Regression Analysis of Censored Data with Applications in Perimetry

Lindgren, Anna

1999

Regression Analysis of Censored Data with Applications in Perimetry

Anna Lindgren

Lund University

Centre for Mathematical Sciences
Mathematical Statistics

Mathematical Statistics
Centre for Mathematical Sciences
Lund University
Box 118
SE-221 00 Lund
Sweden

Doctoral Theses in Mathematical Sciences 1999:1
ISSN 1404-0034

ISBN 91-628-3499-1
LUNFMS-1008-1999

Printed in Sweden by KFS AB
Lund 1999

Contents

List of papers

This thesis is based on five papers referred to in the text by the capital letters A, B, C, D, and E.

A. Lindgren, A.: Quantile regression with censored data using generalized L_1 minimization. *Computational Statistics and Data Analysis*, 23 (1997) 509–524.

B. Lindgren, A.: L_2 regression with censored covariate. Manuscript.

C. Heijl, A., Lindgren, G., Lindgren, A., Olsson, J., Åsman, P., Myers, S., and Patella, M.: Extended empirical statistical package for evaluation of single and multiple fields in glaucoma: Statpac 2. *Perimetry Update* 1990/91, 303–315.

D. Heijl, A., Lindgren, A., Lindgren, G., and Patella, M.: Inter-test threshold variability in glaucoma. Importance of censored observations and general field status. *Perimetry Update* 1990/91, 189–192.

E. Bengtsson, B., Lindgren, A., Heijl, A., Lindgren, G., Åsman, P., and Patella, M.: Perimetric probability maps to separate change caused by glaucoma from that caused by cataract. *Acta Ophthalmologica Scandinavica* 1997:75:184–188.

Preface

When I was a child, my siblings and I used to spend our summers measuring the height of the surf at the beach of Kivik and Vitemölle using two measuring-tapes nailed end-to-end on a board. This caused some consternation among the other holiday-makers who, not wanting to appear inquisitive, instead sent their children to ask us what on earth we were doing. It also puzzled the shop-keeper when my brother, having already bought one measuring-tape, came back a little later to buy one more because "the things he was measuring where longer than 1.5 meters". It should perhaps be noted that our statistician father, who initiated the wave-measuring, is still interested in waves but now gets his measurements by satellite, not measuring-tape, while two of his four children have grown up to become statisticians ourselves, although neither of us is particularly interested in waves.

When I grew up and started reading mathematics at the university I ran into professor Gunnar Blom who taught both the basic and the secondary course in probability in an inspiring way, confirming my suspicions that mathematical statistics was actually quite fun.

I am also indebted to my supervisor, professor Jan Lanke, for valuable suggestions and language corrections, to professor Anders Heijl at the department of Ophthalmology in Malmö for a long cooperation and for supplying the 51 "sick normals", and to Mike Patella at Humphrey Instruments for support, discussions, and patience over the years. Further, to the people at the Department of Community Medicine in Malmö, particularly the division for Health Economics and Biostatistics, for presenting me with a number of intriguing statistical questions.

Finally, I would like to thank Anders, Anders, Anders, Bengt, Fredrik, Hans, Henrik, Lars, Magnus, and Mikael for enlivening work with extended lunches, coffee-breaks, bruises, and feuds over whether applications are really necessary, my brother Finn for showing me how to make my Matlab programs run faster, and anyone else at the department with a reasonably fast computer for not complaining too much when I've kept my simulations running on their machines for weeks on end.

1 Overview

In this thesis we examine some problems arising in regression analysis when data are censored. We will first, in Section 2, give an introduction to some variants of linear regression, in particular quantile regression, and introduce the concept of censoring in this context. Then, in Section 3, we introduce the medical problem that spurred this research, namely determination of normal variability limits for glaucomatous visual fields, and in Section 4 we present a more detailed statistical discussion than is found in the applied Papers C, D, and E. Finally, we suggest some topics for further research in this area.

Paper A proposes a distribution-free parametric solution to the problem of estimating a quantile function when the dependent variable is censored. Paper B adresses the problem of least squares estimation of a linear function when the covariate is censored. A solution is proposed and some variants of it are examined. Papers C, D, and E are all devoted to the glaucoma problem and the practical use of the quantile regression method presented in Paper A. Paper C describes the statistical package into which the variability limits have been incorporated, Paper D demonstrates the importance of taking the censoring into account when examining the variability, and Paper E, finally, describes a more specific use of the variability limits in practice concerning patients with both glaucoma and cataract.

2 Regression and censoring

2.1 Regression

2.1.1 A brief history

When studying the relationship between variables one is often interested in expressing one of the variables as a function of the others. A simple example is the linear relationship $y = \alpha + \beta x + \varepsilon$, where the dependent variable y is described as a linear function of the independent variable, or covariate, x, plus some error ε. The problem then lies in finding the values of the parameters α and β that describe this relationship using n observed data pairs $(x_i,\, y_i)$ where $y_i = \alpha + \beta x_i + \varepsilon_i$.

The errors $\varepsilon_i = y_i - \alpha - \beta x_i$, i.e. the points' deviations from the line, are unknown since the line itself is unknown; we want to find the line that makes these deviations as small as possible. There are several different ways of defining "as small as possible", i.a. the classical method of minimizing the sum of squared deviations, but several other methods have been, and still are, in use. For the fascinating story of the early approaches to regression see Chapter 6 of (Hald, 1998), which has been an important source for the following account.

The earliest method, in use before 1750, consisted of choosing the necessary number of points, two in the case of simple linear regression, and fitting a line between them. The result is, of course, highly dependent on the choice of points and the definition of

"as small as possible" is left entirely to the discretion of the person doing the estimation. In fact, it seems that the general rule was to take only as many measurements as was necessary in order to calculate the estimate.

The method of averages; see (Mayer, 1750). The German cartographer and astronomer Tobias Mayer (1723–1762) found the arbitrariness of the old method unsatisfactory and proposed the following method instead: group the n equations $y_i = \alpha + \beta x_i$, into two groups (the number of unknown parameters), take the average of the equations within each group, and then solve the resulting system of linear equations. There is still some subjectivity left in the grouping of the points but less so than when only choosing certain points. The groups should be chosen with care in order to achieve as much variability as possible, and different choices lead to different estimates. Since the method is computationally easy it was widely used even after the least squares method was invented. In fact, it is still in use today in some medical applications where it is not uncommon to dichotomize a covariate, at the median (which would have been Mayer's choice) or a tertile, before estimating its influence on the dependent variable.

The method of least absolute deviations; see (Boscovich, 1757): Roger Joseph Boscovich (1711–1787), also unsatisfied with the subjectivity of the available methods, instead proposed the use of an estimate satisfying the following conditions: $\sum_{i=1}^{n} \varepsilon_i = 0$, which means that the line should pass through the central mass of the data, i.e. $\bar{y} = \alpha + \beta\bar{x}$, and the estimates of α and β should be chosen to minimize the sum of absolute deviations, $\sum_{i=1}^{n} |\varepsilon_i|$. This method is robust with respect to outliers and does not leave any room for subjective choices on the part of the practitioner. However, since the method is computationally difficult, lacking a closed form for the estimates, it was more or less forgotten for two hundred years until in the 1950s the growing use of computers and the development of programming methods managed to produce a simple solution algorithm. By then the restriction $\sum_{i=1}^{n} \varepsilon_i = 0$ had been dropped. The asymptotic theory, on the other hand, did not appear until fairly recently with (Koenker & Basset Jr., 1978).

The idea of using absolute deviations as a rule for fitting data was not new in 1757, although it had not been used in regression situations. It is described as early as 1632 when Galileo Galilei (1564–1642) used it to determine whether a new star appearing in 1572 was sub-lunar or a fixed star, i.e. how far away it lay; see (Galilei, 1632).

Minimizing the largest absolute residual; see e.g. (Laplace, 1805): Laplace, who also generalized and modified the methods of Mayer and Boscovich, proposed another estimate defined as follows: the estimates of α and β should be chosen to minimize the largest absolute deviation, i.e. minimize $\max_i |\varepsilon_i|$. This method is extremely sensitive to outliers and Laplace used it more as a tool for testing hypotheses than as a method for estimating relationships.

Method of least squares; see (Legendre, 1805): Adrien-Marie Legendre (1752–1833), unhappy with the arbitrariness of the method of averages and the computational difficulties of the method of least absolute deviations, described the following method:

the estimates of α and β should be chosen to minimize the sum of squared deviations, $\sum_{i=1}^{n} \varepsilon_i^2$. This automatically requires the line to pass through the gravitational centre of the data, the same restriction as Boscovich imposed in the least absolute deviation method. The least squares estimates can be given in closed form, and the method is computationally easy. This, along with the probabilistic foundation developed by Gauss and Laplace 1809–1828, leading to well-known properties for the model where the errors follow a Gaussian distribution, has caused the least squares method to become the standard estimation procedure.

One problem with the least squares method is that it is sensitive to outliers and a number of robust regression methods have been developed to take care of this. In addition, the maximum likelihood method, first used in (Lambert, 1760) and thoroughly investigated by Fisher in the 1920s, allows the specification of any parametric relationship, as long as the type of the distribution of the measurements is known. In fact, for a linear model with Gaussian errors the maximum likelihood estimate is the same as the least squares estimate, and when the errors follow a Laplace distribution the maximum likelihood estimate is the same as the least absolute deviations estimate. With the advent of computational power there has also been a development of various non-parametric methods.

2.1.2 Quantiles

The least squares method, L_2, estimates the conditional mean of the dependent variable, y, as a function of the independent ones, x, whereas the least absolute deviations method (without any restriction on $\sum_{i=1}^{n} \varepsilon_i$), L_1, estimates the conditional median of y as a function of x. While these two entities are the same when the distribution of the errors is symmetrical they may differ if the distribution is skewed.

Sometimes it is more appropriate to look at a conditional quantile of y as a function of x. For example, this is the case when we want to derive limits for the natural variation of y for different values of x, such as limits for the intra-uterinal foetus growth as functions of gestational week, telling us what can be considered an unusually small or large foetus. Another example, examined in this thesis, is the variability of the light sensitivity in stable glaucomatous eyes from one week to the next. The purpose of the regression is then not to determine what the average behaviour is, i.e. finding the mean or median, but rather where lies the limit for what is unusual behaviour, e.g. finding the lower 5% quantile; see Figure 1. The 5% quantile is defined here as the value below which 5% of the data will fall.

If we know the error distribution we can estimate the quantile function by estimating the usual mean, by least squares, and then use our knowledge of the distribution to derive the quantiles. For instance, when the errors are normally distributed we know that 5% of the data should lie below the mean minus 1.64 standard deviations, that is, the quantile function can be calculated as $u_{0.05}(x) = \alpha + \beta x - 1.64\sigma$ where $\alpha + \beta x$ is the ordinary regression line and σ is the standard deviation of the errors. However, this method is highly

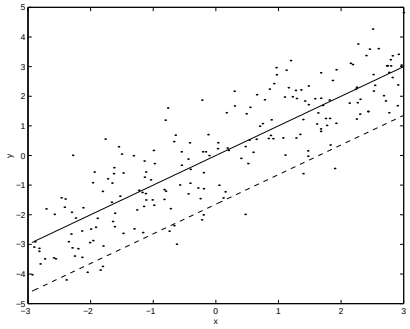

Figure 1: Regression of the mean (solid) and regression of the lower 5% quantile (dashed)

sensitive to mis-specification of the error distribution; if the true distribution has heavier tails than the one we use in the estimation we will underestimate the distance between the quantile and the mean, and if the true distribution is skewed we will underestimate the distance at one end of the distribution and overestimate it at the other.

An advantage of the L_1 method is that it can easily be modified to estimate any quantile function, and not just the median; see (Koenker & Basset Jr., 1978). Instead of minimizing the sum of the absolute deviations we minimize the asymmetrically weighted sum $\sum_{i=1}^{n} |\varepsilon_i|_p$ where

$$|\varepsilon|_p = \begin{cases} (1-p)|\varepsilon| & \text{if } \varepsilon \leq 0, \\ p|\varepsilon| & \text{if } \varepsilon > 0. \end{cases}$$

This means that when we are estimating the 5% quantile and thus want $p = 5\%$ of the data points to lie below the line and the other 95% of the points above the line, we should weigh those few points that fall under the line by $1 - p = 0.95$ and the many above by $p = 0.05$ in order to have the points below the line weigh as much as those above.

2.2 Censoring

In some practical situations it is not possible to measure the value of a variable if this value lies above or below a certain level. For instance, the amount of a substance in the blood may be below the minimum detection level of the measuring instrument, or a patient may still be alive when the study of post-surgical mortality he is part of is closed two years after he underwent surgery. The value is then said to be *censored*: left-censored in the case of falling below the minimum detection level, and right-censored in the case of exceeding the two-year follow-up period. In these two examples all censored values are censored at the same level. In other cases the censoring limits may be different for different subjects, e.g. patients may move abroad before the study is closed, in which case we know that they

were alive when whey moved but not whether they were alive or dead when the study was closed.

When we have censored data we need to adjust our estimation procedures to allow for the fact that some observations should have been smaller (left-censoring) or larger (right-censoring) than what was actually observed. If we disregard the fact that some y-values are left-censored our estimated line will be too low; see Figure 2 and Paper D.

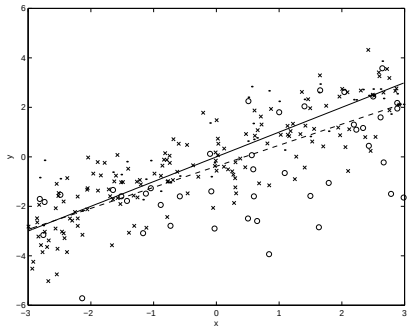

Figure 2: Regression ignoring the fact that the dependent variable, y is censored: true line (solid), estimated line (dashed), uncensored data ($\times$), left-censored data ($\circ$), true value of censored data ($\cdot$)

When the dependent variable is non-negative and right-censored there exist a number of techniques based on estimation of the hazard function under various assumptions; see Chapter VII of (Andersen, Borgan, Gill & Keiding, 1993). One popular method is Cox's proportional hazards model, which assumes that the risk, e.g. of dying, changes proportionally with increasing covariate value, e.g. the patient's age or the severity of the disease, while the baseline hazard itself is allowed to vary freely over time.

In this thesis we are interested in estimating, not the risk of dying as a function of the covariates, but rather the lower and upper 5% limits of the survival times, as functions of the covariates. This calls for a different estimation strategy, based on the ordinary quantile regression methods. There are a number of non-parametric techniques (see, e.g. (Dabrowska, 1992)) but we will pursue a parametric one.

When it is the covariate that is censored we have two choices: either throw away the censored data and use the remainder to estimate means or quantiles in the ordinary way, or use what information there is in the censored data to help improve our estimates. As in the case of censoring in the dependent variable, using the censored values as if they were uncensored will cause the estimates to be biased. Throwing away the censored data points will not cause any systematic errors of the estimates as long as the censoring limits are independent of the actual values. However, if we are forced to throw away a lot of data our estimates will become more uncertain. For the censored data points we have full knowledge of the value of the dependent variable, and we know that the true covariate

value is at least as large as the observed, censored one. It is to be hoped that the use of this information could improve the estimates and decrease the uncertainty; this is the topic of Paper B.

An even more complicated situation occurs when both the dependent and the independent variable are censored. We then have four types of points: uncensored, censored in x but not in y, censored in y but not in x, and censored in both x and y. The solution to this problem will have to wait until later; see the suggestions for further research in Section 5.

3 Perimetry

3.1 Introduction

Quantitative examination of the visual field (perimetry) is used as a tool for diagnosis and follow-up of many diseases affecting the eye and the central nervous system. Typical changes in the visual field occur, e.g. in glaucoma, where the optic nerve fibres in the retina are affected and the visual ability deteriorates in growing areas of the visual field. Computerized perimetry is an important tool in early diagnosis of glaucoma and supervision of the development of the disease.

There are a number of automatic, computerized perimeters in use. Most of them can determine the *threshold value* of the ability to detect light stimuli, for a large number of points in the visual field. By threshold value is usually meant the light intensity needed for the patient to detect, with 50% probability, a stimulus of a certain size and exposure time against an evenly lighted background of constant luminosity.

In this thesis we study the results of a perimetry study performed with the instrument Humphrey Field Analyzer (HFA), developed by Humphrey Instruments, California, in cooperation with professor Anders Heijl at the Department of Ophthalmology at Malmö University Hospital; see (Haley, 1986) and (Heijl, 1985). A statistical package, Statpac, and its extension, Statpac 2, developed in cooperation between Humphrey Instruments, the Department of Ophthalmology and the Department of Mathematical Statistics, Lund University, is included in this instrument; see (Heijl, Lindgren & Olsson, 1987c), (1986, Statpac User's Guide), and (Heijl, Lindgren, Lindgren, Olsson, Åsman, Myers & Patella, 1987a).

This statistical package calculates different indices (regarding i.a. the level of sensitivity and variation), which can be used for classifying a visual field as "normal" or "diseased".

3.2 Glaucoma and cataract

Glaucoma affects the axons in the retinal nerve fibre layer, which causes visual field defects, i.e. areas of the retina with systematically deteriorated ability to detect light stimuli.

One problem is to separate actual defects from the variation that always occurs among visual fields.

Detected visual defects often have a large diagnostic value and one can usually not be certain that a patient has glaucoma until such defects have occurred. This is because the visible changes in the retina are often too unspecific to allow a diagnosis with any certainty. Furthermore, many patients with suspected glaucoma, e.g. with increased eye pressure, never develop glaucoma. Hence, a patient with increased pressure is often not treated until glaucoma has been verified. The treatment is dependent on the result of the perimetry.

Glaucoma is more common in elderly people who also tend to have more cataracts. Cataracts tend to cause a homogeneous loss of visual ability across the entire visual field, rather than the localized losses often seen in connection with glaucoma. When monitoring the progress of the glaucoma in a patient suffering from both glaucoma and cataract, it is thus desirable to be able to separate these two different types of changes.

3.3 Visual field examinations

An examination of the visual field using automatic, static, perimetry is performed as follows: luminous points are shown, one at a time, on a white hemisphere, and the patient presses a button if he/she sees the point in question. The light intensity of the points is changed in steps (increased or decreased) in order to find the value for the light intensity where the patient sees the points with probability 0.5, the so called threshold value.

The light intensity is measured in apostilbs (asb) or in logarithmic units in decibels (dB); see Figure 3.

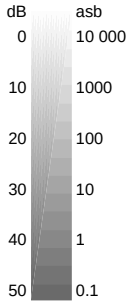

Figure 3: Relationship between apostilbs and decibels

When determining the light sensitivity of a point one starts by showing light of an intensity 2 dB stronger than what the patient is expected to see in that point. If the patient does not see the point the intensity is increased in steps of 4 dB until the patient notices

the stimulus. Then the light is decreased in 2 dB steps until the patient again does not see the point. The threshold value of the point is defined as as that value of the light intensity that the patient last saw.

Within the central part of the visual field (30° from the centre) one measures, in a partly random order, 76 points. Two of these lie in the physiological blind spot where the patient normally does not see anything; see Figure 4.

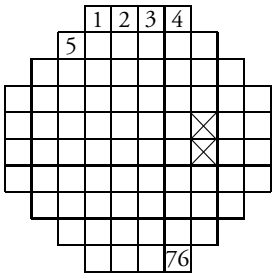

Figure 4: The 76 points of the central visual field of the right eye, including the blind spot (×)

3.4 Normal visual fields

A normal visual field has a certain shape, so that the sensitivity (visual ability) is highest in the middle (in fovea, the yellow spot), and decreases toward the edges, and a certain height, i.e. a certain overall sensitivity. The visual ability, measured in dB, decreases linearly with age in all points. The decrease is larger in some points, e.g. in the periphery of the visual field, than in others.

Hence, what a patient should be able to see can be calculated using *normal values*, adjusting for the age of the specific patient. Normal visual fields and age correction for the HFA have been calculated using a large material consisting of 487 normal visual fields from 239 randomly chosen people; see (Heijl, Lindgren & Olsson, 1987b).

Since the reliability of patients varies, e.g. their tendency to move their eye away from the fixation point or to press the button despite their not seeing anything, it is necessary to have some measurement of the reliability of the examination in order to be able to exclude unreliable results. In the perimeter used three different reliability tests are performed, namely False Negative answers (FN), False Positive answers (FP) and Fixation Losses (FL). FN, FP and FL were used to discard unreliable visual fields when the normal values were calculated.

3.5 Visual field indices

The intention behind visual field examinations is to indicate whether the patient has a normal visual field or not, or to determine whether detected visual field defects are im-

proving or deteriorating. In a normal visual field the threshold value in each point has a certain variability, both between measurements and between individuals. The normal values for different ages, as well as the variation, have been calculated using the aforementioned normal material.

A patient with glaucoma has a lower visual ability than normal and has a larger variation across the visual field. If x_i is the threshold value at point i and N_i is the age dependent normal value at point i then $y_i = x_i - N_i$ is the deviation from normal at point i, also called the *total deviation*.

For the study of point-wise test-retest variability it appears to be best to look at the so called *pattern deviation*, rather than the total deviation; see (Bengtsson, Lindgren, Heijl, Lindgren, Åsman & Patella, 1997). The pattern deviation is calculated by subtracting the seventh largest observed threshold value of the 45 most central points from the total deviations. The reason for this is to avoid influence of cataracts which affect the entire visual field.

The general status of the field is measured by the weighted Mean Deviation (MD) of the total deviations calculated as follows (see (Heijl et al., 1987b)):

$$\text{MD} = \frac{\sum_{i=1}^{n} y_i/\delta_i^2}{\sum_{i=1}^{n} 1/\delta_i^2}$$

where $n = 73$ is the number of points in the visual field used (since three points in or close to the blind spot are excluded in the analysis), and δ_i^2 is the normal variation at point i.

3.6 Censoring

Due to the limited light range of the perimeter used for the testing, some threshold values, and thus the corresponding pattern deviations, are left-censored, i.e. the true deviation exceeds the registered one by an unknown amount. In this case the censoring limit is 10 000 asb or 0 dB, which means that the threshold values are all censored at 0 while the pattern deviations are all censored at different levels determined by point location, patient age, and the seventh largest threshold value in the central part of the field. This is the practical situation that inspired the development of the statistical methods in this thesis.

3.7 Variability of glacomatous eyes

One important use of perimetry on glaucoma patients is monitoring the progress of the disease and determining if the patient is improving or deteriorating over time. To this end it is important to know how large the normal variation is for stable glaucomatous visual fields. Due to the nature of the disease this variation is quite large and it is not unusual for a patient to have an almost normal visual ability in some points of the eye one week and then to be almost blind there the next, and vice versa. In order to determine whether a patient really has deteriorated it is therefore necessary to take this variability into account.

In order to be able to use perimetry for following up the progress of glaucomatous eyes it is necessary to produce limits for the normal visual ability in an individual point in an eye, based on the results of previous examinations of that eye. The variation in light sensitivity at a specific point in an eye with glaucoma depends on a large number of covariates. The most important of these are the mean defect of the eye and the defect depth at the studied point. To find the limits for the normal variability is thus a typical quantile regression problem.

The problem is complicated since the data are affected by censoring in both the dependent variable and in one of the independent ones or, strictly speaking, in both of the independent variables since the mean deviation is itself a weighted mean of censored values. We start to solve this problem by first solving the simpler problem of quantile regression with censoring in the dependent variable only (see Papers A, C, D, and E), and then proceeding to regression when only the covariate is censored (see Paper B). The complete problem, with censoring in both the dependent and the independent variables, is outside the scope of this thesis.

4 Test-retest variability limits

This section gives the details of the limit calculations performed in Papers C, D, and E. The technical details of the method are described in Paper A.

4.1 Data description

We examined 51 glaucomatous eyes from 51 patients tested four times over a one-month period. The eyes ranged from severely ill to almost normal. The light sensitivity of the eye was measured at 74 different points (blind spot excluded) over the eye retina, giving 74 threshold values, i.e. the light that has probability 0.5 to be seen, for each eye on each test occasion.

The threshold values were corrected for the patient's age, giving the total deviation at a certain point as the difference between the threshold value at that point and the normal value for a point at that location, for a patient of that age. The total deviations where in turn corrected for cataracts giving the pattern deviations for each point. The mean deviation, MD, was calculated as a weighted mean of the total deviations, giving a measure of the general status of the eye.

4.2 Problem

We are interested in finding the 5% and 95% quantile limits for the pattern deviation at each point at a follow-up test. The limits should be functions of the observed pattern deviation at the point at one or more baseline tests. Further, the limits should take into

account the overall field status and thus depend on the mean deviation at the baseline test, or tests.

Since the test-retest variability differs over the field, being smaller in the center and larger on the periphery, the points were divided into three zones: inner, middle, and outer; see Figure 5. Limits were calculated separately for each zone, using all points in that zone.

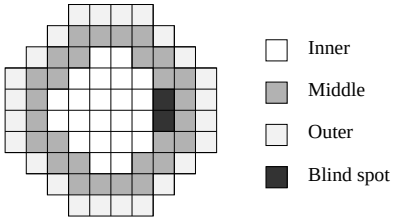

Figure 5: The visual field divided into three eccentricity zones

With this procedure we regard neighbouring points from the same eye as independent, given eye status. This is not actually the case since, due to the nature of the disease, visual loss often occurs in bands along the nerve fibres; see (Heijl, Lindgren & Lindgren, 1988).

We will only deal with the censoring at the follow-up, not at the baseline tests, regarding the censored baseline values as valid observations. We will also ignore the fact that the mean deviation is a weighted average of possibly censored total deviations, and thus itself censored. To get right-censored data rather than left-censored, we multiplied the follow-up pattern deviations by minus one in the calculations. This also gives us (mostly) positive data.

4.3 Solution

Variability limits are designed to contain the normal variation of data, and in the simplest case they can therefore be based on the empirically observed variation in a homogeneous population. In our case, the limits depend on covariates in a non-homogeneous population, and then a more elaborate method needs to be used to calculate the limits.

A standard way to estimate quantiles in such a regression situation is to find the value that minimizes a certain sum of weighted absolute deviations; see (Koenker & Basset Jr., 1978). Due to the censoring, we have not observed the true pattern deviations but rather the minimum of the true pattern deviations and the corresponding censoring limits, and then the standard way does not work. If the censoring limits are independent of the true deviations, we can rewrite the problem of finding the p-quantile function of the true deviations into one of finding a q-quantile function of the observed, censored, deviations, where q is a function of the distribution of the censoring limits; see Paper A for details.

4.3.1 Notation

Let x_{1i} be the mean deviation (MD) at one baseline test or the average of the mean deviations at two or more baseline tests for eye i and let x_{2ij} be the baseline pattern deviation for eye i and point j (a single test or the average of several). Further, let y_{ij} be the reversed pattern deviation at the follow-up for eye i and point j, i.e. y_{ij} is equal to minus the pattern deviation; this reversal is done in order to have right-censored data rather than left-censored. Also, let $u_p(x_{1i}, x_{2ij};\, \theta)$ be the p-quantile limit that we want to estimate. The quantile function, $u_p(x_1, x_2;\, \theta)$, is a parametric function, chosen so that its shape can model that of the data (see Figure 6), while letting the ratio p of the true, uncensored, data fall below it.

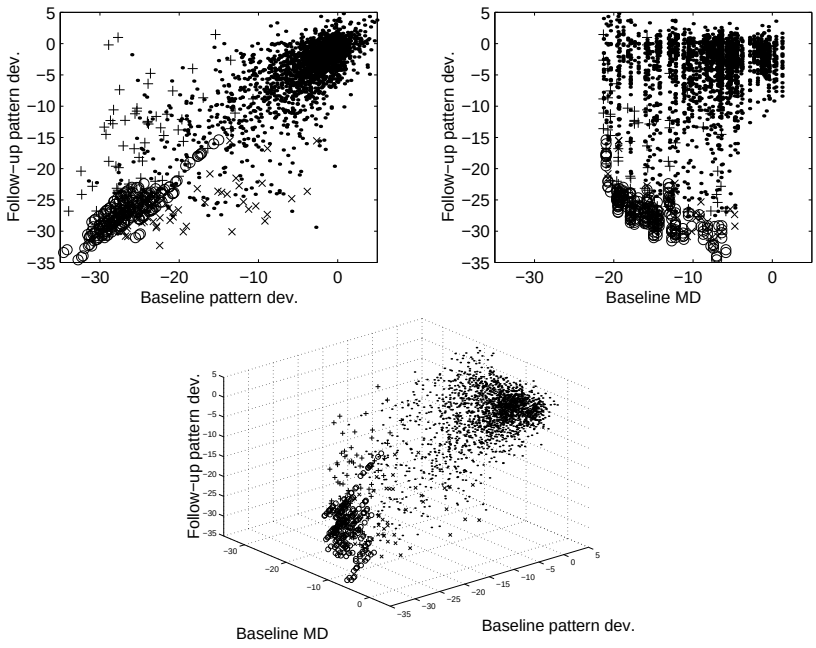

Figure 6: Data in the inner zone: fully observed (·), follow-up censored (×), baseline censored (+), both baseline and follow-up censored (○)

Due to the censoring, an observed pattern deviation (with reversed sign), y_{ij}, is the minimum of the (reversed) true pattern deviation, t_{ij}, i.e. the one we would have obtained if we had had an infinite light-range, and a corresponding censoring limit, z_{ij}, i.e. the largest possible measurement at that point, for that eye. The censoring limits can be regarded as random with the distribution function $F_Z(t;\, x_1, x_2)$.

4.3.2 Minimization

If we assume that the true pattern deviations and the censoring limits are independent, which is reasonable enough, the distribution function for the observed, reversed, follow-up deviations, given the baseline values, $F_Y(t;\, x_1, x_2)$, is a simple function of the distribution functions for the true deviations, $F_T(t;\, x_1, x_2)$, and the censoring limits:

$$F_Y(t;\, x_1, x_2) = 1 - (1 - F_T(t;\, x_1, x_2))(1 - F_Z(t;\, x_1, x_2)).$$

All these distributions are unknown.

Our goal is to calculate a limit for the true pattern deviations, T. Since we are fitting parametric functions this involves estimating an unknown parameter vector, θ, and then finding the limits as

$$u_p(x_1, x_2;\, \theta) = \inf\{t;\, F_T(t;\, x_1, x_2) \geq p\}$$

for a fixed value p. This means that for each combination of baseline values the follow-up limit should be the smallest value that we have probability at least p to fall below. Since the distributions are continuous in this case we simply want $u_p(x_1, x_2;\, \theta)$ to fulfil $F_T(u_p(x_1, x_2;\, \theta);\, x_1, x_2) = p$. Due to the independence of T and Z we can estimate $u_p(x_1, x_2;\, \theta)$ as the $q(x_1, x_2;\, \theta)$-quantile function for Y instead of as the p-quantile function of T, where

$$q(x_1, x_2;\, \theta) = 1 - (1 - p)(1 - F_Z(u_p(x_1, x_2,\, \theta);\, x_1, x_2)).$$

This involves estimating the unknown distribution function of the censoring limits, e.g. using a local Kaplan-Meier function in the neighbourhood of x_1 and x_2. The neighbourhood of a point can be defined in many different ways, e.g. as all points within a circle with a fixed radius.

The regression parameter θ can be estimated iteratively with the estimate at the mth step calculated as

$$\theta^{(m)} = \arg\min_{\theta} \sum_{ij} |y_{ij} - u_p(x_{1i}, x_{2ij};\, \theta)|_{\hat{q}_{ij}(\theta^{(m-1)})}$$

where

$$|r|_p = \begin{cases} |r| \cdot (1 - p) & \text{if } r \leq 0, \\ |r| \cdot p & \text{if } r > 0 \end{cases}$$

is the asymmetric weight function and

$$\hat{q}_{ij}(\theta) = 1 - (1 - p)(1 - \widehat{F}_Z(u_p(x_{1i}, x_{2ij};\, \theta);\, x_{1i}, x_{2ij}))$$

is the estimated quantile value. We have used the MATLAB-function `fmins` to calculate the iterative estimates. The p-values will be 0.05 and 0.95 to give 5% limits for significant improvement and deterioration, respectively.

4.4 Results

As can be seen in Figures 7 and 8, the width of the limits increases with decreasing visual ability; a moderately depressed point varies more than a point with normal visual ability.

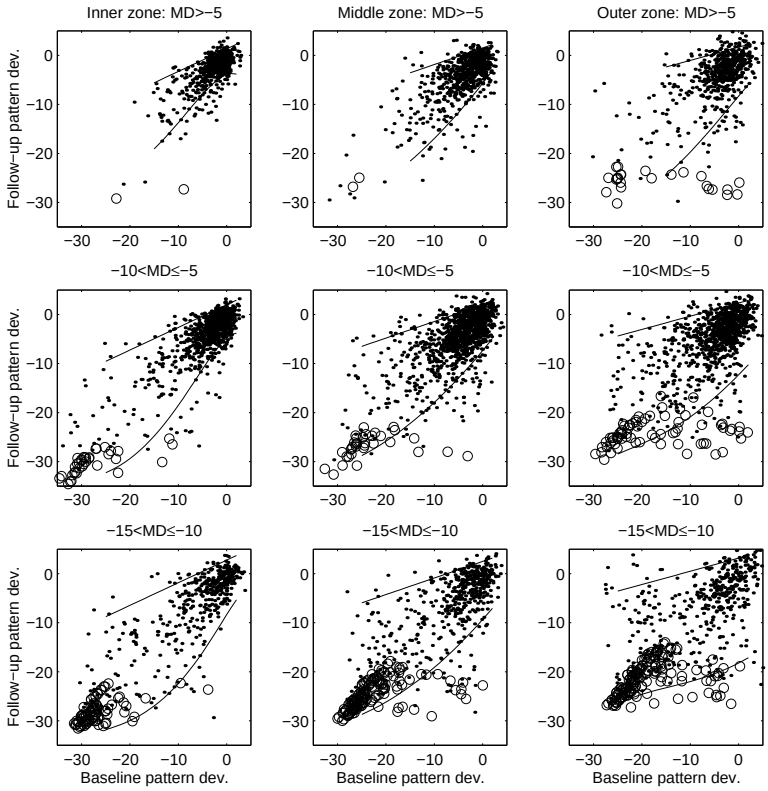

Figure 7: Normal 5% and 95% variability limits for the three eccentricity zones, using one baseline test; uncensored point (·), censored point (○)

Further, the limits become wider with growing peripheriсity so that a point in the outer eccentricity zone varies more than a point with the same pattern deviation but located in the inner zone. Also, the limits grow wider with decreasing overall status (MD) of the eye so that we expect a healthy point in an otherwise moderately depressed field to vary more than a healthy point in a healthy field. We can also note that when we use the average of two baseline tests the limits become slightly more narrow. This should be expected since we are then able to determine the baseline values more accurately, and so we have a better estimate of the actual light sensitivity of the individual points.

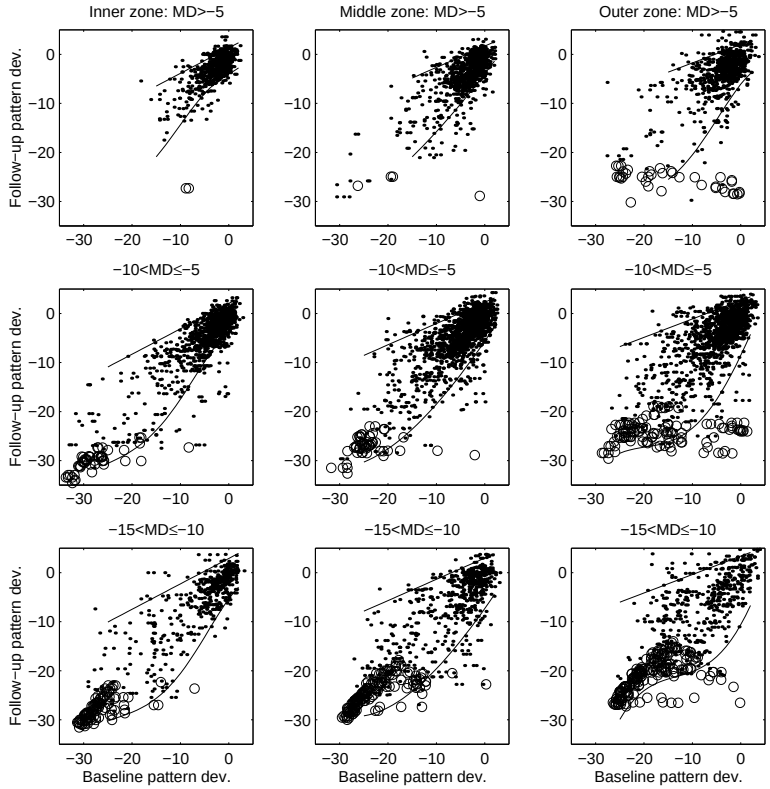

Figure 8: Normal 5% and 95% variability limits for the three eccentricity zones, using the average of two baseline tests; uncensored point (·), censored point (○)

5 Suggestions for further research

We started out with the problem of estimating quantile functions when both the dependent and the independent variables were censored, and so far we have adressed some of the steps toward a solution to this problem. However, there still remain quite a few things to be done. The method of L_2-regression with a censored covariate presented in Paper B needs to be tested on real data; we cannot use the glaucoma data since it has censoring in y as well as in x. We also need to examine the behaviour of the σ^2-estimates closer and the method should be generalized to handle several covariates, in particular a polynomial in x.

Further, we should try the Redistribution technique on quantile regression rather than L_2 regression. This should present no difficulty when using the parametric method since if we know the distribution, we also know how the quantiles are expressed in terms of its parameters and all we need to do is estimate those parameters. The semi-parametric method, on the other hand, relies on the regression of the conditional mean of Y given X when redistributing the censored x-values, and so it might not be directly adaptable to handling quantiles. It could, however, be used as a preliminary step to determine the appropriate redistribution locations of the censored points, which could in turn, as a second step, be used to estimate the quantile function. The non-parametric metod seems to be easy to adapt to estimating quantiles instead of means since it does not rely on any estimates of the mean of Y given X.

Finally, the methods for handling censoring in y and for handling censoring in x should be combined to handle censoring in both x and y simultaneously.

References

Andersen, P. K., Borgan, Ø., Gill, R. D. & Keiding, N. (1993). *Statistical Models Based on Counting Processes*. Springer-Verlag.

Bengtsson, B., Lindgren, A., Heijl, A., Lindgren, G., Åsman, P. & Patella, M. (1997). Perimetric probability maps to separate change caused by glaucoma from that caused by cataract. *Acta Ophthalmologica Scandinavica*, **75**, 184–188.

Boscovich, R. (1757). De litteraria expeditione per pontificiam ditionem, et synopsis amplioris operis. *Bononiesi Scientiarum et Artum Instituto Atque Academia Commentarii*, **4**, 353–396.

Dabrowska, D. M. (1992). Non-parametric quantile regression with censored data. *Sankhyā*, **54**(Series A), 252–259.

Galilei, G. (1632). Dialogo sopra i due massimi sistemi del mondo, Tolemaico, e Copernicano.

Hald, A. (1998). *A History of Mathematical Statistics. From 1750 to 1930*. John Wiley and Sons, Inc.

Haley, M. J. (1986). *The Field Analyzer Primer*. Allergan Humphrey, Los Angeles.

Heijl, A., Lindgren, G., Lindgren, A., Olsson, J., Åsman, P., Myers, S. & Patella, M. (1987a). Extended empirical statistical package for evaluation of single and multiple fields in glaucoma: Statpac 2. In *Proceedings of the 9th International Parimetric Society Meeting, Malmö, Sweden*, pp. 303–315. Kugler, Amsterdam.

Heijl, A., Lindgren, G. & Olsson, J. (1987b). Normal variability of static threshold values across the central field. *Archives of Ophthalmology*, **105**, 1544–1549.

Heijl, A., Lindgren, G. & Olsson, J. (1987c). A package for the statistical analysis of computerized fields. *Documenta Ophthalmologica Proc. Sec.*, 153–168.

Heijl, A., Lindgren, A. & Lindgren, G. (1988). Inter-point correlations of deviation of threshold values in normal and glaucomatous visual fields. In Heijl, A. (Ed.), *Perimetry Update*, pp. 177–183. Kugler & Ghedini, Amsterdam.

Heijl, A. (1985). The Humphrey Field Analyzer: concepts and clinical results. In Greve, E. L., Leydhecker, W. & Raitta, C. (Eds.), *Second European Glaucoma Symposium. Helsinki 1984*, pp. 55–64. Dordrech, W. Junk Publishers.

Koenker, R. & Basset Jr., G. (1978). Regression quantiles. *Econometrica*, **46**(1), 33–50.

Lambert, J. (1760). *Photometria sive de Mensura et Gradibus Lumininx, Colorum et Umbrae*. Detleffsen, Augsburg.

Laplace, P. S. (1799–1805). *Traité de méchanique céleste*. Paris.

Legendre, A.-M. (1805). *Nouvelles méthodes pour la détermination des orbites des comètes*. Courcier, Paris.

Mayer, T. (1750). Abhandlung über die Umwalzung des Monds um seine Axe und die scheinbare Bewegung der Mondsflecten. *Kosmogr. Nachr. Samml. auf des Jahr 1748*, **1**, 52–183.

Statpac User's Guide (1986). Allergan Humphrey, Los Angeles.