Suggestions How to Analyse Highly Episodic Repeated Measurements

Catrin Bergqvist

With Application in Growth Hormone

Abstract

Growth hormone (GH) affects most of the tissues in the body, where the main impact being on growth and metabolism. Infantile over secretion or lack of GH cause medical problems throughout life. To be able to diagnose GH deficiency (GHD) and improve its treatment a great deal of effort has been made and is still being made to explain how GH works.

This thesis discusses methods to describe repeated measures of GH in serum. In the first paper methods to conclude the existence and define the terminal time points of a high intensive interval of either changes of frequency or changes of amplitude are discussed. Both simulations and an application to a real data set are analysed. The second paper discusses some explorative methods to improve the estimation of peak amplitude, how to analyse the standardised curve form of a pulse and how to analyse periodicity.

Keywords:
Episodic data, Growth hormone, Bootstrap simulation, Directional Data Analysis, Harmonic regression, Isotonic regression

Acknowledgements

I am obliged to all who inspired my love of learning throughout the years - from my math teacher in upper secondary school, Karl-Georg 'Rödskägg' Jacobsson, to my supervisor Professor Sture Holm. I would also like to thank my family and friends, who never stop trying to distract me from work. Besides being my friend, Maj-Britt Menday has been a great help by checking the grammar in this thesis.

I am grateful to ITM, Swedish Institute of Applied Mathematics, who has given me financial support. I would also like to thank Kerstin Albertsson-Wikland, Chatarina Löfqvist and Sten Rosberg for introducing me into a fascinating area of research, containing a number of intriguing statistical issues.

It has been few moments of rest, a number of distress and frustration but mostly of great fun. Though the result would had been nothing without a wise, patient and encouraging supervisor, always prepared with a clever suggestion. Thank you, Sture.

Catrin Bergkvist
Göteborg, January 2001

1 Introduction

Growth hormone (GH) affects most of the tissues in the body, where the main impacts being on growth and metabolism. One of the latest achievements is that GH has an affect on the neurogenesis, i.e. new stem cells (Aberg et. al. 2000). More well-established are the stimulating effects of GH on the growth of muscles and bones and that GH mobilises fat cells. Lack of GH during childhood is conducive to an increased risk of diabetes, obesity and hypohysial dwarfism. Infantile over secretion leads on the other hand to acromegaly and gigantism. To be able to diagnose growth hormone deficiency (GHD) and improve its treatment a great deal of effort has been made and is still being made to explain how GH works.

This study discusses methods how to describe repeated measures of GH in serum. The characterisation of GH secretion data is a very high variation, both individually and between subjects. Every subject has a different baseline of the GH level that varies with time. The pulses of secretion give rise to a pulsatile pattern with a high variation of peak frequencies and peak amplitudes within individuals as well as between subjects. This thesis presents two methods to test if data changes significantly in peak frequency or amplitude at a certain interval of time. The methods will also define when these time intervals occur. In addition to that we will discuss methods to reveal if there is any periodicity of GH data and the appearance of the curve form of a standardised pulse.

The first section will give a very brief and 'popular' description of hormones in general and GH in particular. This will hopefully make it easier for the reader to grasp the biological background data. The next section explains how data is collected and discusses possible sources of errors. This section is followed by a description of the papers, paper I and paper II, respectively. Finally, a conclusion is given with a discussion of future topics.

2 Basic theory of growth hormone

The cells in all animals have different tasks. This requires that different cells must be able to communicate with each other. There are at least four ways of communication. In the *the autocrine system* cells are acting on themselves, in *the paracrine system* others are sending their molecules to the cell in the immediate surrounding. Then some cells are able to transport their molecules (often via the blood system) far in the body (*the endocrine system*). The fourth type of communication is *the neuroendocrine system*. A nerve cell releases a molecule directly into the blood to its target cell. Such a molecule or protein carrying information to cells is called a hormone.

The hormone hits its receptor molecule (or binding protein, BP), which is placed on the target cell. It is a rather beautiful concept, as every receptor molecule is adapted to a special hormone. It is the receptor molecule that guarantees that the right hormone ends up in the expected cell. This affinity of the target cells makes it possible for the concentration of hormones in the plasma to stay low.

Growth hormone is one of the hormones that belongs to the endocrine system. Other examples of endocrine regulators that also influence growth are insulin, cortisol and estrogens.
GH is a protein of the pituitary gland and its secretion is regulated by GH-releasing hormone and somatostatin (abbreviated GHRH and SS respectively). These regulators (or peptides) are secreted from the hypothalamus and influenced by a number of factors (such as nutrition, sleep, level of stress).

The GH consists mainly of four isoforms. The most abundant form is 22kDa GH, which constitutes about 75% of the secreted GH (Kriström, 1999). The difference isoforms have different half-times ($t_{\frac{1}{2}}$). Growth et. al. (1994) discern two groups when estimating the half-time; one 'slow' group with a $t_{\frac{1}{2}}$ about 19-22 minutes and one 'fast' group with a $t_{\frac{1}{2}}$ about 12-16 minutes. GH affects growth in three different ways; stimulates the growth of tissues, affects the liver (which stimulates other growth factors) and initiates the mobilisation of fatty acids.

The amount of secreted GH varies to a great extent through life. During foetal life and puberty high levels of GH are secreted compared to the levels secreted during prepuberty and adulthood. There is also a difference between gender. The secretion shows similar behaviour for both sexes during foetal life and infancy, but during puberty the difference between the sexes appears. As Albertsson-Wikland et. al. (1994) conclude, there is a two year difference between the sexes. Both, according to the age at the onset of the growth spurt, the age at the peak maximum and the age at the end of the growth spurt. It is believed that girls have higher peak amplitudes and different baselines of the GH levels compared to boys during puberty. Later, boys catch up at mid puberty. Even after puberty GH continues to be an important factor of metabolism and the difference between gender is believed to sustain, due to estrogen secretions in females.

3 Method of measure

The sampling of the data used in the papers took place at the Queen Silvia Children's Hospital, Göteborg. The data is based on 50 premature children, 37 boys and 13 girls. The children were of Caucasian origin and mainly Swedish. Their ages differed between 5 and 14 years old (the range of the boys was 5-14 and of the girls 8-13). The children were allowed and encouraged to activity and sleep as normally as possible and they were on an ordinary diet.
The blood was continuously drawn with a pump system and sampled every 20 minute for a 24-hour period, that is, each individual generated 72 sampling values. Then the plasma samples were assayed for GH.

We might note that the level of GH measured is the level in the blood, not the level actually affecting the binding proteins. This might cause a measure of error since different humans might have different capacity to absorb the hormones. Another source of error is problems in sampling due to all the different factors influencing GH. For instance it would be natural if the children felt some stress during observation, which is known to affect GH. These possible errors are very difficult to adjust. On the other hand, they are probably small and therefore not important.

4 Outline of the papers

As mentioned above, the data is measured repeatedly and shows a very episodic pattern. This irregularity constitutes the main challenge when analysing data. The aim of this thesis is to develop methods that describe the fluctuations of GH secretion. The thesis consists of two papers. In the first paper we seek methods that divide the data into a low and high intensity interval respectively. That is, we want to find methods which test if peak frequencies or amplitude is changing during a certain period of the day. In addition to that the distribution of data is described and how a possible high intensive interval depends on its starting point. In the second paper we discuss more explorative methods, for instance a method how to calculate the curve form of a standardised pulse is described.

4.1 Paper I

In the first paper we develop two methods to describe how the peak frequencies or amplitude of GH is changing during a 24-hour period. The first method is a criteria-based method, i.e. it is based on a test statistic. Using a bootstrap-like simulation we may test if frequency or amplitude is stationary or not. A procedure to find when the high intensive interval occurs, if the null hypothesis is rejected (i.e. if it is non-stationary), is presented.
The second method is a model-based method, where we first split data into two intervals (a high intensive and a low intensive). Then the technique and theory of directional data are used to evaluate how well the splitting explains our data. We might as well model how the length of the high intensive interval varies with the time point when the interval begins.

We test our methods by simulating four different cases. One deterministic, two with varied degrees of randomness but where there is a high intensive interval and finally a case illustrating stationary, i.e. the probability of peak or high amplitude is the same regardless of time point. Both methods work well when testing the null hypothesis of all four cases. Though, the weakness of the methods appear when determining the terminal points of the high intensive interval. Specially, the procedure to determine the terminal points of the criteria-based method is very indistinct and subjective.

Finally, 50 children's GH level is investigated. It is known that the major secretion of GH appears during night. We could confirm this with our methods.

4.2 Paper II

The second paper presents some explorative methods. First of all, a method for correcting estimates of peaks is given. When analysing continuous data at discrete intervals we are bound to make mistakes, the peak time of occurrences and its amplitude might be misjudged. This will cause the amplitude to be underestimated and the time for peak misjudged. The correction is based on a second order polynomial. We will also estimate the curve form of a standardised pulse. The estimated half-time from this curve is in accordance with the corresponding results of Groth et. al. (1994). This curve may be used for comparative studies.
Then methods to measure both the population and the individual periodicity respectively are discussed. We will present descriptive methods to estimate the presence of periodicity. The analysis of periodicity is based on rather naive assumptions and further development is needed.

5 Conclusions and future developments

The two approaches mentioned in paper I will not be fully comparable with each other, since they are based on different models and partly answer different questions. With the first method (the so called, criteria-based method) we might test the null hypothesis of stationarity, i.e. if frequency or amplitude respectively may be divided into two different intervals of intensity, and determine the terminal points of the high intensive intervals. The second method (the so called, model-based method) is a more refined tool. Besides the above, we might model how length of the high intensive interval varies with its starting point. The last procedure might be of importance when treating growth hormone deficient (GHD) children. Further investigations of the properties of both methods need to be done, for example of strength and robustness.

The measure to compare periodicity, presented in paper II, will not give a complete picture of periodicity. We will only be able to hint the degree of periodicity. We use, for instance a very naive definition of a peak, which may influence the periodicity. That is, we make no attempt to remove possible false peaks, due to noise. It is of interest to study this further, since it is believed that the periodicity of the injection is of importance for the capacity to absorb GH when treating GHD-patients.

The development of the curve form of a standardised pulse in the second paper, paper II, might also be analysed further. It remains to develop methods for comparative studies. A method to compare different populations (such as gender, ages or tall vs short children) would be of interest. It could also be useful when developing an objective tool to classify GHD patients.

References

Aberg M.A.I., Aberg N.D., Hedbacker H., Oscarsson J., Eriksson P.S. (2000). *Peripheral infusion of IGF-I selectively induces neurogenesis in the adult rat hippocampus*, Journal of neuroscience, Vol. 20, No 8, 2896-2903.

Albertsson-Wikland, K., Rosberg, S., Karlberg, J., Groth, T. (1994). *Analysis of 24-Hour Growth Hormone Profiles in Healthy Boys and Girls of Normal Stature:Relation to Puberty*, Journal of Clinical Endocrinology and Metabolism, Vol. 78, No 5,1195-1201.

Eckert, R., Randall, D., Augustine, G. (1938). *Animal physiology, mechanisms and adaptations* , Freeman and Company, New York.

Groth, T., Rosberg, S. and Albertsson-Wikland, K. (1994). *Estimation of Growth Hormone Secretory Patterns in Children with Use of a Numerical Deconvolution Technique: Experimental Design with Use of Computer Simulation*, Horm. Res., 42, 245-252.

Kriström, B. (1999). *Growth Hormone Treatment in Children Prediction of Growth Response*, Umeå University and Göteborg University, Umeå.

Löfqvist, C. (1999). *Growth Hormone Measurements Methodological and Interpretational Aspects*, Göteborg University, Göteborg.

Robinson, I.C.A.F. (1991). *Chronopharmacology of growth hormone and related peptides*, Advance Drug Delievery Reviews, No 6, 57-82.

Appendix A

Below the algorithms for the programs in paper I are presented. First the algorithms for estimating the high intensive interval of the criteria-based method are given. Then the algorithms for estimating the high intensive interval of the model-based method are described. The random seed used for all programs is 143255.

Algorithms of criteria-based method

First the algorithm for calculate frequency and then the corresponding algorithm for the amplitude are given.

Estimating frequency

Frekvens($datav$)	
1.	Initiate an empty matrix.
2.	For every individual and time point ....
2.1	If the sample value is related to a peak, then add up the frequency.
3.	Return the counts divided by the total number of individuals.

Rm($m, datav$)	
1.	Initiate variables.
2.	For every individual $i = 1, 2, .., n$
2.1	Generate an integer, k_i, between 0 and m.
2.2	For every time point $j = 1, 2, .., t$.
2.2.1	Translate the data set ($datav$) k_i step forward.
2.3	Add individual i's contribution to an empty matrix.
3.	Calculate the frequency (using Frekvens($datav$)) of the remixed data set.
4.	Return the test statistic (see equation 1).

Main program	
1.	Initiate constants and variables.
2.	Start a loop through $varv = 1, 2, .., B$.
2.1.	Simulate the distribution by the method described in section 2.2 (Rm($m, datav$)), where $m = t - 1$.
3.	Compare our value of the test statistic of our sample, with the quantiles of the generated distribution.

Estimating amplitude

Amplitud($datav$)	
1.	Initiate two empty matrices.
2.	For every individual and time point ...
2.1	If the response value is related to a peak, then add up the amplitude with that value.
3.	Return the sum divided by the total number of peaks.

The main program is analogously as for frequency.

Algorithms of model-based method

First the algorithm for calculate frequency and then the corresponding algorithm for the amplitude are given.

Estimating frequency

Nollett($pos1, pos2$)	
1.	Initiate the variables, '*noll*' and '*ett*'.
2.	Return the number of noughts (*noll*) and ones (*ett*) in the specified interval ($pos1, pos2$).

Ansattpos(pos)	
1.	Return the position for the peak after the peak at position *pos* in the sequence for individual i.

Berakning($pos1, pos2, tett$)	
1.	Count e_L, n_L and e_H, n_H for the given interval $(pos1, pos2)$ and its complement, via Nollett$(pos1, pos2)$.
2.	Calculate p_L and p_H.
3.	Define the high intensive interval, i.e. the one which correspond to p_H.
4.	Calculate the value of the sum of equation 3.
5.	Return the value of equation 3, p_L, p_H and the end points of the high intensive interval.

Main program	
1.	Initiate constants and variables.
2.	Start a loop through all individuals.
2.1.	Start a loop which passes through all peaks for one individual.
2.1.1.	Calculate the starting point of an interval, via Ansattpos(pos).
2.1.2.	Start a loop which passes through all remaining peaks for one individual.
2.1.2.1.	Calculate the end point of an interval, via Ansattpos(pos).
2.1.2.2.	Calculate the value of equation 3, via Berakning$(pos1, pos2, tett)$.
2.2.	Find the interval which maximises equation 3.
2.3.	Transpose the end points to trigonometric values.

Estimating amplitude

Berakning($pos1, pos2, ytot$)	
1.	Let the variable '$antal$' define the number of elements of the high intensive interval.
2.	Calculate y_L and y_H.
3.	If $\bar{y}_L > \bar{y}_H$ then ...
3.1	Switch positions, such that $(pos1, pos2)$ define the high intensive interval and recalculate y_L and y_H.
4.	Calculate the value of the sum of equation 4.
5.	Return the value of equation 4, y_L, y_H and the end points of the high. intensive interval.

Main program	
1.	Initiate constants and variables.
2.	Start a loop through all individuals.
2.1.	Start a loop which passes through all response values.
2.1.1.	Initiate the starting point of an interval.
2.1.2.	Start a loop which passes through all remaining peaks for one individual.
2.1.2.1.	Initiate the end point of an interval.
2.1.2.2.	Calculate the value of equation 4, via Berakning($pos1, pos2, ytot$).
2.2.	Find the interval which maximises equation 4.
2.3.	Transpose the end points to trigonometric values.

Appendix B

Below the algorithms for the program to estimate the standardised curve form of a pulse are presented (see paper II).

	Posamp($datav$)
1.	Define peaks.
2.	Calculate the adjusted peaks, via ifsatser1($datav, k1, k2, i, j$).

	ifsatser1($datav, k1, k2, i, j$)
1.	Adjust the peaks its corresponding time point according to the formulas in section 2.1 (equation 1 and 2 respectively).

	Kvotfkn($datav, i, avstand$)
1.	Return two vectors containing the logarithmic difference between peaks and its surrounding values (up to distance $avstand$) and the difference of the corresponding time points for individual i.

The subrutins Stigande($xdata1, ydata1$) and Avtagande($xdata2, ydata2$) estimate the isotonic and antitonic regression respectively, according to the flow chart at section 2.2.

	Main program
1.	Denote the number of values surrounding a peak *avstand*, which is considered to have an impact of the peak.
2.	Adjust the peaks according the theory of section 2.1., via Posamp($datav$), for each individual i.
3.	Use Kvotfkn($datav, i, avstand$) to create a matrix containing the logarithm of the ratio between the amplitude and its surrounding values.
4.	Use Kvotfkn($datav, i, avstand$) to create a matrix containing the difference in time of the corresponding time points.
5.	Split the data set, where one containing the values to the left of a peak and the other containing the values to the right of the peak.
6.	Plot raw data.
7.	Use Stigande($xdata1, ydata1$) to estimate the isotonic regression of the data set to the left of the peak.
8.	Use Avtagande($xdata2, ydata2$) to estimate the antitonic regression of the data set to the right of the peak.
9.	Plot the result.

Two Methods to Analyse Time Dependence of Hormone Data in Frequency or Amplitude

1 Introduction

One usually comes across difficulties when analysing hormone data, and growth hormone (GH) data makes no exception. The main challenges consist of individual variations and the fact that responses depend on a complex system of known (and unknown) substances and factors. The researchers in this area are interested in the population as a group as well as individuals. Some of the issues concern how frequencies of the peaks or amplitudes change over time, i.e. if the peaks come more often or are higher during a certain period of a 24-hour period. Another question is if the peaks come regularly, in say some particular 3-hour period. Beside this, there is a need for developing a method to evaluate if the individuals are growing in a 'normal' way or not.

The main issue of interest dealt with in this paper is if and how frequency or amplitude for GH is changing over time (the first question above). The problem has mostly been treated in an ad hoc way before (e.g. Martha et. al., 1989). Though, during the last decade numerous progresses in modelling data have been made. O'Sullivan and O'Sullivan (1988) model a pulse, with both exponential rise and decay, to depict where pulses appear. Their model and analysis is an example of a criteria-based method.
The common way of modelling hormonial data uses model-based methods, especially time series analysis. Diggle and Zeger (1989) introduced a non-Gaussian auto regression model, based on an underlying process. The latter depends on the history of the responses. This technique resembles hidden Markov models. Komaki (1993) adapted a similar pulse-shape as Diggle and Zeger, and developed a method to find the correct position and amplitude of the pulses. After that a state-space model is developed. Diggle and AlWasel (1997) carried on a spectral analysis, where they took the between-individual variation into special account. But in their approach no consideration of the exact location of the peak and its amplitude was made. Gou et. al. (1999) did not just consider that, but also the fact that the circadian rhythm gives

rise to a baseline in the pulse pattern. These effects were jointly modelled by a state-space representation.

This paper will present two methods for testing the above question in slightly different ways. Chapter 2 gives the theoretical background. The first method to be presented is a rather simple and naive criteria-based method. The other is model-based, related to the theory of directional data. These approaches will not be fully comparable, since they are based on different models and partly answer different questions (chapter 3). Hopefully though, the reader will find that every method has its own justification. Chapter 4 will illuminate the different methods in an example investigating the GH of 50 children.

2 The methods

This chapter gives the theoretical background to the two methods investigated. The first method to be described (section 2.2) simulates the distribution under the null hypothesis that the process is stationary (i.e. no major fluctuation of the frequency or the amplitude during the day). A test statistic is developed and used to conclude whether the null hypothesis is significant or not. The second method (section 2.3) divides the data into a low and high intensive interval. Then the technique and theory of directional data are used to test if the starting time of the high intensive interval affects the length of the same. In section 2.1, we first state some overall definitions.

2.1 Some overall definitions

The treated methods deal both with dichotomised data, when analysing peak frequency, and with continuous data, when analysing amplitude. The purpose is to separate the set of individual observations into two intervals, a low intensive, L, and a high intensive, H, interval with respect to peak frequency or amplitude respectively.

Consider a data set of n individuals, from which blood samples were taken every 20 minutes during a 24-hour period. Let

$$X_{i,j} \;=\; \begin{cases} 1 & \text{if } Y_{i,j-1} < Y_{i,j} > Y_{i,j+1} \\ 0 & \text{otherwise} \end{cases} ,$$

where $Y_{i,j}$ denotes the sample value of individual i at time point j for every $i = 1, 2, \ldots, n$, $j = 1, 2, \ldots, t$. Thus $X_{i,j} = 1$ indicates a local maximum at j for individual i. We suppose that observations for different individuals are independent while there may be dependence within individuals. As the reader might notice, the definition of a peak is rather naive, since we are analysing continuous data at discrete intervals. A peak may be missed or its amplitude may be underestimated. Nor do we consider the fact that each individual starts from a different baseline (which furthermore may fluctuate during a 24-hour period). Though, with these drawbacks in mind the main purpose of this paper is to develop methods for analysing growth hormone data (GH) rather than to find algorithms for the optimal position of a peak. A number of algorithms have already been developed to find a good prediction of the exact position when a peak will occur (e.g. Hauffa et. al., 1994 compare two of them). Our definition does not lean on any assumption. For instance we do not consider the size of difference to neighbouring values. This increases the influence of random variation, but the advantage is the objectivity of the definition.

It is also worth noting that the problem is circular, since the data is collected at specific time points following a 24-hour period. That is, the sequence of data does not stop at the last time point but continues with a value estimated by the value of the first time point.

The above definitions will remain throughout the paper. Other notations will be defined under sections of interest. The required computer calculations have been performed in Mathematica, version 4.0.1.0.

2.2 A criteria-based method founded on bootstrap simulations

In this section a criteria-based method to evaluate if frequency or amplitude changes over time is described. Test statistics, V and W, are formulated to test the null hypothesis (no changes over time), where the distribution under the null hypothesis is derived by simulation. Then we describe two possible procedures to find the location of the high intensive interval. Hence the following sequences $\{R_1, R_2, \dots, R_t\}$ and $\{Z_1, Z_2, \dots, Z_t\}$ are calculated, where

$$\begin{aligned} R_j &= \frac{1}{n}\sum_{i}^{n} X_{i,j} \quad \text{when calculating frequency,} \\ \text{or} \quad Z_j &= \frac{1}{N_j}\sum_{i}^{n} Y_{i,j}X_{i,j} \quad \text{when calculating amplitude at peaks,} \\ \text{where} \quad N_j &= \sum_{i}^{n} X_{i,j}. \end{aligned}$$

Appropriate and simple test statistics are given by

$$V = \frac{1}{t}\sum_{j=1}^{t}(R_j - \bar{R})^2 \quad \text{and} \tag{1}$$

$$W = \frac{1}{t}\sum_{j=1}^{t}(Z_j - \bar{Z})^2. \tag{2}$$

The null hypothesis for peak frequency is given by

$$\begin{aligned} H_0 &: f_1 = f_2 = \dots = f_t \quad \text{with the alternative,} \\ H_1 &: f_j \neq f_{j+1} \quad \text{for at least some } j, \end{aligned}$$

where $f_j = E[X_{i,j}] = E[R_j]$.

The null hypothesis for amplitude is given by

$$\begin{aligned} H_0 &: a_1 = a_2 = \dots = a_t \quad \text{with the alternative,} \\ H_1 &: a_j \neq a_{j+1} \quad \text{for at least some } j, \end{aligned}$$

where $a_j = E[Y_{i,j}X_{i,j}] = E[Z_j]$.

If the process, generated by frequencies, is stationary each individual has the same probability to reach a peak from a non-peak position, whatever time point, i.e. if the null hypothesis is true. The corresponding process, generated by the amplitude, is stationary if the probability to reach a high amplitude (compared to the mean) is the same during the 24-hour period. Therefore we need to compare the values of the test statistic with its distribution in a stationary process to test the null hypothesis. The latter is evaluated by a type of bootstrap simulation in the following way;

1. Move all the responses of individual i k_i step, where k_i is a randomly chosen integer between 0 and $t-1$. Accordingly, k_i is independent between individuals but constant within.
2. Calculate the test statistic denoted by v_r, where $r = 1, 2, \ldots, B$.
3. Repeat the above steps B times. In the applications (section 3.2.1 and 3.2.2) B is set to 1000.

The simulated values can be used to calculate the null hypothesis distribution of the test statistic.
The null hypothesis is rejected if the test statistic, V, is lower than the $\frac{\alpha}{2}$ quantile, $v(B, \frac{\alpha}{2})$, or higher than the $1-\frac{\alpha}{2}$ quantile, $v(B, 1-\frac{\alpha}{2})$, of the simulated distribution. Thus we conclude that the probability to peak or getting a high amplitude respectively is altered with respect to time if $V < v(B, \frac{\alpha}{2})$ or $V > v(B, 1-\frac{\alpha}{2})$. The bootstrap procedure for the test of the hypothesis of stationary amplitude is treated analogously.

One might think it is unnecessary to use a bootstrap procedure for the test of stationarity of peak frequency and that the distribution of the test statistic V is determined by a calculation for independent indicator variables. This is not the case. It will always exist a dependence between the $X_{i,j}$ within individuals, since the very definition of peak makes them dependent. If $X_{i,j} = 1$ then the values of both the previous and subsequent peak will equal 0, i.e. $X_{i,j-1} = X_{i,j+1} = 0$.

A rejection of H_0 will not give an answer how to separate the data set into high and low intervals, with respect to peak frequency and amplitude. A rejection just states there is a difference with respect to time. To be able to locate the high intensive interval we will present one possible procedure. The method tries to find the most likely interval and test that one.

2.2.1 The longest high level interval

One possible approach is to find the longest interval where most frequencies are above the mean. This interval is most likely to be the high intensive one. Hence, if we reject the null hypothesis, H_0, we might proceed by deriving the most likely interval, i.e. calculate the longest interval where the contribution of $\sum(R_j^+ - \bar{R})$ is greatest, where

$$(R_j^+ - \bar{R}) \quad = \quad \begin{cases} R_j - \bar{R} & \text{if } R_j > \bar{R} \\ 0 & \text{otherwise} \end{cases}$$

and $\bar{R}$ is the overall mean. The algorithm is given below.

1. Test the hypothesis

$$\begin{aligned} H_0 &: f_1 = f_2 = \ldots = f_t \quad \text{against the alternative} \\ H_1 &: f_j \neq f_{j+1} \quad \text{for at least some } j \in t, \end{aligned}$$

at a significant level α.

1.1. If H_0 is not rejected, then we stop and conclude that it is not possible to separate the data set.

1.2. If H_0 is rejected we calculate s_l, where

$$s_l \quad = \quad \max_j \begin{cases} \left(\sum_{k=1}^{t}(R_k^+ - \bar{R}) - \sum_{k=j}^{j+t-l-1}(R_k - \bar{R})\right)/l & \text{if } j \leq l \\ \left(\sum_{k=j-l}^{j-1}(R_k^+ - \bar{R})\right)/l & \text{if } j > l \end{cases},$$
$$\forall \quad j = 1, 2, \ldots, t \quad and \quad l = t-1, t-2, \ldots, 1.$$

Hence s_l is the maximal average value of all possible subsets with length l, i.e. the average of the total data set except the $j:th$ value and the subsequent $(t-l-1)$ values.

1.2.1. Determine the longest positive interval with a considerable contribution to V (i.e. the interval that corresponds to the maximal s_l with respect to l).

Hence $\mathbf{s}$ is a vector of length $t-1$, containing the maximum average value of the sum over every subinterval of length l, where $l = 1, 2, \ldots, t-1$. The longest interval that contains the relative greatest value of $\mathbf{s}$ with respect to length l, will reveal the high intensive interval. The time points corresponding to this interval are then easily found by a reversed procedure.

Plotting s_l against l is one way to decide which value to choose, see for instance figure 1 (page 16). This type of plot will also hint at other high intensive intervals. Though one drawback of this procedure is its subjectivity.

It is important to note that this last procedure can be used as a descriptive tool only. We might, at a specific level of significance, conclude whether there is a high intensive interval or not. The procedure to describe when this interval occurs is merely descriptive. Though the above procedure is an easy and effective way to deduce when the high intensive interval occurs.

2.3 A model-based method founded on the technique of directional data analysis

Directional data is a well known and an effective method to deal with observations that constitute a direction. There are a number of examples when the theory successfully is used to analyse observations with different geological interpretations, such as analysing wind directions or the homing ability of pigeons and so forth. It can also be used for observations which follow a circular scheme, like endocrine hormones.

The purpose in this section is the same as before, to separate the frequency of peaks or the overall amplitude into two intervals (a low intensive, L, and a high intensive, H), which is done with the least square method. Then the technique of directional data is used to estimate the start of the high intensive interval and the length of the same. This approach will give us a powerful tool to model how the length of a high intensive interval varies with its starting point.

To be able to separate the observations a suitable measure is needed. If we study the change of frequency, the estimate is given by minimising

$$
\begin{aligned}
& \sum_j^t \left\{X_{i,j} - I_{j\in H}\cdot p_H - I_{j\in L}\cdot p_L\right\}^2 = \\
= \; & \sum_j^t \left\{I_{j\in H}[X_{i,j} - p_H] + I_{j\in L}[X_{i,j} - p_L]\right\}^2 = \\
= \; & \sum_j^t I_{j\in H}[X_{i,j} - p_H]^2 + \sum_j^t I_{j\in L}[X_{i,j} - p_L]^2 = \\
= \; & \sum_j^t I_{j\in H}\left\{I_{j:X_{i,j}=1}[1 - p_H]^2 + I_{j:X_{i,j}=0}p_H^2\right\} + \\
& + \sum_j^t I_{j\in L}\left\{I_{j:X_{i,j}=1}[1 - p_L]^2 + I_{j:X_{i,j}=0}p_L^2\right\} = \\
= \; & e_H(1-p_H)^2 + (n_H - e_H)p_H^2 + e_L(1-p_L)^2 + (n_L - e_L)p_L^2 = \\
= \; & n_H p_H(1-p_H) + n_L p_L(1-p_L) = (n_L + n_H)p - n_H p_H^2 - n_L p_L^2,
\end{aligned}
$$

where

$$
\begin{aligned}
p_L &= \frac{e_L}{n_L}, \quad p_H = \frac{e_H}{n_H}, \quad p = \frac{e_H + e_L}{n_H + n_L}, \\
e_L &- \text{number of ones in the low intensive interval,} \\
n_L &- \text{total number of points in the low intensive interval,} \\
e_H &- \text{number of ones in the high intensive interval,} \\
n_H &- \text{total number of points in the high intensive interval.}
\end{aligned}
$$

If studying amplitude, the estimate is analogously given by minimising

$$
\sum_j^t \{Y_{i,j} - I_{j\in H}\cdot \bar{Y}_H - I_{j\in L}\cdot \bar{Y}_L\}^2,
$$

where

$$
\begin{aligned}
\bar{Y}_H &- \text{mean of all the responses for all } j \in H \text{ and} \\
\bar{Y}_L &- \text{mean of all the responses for all } j \in L.
\end{aligned}
$$

This means we want to maximise

$$n_H p_H^2 + n_L p_L^2 \quad \text{or} \tag{3}$$
$$n_H \bar{Y}_H^2 + n_L \bar{Y}_L^2 \quad \text{in the two cases, respectively.} \tag{4}$$

The first method corresponds to Bernoulli variables and the second to normally distributed variables, if assuming independent variables. Our observations are certainly not independent but the procedures seems robust enough. The methods we propose here coincide with the maximum likelihood estimates if we assume independent observations with constant parameters within the interval.

In practice an optimal interval cannot always be uniquely obtained. When we got a multiple choice the first option was chosen. This may lead to a bias, but we will not consider it further. When the optimal interval is found for each individual, the next step is to calculate the mean directions and measure the length of high intensity intervals. Since at least one interval of the individual intervals (either the low or the high interval) usually overlaps the day circle, it is appropriate to transform the data into trigonometric data. This motivates the representation of every sample value as a circular variable, with angle θ_i, where $0 \leq \theta_i \leq 2\pi$. The vector $(cos\theta_i, sin\theta_i)$ therefore represents the direction of an individual point on the circle.

The mean of the individual angles is not an appropriate estimate of the mean direction. This since it does not give an unambiguous measure (see for instance Mardia et. al., 1979). The following estimates and test statistics are used.
Let

$$tan\theta_i \;=\; \frac{C}{S}, \quad \text{where} \quad C = \frac{1}{n}\sum_{i=1}^{n} cos\theta_i \quad \text{and} \qquad S = \frac{1}{n}\sum_{i=1}^{n} sin\theta_i.$$

Then the mean direction of the above vector is defined by the direction of the unit vector

$$\frac{(C,S)}{\bar{R}}, \quad \text{where} \quad \bar{R} = \| (C,S) \| .$$

Thus, the direction is defined by the rescaled mean of the vector sum. The length, $\bar{R}$, can also be used as a measure how well data is clustering around the mean direction, i.e. as a measure of association. Hence it defines a value of the proportion of data that is close to the mean. It is easily shown (after some trigonometric transformations) that $0 \leq \bar{R} \leq 1$. $\bar{R}$ equals 1 if all the individual data lies in a perfect direction agreement.

When the optimal interval is found for each individual the next step is to fit a harmonic regression function. Our model on an individual basis is

$$L_i = a + bcos\phi_i + csin\phi_i + \epsilon_i, \quad (5)$$

where L_i represents the logarithm of the length of the high intensive interval for the individual i and ϕ_i represents the angle of the individual i's circular variable. The coefficient, a, represents the constant length of the interval and b and c reflect how the starting point influences the interval. If b has the main positive influence, then the interval is at its longest if it starts near midnight and shortest if it starts near noon. If c reflects the main positive influence, then the high intensive interval is at its longest if the starting point is near 6 am and at its shortest if the starting point is near 6 pm.

This model is the simplest natural regression model for circular dependence. Simple linear regression is not possible to use if continuity is required since it violates continuity at point $2\pi = 0$. Regression models including shorter periods (i.e. $\frac{2\pi}{2}, \frac{2\pi}{3}, \ldots$) correspond to polynomial regression on the real line and are suitable if a richer model is required.

A matrix notation for our model gives

$$\boldsymbol{L} = \boldsymbol{\Phi\beta} + \boldsymbol{\epsilon},$$

where

$$\begin{aligned} \boldsymbol{\Phi} &\ - \ \text{represents a } n*3 \text{ matrix,} \\ \boldsymbol{\beta} &\ - \ \text{represents a } 3*1 \text{ matrix of parameters and} \\ \boldsymbol{\epsilon} &\ - \ \text{represents a } n*1 \text{ matrix of errors.} \end{aligned}$$

Since this is a parameter linear model the usual estimate of $\boldsymbol{\beta}$ will be valid here. Hence

$$\begin{aligned} \hat{\boldsymbol{\beta}} &= \left(\boldsymbol{\Phi}^T\boldsymbol{\Phi}\right)^{-1}\boldsymbol{\Phi}^T\boldsymbol{L} \quad \text{and} \\ V(\hat{\boldsymbol{\beta}}) &= \left(\boldsymbol{\Phi}^T\boldsymbol{\Phi}\right)^{-1}\sigma^2, \end{aligned}$$

where σ^2 is the variance of the length.

It might be of high interest to estimate $\hat{\boldsymbol{\beta}}$. If only a is significantly nonzero then the length of the interval is constant. It should indicate that the starting point of the high intensive interval does not affect the length of the same. If b is significant then the high intensive interval is longer if it begins near midnight and shorter if it begins at noon. If c is significant, then the high intensive interval is longer if it begins at 6 am, and shorter if it begins at 6 pm. This information might be of importance in studies of GHD (growth hormone deficient) children.

Another way of looking at equation 5 is to estimate the interaction of how the starting point affects the length of interval. That is, we transform the above equation to

$$L_i = a + bcos\phi_i + csin\phi_i + \epsilon_i = a + \sqrt{b^2 + c^2}cos(\phi_i - \varphi) + \epsilon_i, \quad (6)$$

where

$$\varphi = \begin{cases} arctan(\frac{c}{b}) & \text{if } b > 0 \\ \pi + arctan(\frac{c}{b}) & \text{otherwise} \end{cases}.$$

The maximal length of interval is given by $a + \sqrt{b^2 + c^2}$ at the time point when $\phi_i = \varphi$, i.e. $\frac{24}{2\pi}arctan(\frac{c}{b})$ if $b > 0$ and $\frac{24}{2\pi}(\pi + arctan(\frac{c}{b}))$ if $b < 0$. Hence, the latter equation gives a more distinct idea of how length varies with the starting point. We may use it when calculating the maximum length and its corresponding starting point.

A test of the hypothesis of homogeneous circular distribution may be performed by using the estimates of $\bar{R}$ according to the theorem below.

Theorem
Suppose we have t repeated measurements of n independent individuals. Let L_i denote the length of an individual interval and its corresponding starting point expressed with its angle, ϕ_i, where $i = 1, 2, \ldots, n$, according to equation 5. The hypothesis that the distribution contains no direction, i.e. $H_0 : E[L] = a \quad \forall \quad \phi_i \in [0, 2\pi]$, may be rejected if $\bar{R} > \sqrt{\frac{-ln\alpha}{n}}$ at the significance level α.

Proof
Due to the central limit theorem, we note that both C and S are normally

distributed with mean 0 and variance $\frac{1}{2n}$, respectively. The two-dimensional random variable (C, S) has an asymptotically bivariate normal distribution according to Cramer-Wold theorem.
Since C and S are uncorrelated, $2n\bar{R}^2$ has an asymptotically chi-squared distribution with parameter 2, or equivalently an exponential distribution with parameter $\frac{1}{2}$, under the null hypothesis. $\square$

That is, with a sample size of 100 we may conclude a high intensive interval at a 5% level of significance if $\bar{R} > 0.173$. Though, it must be noted that this result is merely asymptotic.

3 A theoretical comparison

The purpose of this chapter is to study the difference between the methods. Some differences are obvious (section 3.1) and others more delicate. In section 3.2 the differences are illustrated by four simple simulations. First, a deterministic case is analysed, where all individuals have exactly the same appearance, i.e. a high intensive interval during a certain period of time. In the second simulation some stochastic effects are added, but the start of the high intensive intervals is still predetermined. The third simulation is the same as the second, but the high intensive interval starts randomly within a certain interval. Finally, the fourth simulation illustrates the case without any structure, i.e. when frequency or amplitude follows a stationary process. The chapter will end with an overall discussion about the results and a comparison between the methods (section 3.3).

3.1 Some obvious differences

The criteria-based method estimates if frequency or amplitude varies during a certain period under a 24-hour period. Then a procedure to find when the high intensive interval occurs is developed.
In the model-based method the data is divided into two intervals (regardless if it is appropriate or not), which is done with an ordinary least square method. With the technique of directional data, the clustering around the

mean direction is measured. In addition we estimate the following harmonic regression equation;

$$L_i = a + bcos\phi_i + csin\phi_i + \epsilon_i,$$

where the parameters indicate how the length of the high intensive interval varies with the time point when the interval begins.

Hence the criteria-based method is a method for testing the null hypothesis of stationarity and finding the high intensive interval. In addition to that, the model-based method may be used as a tool to describe and model the high intensive interval. To give a more detailed illustration of their difference and similarity it might be helpful to observe how they behave in simulations.

3.2 The results of the simulations

In this section the four different cases of simulation are described. Each case simulates 100 individuals, where one individual is to be measured every 20 minutes during a 24-hour period (i.e. in total 72 points of measure per individual and day). Each simulation is thought to be commenced at 8:00 am. First, the criteria-based method is used to analyse frequency (section 3.2.1) and amplitude (section 3.2.2) respectively and then the model-based method is used to analyse the same data set (section 3.2.3 and 3.2.4 respectively).

The first simulation will describe a deterministic case, where every individual gets a peak at the same time during a predetermined interval (between 11:00 pm to 4:00 am). Outside this interval we have a declining (or increasing) sequence, i.e. $X_{i,j} = 0$. Nota bene, we are unable to peak twice in a row, so the sequence of each individual is $\{0, \dots, 0, 1, 0, 1, \dots, 0, 1, 0, \dots, 0\}$. The probability to peak is thus

$$P\big(X_{i,j} = 1 | X_{i,j-1} = 0\big) = \begin{cases} 1 & \forall \quad j = 45, 46, \dots, 60 \\ 0 & \text{otherwise} \end{cases}$$

and this case will be referred to as sim 1.

In the second case (referred to as sim 2) we add some stochastic effects. Still it will be a rather structured sequence, with a good possibility to deduce the data in two intervals (a low and high intense respectively). This time, the probability to peak (if we did not peak the observation before) is 0.6 if we are in the high intensive interval, otherwise it is 0.14. The start and length of the high intensive interval will be the same as in the first case. The probability to peak is thus

$$P\big(X_{i,j} = 1 | X_{i,j-1} = 0\big) = \begin{cases} 0.60 & \forall \quad j = 45, 46, \ldots, 60 \\ 0.14 & \text{otherwise} \end{cases}.$$

The third case (referred to as sim 3) is quite similar to the second case. The only difference is that each individual may start at different time points. The high intensive interval starts randomly somewhere between 9:00 pm and 11:00 pm and the length of the interval is determined to be 5 hours for each individual. The probability to peak is the same as for sim 2 and the probability to enter the high intensive interval follows a sinus-curve, where the cumulative distribution function is $Sin[(j-38) * ArcSin[1]/7] \quad \forall \quad j = 38, 39, \ldots, 45$.

The last case (referred to as sim 4) illustrates the stationary case. The probability to peak is 0.4, if we did not peak the observation before, thus

$$P\big(X_{i,j} = 1 | X_{i,j-1} = 0\big) = 0.40 \quad \forall \quad j = 1, 2, \ldots, 72.$$

The four cases above illustrate how different grades of randomness affect our methods. The first case includes no randomness at all but two levels of intervals, the second case introduces some noise in the simulation. The third case adds some extra noise, with a random starting point of the high intensive interval and finally, the last simulation illustrates the stationary case.

When creating the analogous cases for amplitude the above simulations are used. Every individual has a constant baseline (a simplification since the baseline fluctuates during a 24-hour period), which is uniformly distributed per individual. When $X_{i,j} = 0$, i.e. when not having a peak, we add a value, simulated with a log normal distribution (mean equals 0.3 and variance equals 1), to the baseline. When $X_{i,j} = 1$, i.e. when having a peak, the response value of the amplitude equals the baseline and a value ten times a random value following the same distribution as above.

$$Y_{i,j} = \begin{cases} Z_{i,j} + k_i & \text{if } X_{i,j} = 0 \\ 10Z_{i,j} + k_i & \text{if } X_{i,j} = 1 \end{cases},$$

where $Z_{i,j}$ is log normal distributed, with mean 0.3 and variance 1 and k_i is the constant baseline, for each individual i, uniformly distributed between 0.1 and 0.4.

3.2.1 An analysis of the frequencies with the criteria-based method

In this section the null hypothesis of stationarity of frequency is tested with the criteria-based method. The variable B, i.e. the number of loops, is set to 1000 when simulating the distribution, as described in section 2.2. The results of the four different cases of simulation are presented and analysed. The table 3.2.1 shows the values of the test statistics and the medians, the 2.5% and the 97.5% quantiles, the means and the standard deviations respectively of the simulated distributions for the different cases.

Table 3.2.1. The results of all four simulations (sim 1, sim 2, sim 3 and sim 4 respectively) where V denotes the test statistic of the frequency and the other values describe the distributions in the stationary case.

	sim 1	sim 2	sim 3	sim 4
V	0.0988	0.0144	0.0110	0.00157
Median	0.000915	0.00145	0.000915	0.00202
$v(1000, 0.025)$	0.000432	0.00102	0.000432	0.00138
$v(1000, 0.975)$	0.00189	0.00206	0.00189	0.00287
mean	0.000983	0.00148	0.000983	0.00205
SD	0.000379	0.000271	0.000379	0.000387

We might conclude that the distributions are symmetric, since the means and corresponding medians are almost the same. This should not be a big surprise. The simulations take place to assure stationarity, i.e. the frequencies are equally spaced over time.

The first column of the table 3.2.1 shows the results of the first simulation. There is a significant difference between the test statistic and the stationary state. Hence we might conclude that there is a difference in the frequencies. To discern when the high intensive interval occurs the longest positive interval of V is derived, as suggested in section 2.2.2. Plotting (see the first graph at figure 1) indicates that the length is 5 hours, since the latest peak is at $l = 15$ and we are measuring each 20 minutes (i.e. every third of an hour). The high intensive interval is found to be between 11:00 pm to 4:00 am, which is as expected.

The second column shows the results of sim 2. As expected the result shows that there is a difference in the frequencies. To discern if there is an interval with higher intensity we proceed as before. The figure (the second graph at figure 1) indicates that the length of the interval is 5 hours and 20 minutes and further investigation shows the high intensive interval appearing between 11:00 pm and 4:20 am. The interval is longer compared to the interval of sim 1. This is expected, since it is a nonzero possibility to peak outside the high intensive interval.

The result of the third simulation, sim 3, is shown in the third column of table 3.2.1. As expected, the null hypothesis may be rejected. The figure (figure 1, graph c) is difficult to interpret. One suggestion is that the interval is 4 hours and 20 minutes (i.e. $l = 13$). This would mean that the interval occurs between 10:00 pm to 02:20 am.

The last column shows the results from the forth simulation. The table shows no significant difference between the test statistic and the stationary state. Hence we cannot divide the data into different intervals.

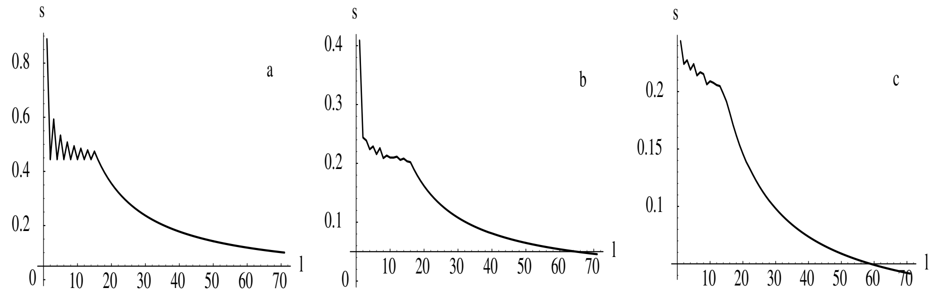

Figure 1: A plot of $\mathbf{s}$ against the length of the subintervals (l) of sim 1 (a), sim 2 (b) and sim 3 (c).

The conclusion is that our method works satisfactorily in all different simulations. Though it might be difficult to establish the terminal points of the interval with our procedure. It turns out to be difficult to interpret the figure, when dealing with a high degree of randomness.

3.2.2 An analysis of the amplitude with the criteria-based method

In this section the null hypothesis of stationarity of amplitude is tested with the criteria-based method. The variable B, i.e. the number of loops, is set to 1000 when simulating the distribution, as described in section 2.2. We analyse the logarithmic values of the amplitude, i.e. the data is expected to follow a normal distribution. The results of all four simulations (sim 1, sim 2, sim 3 and sim 4 respectively) are shown in table 4.3.2. The variable W denotes the value of the test statistic and the other values describe the distribution in the stationary case for each simulation.

Table 4.3.2. The results of all four simulations (sim 1, sim 2, sim 3 and sim 4 respectively) where W denotes the test statistics of the amplitude and the other values describe the distributions in the stationary case.

	sim 1	sim 2	sim 3	sim 4
W	0.555	0.192	0.164	0.0352
Median	0.0332	0.0401	0.0371	0.0351
$w(1000, 0.025)$	0.0220	0.0276	0.0259	0.0248
$w(1000, 0.975)$	0.0510	0.0541	0.0506	0.0477
mean	0.0341	0.0404	0.0373	0.0354
SD	0.00744	0.00689	0.00644	0.00600

As before we might conclude that the simulated distributions are symmetric, since the means and the corresponding medians are almost the same. We might also note that the standard deviations vary to a greater extent, compared to analysing frequency. This is quite natural.

The first column of the table 4.3.2 shows the results of the first simulation. There is a significant difference between the test statistic and the stationary state. Hence we might conclude there is a difference in amplitudes over time. To discern when a high intensive interval occurs, the longest positive interval of W, as suggested in section 2.2.2, is derived. Plotting (see the first graph at figure 11) indicates that the length is 4 hours and 20 minutes, since the latest peak is at $l = 13$. The high intensive interval is found to be between 10:40 pm to 3:00 am. We might have expected a longer interval than the underlying high intensive interval. This since it is difficult to control the simulation of a continuous variable. The way we simulate the amplitude will produce peaks outside our control.

The second column shows the results of sim 2. As expected the result shows that there is a difference in the frequencies. To discern if there is an interval with higher intensity we proceed as before. The figure (the second graph at figure 11) indicates that the length of the interval is 5 hours (i.e. l=15). That is, the high intensive interval should be between 11:00 pm to 04:00 am.

We might reject the null hypothesis of sim 3 (see the third column in table 4.3.2) as expected. The figure (figure 11, graph c) is difficult to interpret as when simulating frequency. Since the line is monotonously decreasing we refrain from interpreting the length of the interval.

The last column shows the results from the forth simulation. The table shows no significant difference between the test statistic and the stationary state. Hence we cannot divide the data into different intervals.

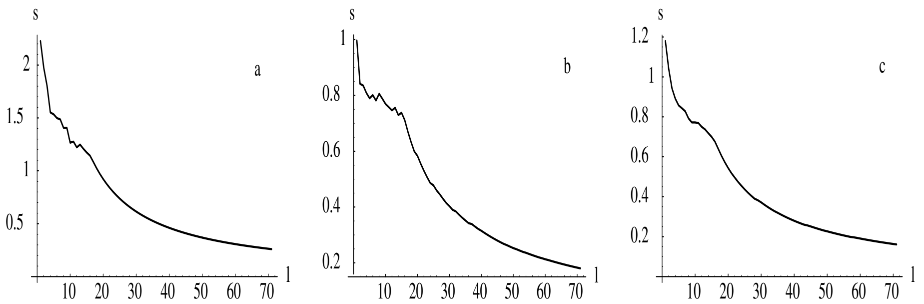

Figure 2: A plot of $\mathbf{s}$ against the length of the subintervals (l) of sim 1 (a), sim 2 (b) and sim 3 (c).

The results when analysing amplitude are analogous to the results when analysing frequency. The test of time dependence seems to work well enough, but the procedure to deduce when the high intensive interval occurs is too subjective and indistinct when we have lots of randomness in data.

3.2.3 An analysis of the frequencies with the model-based method

In this section the frequency is analysed with the model-based method. The results of the method are shown in table 3.2.3. All simulations are given in the same table. In the first row $\bar{R}$ and times when the high intensive intervals begin are presented. In the second row the corresponding values for the end points are given.

Table 3.2.3. The results of all four simulations of the frequency, where the values of $\bar{R}$ and times of the terminal points are given.

	sim 1	sim 2	sim 3	sim 4
$\bar{R}_{beginning}$	1.0	0.834	0.848	0.0256
starting time	-1.0	-1.41	-2.49	12.0
$\bar{R}_{end}$	1.0	0.809	0.838	0.0481
ending time	3.67	3.84	2.32	2.83

The first column of table 3.2.3 shows the result of the first simulation. Its high intensive interval begins at 11:00 pm and ends at 3:40 am the following day. This is as expected, since the underlying interval ends with a nadir. The clustering around mean ($\bar{R}$) for both the beginning and end of the interval is 1.0, since we analyse the same deterministic pattern for all individuals. Hence, the average length of the high intensive interval is approximately 4 hours and 40 minutes for the population. The estimate of β cannot be calculated since the inverse of $\mathbf{\Phi}^T\mathbf{\Phi}$ is singular. Therefore we cannot calculate the average length of the high intensive interval of an individual.

According to the table above, the high intensive interval of the second simulation begins at 10:35 pm and ends at 3:50 am the following day (see table 3.2.3, second column). This interval is expected to be more conservative (i.e. longer) than the underlying high intensive interval, since there is a nonzero possibility to peak outside the predetermined interval. The clustering around mean ($\bar{R}$) for both the beginning and the end of the interval is strong, above 0.8. Figure 3 shows how the trigonometric coordinates of the individuals are distributed around the unit circle.

It might be appropriate to explain how to read the figure 3. Since we are transforming data into trigonometric data, the point (1,0) represents midnight, (0,1) represents 6:00 am and so forth. Hence the figure is rotated compared to the clock and goes anti-clockwise.

The estimates of β are 1.44, -0.169 and -0.732 and variances 0.0245, 0.0231 and 0.0354 respectively. The maximal length of the interval is around 9 hours, when starting at 5:00 pm. The estimated length of the high intensive interval of an individual, when starting at 10:35 pm, is approximately 4 hours and 40 minutes. The length of the high intensive interval of the whole population is around 5 hours). This is expected since all individuals start at the same time point.

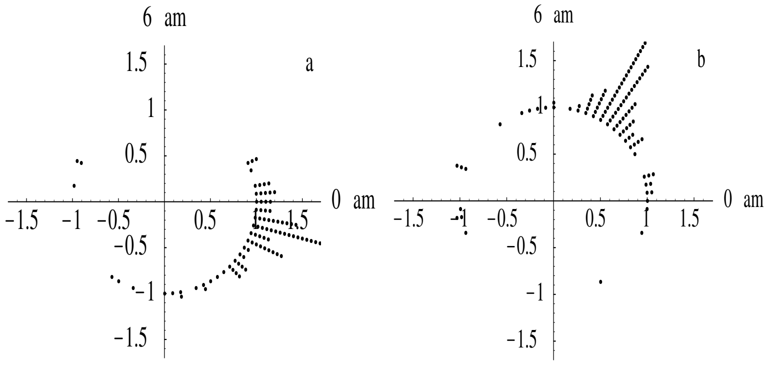

Figure 3: The circular coordinates of the individual's frequencies of sim 2 plotted around the unit circle, where (1,0) represents midnight and goes anti-clockwise. The first graph (a) shows how the beginning of the high intensive interval is distributed for the different individuals. The second graph (b) illustrates the corresponding data of the end point of the high intensive interval.

The high intensive interval of the third simulation (see table 3.2.3, third column) begins at 9:30 pm and ends at 2:20 am the following day. We might have expected that the high intensive interval should be slightly longer than the interval in the second simulation. Still, a conclusion might be that our method is rather robust at randomly chosen starting points. The clustering around mean ($\bar{R}$) for both the beginning and the end of the interval is considerable, $\bar{R} > 0.82$. Figure 4 shows how the data is distributed.

The estimates of β are 1.14, 0.345 and -0.968 and variances 0.0378, 0.0263 and 0.0516 respectively. The maximal length of the interval is 8 hours and 45 minutes, when starting at 7:20 pm. The estimated length of the high intensive interval of an individual, when starting at 9:30 pm, is approximately 7 hours and 20 minutes. The high intensive interval of the whole population is around 5 hours.

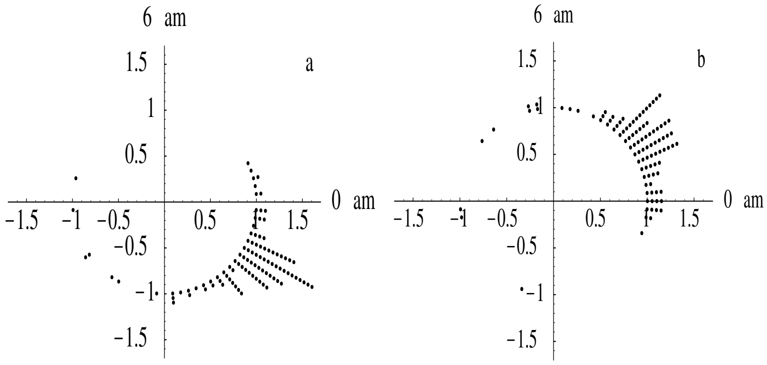

Figure 4: The circular coordinates of the individual's frequencies of sim 3 plotted around the unit circle, where (1,0) represents midnight and goes anti-clockwise. The first graph (a) shows how the beginning of the high intensive interval is distributed for the different individuals. The second graph (b) illustrates the corresponding data of the end point of the high intensive interval.

The result of the last simulation is given in the forth column of table 3.2.3. The high intensive interval begins at 12:00 pm and ends at 2:50 am the following day and a low proportion of data clusters around the mean direction ($\bar{R} = 0.05$ and below for both terminal points). Based on the asymptotic result in the theorem at page 11 and the value of $\bar{R}$, we claim that the interval is non significant. The figure shows how the data is distributed, see figure 5.

Our conclusion is that all simulations turned out as expected. We are able to reject sim 1, sim 2 and sim 3, but not sim 4. The estimated time intervals correspond rather well to to the result of the criteria-based method. The results of the individual lengths should be interpreted with some reservations since we did not check the regression model.

3.2.4 An analysis of the amplitude with the model-based method

In this section the amplitude is analysed with the model-based method. As before, we analyse the logarithmic values of the amplitude. The results of the method are shown in table 3.2.4.

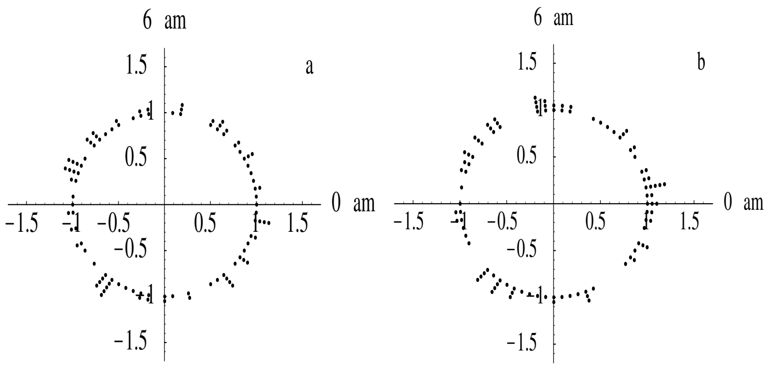

Figure 5: The circular coordinates of the individual's frequencies of sim 4 plotted around the unit circle, where (1,0) represents midnight and goes anti-clockwise. The first graph (a) shows how the beginning of the high intensive interval is distributed for the different individuals. The second graph (b) illustrates the corresponding data of the end point of the high intensive interval.

Table 3.2.4. The results of all four simulations of the amplitude, where the values of $\bar{R}$ and times of the terminal points are given.

	sim 1	sim 2	sim 3	sim 4
$\bar{R}_{beginning}$	0.936	0.383	0.332	0.0618
starting time	-0.920	-1.36	-3.25	6.11
$\bar{R}_{end}$	0.942	0.424	0.421	0.0764
ending time	3.62	3.72	3.00	4.15

The first column of table 3.2.4 shows the result of the first simulation. Its high intensive interval begins at 11:05 pm and ends at 3:37 am the same day. The interval is slightly shorter than what we might expect, the length of the interval is 4 hours and 30 minutes. The clustering around mean ($\bar{R}$) for both the beginning and end of the interval is strong, above 0.90. The reason for not lying in a perfect direction, i.e. not being 1, may be explained by the simulation of the amplitude. We do not have a distinct starting point, since we might get peaks outside the high intensive interval. Figure 6 shows how the trigonometric coordinates of the individuals are distributed around the unit circle.

The estimates of β are 2.26, -1.09 and -0.796 with variances 0.059, 0.0635 and 0.0381 respectively. The maximal length of the interval is around 36 hours and 50 minutes, when starting at 2:25 pm. The estimated length of the high intensive interval of an individual, when starting at 11:03 am, is approximately 4 hours.

According to the table above, the high intensive interval of the second simulation begins at 10:40 pm and ends at 3:42 am (see table 3.2.4, second column). The clustering around mean ($\bar{R}$) for both the beginning and end of interval is considerable less than sim 1, around 0.35. Figure 7 shows how the trigonometric coordinates of the individuals are distributed around the unit circle.

The estimates of β are 0.932, -0.342 and -0.495 and variances and 0.0273, 0.0437 and 0.0533 respectively. The maximal length of the interval is around 4 hours and 40 minutes, when starting at 3:40 pm. The estimated length of the high intensive interval of an individual, when starting at 10:40 pm, is approximately 2 hours and 10 minutes. The high intensive interval of the whole population is around 5 hours.

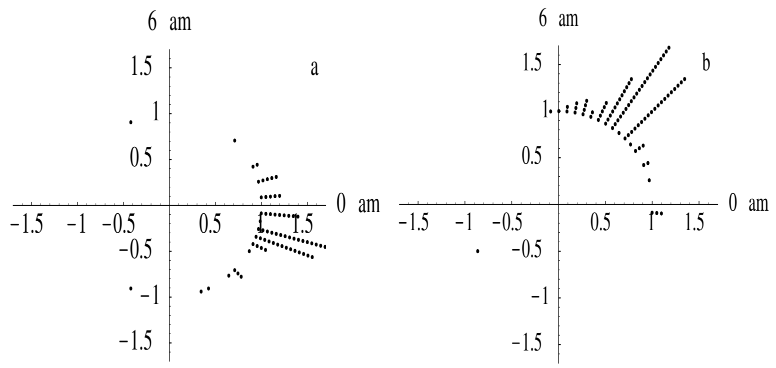

Figure 6: The circular coordinates of the individual's amplitudes for sim 1 plotted around the unit circle, where (1,0) represents midnight and goes anti-clockwise. The first graph (a) shows how the beginning of the high intensive interval is distributed for the different individuals. The second graph (b) illustrates the corresponding data of the end point of the high intensive interval.

The high intensive interval of the third simulation (see table 3.2.4, third column) begins at 8:45 pm and ends at 3:00 am. The interval is as expected compared to the underlying interval. The clustering around mean ($\bar{R}$) for both the beginning and end of interval is low, below 0.43. Figure 8 shows

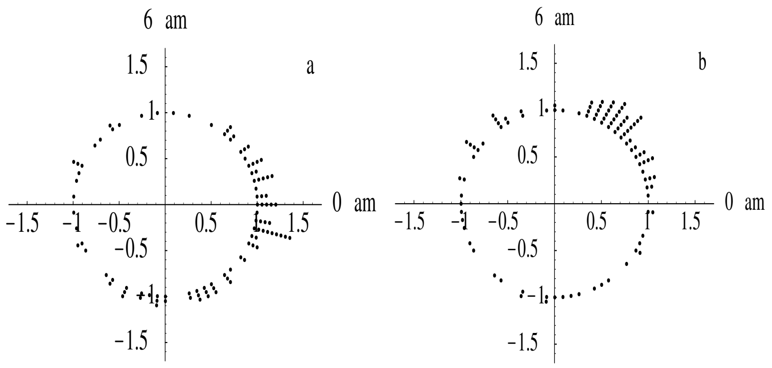

Figure 7: The circular coordinates of the individual's amplitudes for sim 2 plotted around the unit circle, where (1,0) represents midnight and goes anti-clockwise. The first graph (a) shows how the beginning of the high intensive interval is distributed for the different individuals. The second graph (b) illustrates the corresponding data of the end point of the high intensive interval.

how the data is distributed.

The estimates of β are 1.053, -0.457 and -0.463 with variance 0.0215, 0.0342 and 0.0552 respectively. The maximal length of the interval is around 5 hours and 30 minutes, when starting at 3:00 pm. The estimated length of the high intensive interval of an individual, when starting at 8:45 pm, is around 3 hours. The high intensive interval of the whole population is around 5.

The result of the last simulation is given in the forth column of table 3.2.4. The high intensive interval begins at 6:06 am and ends at 4:09 am and a low proportion of data clusters around the mean direction ($\bar{R} = 0.08$ and below for both terminal points). The figure also shows that the data is evenly distributed around the day, see figure 9.

Hence we may reject the null hypothesis (i.e. no high intensive interval) for sim 1, sim 2 and sim 3, but not for sim 4. This result is reassuring. It seems that the method to estimate the terminal points works satisfactorily as well and is in accordance with the result of the criteria-based method. The results of the harmonic regression analysis may be interpreted with some reservations, since we did not check our model. It is worth noting that the estimated individual lengths correspond to the geometric mean, while the length of the whole population corresponds to the arithmetic mean.

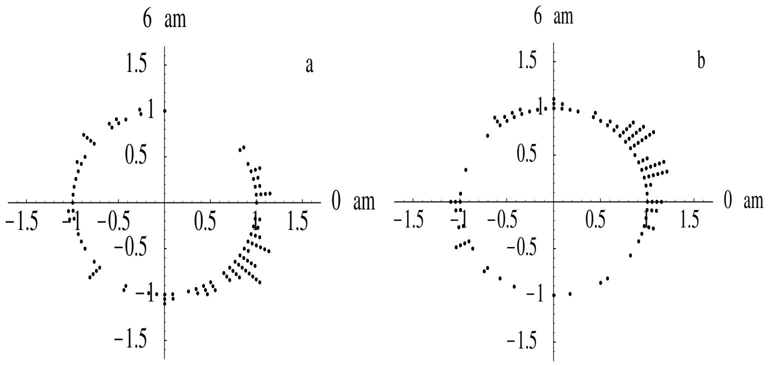

Figure 8: The circular coordinates of the individual's amplitudes for sim 3 plotted around the unit circle, where (1,0) represents midnight and goes anti-clockwise. The first graph (a) shows how the beginning of the high intensive interval is distributed for the different individuals. The second graph (b) illustrates the corresponding data of the end point of the high intensive interval.

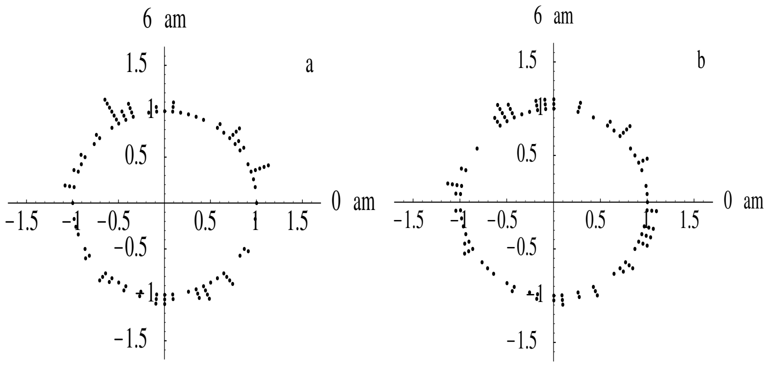

Figure 9: The circular coordinates of the individual's amplitudes for sim 4 plotted around the unit circle, where (1,0) represents midnight and goes anti-clockwise. The first graph (a) shows how the beginning of the high intensive interval is distributed for the different individuals. The second graph (b) illustrates the corresponding data of the end point of the high intensive interval.

3.3 An overall discussion

The criteria-based method seems to predict the high intensive interval rather well. Even when analysing amplitude we get satisfactory results. Though it may be difficult to deduce the terminal points in an objective manner, specially when dealing with data with a high degree of variation. The model-based method works well analysing both frequency and amplitude.

The model-based method may also be used when examining data further, for instance harmonic regression analysis of individual length. It should be noted that we did not check the validity of the harmonic regression. It was merely mentioned as a possible approach. Further investigation regarding strength of both methods also need to be done.

4 An example

To illuminate our methods even further, we will investigate a real data set. The data set consists of children of both sexes in early puberty. A more detailed description of the individuals will be given in section 4.2. First (section 4.1) the method of measurements will be described. In section 4.3 a presentation of the statistical analysis and some comments of the results will be given.

4.1 Method of measurement

The sampling took place at the Children's Hospital of Queen Silvia, Göteborg. The children were allowed and encouraged to activity and sleep as normally as possible and they were on an ordinary diet.
The blood sampling commenced between 8:00 am and 9:00 am. A constant withdrawal pump (Swemed, Göteborg, Sweden) with a catheter (Carmeda AB, Stockholm, Sweden) was used. The tubes were changed every 20 minutes for a 24-hour period, i.e. each individual generated 72 sampling values.

GH was measured in samples of plasma by a Pharmacia GH radioimmuno assay using plyclonical antibodies and all samples except one used the WHO First International Reference Preparation for human GH (WHO 66/217). The latter sample was analysed with WHO 80/505 standard, which has been adjusted for, according to Löfqvist (1999).

4.2 Patients

The study consisted of 50 premature children, 37 boys and 13 girls. The children were of Caucasian origin and mainly Swedish. Their ages differed between 5 and 14 years old (the range of the boys were 5-14 and of the girls 8-13). Even if gender might be an important factor during prepuberty we did not split the data set. This since the data set is a rather small as it is and since this example is merely meant as an illustration of the methods.

4.3 Analysis of data

In the raw data a missing observation is indicated as -1. If no GH was detected the observation was indicated as 0. In the analysis an undetectable value was set to the lowest detectable value of that individual and a missing observation was imputed by the mean of the nearest surrounding non-missing values. This means that a missing value will not appear as a peak in the analysis.

In the first section (section 4.3.1) the frequency is analysed. First with the criteria-based method and then with the model-based method. Then (section 4.3.2) the logarithmic of the amplitude is analysed analogously.

4.3.1 Analysis of frequency

The result of the analysis of the frequency using the criteria-based method are presented in table 4.3.1. In the first column the value of the test statistic is given. Then the simulated distribution is described with median, the 2.5% and the 97.5% quantiles, the mean and the standard deviation.

Table 4.3.1. The result for the 50 children where V denotes the test statistic and the other values describe the distribution in the stationary case.

V	median	$v(1000, 0.025)$	$v(1000, 0.975)$	mean	sd
0.00489	0.00374	0.00257	0.00511	0.00377	0.000625

As the table indicates, we cannot reject the null hypothesis. Hence we cannot divide the data set into a low and high peak intensive interval. This is in good conformity with previous results.

The result of the corresponding analysis using the model-based method is given in table 4.3.1. In the first and second columns $\bar{R}$ and the time when the high intensive intervals begins are shown. In the third and fourth columns the corresponding values of the end point are shown.

Table 4.3.1. The result for the 50 children where the values of $\bar{R}$ and times of the terminal points are given.

$\bar{R}_{beginning}$	starting time	$\bar{R}_{end}$	ending time
0.107	12.2	0.0762	13.9

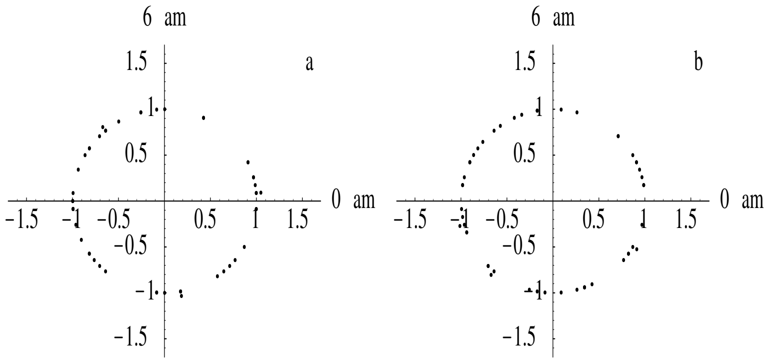

Figure 10: The circular coordinates of the individual's frequency plotted around the unit circle, where (1,0) represents midnight and goes anti-clockwise. The first graph (a) shows how the beginning of the high intensive interval is distributed for the different individuals. The second graph (b) illustrates the corresponding data of the end point of the high intensive interval.

The result of the model-based method (see table 4.3.1) confirms the result of the criteria-based method. The clustering around the mean direction is low, which is illustrated in table 4.3.1 as well as in figure 10. Note as well that the value of $\bar{R}$ is very small, not even significant according to the theorem at page 11.

4.3.2 Analysis of amplitude

The result of the criteria-based method, when analysing the logarithmic of the amplitude, is presented in table 4.3.2. In the first column the value of the test statistic is given. Next the simulated distribution is described with median, the 2.5% and the 97.5% quantiles, the mean and the standard deviation.

Table 4.3.2. The result of the 50 children where W denotes the test statistic and the other values describe the distribution in the stationary case.

W	median	$w(1000, 0.025)$	$w(1000, 0.975)$	mean	sd
0.638	0.322	0.223	0.449	0.326	0.0579

According to previous results, there is an increased amplitude during night. This we can confirm, since we reject the null hypothesis. The high intensive interval is about 5 hours (i.e. $l = 15$) and occurs between 23:20-4:20.

Figure 11: A plot of $\mathbf{s}$ against the length of the subintervals.

The result of the corresponding analysis using the model-based method is given in table 4.3.2. In the first and second columns $\bar{R}$ and the time when the high intensive intervals begins are presented. In the third and fourth columns the corresponding values of the end points are shown.

Table 4.3.2. The results of the analyses, where the values of $\bar{R}$ and times of the terminal points are given.

$\bar{R}_{beginning}$	starting time	$\bar{R}_{end}$	ending time
0.484	-1.87	0.624	7.17

The estimates of β are 2.19, -0.385 and -0.512, with variances 0.0203, 0.0271 and 0.0508 respectively. The maximal length of the interval is around 17 hours, when starting at 3:30 pm. The estimated length of the high intensive interval of an individual, when starting at 11:30 pm, is around 8 hours. The high intensive interval of the whole population is around 9 hours.

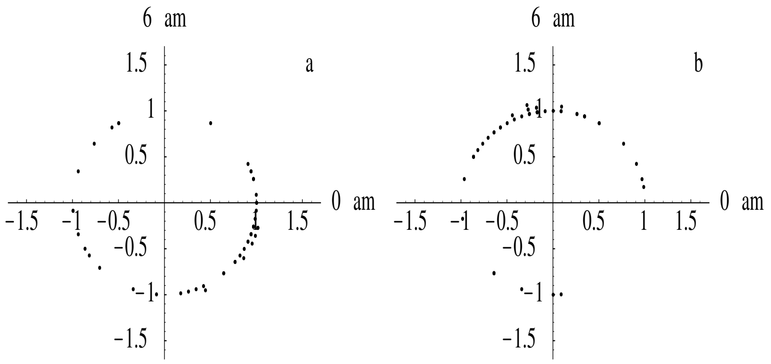

Figure 12: The circular coordinates of the individual's amplitude plotted around the unit circle, where (1,0) represents midnight and goes anti-clockwise. The first graph (a) shows how the beginning of the high intensive interval is distributed for the different individuals. The second graph (b) illustrates the corresponding data of the end point of the high intensive interval.

The result of the model-based method indicates a rather high proportion of the observations clustering around the mean direction. Table 4.3.2 indicates a high intensive interval between 10:07 pm to 7:10 am. The distribution of the data is shown in figure 12.

Both methods reject the null hypothesis, but indicate different intervals. The criteria-based method indicates a much shorter interval than the model-based method. Both intervals may be correct though. The first might show the interval of extreme peaks, while the second captures a more 'overall' high intensive interval. Both results are in accordance with the presumptions. The high intensive interval is believed to occur at night and going on until morning.

5 Conclusions

This paper presents two methods to test time dependence of peak frequency and amplitude of hormonial data. The first method is based on a test statistic. This method allows us to test the hypothesis that data is stationary. Then a procedure is developed to deduce when the high intensive interval will occur. The second method first divides the data into two intervals, a high intensive and low intensive respectively, with an ordinary least square method. Then the technique of directional data is used to estimate the mean direction and the concentration of data in that direction. After that we use a harmonic regression equation to describe how the length varies with its starting point.

The criteria-based method tests the hypothesis of stationarity with a bootstrap-like simulation. The model-based method gives a asymptotic test of the same null hypothesis. Both methods seem to work well, when dealing with dichotomous data as well as with continuous data. The model-based method may also be effectively used as a descriptive method characterising how GH behaves during its high intensive interval.

References

Diggle, P.J., AlWasel, I., (1997). *Spectral analysis of replicated biomedical time series*, Applied statistical journal of Royal Society Serie C, Vol. 46, No. 1, pages 31-60.

Diggle, P.J., Zeger, S.L., (1989). *A Non-Gaussian Model for Time Series With Pulses*, American Statistical Association, Vol. 84, No. 406, pages 354-359 .

Gou, W., Wang, Y. and Brown, M.B., (1999). *A Signal Extraction Approach to Modelling Hormone Time Series with Pulses and a Changing Baseline*, Journal of the American Statistical Association, Vol. 94, No 447, pages 746-755.

Hauffa, B.P. and Stolecke, H., 1994. *Receiver-Operated Characteristic Curve Analysis of Two Algorithms Assessing Human Growth Hormone Pulsatile Secretion (Pulsar, Cluster): Comparison of Peak Detection Efficacy*, Horm Res, 41, pages 169-176.

Jenkins, G.M., Watts, D.G., (1968). *Spectral analysis and its applications*, Holden-Day, California.

Komaki, F.,(1993). *State-space modelling of time series sampled from continuous processes with pulses*, Biometrika, 80, 2, pages 417-429.

Löfqvist, C. (1999). *Growth Hormone Measurements Methodological and Interpretational Aspects*, Göteborg University, Göteborg.

Mardia, K.V., Kent, J.T., Bibby J.M., (1979). *Mulitvariate Analysis*, Academic Press, San Diego.

Martha, P.M., Rogol, A.D., Veldhuis, J.D., Kerrigan, J.R., Goodman, D.W. and Blizzard R.M., (1989). *Alterations in the Pulsatile Properties of Circulating Growth Hormone Concentrations during Puberty in Boys*, JCE & M, Vol 69, No 3, pages 563-569.

O'Sullivan, F. and O'Sullivan, J., (1988). *Deconvolution of Episodic Hormone Data: An Analysis of the Role of Season on the Onset of Puberty in Cows*, Biometrics, 44, pages 339-353.

Some Explorative Methods Useful when Analysing Hormone Data

1 Introduction

There are a lot of challenges when approaching hormone data, because of the wide individual variations and the fact that the responses depend on a complex system of known (and unknown) substances and factors. The researchers have a need for explaining the behaviour, both at a population level and at an individual level. One example of issues of interest is whether the frequency of peaks changes during a 24-hour period or not. The same investigation can be done with amplitude. Another issue is how to estimate the exact amplitude and its corresponding time point. A third issue is to define if there is some periodicity or not.

In this paper we are going to discuss some explorative methods how to analyse growth hormone data. A method for adjusting peaks is described in chapter 2. When analysing continuous data at discrete intervals, we are bound to make mistakes. Some peaks are recorded before and some after their real appearance. This will cause the amplitude to be underestimated. In addition chapter 2 will present a method to calculate the curve form of a standardised pulse. Looking at data from growth hormone (GH) secretion one might notice that the individual level of amplitude varies substantially. The curve form of the pulses all show the same appearance, i.e. a sharp rise to the peak and then a slower decline. An estimation of this underlying curve might be of interest when comparing different individuals.
Another task to analyse is whether the frequencies follow a certain periodicity or not. We will discuss both population and individual periodicity. In reality periodicity is a subjective estimate, since the noise of measurements makes it very vague and indistinct. It is therefore difficult to quantify. A proposal how to analyse this will be given in chapter 3.

The results of previous investigations of periodicity give a somewhat shattered picture, some authors indicate a periodicity and others cannot find this. Pincus et. al. (1996) quantify the entropy with ApEn. ApEn is a statistic

that measures how well two patterns agree. Pincus et. al. conclude a gender difference, where females show a greater irregularity compared to males. Others (see for instance Hindmarsh et. al., 1988) cannot find this difference, but do confirm a periodicity within individuals.
In a well-written overview, Robinson (1990) points out the importance of periodicity, when treating growth hormone deficient (GHD) patients. He claims that an intermittent infusion of GH stimulated growth better compared to continuous infusions.

2 Estimation of the standardised curve form

To be able to adjust the peaks, it is assumed the curve follows a second order polynomial near its peaks (section 2.1). Then both the amplitude and its corresponding time point are re-estimated. Next, isotonic regression is used to estimate the curve form of the standardised pulse. A brief recapture of the theory of ordered restricted inference will be given in section 2.2 and in section 2.3 the curve form of the standardised pulse for 50 healthy children is calculated. It might be helpful to define some variables that will be used throughout this paper.

The treated methods deal both with dichotomised data, when analysing peak frequency and continuous data, when analysing amplitude. Consider a data set of n individuals, from which blood samples were taken every 20 minutes during a 24-hour period. We define every local maximum as a peak, where y_p describes the response value pth steps from the nearest peak and t_p describes the corresponding time point (measured in hours). An index over the n individuals denotes $i = 1, 2, \dots, n$.

We further suppose that observations for different individuals are independent while there may be dependence within individuals. It is also worth noting that the problem is circular, since the data is collected at specific time points during a 24-hour period. That is, the sequence of data does not stop at the last time point, but continues with a value estimated by the value of the first time point.

The above definitions will remain throughout the paper. Other notations will be defined under sections of interest. The required computer calculations have been performed in Mathematica, version 4.0.1.0.

2.1 Adjusting the peak estimates

As explained above, we have to consider misjudgement derived from the fact that we are dealing with continuous data at discrete time observations. Mauger et. al. (1995) define two types of error that might appear.
"False positive pulse" (FPP) defines the error when we have too tight measurements per half-life of a pulse. Hence the addition of noise will increase the number of peaks, which does not portray reality. "False negative pulse" (FNP) defines the error when we have been too sparse with measurements per half-life of a pulse. Hence we might miss a real pulse between two points of measurements. According to Mauger et. al. (1995), 3-5 points per half-life is the proper observational rate to reduce both FPP and FNP. A third source of error is misjudging the time when a peak appears, which causes the amplitude to be underestimated.

The two first errors may be restrained by choosing a sampling interval with respect to the half-time ($t_{\frac{1}{2}}$) of GH. GH consists of different isoforms, with different half-times, where we might discern two main groups (Growth et. al., 1994). One 'slow' group with a $t_{\frac{1}{2}}$ about 19-22 minutes and one 'fast' group with a $t_{\frac{1}{2}}$ about 12-16 minutes. Groth et. al. found that a sampling interval every 20 minutes gave a desirable result for the estimation of the total secretion rate as well as for determining the pulse pattern.

This contradicts the result of Mauger et. al. (1995) and their claim that 3-5 points per half-life is the proper rate to reduce both FPP and FNP. In spite of that, we will ignore the problem of FPP and FNP further, leaning on the results of Groth et. al.

One way to adjust for the third short-coming is to assume that the curve form at the top of a pulse follows a second order polynomial, defined by three points ($p = -1, 0$ and 1). That is, the second order polynomial is defined by its peak and the two surrounding values. The above model is a simplification that introduces a bias, since the real pulse form probably is asymmetric. However, the yield of a more complex model is questionable.

We estimate

$$\begin{aligned}
y_p &= a + bt_p + c((t_p - \bar{t})^2 - Q), \quad \text{where} \\
Q &= \frac{1}{3}\sum_{p=-1}^{1}(t_p - \bar{t})^2 = \frac{2}{3} \\
\hat{a} &= \bar{y}_p \\
\hat{b} &= \frac{\sum_{p=-1}^{1}(t_p - \bar{t})y_p}{\sum_{p=-1}^{1}(t_p - \bar{t})^2} = \frac{y_1 - y_{-1}}{2} \\
\hat{c} &= \frac{\sum_{p=-1}^{1}((t_p - \bar{t})^2 - Q)y_p}{\sum_{p=-1}^{1}((t_p - \bar{t})^2 - Q)^2} = \frac{y_1 + y_{-1} - 2y_0}{2}
\end{aligned}$$

The distance to the actual time point at the curve's maximum becomes;

$$t_{opt} = \frac{y_1 - y_{-1}}{6(2y_0 - y_1 - y_{-1})}, \tag{7}$$

where we assume that the equivalent distance between the observations is $\frac{1}{3}$ of an hour, which corresponds to the measurements taken every 20 minutes in our application. Thus the observation to the left of the peak is $t_{-1} = -\frac{1}{3}$, the observation at the peak is $t_0 = 0$ and finally the observation to the right of the peak is $t_1 = \frac{1}{3}$.

Hence, if our measured peak is greater than t_{opt} we assume that the actual peak is rising at the observed time. The actual peak is then to the right of the observed peak. The new time point is calculated by adding t_{opt} to the measured time point.

The new amplitude is received by adding

$$y_{opt} = \frac{(y_1 - y_{-1})^2}{8(2y_0 - y_1 - y_{-1})} \tag{8}$$

to the measured peak, y_0.

When estimating the standardised curve form of the pulse, all individual pulses are first adjusted according to the above method. Then the curve form is estimated by the technique of isotonic regression. The next section recapitulates some of the theory about isotonic regression.

2.2 Ordered restricted inference

The method to analyse ordered restricted inference is called isotonic regression. In this section we will only recapitulate those parts of the theory that are required for our future extension. Isotonic regression in itself constitutes a very fascinating area, whose entire usefulness will not be fully described here. The interested reader might consult Barlow et. al. (1972) or Robertson et. al. (1988) for a more detailed introduction.
The benefit of isotonic regression is increased flexibility compared to ordinary regression analysis. This is accomplished by taking into account that the response values are ordered. Thus the term isotonic is to be interpreted as "order preserving". Before describing the method further, we need to postulate two definitions;

Definition 1
Let X be a finite set $\{x_j : j = 1, 2, \ldots, k\}$. A nominal relation (after which we will use $\preceq$ instead of $\leq$) on X is a simple order on X if

1.	it is *reflexive*	:	$x \preceq x \quad \forall \quad x \in X$
2.	it is *transitive*	:	$x, y, z \in X, x \preceq y, y \preceq z \Rightarrow x \preceq z$
3.	it is *antisymmetric*	:	$y, x \in X, x \preceq y, x \succeq y \Rightarrow x = y$
4.	two elements are *comparable*	:	$x, y \in X \Rightarrow x \preceq y \quad \text{or} \quad x \succeq y$

Definition 2
Let X be a finite set $\{x_j : j = 1, 2, \ldots, k\}$ with the simple order $x_1 \preceq x_2 \preceq \ldots \preceq x_k$, where k is an arbitrary integer. In addition, let $g(x)$ be a real valued function on X and $f(x) \in \mathcal{F}$, where $\mathcal{F} = \{x \to f(x) : f(x)$ increases with $x\}$. Then $g^*(x)$ on X is an isotonic regression of g if

$$g^*(x) = \underset{\mathcal{F}}{min} \Big(\sum_{x \in X} \big(g(x_i) - f(x_i)\big)^2 \Big).$$

Hence the solution is the isotonic function, $g^*(x)$, that minimises the least square error of the difference between $g(x)$ and the class of all isotonic functions, $f(x)$, on X. When x is decreasing instead of increasing the class of functions, $f(x)$, is called antitonic instead of isotonic.

We may use the theory of ordered restricted inference to estimate the curve form of a standardised pulse if we split the curve into a rising and declining part. When a pulse is rising we get the isotonic case and when it is declining we get the antitonic case. The usual way to compute $g^*(x)$ is by PAVA (Pool Adjactive Violated Algorithm). The algorithm was deduced by Ayer et. al. (1955).

Figure 13: A flow chart over PAVA. The active block is abbreviated AB, the consequent block B_+ and PU is just a control variable.

The algorithm is illustrated in a flow chart at figure 13 and a brief description is given below.
Let us define a *block* as a set of consecutive elements, which is a subset of X. The active block is abbreviated AB, the consequent block B_+. As a first step, every individual element forms a block of its own. These blocks may then be joined if the criterion for pooling is reached. The set is then re-examined to see if the final partition is reached. If not, the procedure is repeated. The criterion for pooling is given by

$$AB_{new} = \begin{cases} AB & \text{if } Av(AB) \leq Av(B_+) \\ B_+ & \text{otherwise} \end{cases},$$

where $Av(x, y)$ is the weighted mean of x, y.

This algorithm might then be used to estimate the regression which describes how the ordered data increases or declines. In our case, we want to estimate how a pulse increases and declines respectively to be able to develop the curve form of a standardised pulse.

2.3 The general pattern of the curve form of a pulse

In this section we will give an example how to calculate the curve form of the standardised pulse. We will investigate a real data set. The data set consists of children of both sexes in prepuberty but it will be analysed as one dataset. It is not fully made clear whether sex is an important factor to explain the level of GH during the prepuberty phase. This factor will be ignored though, since the object is to exemplify the methods rather than claiming facts.

2.3.1 Patients

The sampling of the data used in the paper took place at the Queen Silvia Children's Hospital, Göteborg. The children were allowed and encouraged to activity and sleep as normally as possible and they were on an ordinary diet. The blood sampling commenced between 8:00 am and 9:00 am. The blood was continuously drawn with a pump system and sampled every 20 minute for a 24-hour period, that is each individual generated 72 sampling values.

GH was measured in samples of plasma by a Pharmacia GH radioimmuno assay using plyclonical antibodies and almost all samples used the WHO First International Reference Preparation for human GH (66/217). One sample used the WHO First International Reference Preparation for human GH (80/505). The latter was adjusted to the former standard according to Löfqvist (1999).

The study consisted of 50 premature children, 37 boys and 13 girls. The children were of Caucasian origin and mainly Swedish. Their ages differed between 5 and 14 years old (the range of the boys was 5-14 and of the girls 8-13). In the analysis of the data undetectable values were set to the lowest detectable and missing observations were imputated by the mean of the nearest non-missing values.

2.3.2 Results

Before applying the so called PAVA algorithm to the data set, some adjustments are made. First the amplitude and its time point (i.e. when the peak occurs) are adjusted according to the method described in section 2.1. Then we proceed by calculating the ratios between the response values of the a peak and its surrounding values. After that the corresponding difference in time (measured in hours) is calculated. The influence of a peak is assumed to have diminished after three previous and three subsequent values respectively. That is, we calculate the difference up to the six points that surround a peak.
To minimise variance, the logarithm of the ratio of the amplitude values is taken. The difference between times is not transformed. Finally the mean is calculated by means of isotonic regression on the sorted values with respect

to the difference of time.

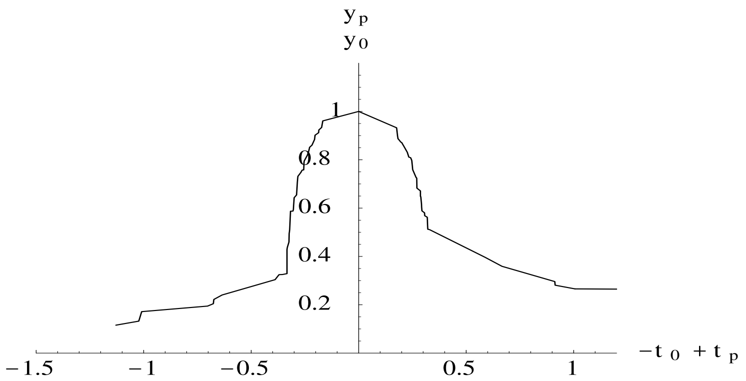

Figure 14: The curve form of a standardised pulse, based on 50 children.

The result is given in figure 14 after a 'back transformation' of the logarithmic data. The characteristic form of the pulse is a rather sharp rise followed by a slower decline. We might note that the half-time for the declining part is approximately 20 minutes (i.e. the time difference is approximately 0.3). This is in accordance with previous results (Growth el. al., 1994).

3 Investigation of periodicity

When analysing periodicity, we must first define if we mean between individuals or within individuals. The first is called population periodicity (see section 3.1) and the latter individual periodicity (see section 3.2). In section 3.3 an alternative procedure to evaluate the population periodicity is presented. Graphical procedures will be discussed only, since periodicity is a quantity difficult to define mathematically.

We have a very naive definition of peaks, which is sensitive to a number of errors. The first type of error is 'double peaks', i.e. two peaks may appear very close to each other, where one is a true peak and the other one is due to noise. Another error may be termed 'baseline peaks', i.e. some noise at the baseline may be interpreted as a peak. The third type of error is a misplaced peak, i.e. a true peak is misplaced due to some noise. The latter may be adjusted by the method mentioned in section 2.1. This type of error will not affect the estimation of periodicity since we are measuring discrete intervals (every 20 minutes), but the two other errors do have an impact on the procedures presented.

3.1 Periodicity at a population level

A simple method is to analyse the number of points of measure until a new peak is reached. This will result in an array of numbers. Let $N_{i,j}$ denote the number of measurements between the $(j-1)$:th and the j:th peak for individual i.

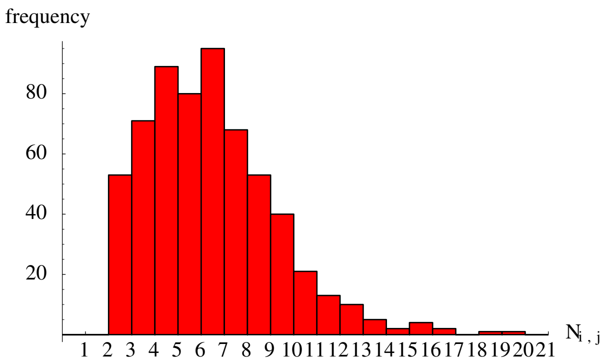

Figure 15: Histogram of the number of measurements between two peaks ($N_{i,j}$), calculated on the whole population.

We might illustrate the method with the same individuals as in section 2.3. Figure 16 shows the histogram of the responses (the number of intervals between two peaks). Straight from the figure we see that the probability to get, for instance $N_{i,j} = 4$ (i.e. corresponding to a periodicity of 1 hour and 20 minutes), is almost as big as getting $N_{i,j} = 6$ (i.e. corresponding to a periodicity of 2 hours).

That is, a simple histogram reveals how the variables of the $N_{i,j}$ are distributed. The sharpness of the distribution indicates the strength of periodicity. If we have a deterministic periodicity, that is a peak every certain time point then all the variables of the $N_{i,j}$ are equal. If not, the values will differ to a different extent depending on the strength of the periodicity.

3.2 Periodicity at an individual level

In this section we will use the same method at an individual level. Even if we are unable to distinguish a periodicity between individuals it is possible that each individual follows a certain rhythm. Figure 16 shows the histograms of the variables of the $N_{\cdot,j}$ of three of the individuals.

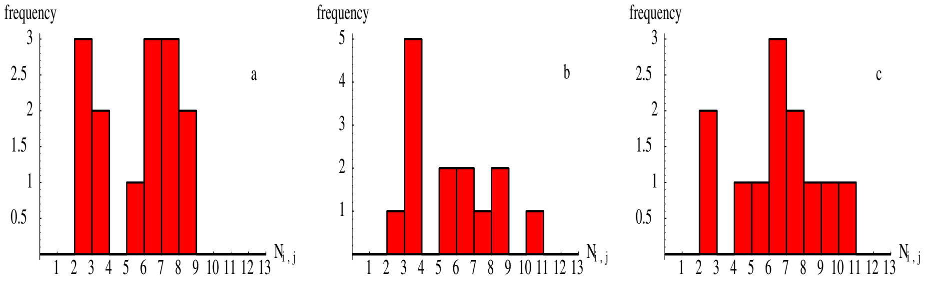

Figure 16: The histograms over $N_{\cdot,j}$ for some of the individuals from the same data set as above.

In the first figure it (see figure a) appears to be two periodicity levels, one at every 40 minutes and one at every two hours to two hours and 20 minutes respectively. The individual represented in the next figure (see figure b) shows a periodicity every one hour and the third individual (see figure c) gives a very vague appearance in respect to periodicity. It must be noted that the number of observations per individual are low, but a bold conclusion of the above graphs might be that there is no periodicity in data. Though our naive definition of a peak may have a major impact on the result. We define a peak as every local maximum in the sample. A false peak may easily affect the result. Further studies are therefore needed.

To give some examples how the distribution of the variables of the $N_{\cdot,j}$ varies with periodicity, we simulate different possibilities. Below three cases are described.

In the first case (figure a and b) there is an equal possibility (1/2) to peak every second or third time. In the second case (figure c and d) a peak may

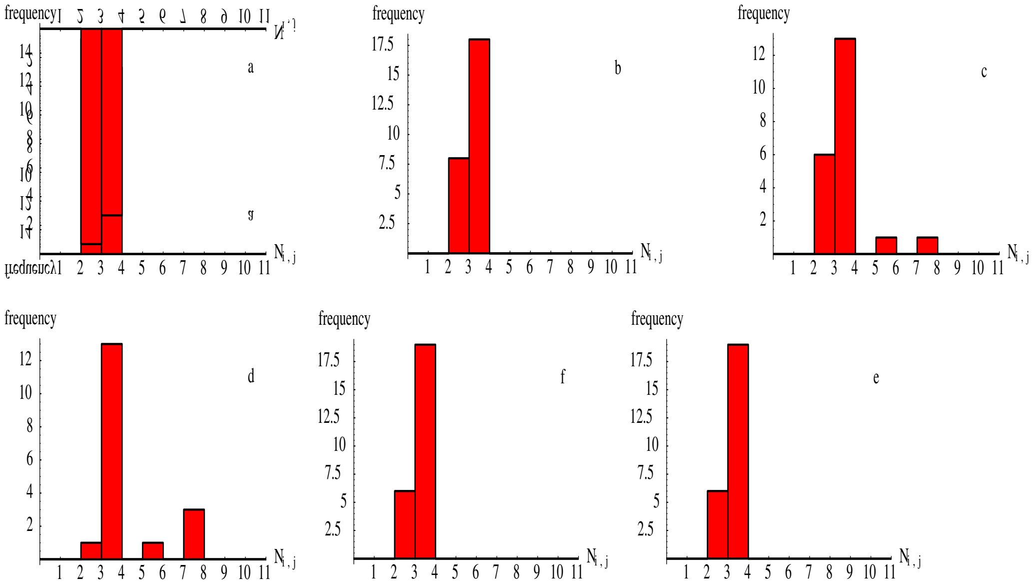

Figure 17: The histograms over $N_{.,j}$ for the simulated values.

occur every second time (with possibility of 1/5), every third time (with possibility of 3/5), every fifth time (with possibility of 1/10) and every seventh time (with possibility of 1/10). In the third case (figure e and f) a peak occurs every second time with a possibility of 1/3 and every third time with a possibility of 2/3.

The figures show very sharp distributions, which might be expected in an ideal world as above. The main conclusion of the simulations is that the above method to discover periodicity might be an interesting tool to develop further.

3.3 A complement to investigating the periodicity

A drawback with the previous method is its sensitivity to false peaks. The procedure makes it possible for a false peak to conceal a true periodicity easily. One result might be that the true underlying periodicity is underestimated or undiscovered.

Another descriptive way to analyse periodicity is to calculate the number of neighbouring peaks at a certain distance, τ, to a peak. That is, we count the number of neighbouring peaks at distance τ and plot these values v.s. τ, where $\tau = 1, 2, \dots, t$. Even if the resulting figure will be interfered by false peaks we will not miss or underestimate an underlying periodicity. We analyse the same data set as before.

The result is illustrated in figure 18, where the relative frequency is plotted v.s. the distance between peaks. Since noise increases with distance we have chosen to neglect data corresponding to a distance above $\tau = 36$.

Figure 18: The plot of the relative frequency of $N_{i,j} = \tau$ v.s. τ of the 50 children investigated.

The result is in accordance with the previous method. If there is a periodicity, the best estimate is around every second hour (i.e. $\tau = 6$), though the slow rate of decline is an indication of no periodicity.

4 Conclusions

This paper has been dealing with explorative methods analysing GH data. The first chapter describes how to adjust for the errors which appear when analysing continuous data at discrete intervals. The chapter also describes how to calculate the curve form of a standardised pulse. This can be used for comparative studies, i.e. when analysing if there are any differences between puberty phases, gender, GHD children vs normal children and so forth.

The third chapter suggests a method to test whether data contains periodicity or not. There is a supposition that such a periodicity exist. The method discussed is merely a descriptive tool to measure periodicity. This is evaluated by a transformation of data. Instead of analysing data, we observe the number of steps needed to reach a peak, termed $N_{i,j}$. The distribution of the variables of the $N_{i,j}$ are then analysed. A drawback is that this method may underestimate the periodicity if there are a lot of 'double peaks', i.e. where one peak is due to noise. Another procedure, not so sensitive to that type of errors, is to plot the number of the variables of the $N_{i,j}$ equal τ vs τ, where τ

is the distance between two peaks. We cannot, with our methods, distinguish any periodicity. This may still be due to our naive definition of a peak as a local maximum. The peaks due too noise might disturb the distribution of $N_{i,j}$ to much. Further developments are needed.

References

Ayer M., Brunk, H.D, Ewing, G.M., Reid W.T. and Silverman, E. (1955). *An empirical distribution function for sampling with incomplete information*, Annuals of Mathematical Statistics, 26, 641-647.

Barlow, R.E., Bartholomew, D.J., Bremner, J.M. and Brunk, H.D. (1972). *Statistical Inference under Order Restrictions. The theory and Application of Isotonic Regression*, John Wiley and Sons, The Pitman Press, Bath.

Groth, T., Rosberg, S. and Albertsson-Wikland, K. (1994). *Estimation of Growth Hormone Secretory Patterns in Children with Use of a Numerical Deconvolution Technique: Experimental Design with Use of Computer Simulation*, Horm. Res., 42, 245-252.

Hindmarsh, P.C., Matthews, D.R., Brook, C.G.D. (1988). *Growth hormone secretion in children determined by time series analysis*, Clinical Endocrinology., 29, 35-44.

Löfqvist, C. (1999). *Growth Hormone Measurements Methodological and Interpretational Aspects*, Göteborg University, Göteborg.

Mauger, D.T. and Brown, M.B. (1995). *A comparison of methods that characterise pulses in a time series*, Statistics in medicine, Vol 14, 311-325.

Pincus, S.M., Gevers, E.F., Robinson, I.C.A.F., Van Den Berg, G., Roelfsema, F., Hartman, M.L. and Veldhuis, J.D. (1996). *Females secrete growth hormone with more process irregularity than males in both humans and rats*, Am J Physiol, No 270, 107-114.

Robertson, T., Wright, F.T. and Dykstra, R.L. (1988). *Order Restricted Statistical Inference*, John Wiley and Sons, Tiptree, Essex.

Robinson, I.C.A.F. (1991). *Chronopharmacology of growth hormone and related peptides*, Advanced Drug Delivery Reviews, No 6, 57-82.

Strömberg, U. (1992). *Isotontic regression based on asymmetric distance functions with applications to biological data*, Dissertation, Lund, Lund university.