed models, and Generalized Linear Models, with their extension Generalized Estimation Equations (see Vonesh and Chinchilli, 1997, Davis, 2002, Diggle *et al.*, 2002 or Hedeker and Gibbons, 2006 for details about the different approaches). A common feature of these approaches is that the mean usually is the measure of centrality. However, as is well known, the mean has some disadvantages. Specifically, it is a bad measure of centrality for skewed data. In such cases, besides often having low efficiency, it is difficult, especially for non-statisticians, to interpret what the mean measures. An alternative to the mean is to use the median as the measure of centrality. In addition to often having higher efficiency than the mean for skewed data (e.g., Koenker and Bassett, 1978), the median always has an easy interpretation.

For the estimation of longitudinal data, regression analysis is a natural choice to use, with time as one of the independent variables and some variable of interest as the dependent variable. Then, the median regression could be used as an alternative to the usual mean regression. The median regression, in turn, is the special case of the 50th quantile for the general quantile regression, which was first introduced in the seminal paper by Koenker and Bassett (1978). This can be used to calculate regression curves for any arbitrary quantile.

By calculating regression curves for several different quantiles of a longitudinal data set, one gets a distribution of quantile regression curves that shows the distribution of the longitudinal data for each time point, conditional on the specific time points. Such a procedure makes it possible to study any changes over time in the shape of the entire conditional distribution of longitudinal data, not only the change over time in the conditional mean or median. Thus, one can study the development over time for different parts of the conditional distribution. For example, the slope of a curve need not be the same for the subjects in the lower and upper 10 percent of the distribution, a difference that can be studied by comparing the 10th and 90th quantiles. By comparing the values of an individual subject with the quantile regression curves, it is also easily seen how

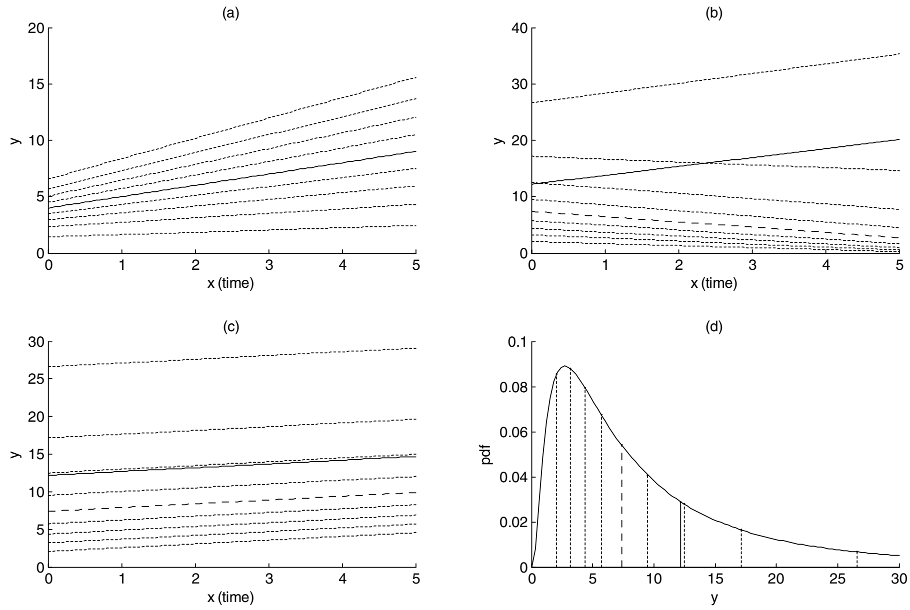

Figure 1.1: (a) Linear regression model with normally distributed heteroskedastic error terms. (b) Linear regression model with lognormal(2,1) distribution for $x = 0$ and lognormal(1,2) distribution for $x = 5$. (c) Linear regression model with lognormal(2,1) distributed error terms for all x. (d) Cross section of the conditional distribution of y for (c). Dotted lines are quantile regression lines for $q = 0.1, ..., 0.4, 0.6, ..., 0.9$; dashed lines are median regression lines using $q = 0.5$; and solid lines are mean regression lines.

the individual subject performs as compared with the overall performance of the sample. This gives the analyzer a much more complete picture of the data set than would be given by only studying the mean or the median, which can be very useful in some instances. In medical applications, for instance, a subject with values above or below a certain quantile regression curve may be identified as possibly suffering from some pathological condition. For overviews of quantile regression, see the monograph by Koenker (2005) as well as the articles by Buchinsky (1998) and Yu *et al.* (2003).

1.1.1 An example

Figure 1.1 illustrates some of the advantages of using quantile and median regression for longitudinal data compared with only using mean regression. It gives the quantile regression lines for quantiles $q = 0.1, ..., 0.4, 0.6, ..., 0.9$ (dotted lines), the median regression line $q = 0.5$ (dashed lines) and the mean regression line (solid lines) for the linear regression model

$$y = \alpha + \beta x + \varepsilon, \tag{1.1}$$

where ε follows an unknown distribution. Panel (a) of Figure 1.1 shows a case where ε follows a normal distribution with increasing variance, i.e. ε is heteroskedastic. By looking at the quantile regression lines, it is easily noted that the variance is increasing over time, and the assumption of homoskedasticity in the classical normal linear regression model is thus not fulfilled. It is also seen that, e.g., the slope for $q = 0.9$ is much steeper than the slope for $q = 0.1$, the latter being almost flat. The impact of x on y is therefore much larger for the upper 10 percent of the conditional distribution of y than for the lower 10 percent. Since the median regression line is the same as the mean regression line, it can also be concluded that the distribution is symmetric.

Panel (b) of Figure 1.1 shows an interesting case in which the distribution of ε is not the same over time, but changes from one time point to another. For $x = 0$, ε follows a lognormal(2,1) distribution, whereas for $x = 5$ it is distributed according to a lognormal(1,2) distribution. Looking only at the mean regression line, one would conclude that the impact of x is to increase the value of y. However, looking at the quantile regression lines, it is seen that for about 85 percent of the observations, the impact of x is to *decrease* the value of y, while y is increasing for only about 15 percent of the observations. This casts serious doubt about the usefulness of the mean regression for this case. By looking at the quantile regression lines, one can also see how the conditional distribution of y changes for one value of x to another.

A linear regression model in which ε follows a lognormal(2,1) distribution for all observations is plotted in panel (c) of Figure 1.1. This plot shows how a plot of several equally spaced quantiles q for the same data set forms a contour map of all the conditional distributions of y over time. The lower quantile regression lines are quite close, indicating that the conditional distribution is steep at this section, but increasing values of q lead to increasing distances between the quantile regression lines, indicating that the distribution is flattening out. Panel (d) shows a cross section of this distribution. Since the median regression line is not the same as the mean regression line, it can be concluded that the distribution is not symmetric. Looking at the mean regression line, which is a little below the quantile regression line for $q = 0.7$, it is obvious that it would be hard for a non-statistician to understand how this line should be interpreted.

1.2 Inference for quantile regression of longitudinal data

A fundamental goal of statistics is to make inference about parameters by, for example, calculating confidence intervals and conducting hypothesis tests. A drawback of using median regression instead of mean regression is that the statistical properties are less tractable, making it harder to make statistical inferences about the parameters. This characteristic is, of course, also transferred from the median regression case to the general quantile regression case. For longitudinal data, the observations from different subjects are usually assumed to be independent, whereas the different observations from the same subject are correlated.

This special structure of the longitudinal data, with partly dependent observations, makes it harder to make statistical inference. Consequently, making inference for the quantile regression case as applied to longitudinal data is even harder. Moreover, the fact that one in applied sciences neither knows the distribution of the underlying population nor the correlation structure of it implies that one needs to use robust methods for making inference about the quantile regression parameters. The bootstrap (Efron, 1979) is a powerful robust method that can be applied to complicated data sets on which the statistical properties of the applied estimator are hard to derive. It is therefore a natural candidate to use for making inference in the case of quantile regression for longitudinal data.

2 Quantile regression model and estimation methods

To start with, some notation needs to be introduced (cf. Diggle *et al.*, 2002). Let git denote observation $t = 1, ..., T_i$ of subject $i = 1, ..., n_g$ in group $g = 1, ..., G$, observed at time point τ_{git}. Each subject i thus has T_i observations from different time points, and each group g has n_g subjects. The number of observations T_i may vary between subjects, as may the number of subjects n_g between groups. For the case when all subjects i for all groups g have equally many observations, the number of observations will be denoted T, i.e. in this case $T = T_1 = \cdots = T_{n_g}$. In the same way, when all the G groups have equally many subjects i, the number of subjects will be denoted n, i.e. in this case $n = n_1 = \cdots = n_G$.

For a general regression model, let y_{git} denote the response variable for observation t of subject i in group g, with x_{gitk} denoting the explanatory variable $k = 1, ..., K$ for y_{git}, and ε_{git} denoting the corresponding error term. Further, let β_{gh}, with $h = 1, ..., H$, denote the hth parameter for group g of the known response function $f(\cdot)$. Furthermore, let $\mathbf{x}_{git}$ denote the $K \times 1$ vector that contains all the K explanatory variables for y_{git}, let $\mathbf{y}_{gi}$ denote the $T_i \times 1$ vector that contains all the T_i response variables y_{git} for subject i of group g, let $\boldsymbol{\varepsilon}_{gi}$ denote the corresponding $T_i \times 1$ vector that contains all the T_i unknown error terms ε_{git}, and let $\boldsymbol{\beta}_g$ denote the $H \times 1$ vector containing the H unknown parameters β_{gh} for group g.

This results in the following column vectors:

$$\mathbf{y}_{gi} = \begin{pmatrix} y_{gi1} \\ \vdots \\ y_{git} \\ \vdots \\ y_{giT_i} \end{pmatrix}_{T_i \times 1}, \ \mathbf{x}_{git} = \begin{pmatrix} x_{git1} \\ \vdots \\ x_{gitk} \\ \vdots \\ x_{gitK} \end{pmatrix}_{K \times 1}, \ \boldsymbol{\varepsilon}_{gi} = \begin{pmatrix} \varepsilon_{gi1} \\ \vdots \\ \varepsilon_{git} \\ \vdots \\ \varepsilon_{giT_i} \end{pmatrix}_{T_i \times 1}, \ \boldsymbol{\beta}_g = \begin{pmatrix} \beta_{g1} \\ \vdots \\ \beta_{gh} \\ \vdots \\ \beta_{gH} \end{pmatrix}_{H \times 1}.$$

Moreover, let $\mathbf{X}_{gi}$ denote the $T_i \times K$ matrix in which the rows consists of the transposed vectors $\mathbf{x}'_{git}$ of explanatory variables for y_{git},

$$\mathbf{X}_{gi} = \begin{pmatrix} \mathbf{x}'_{gi1} \\ \vdots \\ \mathbf{x}'_{git} \\ \vdots \\ \mathbf{x}'_{giT_i} \end{pmatrix}_{T_i \times K} = \begin{pmatrix} x_{gi11} & \cdots & x_{gi1k} & \cdots & x_{gi1K} \\ \vdots & & \vdots & & \vdots \\ x_{git1} & \cdots & x_{gitk} & \cdots & x_{gitK} \\ \vdots & & \vdots & & \vdots \\ x_{giT_i1} & \cdots & x_{giT_ik} & \cdots & x_{giT_iK} \end{pmatrix}_{T_i \times K}. \tag{2.1}$$

Finally, let $f(\mathbf{x}_{git}, \boldsymbol{\beta}_g)$ denote a known response function of $\mathbf{x}_{git}$ and $\boldsymbol{\beta}_g$, with $\mathbf{f}(\mathbf{X}_{gi}, \boldsymbol{\beta}_g)$ denoting the $T_i \times 1$ vector

$$\mathbf{f}(\mathbf{X}_{gi}, \boldsymbol{\beta}_g) = \begin{pmatrix} f(\mathbf{x}_{gi1}, \boldsymbol{\beta}_g) \\ \vdots \\ f(\mathbf{x}_{git}, \boldsymbol{\beta}_g) \\ \vdots \\ f(\mathbf{x}_{giT_i}, \boldsymbol{\beta}_g) \end{pmatrix}_{T_i \times 1}. \tag{2.2}$$

A general regression model, linear or nonlinear, for longitudinal data can then be written as

$$y_{git} = f\left(\mathbf{x}_{git}, \boldsymbol{\beta}_g\right) + \varepsilon_{git}, \quad g = 1, .., G,\ i = 1, ..., n_g,\ t = 1, ..., T_i, \tag{2.3}$$

or, in matrix form,

$$\mathbf{y}_{gi} = \mathbf{f}(\mathbf{X}_{gi}, \boldsymbol{\beta}_g) + \boldsymbol{\varepsilon}_{gi}, \quad g = 1, .., G,\ i = 1, ..., n_g. \tag{2.4}$$

As can be seen, this is a form of the marginal analysis model for longitudinal data discussed in Diggle *et al.* (2002).

Since the data sets used are longitudinal, it is assumed that the values of the error terms $\varepsilon_{gi1}, ..., \varepsilon_{giT_i}$ for the different observations of a subject i are dependent to an unknown degree, but the error terms ε_{git} and ε_{gjt} for observations of different subjects i and j are independent. If there is only one group in the longitudinal data set, i.e. $G = 1$, the subscript g may be dropped.

2.1 Quantile regression estimation

The seminal paper by Koenker and Bassett (1978) that introduced quantile regression used the case of cross-sectional data with a linear response function. With the notation of this thesis, this corresponds to the model

$$y_i = \mathbf{x}_i'\boldsymbol{\beta} + \varepsilon_i, \quad i = 1, ..., n. \tag{2.5}$$

Now, for $0 < q < 1$, let $\rho_q\left(y_i - \mathbf{x}_i'\boldsymbol{\beta}\right)$ denote the check function

$$\rho_q\left(y_i - \mathbf{x}_i'\boldsymbol{\beta}\right) = \left(y_i - \mathbf{x}_i'\boldsymbol{\beta}\right)\left(q - I\left\{y_i - \mathbf{x}_i'\boldsymbol{\beta} < 0\right\}\right), \tag{2.6}$$

where $I\{\bullet\}$ denotes the indicator function

$$I\left\{y_i - \mathbf{x}_i'\boldsymbol{\beta} < 0\right\} = \begin{cases} 1 \text{ if } y_i - \mathbf{x}_i'\boldsymbol{\beta} < 0 \\ 0 \text{ if } y_i - \mathbf{x}_i'\boldsymbol{\beta} \geq 0 \end{cases}. \tag{2.7}$$

The quantile regression estimate for quantile q given by Koenker and Bassett (1978) is then

$$\mathbf{b}\left(q\right) = \min_{\boldsymbol{\beta}} \sum_{i=1}^{n} \rho_q\left(y_i - \mathbf{x}_i'\boldsymbol{\beta}\right). \tag{2.8}$$

2.1.1 Weighted quantile regression for longitudinal data

Since the longitudinal data in Model (2.3) are correlated, a natural idea to consider when extending Estimator (2.8) to the case of longitudinal data is to use a weighted quantile regression estimator that could take this correlation into account. To estimate a weighted quantile regression for longitudinal data, let w_{git} denote the weight for observation git and $\mathbf{W}_{gi}$ a symmetric $T_i \times T_i$ matrix of weights for subject i of group g,

$$\mathbf{W}_{gi}=\begin{pmatrix} w_{gi11} & \cdots & w_{gi1t'} & \cdots & w_{gi1T_i} \\ \vdots & \ddots & \vdots & & \vdots \\ w_{git1} & \cdots & w_{gitt'} & \cdots & w_{gitT_i} \\ \vdots & & \vdots & \ddots & \vdots \\ w_{giT_i1} & \cdots & w_{giT_it'} & \cdots & w_{giT_iT_i} \end{pmatrix}_{T_i\times T_i}, \quad \mathbf{W}'_{gi}=\mathbf{W}_{gi}, \tag{2.9}$$

where the element $w_{gitt'}$, with $t,t'=1,...,T_i$, denotes the weight for observation t of subject i belonging to group g. Further, for $0<q<1$, let $\rho_q\left(y_{git}-f\left(\mathbf{x}_{git},\boldsymbol{\beta}_g\right)\right)$ denote the check function

$$\rho_q\left(y_{git}-f\left(\mathbf{x}_{git},\boldsymbol{\beta}_g\right)\right)=\left(y_{git}-f\left(\mathbf{x}_{git},\boldsymbol{\beta}_g\right)\right)\left(q-I\left\{y_{git}-f\left(\mathbf{x}_{git},\boldsymbol{\beta}_g\right)<0\right\}\right), \tag{2.10}$$

with $I\{\bullet\}$ denoting the indicator function

$$I\left\{y_{git}-f\left(\mathbf{x}_{git},\boldsymbol{\beta}_g\right)<0\right\}=\begin{cases} 1 \text{ if } y_{git}-f\left(\mathbf{x}_{git},\boldsymbol{\beta}_g\right)<0 \\ 0 \text{ if } y_{git}-f\left(\mathbf{x}_{git},\boldsymbol{\beta}_g\right)\geq 0 \end{cases}. \tag{2.11}$$

Finally, let $\boldsymbol{\rho}_q\left(\mathbf{y}_{gi}-\mathbf{f}\left(\mathbf{X}_{gi},\boldsymbol{\beta}_g\right)\right)$ denote the corresponding column vector of length T_i,

$$\boldsymbol{\rho}_q\left(\mathbf{y}_{gi}-\mathbf{f}\left(\mathbf{X}_{gi},\boldsymbol{\beta}_g\right)\right)=\begin{pmatrix} \rho_q\left(y_{gi1}-f\left(\mathbf{x}_{gi1},\boldsymbol{\beta}_g\right)\right) \\ \vdots \\ \rho_q\left(y_{git}-f\left(\mathbf{x}_{git},\boldsymbol{\beta}_g\right)\right) \\ \vdots \\ \rho_q\left(y_{giT_i}-f\left(\mathbf{x}_{giT_i},\boldsymbol{\beta}_g\right)\right) \end{pmatrix}_{T_i\times 1}. \tag{2.12}$$

The weighted quantile regression estimator for longitudinal data is then given by

$$\widehat{\boldsymbol{\beta}}_g(q)=\min_{\boldsymbol{\beta}_g}\sum_{i=1}^{n_g}\sum_{t=1}^{T_i} w_{git}\rho_q\left(y_{git}-f\left(\mathbf{x}_{git},\boldsymbol{\beta}_g\right)\right), \tag{2.13}$$

with the corresponding matrix case being

$$\begin{aligned} \widehat{\boldsymbol{\beta}}_g(q) &= \min_{\boldsymbol{\beta}_g}\sum_{i=1}^{n_g}\mathbf{1}'_{T_i}\mathbf{W}_{gi}\boldsymbol{\rho}_q\left(\mathbf{y}_{gi}-\mathbf{f}\left(\mathbf{X}_{gi},\boldsymbol{\beta}_g\right)\right) \\ &= \min_{\boldsymbol{\beta}_g}\sum_{i=1}^{n_g}\sum_{t=1}^{T_i}\sum_{t'=1}^{T_i} w_{gitt'}\rho_q\left(y_{git}-f\left(\mathbf{x}_{git},\boldsymbol{\beta}_g\right)\right), \end{aligned} \tag{2.14}$$

where $\mathbf{1}_{T_i}=(1,...,1)'$ denotes a column vector of 1s with length T_i. With $\boldsymbol{\beta}_g(q)$ denoting the column vector $\boldsymbol{\beta}_g$ containing the H unknown parameters of group g for quantile q, $\widehat{\boldsymbol{\beta}}_g(q)$ is thus the column vector of estimates of the unknown values of $\boldsymbol{\beta}_g(q)$, i.e.

$$\boldsymbol{\beta}_g(q) = \begin{pmatrix} \beta_{g1}(q) \\ \vdots \\ \beta_{gh}(q) \\ \vdots \\ \beta_{gH}(q) \end{pmatrix}_{H\times 1}, \quad \widehat{\boldsymbol{\beta}}_g(q) = \begin{pmatrix} \widehat{\beta}_{g1}(q) \\ \vdots \\ \widehat{\beta}_{gh}(q) \\ \vdots \\ \widehat{\beta}_{gH}(q) \end{pmatrix}_{H\times 1}. \tag{2.15}$$

As is obvious, Estimator (2.13) is the special case of Estimator (2.14) where all matrices $\mathbf{W}_{gi}$ are diagonal with weights w_{git} as diagonal elements $w_{gitt'}$, and the special case $\widehat{\boldsymbol{\beta}}_g(0.5)$ is the estimate for the median regression case. Moreover, as can be seen, Estimators (2.13) and (2.14) are weighted generalizations to the case of longitudinal data of the Koenker and Bassett quantile regression estimator (2.8).

3 Bootstrapping quantile regression for longitudinal data

Let $\widehat{\boldsymbol{\beta}}^*_{gb}(q) = \left(\widehat{\beta}^*_{g1}(q), \ldots, \widehat{\beta}^*_{gh}(q), \ldots, \widehat{\beta}^*_{gH}(q)\right)'$ denote bootstrap estimate b, where $b = 1, ..., B$, of $\boldsymbol{\beta}_g(q)$. The B bootstrap estimates together form a bootstrap distribution, which is an estimate of the sampling distribution of the estimated quantile regression parameter vector $\widehat{\boldsymbol{\beta}}_g(q)$. Thus, by estimating the characteristics of this bootstrap distribution bootstrap estimates of the characteristics of $\widehat{\boldsymbol{\beta}}_g(q)$ are obtained. Similarly, inferences for the quantile regression parameter vector $\boldsymbol{\beta}_g(q)$ can be based on this bootstrap distribution.

When using bootstrap methods with a regression model, one could choose to resample either the residuals e_{git}, i.e. error or residual bootstrapping, or the pair $(y_{git}, \mathbf{x}'_{git})$ of observations with response and explanatory variables, i.e. paired or design matrix bootstrapping, sometimes called case bootstrapping. In this thesis only the latter method will be used. Buchinsky (1995) showed that, for quantile regressions, paired bootstrapping was to be preferred over the error bootstrapping method. Paired bootstrapping was also recommended by Koenker (2005, Chapter 3.9). For details about the use of bootstrapping with regression models, see Davison and Hinkley (1997).

3.1 Paired bootstrap with independent resampling

The simplest bootstrap method for longitudinal data is to disregard the dependence between the observations for a subject and proceed as if all observations and subjects were independent. In this thesis this method is called *paired bootstrap with independent resampling*, PB_I and is performed according to the following algorithm:

Algorithm 1 (PB_I) *For $b = 1, ..., B$*

1. *Generate the paired bootstrap sample $\left(y^*_{g1}, \mathbf{x}'^*_{g1}\right), ..., \left(y^*_{gM_g}, \mathbf{x}'^*_{gM_g}\right)$ with M_g pairs by resampling with replacement from the longitudinal data set consisting of $N_g = \sum_{i\in g} T_i$ pairs $\left(y_{gi}, \mathbf{x}'_{gi}\right)$.*

2. *Calculate a weight w^*_{gj} for each pair $\left(y^*_{gj}, \mathbf{x}^{*\prime}_{gj}\right)$, $j = 1, ..., M_g$.*

3. *Compute the* b*th paired bootstrap estimate* $\widehat{\boldsymbol{\beta}}^*_{gb}(q)$ *by the formula*

$$\widehat{\boldsymbol{\beta}}^*_{gb}(q) = \min_{\boldsymbol{\beta}_g} \sum_{j=1}^{M_g} w^*_{gj} \rho_q \left(y^*_{gj} - f\left(\mathbf{x}^*_{gj}, \boldsymbol{\beta}_g\right)\right), \quad b = 1, ..., B. \tag{3.1}$$

This algorithm produces B bootstrap estimates $\widehat{\boldsymbol{\beta}}^*_{gb}(q)$ of $\boldsymbol{\beta}_g(q)$. Any value for M_g can be used, but in this thesis $M_g = N_g$ is used.

3.2 Paired bootstrap with dependent resampling

Since a longitudinal data set has a special dependence structure, a bootstrap method that makes use of this special structure could possibly give better bootstrap estimates than the simple bootstrap method PB_I, which disregards this dependence. Specifically, since longitudinal data are correlated, it would possibly be a good idea to use bootstrap methods that preserve the correlation structure of the data. Such a method is the block bootstrap method in which blocks of consecutive observations are resampled, instead of resampling single observations. By using this method, the dependence structure in each block is retained (see Lahiri, 2003 for details about different block bootstrap methods). Fitzenberger (1997) considers the moving blocks bootstrap method.

For the case of longitudinal data considered in this thesis, in which it is assumed that the subjects i are independent but the T_i consecutive observations for an individual subject are dependent, the natural block to resample is the pair $(\mathbf{y}_{gi}, \mathbf{X}_{gi})$, consisting of the T_i consecutive observations t for subject i of group g, i.e. all observations for the response and explanatory variables for a single subject. By resampling the n_g pairs or blocks $(\mathbf{y}_{gi}, \mathbf{X}_{gi})$ the correlation structure in the longitudinal data set is retained. Let b denote a bootstrap resample and B the total number of bootstrap resamplings. In this thesis, this method is called *paired bootstrap with dependent resampling*, PB_D and is performed according to the following algorithm:

Algorithm 2 (PB_D) *For* $b = 1, ..., B$

1. *Generate the paired bootstrap sample* $\left(\mathbf{y}^*_{g1}, \mathbf{X}^*_{g1}\right), ..., \left(\mathbf{y}^*_{gm_g}, \mathbf{X}^*_{gm_g}\right)$ *with* m_g *pairs by resampling with replacement from the longitudinal data set consisting of* n_g *pairs* $(\mathbf{y}_{gi}, \mathbf{X}_{gi})$.

2. *Calculate a weight matrix* $\mathbf{W}^*_{gj}$ *for each pair* $\left(\mathbf{y}^*_{gj}, \mathbf{X}^*_{gj}\right)$, $j = 1, ..., m_g$.

3. *Compute the* b*th paired bootstrap estimate* $\widehat{\boldsymbol{\beta}}^*_{gb}(q)$ *by the formula*

$$\widehat{\boldsymbol{\beta}}^*_{gb}(q) = \min_{\boldsymbol{\beta}_g} \sum_{j=1}^{m_g} \mathbf{1}'_{T_j} \mathbf{W}^*_{gj} \boldsymbol{\rho}_q \left(\mathbf{y}^*_{gj} - \mathbf{f}\left(\mathbf{X}^*_{gj}, \boldsymbol{\beta}_g\right)\right), \quad b = 1, ..., B. \tag{3.2}$$

This algorithm produces B bootstrap estimates $\widehat{\boldsymbol{\beta}}^*_{gb}(q)$ of $\boldsymbol{\beta}_g(q)$. Any value for m_g can be used, but in this thesis $m_g = n_g$ is used.

4 Summary of the papers

The overall objective of the four papers is to examine the properties and usefulness of the weighted nonlinear quantile regression estimator in the analysis of longitudinal data. To this end, the questions of which weights to be used, the bias of the estimator, the possibility to calculate confidence intervals and the performance of hypothesis tests have to be examined. This is done in the first three papers, focusing on small samples, as this is the most common case for real world applications in biostatistics. The fourth paper uses the results of the first three papers to demonstrate the new insights that can be obtained regarding the properties of longitudinal biostatistics data from using quantile regression methods.

4.1 Paper I - Nonlinear quantile regression estimation of longitudinal data

This paper examines the properties of some weights used with the weighted nonlinear quantile regression estimators (2.13) and (2.14) for estimating Model (2.3) using the four-parameter logistic growth function

$$f\left(\mathbf{x}_{git}, \boldsymbol{\beta}_g\right) = \beta_{g1} + \frac{\beta_{g2} - \beta_{g1}}{1 + \exp\left(\beta_{g4}\left(x_{git} - \beta_{g3}\right)\right)}, \tag{4.1}$$

and the error terms follow the AR(1) model

$$\varepsilon_{git} = \rho\varepsilon_{gi,t-1} + u_{git}, \quad g = 1..., G,\ i = 1, ..., n_g,\ t = 1, ..., T. \tag{4.2}$$

Only one group is used for this paper, i.e. $G = 1$, so the subscript g is dropped from all variables and parameters in the paper, although it is retained here for reasons of continuity. Further, note that $T = T_i$ is used here since all subjects i have equally many observations.

The interpretation of the parameters in (4.1) for group g is as follows (cf. Davidian and Giltinan 1995, p. 10): β_{g1} gives the value of the dependent variable y_{git} at $-\infty$ and β_{g2} gives its value at ∞. The parameter β_{g3} gives the EC_{50} value, i.e. the value of x_{git} for which y_{git} reaches the midpoint $\left|\beta_{g2} - \beta_{g1}\right|/2$, halfway between the two asymptotes β_{g1} and β_{g2}. The value of y_{git} at this point equals $\left(\beta_{g2} - \beta_{g1}\right)/2$. Further, β_{g4} is a slope parameter governing the steepness of the growth curve. Figure 4.1 shows (4.1) for $\beta_{g1} = 10$, $\beta_{g2} = 20$, $\beta_{g3} = 5$ and $\beta_{g4} = -0.5$.

The performance of the weights is studied by the use of simulation methods for the quantiles $q = 0.5$, 0.75 and 0.9, mainly by looking at the mean absolute percent error (MAPE). The comparison is made for different degrees of autocorrelation, distributions, number of observations, number of subjects and types of time point.

Using simulations, a comparison is also made between the median regression case of the quantile regression estimator and the corresponding mean regression estimator, the weighted least squares (WLSQ) regression estimator, given by

$$\widetilde{\boldsymbol{\beta}}_{g\,\min} = \min_{\boldsymbol{\beta}_g} \sum_{i=1}^{n_g} \sum_{t=1}^{T_i} w_{git}\left(y_{git} - f\left(\mathbf{x}_{git}, \boldsymbol{\beta}\right)\right)^2, \tag{4.3}$$

or in the matrix case

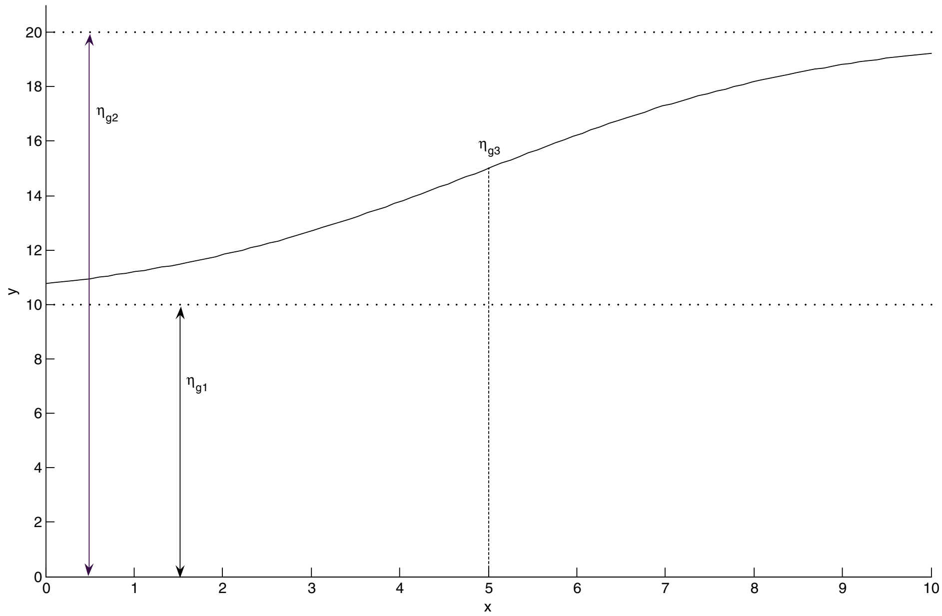

Figure 4.1: A four-parameter logistic growth function with $\beta_{g1} = 10$, $\beta_{g2} = 20$, $\beta_{g3} = 5$ and $\beta_{g4} = -0.5$.

$$\widetilde{\boldsymbol{\beta}}_{g\min} = \min_{\boldsymbol{\beta}_g} \sum_{i=1}^{n_g} \left(\mathbf{y}_{gi} - f\left(\mathbf{X}_{gi}, \boldsymbol{\beta}_g\right)\right)' \mathbf{W}_{gi} \left(\mathbf{y}_{gi} - f\left(\mathbf{X}_{gi}, \boldsymbol{\beta}_g\right)\right), \tag{4.4}$$

where $\mathbf{W}_{gi}$ for the weight specifications i-v are diagonal matrices with the weights w_{git}, $t = 1, ..., T_i$, on the diagonals.

4.1.1 Construction of weights

The weights are based on residuals e_{git} computed by first estimating $\widehat{\boldsymbol{\beta}}_g(0.5)$ when $w_{git} = 1$, i.e.

$$\widehat{\boldsymbol{\beta}}_g(0.5) = \min_{\boldsymbol{\beta}_g} \sum_{i=1}^{n_g} \sum_{t=1}^{T_i} \rho_{0.5}\left(y_{git} - f\left(\mathbf{x}_{git}, \boldsymbol{\beta}_g\right)\right), \tag{4.5a}$$

and then with the help of this $\widehat{\boldsymbol{\beta}}_g(0.5)$ value calculate the residuals as

$$e_{git} = y_{git} - f\left(\mathbf{x}_{git}, \widehat{\boldsymbol{\beta}}_g(0.5)\right), \quad g = 1..., G,\ i = 1, ..., n,\ t = 1, ..., T_i. \tag{4.6}$$

To construct the weights, let $\mathbf{1}_{n_g} = (1, ..., 1)'$ denote a column vector of 1s with length n_g, $\mathbf{E}_g$ denote the $n_g \times T$ matrix of residuals

$$\mathbf{E}_g = \begin{pmatrix} \mathbf{e}_{g1} & \cdots & \mathbf{e}_{gi} & \cdots & \mathbf{e}_{gn_g} \end{pmatrix}_{n_g \times T} = \begin{pmatrix} e_{g11} & \cdots & e_{gi1} & \cdots & e_{gn_g1} \\ \vdots & & \vdots & & \vdots \\ e_{g1t} & \cdots & e_{git} & \cdots & e_{gn_gt} \\ \vdots & & \vdots & & \vdots \\ e_{g1T} & \cdots & e_{giT} & \cdots & e_{gn_gT} \end{pmatrix}_{n_g \times T}, \quad (4.7)$$

and $\mathbf{S}_g$ denote the $T \times T$ estimated variance-covariance matrix

$$\mathbf{S}_g = \frac{1}{n_g}\left(\mathbf{E}_g'\mathbf{E}_g - \frac{1}{n_g}\mathbf{E}_g'\mathbf{1}_{n_g}\mathbf{1}_{n_g}'\mathbf{E}_g\right) = \begin{pmatrix} s_{g11} & \cdots & s_{g1t'} & \cdots & s_{g1T} \\ \vdots & \ddots & \vdots & & \vdots \\ s_{gt1} & \cdots & s_{gtt'} & \cdots & s_{gtT} \\ \vdots & & \vdots & \ddots & \vdots \\ s_{gT1} & \cdots & s_{gTt'} & \cdots & s_{gTT} \end{pmatrix}_{T \times T}, \quad \mathbf{S}_g' = \mathbf{S}_g. \quad (4.8)$$

Furthermore, define s_{gi}^2, s_{gt}^2, u_{gi} and u_{gt} as

$$s_{gi}^2 = \frac{1}{T_i}\sum_{t=1}^{T_i}\left(e_{git} - \overline{e}_{gi}\right)^2, \quad g = 1..., G, \; i = 1, ..., n, \quad (4.9)$$

$$s_{gt}^2 = \frac{1}{n_g}\sum_{i=1}^{n_g}\left(e_{git} - \overline{e}_{gt}\right)^2, \quad g = 1..., G, \; t = 1, ..., T, \quad (4.10)$$

$$u_{gi} = \frac{1}{T_i}\sum_{t=1}^{T_i}\left|e_{git} - \overline{e}_{gi}\right|, \quad g = 1..., G, \; i = 1, ..., n, \quad (4.11)$$

and

$$u_{gt} = \frac{1}{n_g}\sum_{i=1}^{n_g}\left|e_{git} - \overline{e}_{gt}\right|, \quad g = 1..., G, \; t = 1, ..., T, \quad (4.12)$$

where

$$\overline{e}_{gi} = \frac{1}{T_i}\sum_{t=1}^{T_i} e_{git}, \quad g = 1..., G, \; i = 1, ..., n, \quad (4.13)$$

and

$$\overline{e}_t = \frac{1}{n}\sum_{i=1}^{n} e_{it}, \quad t = 1, ..., T. \quad (4.14)$$

The following weights are then used in this paper:

i. $w_{git} = 1$,

ii. $w_{git} = \frac{1}{u_{gt}}$,

iii. $w_{git} = \frac{1}{u_{gi}}$,

iv. $w_{git} = \frac{1}{\sqrt{s_{gt}^2}}$,

v. $w_{git} = \frac{1}{\sqrt{s_{gi}^2}}$,

vi. $\mathbf{W}_{gi} = \left(\frac{1}{|s_{gtt'}|}\right)$, i.e. $w_{gitt'} = \frac{1}{|s_{gtt'}|}$,

vii. $\mathbf{W}_{gi} = \left|\mathbf{S}_g^{-1}\right|$.

4.1.2 Results

The results indicated that the differences between the seven weights used were small, but that weights iii and v (i.e. those based on the dispersion of the estimated regression errors over subjects) generally perform somewhat better than the other weights. A further finding was that the quantile regression estimator performs quite well in terms of bias, especially for the median regression case, but the further away from the median the quantiles q are, the less well the estimator works. The latter is also true for the MAPE.

When the autocorrelation ρ in the AR(1) model is increased, all weights are performing better, which is somewhat surprising; however, these results are clear. Increasing the number of subjects or observations generally leads to better estimates. Overall, there are quite small differences for the performances between the different distributions used, but the estimator is generally performing better for nonrandom than for random time points.

When comparing the median regression case of the quantile regression estimator and the corresponding mean regression estimator, the weighted least squares (WLSQ) regression estimator, the WLSQ estimator was found to be less robust.

Finally, the quantile regression estimator is, together with the WLSQ estimator, applied to a real data set in which the growth patterns of two genotypes of soybean are compared. This comparison provides some insights into how the quantile regressions give a more complete picture of the data than the mean regression does.

4.2 Paper II - Bootstrap methods for bias correction and confidence interval estimation for nonlinear quantile regression of longitudinal data

This paper examines the use of bootstrapping for (a) reducing the bias of the estimates $\widehat{\boldsymbol{\beta}}_g(q)$ from the weighted nonlinear quantile regression estimators (2.13) and (2.14), and (b) calculating confidence intervals for the nonlinear quantile regression parameter vector $\boldsymbol{\beta}_g(q)$. The examination is performed through Monte Carlo simulations of Model (2.3) using response function (4.1), with the error terms following the AR(1) model (4.2). Note that only one group is used for this paper, so the subscript g is dropped from all variables and parameters in the paper, although it is retained here for continuity reasons.

Weights i, v and vi are used in this study in (2.13) and (2.14) for the quantiles $q = 0.5$, 0.75 and 0.9, as well as different distributions, degrees of autocorrelation, numbers of observations and numbers of subjects. Of the three weights, all three were used with bootstrap method PB_D but only weight i with bootstrap methods PB_I since the other weights could not be used with this bootstrap method.

To examine the use of bootstrapping for bias reduction the bias of the bias-corrected parameter estimate

$$\widehat{\boldsymbol{\beta}}_g^c(q) = \begin{pmatrix} \widehat{\beta}_{g1}^c(q) \\ \vdots \\ \widehat{\beta}_{gh}^c(q) \\ \vdots \\ \widehat{\beta}_{gH}^c(q) \end{pmatrix}_{H\times 1}, \tag{4.15}$$

where

$$\widehat{\beta}_{gh}^c(q) = 2\widehat{\beta}_{gh}(q) - \frac{1}{B}\sum_{b=1}^{B}\widehat{\beta}_{ghb}^*(q), \tag{4.16}$$

is compared with the bias of the original quantile egression estimate $\widehat{\boldsymbol{\beta}}_g(q)$. For the calculation of bootstrap-based confidence intervals, four methods are used: the *normal approximation method* CI_N, the *bias-corrected normal approximation method* CI_{NBC}, the *percentile method* CI_P, and the *bias-corrected percentile method* CI_{PBC}. These are compared by examining the coverage and length of confidence intervals when the nominal coverage is 95 percent.

4.2.1 Calculation of confidence intervals

Using CI_N and CI_{NBC}, $100(1-2\alpha)$ percent confidence intervals for $\boldsymbol{\beta}_g(q)$ are given by

$$\widehat{\boldsymbol{\beta}}_g(q) \pm Z_\alpha\sqrt{\frac{1}{B-1}\sum_{b=1}^{B}\left(\widehat{\boldsymbol{\beta}}_{gb}^*(q) - \frac{1}{B}\sum_{b=1}^{B}\widehat{\boldsymbol{\beta}}_{gb}^*(q)\right)^2} \tag{4.17}$$

and

$$\widehat{\boldsymbol{\beta}}_g^c(q) \pm Z_\alpha\sqrt{\frac{1}{B-1}\sum_{b=1}^{B}\left(\widehat{\boldsymbol{\beta}}_{gb}^*(q) - \frac{1}{B}\sum_{b=1}^{B}\widehat{\boldsymbol{\beta}}_{gb}^*(q)\right)^2}, \tag{4.18}$$

respectively. For CI_P, a $100(1-2\alpha)$ percent confidence interval is given by

$$\widehat{\boldsymbol{\beta}}_{g(\alpha(B+1))}^*(q) \leq \boldsymbol{\beta}_g(q) \leq \widehat{\boldsymbol{\beta}}_{g((1-\alpha)(B+1))}^*(q), \tag{4.19}$$

where

$$\widehat{\boldsymbol{\beta}}_{g(b)}^*(q) = \begin{pmatrix} \widehat{\beta}_{g1(b)}^*(q) \\ \vdots \\ \widehat{\beta}_{gh(b)}^*(q) \\ \vdots \\ \widehat{\beta}_{gH(b)}^*(q) \end{pmatrix}_{H\times 1}, \tag{4.20}$$

and $\widehat{\beta}_{gh(b)}^*(q)$ is the bth order statistic of the B bootstrap replicates $\widehat{\beta}_{gh(1)}^*(q) \leq \cdots \leq \widehat{\beta}_{gh(B)}^*(q)$ of $\widehat{\beta}_{gh}(q)$. Finally, let $P^*\left(\widehat{\beta}_{gh}^*(q) \leq \widehat{\beta}_{gh}(q)\right)$ denote the proportion of $\widehat{\beta}_{gh}^*(q)$s in the bootstrap sample that have a value lower than the value of the parameter estimate $\widehat{\beta}_{gh}(q)$, and define z_0 as

$$z_0 = \Phi^{-1}\left(P^*\left(\widehat{\beta}^*_{gh}(q) \leq \widehat{\beta}_{gh}(q)\right)\right), \tag{4.21}$$

where Φ denotes the cumulative distribution function (CDF) of the normal distribution. A $100(1-2\alpha)$ percent confidence interval using CI_{PBC} is then given by

$$\widehat{\boldsymbol{\beta}}^*_{g(\Phi(2z_0+z_\alpha)(B+1))}(q) \leq \boldsymbol{\beta}_g(q) \leq \widehat{\boldsymbol{\beta}}^*_{g(\Phi(2z_0+z_{1-\alpha})(B+1))}(q). \tag{4.22}$$

4.2.2 Results

For the bootstrap-based bias correction, the bootstrap method PB_I was observed to perform badly in improving bias; in fact, overall the bias deteriorated when using PB_I. However, the bootstrap method PB_D performed well, with generally somewhat improved biases for all weights used. Overall, weight vi had the lowest bias, followed by the unweighted alternative weight i. Another finding was that, in general, the bias improvements increase with values of q further away from the median and that the bias improvements decrease with increasing ρ.

For the calculations of 95 percent confidence intervals, the bias-corrected methods CI_{NBC} and CI_{PBC} performed bad for all weights and both bootstrap methods PB_I and PB_D, with true coverage percentages usually below the nominal 95 percent level. But weight i, using PB_I or PB_D, and weight vi performed well for both the normal approximation method CI_N and the percentile method CI_P, with all or nearly all confidence intervals having a true coverage of at least the nominal 95 percent level. Overall, weight i using PB_D for percentile method CI_P performed best in terms of coverage and length of confidence interval, followed by weight vi for percentile method CI_P. A further finding was that values of q further away from the median led to longer confidence intervals, whereas increasing the correlation ρ led to shorter confidence intervals.

The use of bootstrapping for reducing bias and calculating confidence intervals for a weighted nonlinear quantile regression estimator is also applied to a real longitudinal data set with growth patterns, measured as leaf weight per plant, of two genotypes of soybean, where nonlinear quantile regressions are calculated for the nine quantiles $q = 0.1, 0.2, ..., 0.9$. For this application, weight i is used with bootstrap method PB_D, and confidence intervals calculated using the percentile method CI_P and the normal approximation method CI_N. It was found that using bias correction led to some problems for this case. One problem is that when using bias-corrected parameter estimates for the quantile regression crossing quantile regression curves sometimes occur. Another problem is that the bias-corrected parameter estimates are sometimes outside the limits of the confidence intervals for the percentile method CI_P.

In conclusion regarding the performance for both bias correction and confidence interval calculation, the following recommendations can be offered when performing nonlinear quantile regression estimation for longitudinal data: use either weight i or weight vi without bias correction and calculate the confidence intervals with the percentile method CI_P using the bootstrap method PB_D.

4.3 Paper III - Bootstrap-based hypothesis tests for nonlinear quantile regression of longitudinal data

This paper examines the use of bootstrap-based hypothesis tests for the weighted nonlinear quantile regression estimators (2.13) and (2.14) using weights i and vi. The focus is on testing the equality between two groups of one or more of the regression parameters in the four-parameter logistic growth function (4.1) for different quantiles q. The performance of the different tests is compared with simulation methods. The comparison is focused on evaluating the performance of the tests for the empirical significance level attained, the estimated nominal significance level, the empirical power, and the estimated power achieved.

The equality of two regression parameters $\beta_{1h}(q)$ and $\beta_{2h}(q)$ for some quantile q is tested with the null hypothesis

$$H_0 : \beta_{1h}(q) - \beta_{2h}(q) = 0, \tag{4.23}$$

against two separate alternative hypotheses, the one-sided alternative hypothesis

$$H_{a1} : \beta_{1h}(q) - \beta_{2h}(q) > 0 \tag{4.24}$$

and the two-sided alternative hypothesis

$$H_{a2} : \beta_{1h}(q) - \beta_{2h}(q) \neq 0. \tag{4.25}$$

The simultaneous equality between two groups of all H parameters $\beta_{1h}(q)$ and $\beta_{2h}(q)$ for some quantile q is tested with the null hypothesis

$$\mathbf{H}_0 : \mathbf{R}\boldsymbol{\beta}_0(q) = \mathbf{0} \tag{4.26}$$

against the alternative hypothesis

$$\mathbf{H}_a : \mathbf{R}\boldsymbol{\beta}_0(q) \neq \mathbf{0}, \tag{4.27}$$

where

$$\mathbf{R} = \begin{pmatrix} 1 & -1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & -1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & -1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & -1 \end{pmatrix}, \tag{4.28}$$

$$\boldsymbol{\beta}_0(q) = \begin{pmatrix} \beta_{11}(q) & \beta_{21}(q) & \beta_{12}(q) & \beta_{22}(q) & \beta_{13}(q) & \beta_{23}(q) & \beta_{14}(q) & \beta_{24}(q) \end{pmatrix}', \tag{4.29}$$

and $\mathbf{0}$ denotes a column vector of 0s with length H.

4.3.1 Bootstrap-based hypothesis tests

The test procedures that are used can be divided into two variants, namely a normal approximation-based testing procedure in which the critical values for the tests are obtained using the assumption of asymptotic normality and a percentile-based testing procedure in which the critical values for the tests are obtained from the percentiles of the bootstrap distributions of the tests. For both variants, both bias-corrected and non-bias-corrected versions of the tests are used. The bootstrap method PB_D is used for all cases.

Single restriction tests For the single restriction (4.23) of equality of one of the parameters between the two groups, the basic non-bias-corrected test statistic used is

$$Z^* = \frac{\left(\widehat{\beta}_{1h}(q) - \widehat{\beta}_{2h}(q)\right) - \left(\beta_{1h}(q) - \beta_{2h}(q)\right)}{\sqrt{\widehat{Var}^*\left(\widehat{\beta}_{1h}(q)\right) + \widehat{Var}^*\left(\widehat{\beta}_{2h}(q)\right)}}, \tag{4.30}$$

where $\beta_{1h}(q) - \beta_{2h}(q) = 0$ under the null hypothesis (4.23). The bootstrap estimator $\widehat{Var}^*\left(\widehat{\beta}_{gh}(q)\right)$ of $Var\left(\widehat{\beta}_{gh}(q)\right)$ is given by

$$\widehat{Var}^*\left(\widehat{\beta}_{gh}(q)\right) = \frac{1}{B-1}\sum_{b=1}^{B}\left(\widehat{\beta}^*_{ghb}(q) - \frac{1}{B}\sum_{b=1}^{B}\widehat{\beta}^*_{ghb}(q)\right)^2, \quad g = 1, 2. \tag{4.31}$$

Since $\widehat{\beta}_{1h}(q)$ and $\widehat{\beta}_{2h}(q)$ are estimated separately for the two groups, they are independent. Thus, the covariance term has been cancelled from (4.30).

Let Z^*_{obs} denote the observed value of Z^* and Z_α the critical value obtained from the standard normal distribution using significance level α, i.e. Z_α is the value of the standard normal variate Z for which $P(Z > Z_\alpha) = \alpha$. For the normal approximation tests, if the one-sided alternative hypothesis (4.24) is used, the null hypothesis (4.23) is rejected on significance level α if $Z^*_{obs} > Z_\alpha$; if the two-sided alternative hypothesis (4.25) is used, the null hypothesis (4.23) is rejected on significance level α if $|Z^*_{obs}| > Z_{\alpha/2}$. This normal approximation test is denoted Z^*_n in this paper.

For the percentile based testing procedure, the critical values for Z^* are given by the order statistics of the bootstrap test statistic

$$Z^*_b = \frac{\left(\widehat{\beta}^*_{1hb}(q) - \widehat{\beta}^*_{2hb}(q)\right) - \left(\widehat{\beta}_{1h}(q) - \widehat{\beta}_{2h}(q)\right)}{\sqrt{\widehat{Var}^*\left(\widehat{\beta}_{1h}(q)\right) + \widehat{Var}^*\left(\widehat{\beta}_{2h}(q)\right)}}, \tag{4.32}$$

where the reason for replacing the $\beta_{gh}(q)$s with their estimates $\widehat{\beta}_{gh}(q)$ is given by the first guideline in Hall and Wilson (1991). After calculating B bootstrap estimates Z^*_b, let $Z^*_{(b)}$ denote the bth order statistic of the B bootstrap replicates $Z^*_{(1)} \leq \cdots \leq Z^*_{(B)}$ from (4.32). When the one-sided alternative hypothesis (4.24) is used, the null hypothesis (4.23) is rejected on significance level α if $Z^*_{obs} > Z^*_{((1-\alpha)(B+1))}$. This percentile-based test is denoted Z^*_p in this paper.

When the two-sided alternative hypothesis (4.25) is used, two rejection methods can be used: a symmetrical or a non-symmetrical rejection method (cf. Cameron and Trivedi, 2005, Chapter 11.2.6). Using the non-symmetrical rejection method, the null hypothesis (4.23) is rejected on significance level α if $Z^*_{obs} < Z^*_{((\alpha/2)(B+1))}$ or $Z^*_{obs} > Z^*_{((1-\alpha/2)(B+1))}$; using the symmetrical rejection method, the null hypothesis (4.23) is rejected on significance level α if $|Z^*_{obs}| > \left|Z^*_{((1-\alpha)(B+1))}\right|$, where $\left|Z^*_{(b)}\right|$ denotes the bth ordered absolute value of the B bootstrap replicates from (4.32), i.e. $\left|Z^*_{(1)}\right| \leq \cdots \leq \left|Z^*_{(B)}\right|$. The test procedure using the non-symmetrical rejection method is denoted Z^*_{pns} while the test procedure using the symmetrical rejection method is denoted Z^*_{ps}.

The bias-corrected versions of (4.30) and (4.32) are given by replacing the parameter estimates $\widehat{\beta}_{gh}(q)$ with the corresponding bias-corrected parameter estimates $\widehat{\beta}^c_{gh}(q)$, resulting in the formulas

$$Z^{c*} = \frac{\widehat{\beta}^{c}_{1h}(q) - \widehat{\beta}^{c}_{2h}(q)}{\sqrt{\widehat{Var}^{*}\left(\widehat{\beta}_{1h}(q)\right) + \widehat{Var}^{*}\left(\widehat{\beta}_{2h}(q)\right)}}. \tag{4.33}$$

and

$$Z^{c*}_{b} = \frac{\left(\widehat{\beta}^{*}_{1hb}(q) - \widehat{\beta}^{*}_{2hb}(q)\right) - \left(\widehat{\beta}^{c}_{1h}(q) - \widehat{\beta}^{c}_{2h}(q)\right)}{\sqrt{\widehat{Var}^{*}\left(\widehat{\beta}_{1h}(q)\right) + \widehat{Var}^{*}\left(\widehat{\beta}_{2h}(q)\right)}}. \tag{4.34}$$

The rejection regions for these are obtained in an obvious way by replacing Z^{*} with Z^{c*} and Z^{*}_{b} with Z^{c*}_{b}. The resulting test procedures are denoted Z^{c*}_{n}, Z^{c*}_{p}, Z^{c*}_{pns}, and Z^{c*}_{ps}.

Joint restriction tests For the joint restriction (4.26) of equality of all of the parameters between the two groups, the basic non-bias-corrected test statistic used is the Wald-like

$$W^{*} = \left[\mathbf{R}\widehat{\boldsymbol{\beta}}_{0}(q) - \mathbf{R}\boldsymbol{\beta}_{0}(q)\right]' \left[\mathbf{C}\widehat{\mathbf{Var}}^{*}\left(\widehat{\boldsymbol{\beta}}_{\mathbf{0}}(\mathbf{q})\right)\mathbf{C}'\right]^{-1} \left[\mathbf{R}\widehat{\boldsymbol{\beta}}_{0}(q) - \mathbf{R}\boldsymbol{\beta}_{0}(q)\right], \tag{4.35}$$

where $\mathbf{R}\boldsymbol{\beta}_{0}(q) = \mathbf{0}$ under the null hypothesis (4.26),

$$\widehat{\boldsymbol{\beta}}_{0}(q) = \begin{pmatrix} \widehat{\beta}_{11}(q) & \widehat{\beta}_{21}(q) & \widehat{\beta}_{12}(q) & \widehat{\beta}_{22}(q) & \widehat{\beta}_{13}(q) & \widehat{\beta}_{23}(q) & \widehat{\beta}_{14}(q) & \widehat{\beta}_{24}(q) \end{pmatrix}', \tag{4.36}$$

$$\mathbf{C} = \left.\frac{\partial \mathbf{R}\boldsymbol{\beta}_{0}(q)}{\partial \boldsymbol{\beta}'_{0}(q)}\right|_{\widehat{\boldsymbol{\beta}}_{0}(q)} = \begin{pmatrix} 1 & -1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & -1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & -1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & -1 \end{pmatrix}, \tag{4.37}$$

and

$$\widehat{\mathbf{Var}}^{*}\left(\widehat{\boldsymbol{\beta}}_{\mathbf{0}}(\mathbf{q})\right) = \begin{pmatrix} \widehat{\sigma}^{*}_{11,11} & 0 & \widehat{\sigma}^{*}_{11,12} & 0 & \widehat{\sigma}^{*}_{11,13} & 0 & \widehat{\sigma}^{*}_{11,14} & 0 \\ 0 & \widehat{\sigma}^{*}_{21,21} & 0 & \widehat{\sigma}^{*}_{21,22} & 0 & \widehat{\sigma}^{*}_{21,23} & 0 & \widehat{\sigma}^{*}_{21,24} \\ \widehat{\sigma}^{*}_{12,11} & 0 & \widehat{\sigma}^{*}_{12,12} & 0 & \widehat{\sigma}^{*}_{12,13} & 0 & \widehat{\sigma}^{*}_{12,14} & 0 \\ 0 & \widehat{\sigma}^{*}_{22,21} & 0 & \widehat{\sigma}^{*}_{22,22} & 0 & \widehat{\sigma}^{*}_{22,23} & 0 & \widehat{\sigma}^{*}_{22,24} \\ \widehat{\sigma}^{*}_{13,11} & 0 & \widehat{\sigma}^{*}_{13,12} & 0 & \widehat{\sigma}^{*}_{13,13} & 0 & \widehat{\sigma}^{*}_{13,14} & 0 \\ 0 & \widehat{\sigma}^{*}_{23,21} & 0 & \widehat{\sigma}^{*}_{23,22} & 0 & \widehat{\sigma}^{*}_{23,23} & 0 & \widehat{\sigma}^{*}_{23,24} \\ \widehat{\sigma}^{*}_{14,11} & 0 & \widehat{\sigma}^{*}_{14,12} & 0 & \widehat{\sigma}^{*}_{14,13} & 0 & \widehat{\sigma}^{*}_{14,14} & 0 \\ 0 & \widehat{\sigma}^{*}_{24,21} & 0 & \widehat{\sigma}^{*}_{24,22} & 0 & \widehat{\sigma}^{*}_{24,23} & 0 & \widehat{\sigma}^{*}_{24,24} \end{pmatrix}_{8\times 8}, \tag{4.38}$$

with

$$\begin{aligned}\widehat{\sigma}^{*}_{ef,gh} &= \widehat{Cov}^{*}\left(\widehat{\beta}_{ef}(q), \widehat{\beta}_{gh}(q)\right) \quad (4.39)\\ &= \frac{1}{B-1}\sum_{b=1}^{B}\left(\widehat{\beta}^{*}_{efb}(q) - \frac{1}{B}\sum_{b=1}^{B}\widehat{\beta}^{*}_{efb}(q)\right)\left(\widehat{\beta}^{*}_{ghb}(q) - \frac{1}{B}\sum_{b=1}^{B}\widehat{\beta}^{*}_{ghb}(q)\right),\end{aligned}$$

where e,g denotes the group and f,h denotes the parameter. Making use of the assumption of independence and thus zero covariance between $\widehat{\beta}_{1h}(q)$ and $\widehat{\beta}_{2h}(q)$ results in $\widehat{\sigma}^{*}_{1f,2h} = \widehat{\sigma}^{*}_{2f,1h} = 0$.

Let W^{*}_{obs} denote the observed value of W^{*}. For the normal approximation tests, the null hypothesis (4.26) is rejected on significance level α if $W^{*}_{obs} > \chi^{2}_{4,\alpha}$, where $\chi^{2}_{4,\alpha}$ denotes the value of the χ^{2}-distribution with 4 degrees of freedom for which $P\left(\chi^{2}_{4} > \chi^{2}_{4,\alpha}\right) = \alpha$. This test procedure is denoted W^{*}_{n}.

For the percentile-based tests, the critical values for W^{*} are given by the order statistics of the bootstrap test statistic

$$W^{*}_{b} = \left[\mathbf{R}\widehat{\boldsymbol{\beta}}^{*}_{0b}(q) - \mathbf{R}\widehat{\boldsymbol{\beta}}_{0}(q)\right]'\left[\mathbf{C}\widehat{\mathbf{Var}}^{*}\left(\widehat{\boldsymbol{\beta}}_{\mathbf{0}}(\mathbf{q})\right)\mathbf{C}'\right]^{-1}\left[\mathbf{R}\widehat{\boldsymbol{\beta}}^{*}_{0b}(q) - \mathbf{R}\widehat{\boldsymbol{\beta}}_{0}(q)\right], \quad (4.40)$$

where

$$\widehat{\boldsymbol{\beta}}^{*}_{0b}(q) = \begin{pmatrix} \widehat{\beta}^{*}_{11b}(q) & \widehat{\beta}^{*}_{21b}(q) & \widehat{\beta}^{*}_{12b}(q) & \widehat{\beta}^{*}_{22b}(q) & \widehat{\beta}^{*}_{13b}(q) & \widehat{\beta}^{*}_{23b}(q) & \widehat{\beta}^{*}_{14b}(q) & \widehat{\beta}^{*}_{24b}(q)\end{pmatrix}'. \quad (4.41)$$

After calculating B bootstrap estimates W^{*}_{b}, let $W^{*}_{(b)}$ denote the bth order statistic of the B bootstrap replicates $W^{*}_{(1)} \leq \cdots \leq W^{*}_{(B)}$ from (4.40). The null hypothesis (4.26) is then rejected on significance level α if $W^{*}_{obs} > W^{*}_{((1-\alpha)(B+1))}$. This test procedure is denoted W^{*}_{p}.

The bias-corrected versions of (4.35) and (4.40) are given by replacing the parameter estimates $\widehat{\boldsymbol{\beta}}_{0}(q)$ with the corresponding bias-corrected parameter estimates

$$\widehat{\boldsymbol{\beta}}^{c}_{0}(q) = \begin{pmatrix} \widehat{\beta}^{c}_{11}(q) & \widehat{\beta}^{c}_{21}(q) & \widehat{\beta}^{c}_{12}(q) & \widehat{\beta}^{c}_{22}(q) & \widehat{\beta}^{c}_{13}(q) & \widehat{\beta}^{c}_{23}(q) & \widehat{\beta}^{c}_{14}(q) & \widehat{\beta}^{c}_{24}(q)\end{pmatrix}', \quad (4.42)$$

resulting in the formulas

$$W^{c*} = \left[\mathbf{R}\widehat{\boldsymbol{\beta}}^{c}_{0}(q)\right]'\left[\mathbf{C}\widehat{\mathbf{Var}}^{*}\left(\widehat{\boldsymbol{\beta}}_{\mathbf{0}}(\mathbf{q})\right)\mathbf{C}'\right]^{-1}\left[\mathbf{R}\widehat{\boldsymbol{\beta}}^{c}_{0}(q)\right] \quad (4.43)$$

and

$$W^{c*}_{b} = \left[\mathbf{R}\widehat{\boldsymbol{\beta}}^{*}_{0b}(q) - \mathbf{R}\widehat{\boldsymbol{\beta}}^{c}_{0}(q)\right]'\left[\mathbf{C}\widehat{\mathbf{Var}}^{*}\left(\widehat{\boldsymbol{\beta}}_{\mathbf{0}}(\mathbf{q})\right)\mathbf{C}'\right]^{-1}\left[\mathbf{R}\widehat{\boldsymbol{\beta}}^{*}_{0b}(q) - \mathbf{R}\widehat{\boldsymbol{\beta}}^{c}_{0}(q)\right]. \quad (4.44)$$

The rejection regions for these are obtained in an obvious way by replacing W^{*} with W^{c*} and W^{*}_{b} with W^{c*}_{b}, with the resulting test procedures denoted as W^{c*}_{n} and W^{c*}_{p}.

4.3.2 Results

The empirical significance level for a test gives the actual probability of rejecting a true null hypothesis for a data set when a particular nominal significance level is used. In the evaluation of the empirical significance level in this paper the empirical significance level of the test is examined when the nominal significance level is 5 percent, which is calculated from the simulated data set as the proportion of the simulations in which the true null hypothesis is rejected when critical values corresponding to a nominal significance level of 5 percent are used.

In applied statistics a prerequisite for a test to be useful is that it is conservative, i.e. that the empirical significance level is never larger than the nominal significance level. From the evaluation in this paper, it was found that for a single restriction, the normal approximation test Z_n^* is the only conservative test when a one-sided alternative hypothesis is used, whereas the normal approximation test Z_n^*, the symmetrical percentile test Z_{ps}^*, and the bias-corrected symmetrical percentile test Z_{ps}^{c*} are all conservative when a two-sided alternative hypothesis is used. For joint restrictions, both the Wald-like percentile test W_p^* and the bias-corrected Wald-like percentile test W_p^{c*} were found to be conservative. Moreover, it should be noted that for all these tests, they are conservative for both weights i and vi.

Given that a test is conservative, deciding which test is best overall depends on the importance that is placed on the size and power performances of the tests. If size considerations are paramount, a conservative test is preferred in which the empirical significance level is as close as possible to the nominal significance level. This closeness is measured in this paper by the estimated nominal significance level, which is calculated from a second-order polynomial regression when the empirical significance level is 5 percent and shows how large a nominal significance level has to be for a test to attain the empirical significance level of 5 percent. It was found that Z_n^* (using weight i) is the best test for a single restriction with a one- or two-sided alternative hypothesis, whereas W_p^{c*} (using weight i) is the best test for joint restrictions.

If power considerations are the most important aspects of a test, one would in applied statistics want to have power as high as possible given the nominal significance level used. In this paper this is measured by the empirical power, which gives the actual probability of rejecting a false null hypothesis for a data set, given a particular nominal significance level. Here, the nominal significance level used is 5 percent and the empirical power is calculated from the simulated data set as the proportion of the simulations where the false null hypothesis is rejected when critical values corresponding to a nominal significance level of 5 percent is used. It was found that Z_n^* (using weight i) is the best test for a single restriction with a one- or two-sided alternative hypothesis, whereas W_p^{c*} (using weight i) is the best test for joint restrictions. Note that these are the same tests and weights that are preferable when size considerations are paramount.

From a theoretical point of view, it is of interest to know how large the power of a test is if one somehow could ensure that the nominal and empirical significance levels are always equal in that this gives the intrinsic maximum possible power of a test. This estimated power, which is size corrected, is calculated with the help of a second-order polynomial regression by estimating the power when the empirical significance level is 5 percent. Here, it does not matter if the test is conservative since the focus is theoretical and the estimated power is calculated for the same value of empirical significance level for all tests. The conclusions from this paper are that, for testing a single restriction, Z_n^*

(using weight vi) is the preferred test if a one-sided alternative hypothesis is used, whereas Z^*_{ps} (using weight vi) is the preferred test if a two-sided alternative hypothesis is used. If a joint restriction is used, W^*_p (using weight vi) is the preferred test. It is worthwile noting that these tests are also conservative tests. In fact, for both single and joint restrictions, and using both one- and two-sided alternative hypotheses, all conservative tests perform better than the non-conservative tests.

It can be noted that while non-bias-corrected normal approximation tests work well for the empirical significance level of a single restriction, they fail for joint restrictions. A reason for this failure could be that the normality assumption works better for small samples when a single restriction is used. Further, the bias-corrected versions of the tests are often performing worse than the non-bias-corrected tests for the empirical significance level and the estimated power. This observation agrees with the results of Paper II, which also found that bias correction worked badly. A reason for this could be that the bias estimation that the bias correction is based on is bad.

In this paper the normal approximation test Z^*_n (using weight i) is also applied to a real data set with growth patterns of two genotypes of soybean, testing the null hypothesis of equal asymptotic average leaf weight per plant. These results demonstrate the usefulness of quantile regression and how it can be used to give a more complete picture of a data set.

In conclusion, from the evaluation of the tests in this paper, the following recommendation can be suggested in an applied situation, regardless of whether size or power considerations are the most important aspect of the test: Use Z^*_n with weight i for a single restriction with a one- or two-sided alternative hypothesis and W^{c*}_p with weight vi for joint restrictions.

4.4 Paper IV - New insights on longitudinal biostatistics data from using quantile regressions

In this paper seven longitudinal biostatistics data sets are analyzed by quantile regression methods in order to demonstrate the new insights that can be obtained regarding the properties of longitudinal data from using quantile regression methods. The quantile regression estimates are also compared and contrasted with the least squares mean regression estimates. The estimates are calculated using weight i; confidence intervals are constructed using the bootstrap-based percentile method CI_P, and hypothesis tests are performed using the normal approximation test Z^*_n for single restrictions and the percentile test W^{c*}_p for joint restrictions. The bootstrap method PB_D is used for all cases.

4.4.1 Calculation of P-values

The hypothesis tests used are the single restriction (4.23) with the two-sided alternative hypothesis (4.25) for Z^*_n and the joint restriction (4.26) with alternative hypothesis (4.27) for W^{c*}_p. These are evaluated using the P-value

$$P = \Pr\left(Z > Z^*_{obs} | H_0\right) \tag{4.45}$$

for Z^*_n and the bootstrap-based P-value

$$P^{c*} = \frac{\#\left\{W_b^{c*} > W_{obs}^{c*}\right\} + 1}{B + 1}, \tag{4.46}$$

where $\#\{A\}$ denotes the number of elements in set A, for W_p^{c*}.

4.4.2 Results

The first data set analyzed consists of data from five orange trees, originally given by Draper and Smith (1981, p. 524), in which the response variable is the trunk circumference of the trees and the model used is a three-parameter logistic growth function. A simultaneous plot of quantile and mean regression curves suggests that the data set is heteroskedastic, although the confidence intervals for the regression parameters do not give statistically significant support for this conclusion.

The second data set, given by Kwan et al. (1976), is the result of a study of the pharmacokinetics of the drug indomethacin after bolus intravenous injections in six persons. Here, the response variable is the plasma concentrations of indomethacin for the six persons and the model used is a biexponential function. A simultaneous plot of quantile and mean regression curves suggests that the decrease of indomethacin concentration for some time is faster for the lower 25 percent of the conditional distribution than for the other parts of the distribution. However, the confidence intervals obtained for the regression parameters do not give statistically significant support for this conclusion.

The third data set analyzed is obtained from a study of the heights of schoolgirls with short or tall mothers, and was reported by Goldstein (1979, p. 101). A linear model is applied to this data set. Hypothesis tests show that there is a statistically significant difference in growth between the girls with short and tall mothers for the upper 50 percent of the conditional distribution, but no difference can be found for the lower 25 percent of the distribution.

The fourth data set studied is the classical dental study data set of Potthoff and Roy (1964) in which the distance from the centre of the pituitary to the pteryomaxillary fissure was measured for a group of girls and boys. Here, it was shown how plotting lines connecting the observations for one individual together with the quantile regression lines makes it easy to follow the development of this individual over time compared with the general distribution. Hypothesis tests of equal slopes between the boys and the girls result in the conclusion that the slopes can be shown to be different only for the mean and the 10th percentile. A plot of this data set is given in 4.2.

The fifth data set, reported by Zerbe (1979), is the result of a study of the plasma inorganic phosphate levels for two groups of patients (one with obese patients and the other a control group) that were given an oral glucose challenge. This data set is analyzed with a quadratic model. The joint equality of all parameters between the two groups with control and obese patients is tested to determine whether the regression curves are the same for the two groups. The results indicate that only the mean and the 50th and 90th percentiles can be shown to have statistically significant different regression curves.

The sixth data set analyzed is obtained from a longitudinal epidemiological study of the lung function of current and former smokers, with the response variable being a measure of forced expiratory volume (FEV_1). This data set has earlier been analyzed by Fitzmaurice et al. (2004, Chapter 6.5). In contrast to the previous data sets analyzed, this data set is unbalanced since every individual is not measured on all occasions. A linear model is used in the analysis, and the conclusion is that the forced expiratory

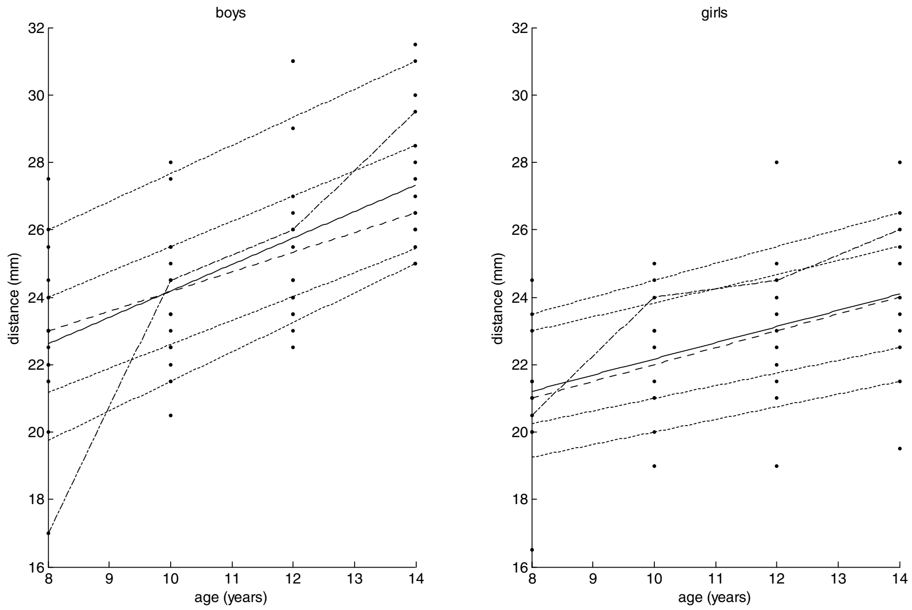

Figure 4.2: Mean and quantile regression lines for the dental study data in which a linear model is used separately for boys and girls. The mean regression line is solid, quantile regression lines ($q = 0.10, 0.25, 0.75, 0.90$) are dotted, and the median regression line ($q = 0.5$) is dashed. Dash-dotted lines connecting the four observations for one boy (individual no. 13) and one girl (individual no. 3), respectively, are also shown.

volumes for the two groups are statistically different at the entry of the study on average and for the 25th and 95th percentiles, but the change in the forced expiratory volume over time is not different between the two groups for any part of the conditional distribution tested, nor is it different on average.

The seventh and final data set analyzed is taken from a study of kidney function and consisted of 619 subjects classified into four separate groups according to whether they had kidney disease and/or hypertension. This data set is from Jones and Boadi-Boateng (1991). The response variable used is the reciprocal of serum creatinine (1/SCR), which is an indicator of kidney function. In contrast to the data sets previously considered in this paper, the measures of 1/SCR for the different subjects are not taken at the same time points, but at arbitrary times. Furthermore, the number of observations varies from one subject to another.

A linear model with separate intercepts and slopes for the four groups is used in the analyses of this data set, and the equality of the slope parameter for group 1 (with kidney disease and hypertension) against each of the other three groups is tested separately. The results show a strongly significant difference in slope between group 1 and group 2 (with kidney disease but without hypertension) for the 5th percentile, whereas the other parts of the distribution do not have significant differences. The difference between group 1 and group 3 (without kidney disease but with hypertension) as well as between group 1 and group 4 (without both kidney disease and hypertension) is shown to be statistically

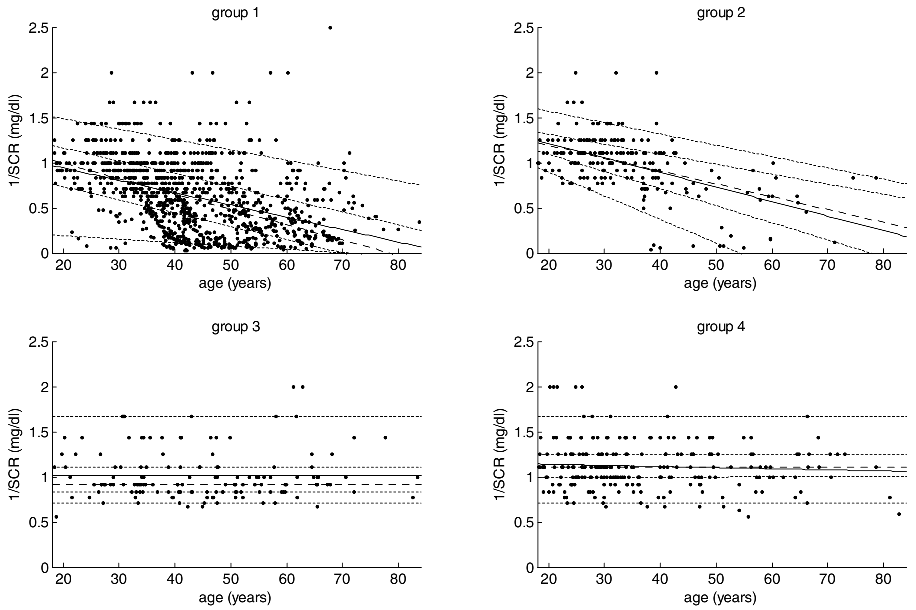

Figure 4.3: Mean and quantile regression curves for the serum creatinine reciprocals data using a linear model separately for the four groups. Mean regression curve is solid, quantile regression curves ($q = 0.05$, 0.25, 0.75, 0.95) are dotted, and median regression curve ($q = 0.5$) is dashed.

significant for the mean and for all parts of the distribution tested, except for the 5th percentile. A plot of this data set is given in 4.3.

One conclusion of this paper is that quantile regressions are very useful for visualizing changes in the conditional distributions of longitudinal data sets over time (e.g., heteroskedasticity in the data set). They are especially useful for longitudinal data sets with two or more groups, making it possible to visually compare the changes for the two groups. Hypothesis tests can find statistically significant differences between two groups for some quantile regression of a longitudinal data set, even though the mean regressions for the two groups are not significantly different. When the mean regressions are in fact significantly different, the quantile regressions can suggest which parts of the conditional distributions differ and thus make the mean regressions differ between the two groups.

5 Conclusions

The main conclusions that can be drawn from the results of the four papers regarding the weighted estimation of nonlinear quantile regression models for longitudinal data and using bootstrap-based methods for bias correction, confidence interval calculations, and hypothesis testing for this case are as follows:

- The differences between which weight is used in the weighted nonlinear quantile regression estimator are small, but weights that take the correlation in the data set into account are preferable for parameter estimation.

- If the interest extends beyond only performing parameter estimates and the researcher is interested in calculating confidence intervals and performing hypothesis tests, weight i, which does not take the correlation in the data set into account, is usually preferable.
- The nonlinear quantile regression estimator is more robust than the mean regression estimator.
- The larger the autocorrelation ρ, the better is the nonlinear quantile regression estimation working in terms of bias and MAPE, and the shorter are the confidence intervals.
- The further away from the median that the quantile value q is, the less well is the nonlinear quantile regression estimation working in terms of bias and MAPE, and the longer are the confidence intervals.
- Bias correction reduces the bias, but has the disadvantage of increasing the risk of getting crossing quantile regression curves.
- The bootstrap method PB_D is preferred over PB_I.
- For confidence interval calculation, the percentile method CI_P is preferred, but the normal approximation method CI_N could also be used if seeking a symmetric confidence interval.
- For hypothesis testing in an applied situation, regardless of whether size or power considerations are the most important aspect of the test, Z_n^* with weight i is preferred for a single restriction with a one- or two-sided alternative hypothesis and W_p^{c*} with weight vi is preferred for joint restrictions.
- Quantile regressions are very useful for visualizing changes in the conditional distributions of longitudinal data sets over time, especially when comparing two or more groups.
- Hypothesis tests can detect statistically significant differences between two groups for some quantile regression of a longitudinal data set, even though the mean regressions for the two groups are not significantly different. When the mean regressions are in fact significantly different, the quantile regressions can suggest which parts of the conditional distributions differ and thus make the mean regressions differ between the two groups.

References

BUCHINSKY, M. (1995). Estimating the asymptotic covariance matrix for quantile regression models: A Monte Carlo study. *Journal of Econometrics* **68**, 303-338.

BUCHINSKY, M. (1998). Recent Advances in Quantile Regression Models: A Practical Guideline for Empirical Research. *The Journal of Human Resources* **33**, 88-126.

CAMERON, A. C. AND TRIVEDI, P. K. (2005). *Microeconometrics. Methods and Applications.* Cambridge University Press, Cambridge.

DAVIDIAN, M. AND GILTINAN, D. M. (1995). *Nonlinear Models for Repeated Measurement Data.* Chapman & Hall, London.

DAVIS, C. S. (2002). *Statistical Methods for the analysis of Repeated Measurements.* Springer, Berlin.

DAVISON, A. C. AND HINKLEY, D. V. (1997). *Bootstrap Methods and their Application.* Cambridge University Press, Cambridge.

DIGGLE, P. J., HEAGERTY, P., LIANG, K.-Y., AND ZEGER, S. L. (2002). *Analysis of Longitudinal Data.* Oxford University Press, Oxford.

DRAPER, N. R. AND SMITH, H. (1981). *Applied Regression Analysis*, 2nd edition. Wiley, New York.

EFRON, B. (1979). Bootstrap Methods: Another Look at the Jackknife. *Annals of Statistics* **7** (1), 1-26.

FITZENBERGER, B. (1997). The moving blocks bootstrap and robust inference for linear least squares and quantile regressions. *Journal of Econometrics* **82**, 253-287.

FITZMAURICE, G. M., LAIRD, N. M., AND WARE, J. H. (2004). *Applied Longitudinal Analysis.* Wiley, New York.

GOLDSTEIN, H. (1979). *The Design and Analysis of Longitudinal Studies.* Academic Press, London.

HALL, P. AND WILSON, S. R. (1991). Two Guidelines for Bootstrap Hypothesis Testing. *Biometrics* **47**, 757-762.

HEDEKER, D. AND GIBBONS, R. D. (2006). Longitudinal Data Analysis. Wiley, New York.

JONES, R. H. AND BOADI-BOATENG, F. (1991). Unequally spaced Longitudinal Data with AR(1) Serial Correlation. *Biometrics* **47**, 161-175.

KOENKER, R. (2005). *Quantile Regression.* Cambridge University Press, Cambridge.

KOENKER, R. AND BASSETT, G. (1978). Regression quantiles. *Econometrica* **46**, 33-50.

KWAN, K. C., BREAULT, G. O., UMBENHAUER, E. R., MCMAHON, F. G., AND DUGGAN, D. E. (1976). Kinetics of Indomethacin Absorption, Elimination, and Enterohepatic Circulation in Man. *Journal of Pharmacokinetics and Biopharmaceutics* **4** (3), 255-280.

LAHIRI, S.N. (2003). *Resampling Methods for Dependent Data.* Springer, Berlin.

POTTHOFF, R. F. AND ROY, S. N. (1964). A generalized multivariate analysis of variance model useful especially for growth curve problems. *Biometrika* **54** (3/4), 313-326.

VONESH, E. F. AND CHINCHILLI, V. M. (1997). *Linear and nonlinear models for the analysis of repeated measurements.* Marcel Dekker, New York.

YU, K., LU, Z. AND STANDER, J. (2003). Quantile regression: applications and current research areas. *The Statistician* **52** (3), 331-350.

ZEGER, S. L. AND LIANG, K.-Y. (1992). An overview of methods for the analysis of longitudinal data. *Statistics in Medicine* **11**, 1825-1839.

Zerbe, G. O. (1979). Randomization Analysis of the Completely Randomized Design Extended to Growth and Response Curves. *Journal of the American Statistical Association* **74**, 215-221.

Acta Universitatis Upsaliensis

Digital Comprehensive Summaries of Uppsala Dissertations from the Faculty of Social Sciences 18

Editor: The Dean of the Faculty of Social Sciences

ACTA
UNIVERSITATIS
UPSALIENSIS
UPPSALA
2006

Distribution: publications.uu.se
urn:nbn:se:uu:diva-7186