

LONGITUDINAL AND GESTATIONAL EFFECTS
OF MINERALS IN HUMAN MILK

CENTRE FOR NEWFOUNDLAND STUDIES

**TOTAL OF 10 PAGES ONLY
MAY BE XEROXED**

(Without Author's Permission)

EDWARD MAXWELL HOYLES



LONGITUDINAL AND GESTATIONAL EFFECTS
OF
MINERALS IN HUMAN MILK

by

©Edward Maxwell Hoyles, B.Sc.

*A Practicum report submitted to the School of
Graduate Studies in partial fulfillment of
the requirement for the Degree of Master
of Applied Statistics*

Department of Mathematics and Statistics
Memorial University of Newfoundland

August, 1998

St. John's, Newfoundland

Abstract

Human milk is often fortified with the appropriate nutrients including minerals to allow pre-term infants and their families to enjoy the benefits conveyed by the feeding of breast milk while delivering an optimal nutrient supply to the baby. But, the amount of nutrients needed for fortification depends in general on the recommendations made by the pediatric societies based on the information available for full-term mothers. This calls for the investigation to examine the difference that may exist between the mineral contents in the milk of the full-term and pre-term mothers. Also it is necessary to check the longitudinal patterns of mineral contents to examine when and how mothers milk needs fortification. It is generally known that a good number of minerals do not follow a symmetric distribution. This asymmetric nature is needed to be carefully accommodated in order to develop statistical test for the gestation effects.

Acknowledgments

Special thanks to my supervisor, Dr. Brajendra Sutradhar, for his guidance and support. Thanks also to Dr. James Friel from the Department of Biochemistry who provided the data set. The encouragement of fellow graduate students cannot go without mention. Finally, I would especially like to thank my parents for continued support and encouragement throughout my graduate program.

Contents

1	Background of the Problem	1
1.1	Introduction	1
1.2	Objectives of the Study	6
2	Exploratory Analysis of the Trace and Ultratrace Elements	8
2.1	Exploratory Analysis of the Trace Elements	9
2.1.1	Distributional Patterns	9
2.1.2	Longitudinal Patterns	11
2.2	Exploratory Analysis of the Ultratrace Elements	13
2.2.1	Distributional Patterns	13
2.2.2	Longitudinal Patterns	14
3	Gestational Effects: TIME AS A SPECIFIC FACTOR	15
3.1	Tests for Gestation Effects at a Given Week	17
3.1.1	An Ad-hoc Median Test	17
3.1.2	Median Test Based on Jackknifed Dispersion Estimate	21
3.1.3	A Permutation Test	25
3.2	Tests for Gestation Effects for all Weeks Combined: A “Work- ing” Independence Approach	31
3.2.1	Ad-hoc Test	32
3.2.2	Weighted Ad-hoc Test	34
3.2.3	Comparison of Median When Dispersion is Computed by Delete-1 Group Jackknife Technique.	38

4 Gestational Effects: TIME AS A NON-SPECIFIC FACTOR	41
4.1 Delete-1 Cluster/Mother Jackknifing	43
4.2 Ad-hoc Median and Mean Tests: TIME AS A NON-SPECIFIC FACTOR	45
4.2.1 Modeling Gestation Effects	45
5 Gestation Effects of Macrominerals: Time as a Non-specific Factor	57
5.1 AR(1) Model and Estimates of the Nuisance Parameters . . .	59
5.2 Score Test for Mean Gestation Levels	62
6 Concluding Remarks	65

Chapter 1

Background of the Problem

1.1 Introduction

Low birth weight infants, defined as weighing less than 2500 grams at birth, comprise 5 to 8 percent of all live born infants in Canada. A sizeable proportion weigh less than 1500 grams and these infants are referred to as very low birth weight infants. The majority of these infants are born prematurely within less than 37 weeks of gestation.

Human milk has been widely used as the sole nutritional source for feeding low birth weight premature infants. While it may be desirable to feed human milk to low birth weight premature infants, it may not be desirable as the

only nutritional source. In other words, it is generally recognized that breast milk cannot support the full range of nutritional requirements for the low weight premature infants. Furthermore, breast milk alone may cause the infants tissue to be deposited at a different rate as would occur if the infants were carried to term.

A viable solution to the above problem is to fortify human milk with the appropriate nutrients to allow the premature infants to enjoy the benefits conveyed by the feeding of breast milk while delivering an optimal nutrient supply to the infant (cf. Modanlou *et. al.* (1986), Kashyap *et. al.* (1990)). It may be acceptable medical practice to feed the premature infant with his/her own mother's milk alone at high volume while monitoring closely for and appropriately treating nutritional inadequacies identified. It is, however, preferable to routinely supplement human milk fed to very low birth weight infants to achieve a more optimal nutritional profile while avoiding the effects of dietary deficiency (cf. Fomon *et. al.* (1977)). In this regard, there seems to be little debate about whether human milk should be fortified if it is to be the sole source of enteral nutrition for very low birth weight infants; there is considerable controversy, however, about when, how and with what. This stems from uncertainty about the long-term consequence of nutritional "in-

adequacy" identified in the perinatal period, potential interactions between components contained within breast milk and fortifiers, extreme variability in the composition of expressed breast milk and marked clinical heterogeneity between infants (cf. Atkinson *et. al.* (1990)).

A solution is to fortify human milk with the appropriate nutritional requirements recommended by the Canadian pediatric Society (CPS). When differences exist between the recommended nutrient intakes for a stable, growing premature infant and the amount of nutrient that would be received if human milk alone were fed at standard fluid intakes, fortification may be recommended. The nutrient level provided by human milk is derived from the mean value determined when expressed breast milk from mothers delivering prematurely is analyzed. When used according to manufacturer's instructions, the sum of the amounts of any nutrient generally fed to the premature infant should come within the range of the recommended nutrient intakes and not exceed the maximum "safe" level of intake. The amount of fortification can be derived from the difference between the recommended CPS nutritional requirements (which is based on the nutrient levels in the milk of full-term mothers) and the mean/median nutritional levels of the milk of pre-term mothers.

Note that as the nutrient level of the milk from mothers of premature infants may vary mainly during the early stage of lactation, the use of simple mean/median nutrient level (determined from the expressed breast milk of the mothers of premature infants) to assess the quantity of human milk fortifier needed over time, may not be valid. Consequently, the mean/median level based prescribed quantity of human milk fortifier may affect the growth of the premature infants negatively. Further since the importance of minerals in the nutritional management of infants is now widely recognized, it seems appropriate to study the longitudinal changes in the milk composition during the first three months of lactation, and to evaluate the effect of gestational length on levels of macro, trace and ultratrace elements in human milk. The macrominerals in the human milk are: calcium (Ca), magnesium (Mg), rubidium (Rb), and strontium (Sr). These macrominerals are measured in micrograms per milliliter ($\mu g/ml$ or PPM). The trace elements in the human milk are: zinc (Zn), copper (Cu), manganese (Mn), nickel (Ni), molybdenum (Mo), and cobalt (Co); and they are measured in $\mu g/ml$ or PPM. The other minerals in human milk, namely, tin (Sn), lead (Pb), cadmium (Cd), cesium (Cs), barium (Ba), lanthanum (La), and cerium (Ce) are known as ultratrace minerals and they are measured in nanograms per millilitre (ng/ml or PPB).

Recently, Sutradhar, Dawson and Friel (1998) studied this problem and discussed, in details, the longitudinal changes and gestation effects for all macrominerals. They have also studied to some extent the longitudinal and gestational behavior of some trace and ultratrace minerals. These trace and ultratrace minerals were found to be nonnormally distributed. Consequently, they were analyzed based on some ad-hoc correlation computational procedures and ad-hoc normal tests. In this practicum, we study these non-normally distributed trace and ultratrace minerals in details.

More specifically, we discuss the pattern of mineral concentration for all non-normal trace elements. We also consider non-normal ultratrace elements, but exclude cadmium (Cd), cesium (Cs), barium (Ba), lanthanum (La), and cerium (Ce), as the elements are known to contribute very little. Finally, for the sake of completeness, we discuss in brief about some macrominerals, specifically, to examine their gestation effects based on certain optimal statistical test procedures.

A total of 288 observations (breast milk) for each of the above minerals were collected by James Friel and his group at the Department of pediatrics at the Janeway Hospital. The data we analyze are from 43 lactating mothers from St. John's, Newfoundland. The breast milk was collected for the first 8

consecutive weeks following birth and a final 12_{th} week after birth. Of the 43 women 19 were mothers of pre-term infants and 24 were mothers of full-term infants. There were 136 observations from the full-term mothers and 152 observations from the pre-term mothers. Samples were not available from all mothers for every week during the duration of the study, which resulted in unbalanced sample sizes for different mothers. As previously mentioned, there every week there were approximately 19 and 24 mothers under the pre-term and full-term mothers respectively. So, as the pbservations are independent, the use of normal test statistics based on such sample sizes is reasonable. A simulation study could be done to validate the normality of different test statistics used in the practicum for small sample sizes, but it was not chosen in the present study.

1.2 Objectives of the Study

The objective of the practicum is to examine the longitudinal patterns of the minerals present in mothers milk. Also, we examine the gestation effects taking the longitudinal correlations into account.

The specific plans of the practicum is as follows. In Chapter 2, we analyze

graphically the trace elements and ultratrace elements. The graphical procedures adopted in Chapter 2 are standard boxplots, histograms, q-q plots, line graphs and smoothed line graphs. In Chapter 3, we study the gestational effects using time as a specific factor. The fourth chapter studies gestational effects using a cluster approach with time as a non-specific factor. In Chapter 5, we reconsider some normal minerals studied in detail by Sutradhar, Dawson and Friel (1998). This is mainly done to verify their results about gestational effects but based on certain optimal statistical test procedures. The final chapter contains the concluding remarks.

Chapter 2

Exploratory Analysis of the Trace and Ultratrace Elements

As mentioned earlier, there are six trace elements in mothers' milk. These minerals are Zn, Cu, Mn, Ni, Mo and Co. Among these elements, the longitudinal and gestational effects of Zn and Cu were studied in details by Sutradhar *et. al.* (1998). These minerals are therefore, not included in our analysis. Among the 7 ultratrace elements, as mentioned earlier, we include two elements, namely Sn and Pb. The distributional patterns of the trace elements are discussed first in Section 2.1.1 and their longitudinal patterns of gestational effects are discussed in Section 2.1.2. Similarly, the distributional

patterns of the trace and ultratrace elements Sn and Pb and their longitudinal patterns are discussed in Section 2.2.1 and 2.2.2 respectively. The exploratory analyses are done through various graphical displays such as standard boxplots, histograms and q-q plots for the distributional patterns, and line graphs and smoothed line graphs, for the longitudinal pattern. The graphs for distributional and gestational patterns are located in Figures 2.1.1 to 2.6.5. For convenience, the figures are numbered in a systematic way to reflect the mineral and the graphical pattern of the mineral. For example, Figures 2.1.1 to 2.1.5 will correspond to five different graphical patterns of the first mineral, Mn. Similarly Figures 2.2.1 to 2.2.5 will correspond to five different graphical patterns for the second mineral, Ni.

2.1 Exploratory Analysis of the Trace Elements

2.1.1 Distributional Patterns

As mentioned above, for each mineral we display five different graphical patterns. First three graphs will show distributional patterns and the remaining

graphs will exhibit longitudinal patterns.

More specifically, for a given mineral, the five different graphs will report five different patterns as follows. The first graph will show boxplots for both full-term and pre-term mothers for any given week. The second graph displays the boxplots, histograms and q-q plots for the same mineral for both full-term and pre-term mothers for the combined weeks. In the third graph, histogram and q-q plot are given for the combined groups and combined weeks. The last two graph will exhibit the median plots and smoothed median plots over the weeks, for the given mineral.

The first boxplots graph for each of the four mineral (Mn, Ni, Mo and Co) shows that the median is generally located either closer to the first or third quantile value. The distributional patterns for trace elements for almost all weeks for both full-term and pre-term groups appear positively skewed. We further notice that the median of the minerals in pre-term and full-term mothers milk are generally different. The second graph for each mineral indicate the non-normality for both pre-term and full-term groups for combined weeks. The boxplot, histogram as well as the q-q plot in the second graph for each of the four minerals show that the distributions for pre-term and full-term are positively skewed. Similarly, overall distributional

behavior for combined groups and combined weeks are also positively skewed for each mineral as indicated by the third graph for each of them.

2.1.2 Longitudinal Patterns

As the distribution of the trace elements considered above are generally positively skewed, in this section we exhibit the median concentrations of these minerals instead of mean concentrations. For this purpose, we display two types of median plots. First, we plot the ordinary medians against every week. Second, we compute the smoothed medians by S-plus and provide them in the second graph. More specifically, to compute the smoothed medians, a method of running medians known as 4(3RSR)2H is used (cf. Venables and Ripley (1996)). Within this procedure twicing is performed. This is a process of smoothing, computing the residuals from the smoothed values, smoothing these residuals and adding the two smoothed series together. The fourth and fifth graph for each mineral displays these median concentrations, and smoothed median concentrations of the particular mineral.

The median concentration for Mn, as indicated by figure 2.1.4, for both gestation periods appear to be relatively constant over all time periods with the concentration being higher for the full-term groups. This constant pat-

tern is more explicit from the smoothed median plot for Mn, as shown in figure 2.1.5. The median concentrations of Ni for both pre-term and full-term as displayed in figure 2.2.4 are not constant over the period of weeks and they do not apparently follow any pattern. The smoothed concentration in figure 2.2.5, however, shows that the Ni concentrations are generally decreasing for the pre-term group whereas the concentration is increasing for the full-term. Now figures 2.3.4 and 2.3.5 for Mo exhibit that the concentration for the full-term group decreases significantly over time while for the pre-term group decreases sharply for the first 4 weeks and then appear to increase slowly for the rest of the time. The pre-term group has an higher concentration at each time period as compared to the full-term group. The plot of the median concentration in figure 2.4.4 for Co is difficult to interpret. The median concentrations for full-term appear to be slightly lower for time periods 1 through 7 but higher for time periods 8 and 9. Figure 2.4.5 for smoothed median concentration, however, shows clearly that Co concentration is slowly decreasing over time, the pre-term concentrations are being higher, as compared to the full-term.

2.2 Exploratory Analysis of the Ultratrace Elements

2.2.1 Distributional Patterns

In this Section, we examine the distributional pattern of two ultratrace elements Pb and Sn for both the pre-term and full-term mothers. These two elements are selected in the study, as among the 7 ultratrace elements, they are generally considered more important than the remaining.

The first boxplot graph for each mineral indicates that similar to the trace elements, the median concentrations for both pre-term and full-term groups for the ultratrace elements are generally different for a given week. The weekly distributional patterns for the ultratrace elements for both full-term and pre-term groups appear positively skewed. The second graph for both minerals indicate that the mineral concentrations are non-normality distributed for both pre-term and full-term groups for combined weeks. Overall distribution of the minerals for the combined groups and combined weeks are also seen to be asymmetric.

2.2.2 Longitudinal Patterns

The median concentration for Sn is very difficult to interpret with no apparent pattern. The smoothed median plot, however, shows that median concentration for the pre-term group remained constant for time periods 1 through 4 then increases until time period 6, followed by a constant pattern over the remaining weeks. For the full-term group, however, there is an increase until time period 6, and then it abruptly decreases. The median concentration for Pb indicates that for the pre-term group, there is a non detectable amount of this mineral until week 6 where it slightly increases. For Pb, it is clear from both figures 2.6.4 and 2.6.5 that there is no Pb in the milk of pre-term mothers, whereas, for the full-term group, the Pb concentrations appear to increase sharply up to week six and then decrease sharply over the remaining periods.

Chapter 3

Gestational Effects: TIME AS A SPECIFIC FACTOR

For the purpose of fortification, it is important to see the difference between the mineral concentrations in the milk of two groups of mothers. In this chapter, we first examine the significance of the difference of mineral concentration for each week for a given mineral. For this, we test the equality of the median concentration of minerals by employing some nonparametric tests. The reason for using the nonparametric tests is this that all the trace and ultratrace elements as discussed in chapter two were found to be asymmetrically distributed. Note however that, the usual median tests for two

samples are generally based on certain regularity conditions which may not be satisfied always for all the above trace and ultratrace elements. With this in view, we have modified some of the widely used nonparametric tests and applied them to test the difference between the medians. These modified tests are: (1) an ad-hoc median test; (2) a median test based on jackknifed dispersion estimate and (3) a permutation median test.

Second, in Section 3.2, we combine the minerals of all weeks and then test the difference between the medians of two combined groups. In developing this type of test, we make a “working” independence assumption between the minerals of mothers at any two points of time. Further, in some cases we assume that the minerals under each time point constitute a stratum. Finally, based on either or both of the above assumptions we develop the following three tests for testing the median differences. The tests are: (1) an ad-hoc median test; (2) a weighted ad-hoc test and (3) comparison of median when the dispersion is computed by delete-1 group jackknife technique.

3.1 Tests for Gestation Effects at a Given Week

3.1.1 An Ad-hoc Median Test

The traditional median tests, for example, Tukey's quick test, median test, and Mann-Whitney test are based on the signs and/or ranks of the observations. These tests, therefore, ignore the dispersion of the data in testing the difference of the median. In order to take the dispersion of the data into account, we, in this section, derive a modified ad-hoc median test, which is quite similar to the two sample z or t -test but tests the difference in the medians instead of the means. The dispersion of the difference between two sample medians are computed in an ad-hoc way. Also, as the distributions involved are asymmetric, the suggested test statistic (which is quite similar to the z statistic) is treated as a normal test statistic, mainly under the assumption that the samples are large. We now formulate this test as follows.

Let y_{ij} be the mineral concentration in the milk of the $j_{th}(j = 1, 2, \dots, n_i)$ mother of the $i_{th}(i = 1, 2)$ group. Here group 1 will refer to the full-term mothers and group 2 will refer to the pre-term mothers. Also let m_i be the median (Med) of $\{y_{ij}\}$, that is, $m_i = \text{Med}\{y_{ij}\}$. Next, suppose that s_i

measures the dispersion of the minerals $\{y_{ij}\}$ for the i_{th} group. A suitable formula for s_i (cf. Hill and Holland (1977)), for the asymmetric data, is given by

$$s_i = \frac{\text{Med}(|y_{ij} - m_i|)}{0.6745}, \quad j = 1, \dots, n_i; \quad i = 1, 2. \quad (3.1)$$

We now construct the ad-hoc test which takes the dispersion of the minerals of each mother into account, as opposed to the traditional median test. The test statistic for testing whether the population median, M_1 , of the full-term group is equal to the population median, M_2 , of the pre-term group, i.e. $H_0 : M_1 = M_2$, is given by

$$z^* = \frac{m_1 - m_2}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}. \quad (3.2)$$

The test based on this z^* statistic may be treated as a normal test, where z^* has mean 0 and variance 1.

We apply this ad-hoc test procedure to all trace and ultratrace elements and provide the z^* values in Tables 3.1(a) and 3.1(b). The results shown in Table 3.1(a) are calculated using all observations for a given mineral for

a particular week, while results in Table 3.1(b) are calculated with outliers removed. Here an observation is determined as an outlier based on the diagnostics done in Splus, by comparing its magnitude with 1.5 times the inter-quartile range (IQR) of the data.

It is clear from Table 3.1(a) that a significant difference exists between full-term and pre-term groups for Mn in the first five weeks. For the remaining weeks, the differences appear to be insignificant. For Co, there exists a significant difference at the beginning (for week 1). The differences appear to be insignificant for weeks 2 through 5, followed by significant differences for the weeks 6 through 8. There does not appear any significant differences for Mn for the first few weeks but the medians between the groups appear to be different for the latter weeks. The behavior of Ni is similar to that of Co. For Pb, the differences between the groups appear to be significant at each week except week 1 where the test statistic was indeterminate ($\frac{0}{0}$). For Sn, there is no apparent difference at any time period.

Table 3.1(a): The values of the z^* test statistic (3.2) for testing $H_0 : M_1 = M_2$ for a given week for trace and ultratrace elements (based on complete data).

	Trace Elements				Ultratrace Elements	
	Mn	Co	Mo	Ni	Pb	Sn
wk 1	3.22202	-3.304362	1.616899	2.484899	ind	-0.4289783
wk 2	2.608291	-1.341041	1.925039	-0.2680191	2.698	0
wk 3	2.049371	-0.7773048	0.6181895	0.5710875	2.698	0.553014
wk 4	2.289925	0	0	-0.6487321	2.612327	-0.7904583
wk 5	2.726551	0	0	0.7407319	3.483103	0.4328899
wk 6	1.710021	-2.698	3.753503	-2.122301	2.612327	0.1321111
wk 7	1.372717	-2.431944	1.102397	0.2715604	2.459425	-1.462838
wk 8	1.908318	2.612327	2.431944	-2.612327	2.612327	0.442415
wk 12	1.815556	0	3.504805	-1.739196	2.336537	-1.386031

When possible outliers are removed, the median differences appear to be similar to the case described above, where all data points remain in the analysis. For the case when outliers are excluded the test statistic has, however, 2 indeterminate values for Pb.

Table 3.1(b): The values of the z^* test statistic (3.2) for testing $H_0 : M_1 = M_2$ for a given week for trace and ultratrace elements (outliers excluded).

	Trace Elements				Ultratrace Elements	
	Mn	Co	Mo	Ni	Pb	Sn
wk 1	3.25088	-3.234788	2.089862	1.513214	ind	-0.8850812
wk 2	2.194718	-2.940077	1.25806	-2.344525	ind	0
wk 3	1.83414	0.767158	2.190439	0.2327054	2.612327	0
wk 4	2.457447	0	0	1.664419	2.612327	0
wk 5	2.631189	-1.160012	2.612327	-0.6820365	3.647917	0.3625329
wk 6	1.99087	-2.698	3.385963	-2.736793	2.523748	0.1315045
wk 7	1.719475	-2.237063	1.773229	0	2.459425	-0.7920271
wk 8	1.640973	2.612327	2.336537	-2.336537	2.523748	-1.858703
wk 12	1.815556	0	3.504805	-1.653222	2.237063	-1.386031

3.1.2 Median Test Based on Jackknifed Dispersion Estimate

In order to construct the ad-hoc median test in the previous section, we computed the dispersion for the difference in medians in an ad-hoc fashion. While this computation for the dispersion appears to be reasonable, there are other methods for calculating the dispersion of this type of asymmetric data. The jackknife dispersion estimate is a well known technique to compute such dispersions.

In this section we use the jackknife procedure to calculate the dispersion of the median m_i for the i_{th} ($i = 1, 2$) group as follows where $i = 1$ refers to the

full-term mothers and $i = 2$ refers to the pre-term mothers. This procedure involves deleting the j_{th} observation from the data and recalculating the parameter of interest (the median for the non-normal minerals) based on the remaining $n_i - 1$ observations. Let $m_{i(j)}$ denote this median yielding the average of the n_i recalculated medians for the i_{th} group as

$$m_{i(.)} = \frac{\sum_{j=1}^{n_i} m_{i(j)}}{n_i}. \quad (3.3)$$

Next, for $j = 1, \dots, n_i$, construct the n_i pseudo-values \tilde{m}_{ij} given by

$$\tilde{m}_{ij} = m_i - (n_i - 1)(m_{i(j)} - m_i) \quad (3.4)$$

where m_i is the median of the complete data set of the i_{th} group, yielding the average of \tilde{m}_{ij} 's as

$$\begin{aligned} \bar{\tilde{m}}_i &= \frac{\sum_{j=1}^{n_i} \tilde{m}_{ij}}{n_i} \\ &= m_i - \frac{n_i - 1}{n_i} \left(\sum_{j=1}^{n_i} m_{i(j)} - n_i m_i \right) \\ &= n_i m_i - (n_i - 1) m_{i(.)}. \end{aligned} \quad (3.5)$$

Now the jackknife dispersion estimate (cf. Effron and Gong (1983)) of m_i is

defined as

$$\begin{aligned}
\hat{V}(m_i) &= \frac{1}{n_i(n_i-1)} \sum_{j=1}^{n_i} (\tilde{m}_{ij} - \tilde{m}_i)^2 \\
&= \frac{1}{n_i(n_i-1)} \sum_{j=1}^{n_i} (m_i - (n_i-1)(m_{i(j)} - m_i) - n_i m_i + (n_i-1)m_{i(\cdot)})^2 \\
&= \frac{1}{n_i(n_i-1)} \sum_{j=1}^{n_i} ((1-n_i)(m_{i(j)} - m_{i(\cdot)}))^2 \\
&= \frac{n_i-1}{n_i} \sum_{j=1}^{n_i} (m_{i(j)} - m_{i(\cdot)})^2.
\end{aligned} \tag{3.6}$$

In the manner similar to that of (3.2), the test statistic for testing whether the population median, M_1 , of the full-term group is equal to the population median, M_2 , of the pre-term group, i.e. $H_0 : M_1 = M_2$, is given by

$$z^* = \frac{m_1 - m_2}{\sqrt{\hat{V}(m_1) + \hat{V}(m_2)}} \tag{3.7}$$

where $\hat{V}(m_i)$ for $i = 1, 2$ is given by (3.6). This statistic is treated as a normal test statistic, that is, $z^* \sim N(0, 1)$.

We apply this procedure to all trace and ultratrace elements and report the results in Tables 3.2(a) and 3.2(b). The results in Table 3.2(a) are based on all data values while the results in Table 3.2(b) are based on the data after the removal of the outliers.

It is clear from Table 3.2(a) that in general, there exist significant differences between the gestation groups for Mn, while there is no difference between gestation groups for Co, Mo and Sn. For Ni, there is initially a significant difference between the groups but as time progresses, the difference becomes insignificant. For Pb, no differences between the full-term and pre-term groups are apparent while differences become significant during the latter weeks. As before 'ind' indicates an indeterminate value, $\frac{0}{0}$, while 'inf' indicates an infinite value.

Table 3.2(a): The values of the z^* test statistic (3.7) for testing $H_0 : M_1 = M_2$ for a given week for trace and ultratrace elements (based on complete data).

	Trace Elements				Ultratrace Elements	
	Mn	Co	Mo	Ni	Pb	Sn
wk 1	10.01735	-0.6255432	0	3.365081	ind	-0.4581362
wk 2	1.3789	-0.2581989	inf	-0.06793129	0.2581989	0
wk 3	2.460372	-0.2236702	0.895443	0.3669207	inf	1.290994
wk 4	2.45385	ind	0.2581989	-0.2254907	1.071429	-0.2314998
wk 5	4.426267	0	0	0.2223504	4.285714	0.5576846
wk 6	4.426267	-inf	inf	-0.8906409	3.214286	0.06443289
wk 7	0.689644	-inf	1.071429	0.1457983	inf	-1.535036
wk 8	4.718893	inf	1.158132	-1.143005	3.214286	0.2778941
wk 12	2.412091	ind	inf	-1.842813	inf	-0.9438798

When possible outliers are removed, the median differences appear to be similar to the case with all data points, except that for Ni now the difference

appears to be insignificant at week 1.

Table 3.2(b): The values of the z^* test statistic (3.7) for testing $H_0 : M_1 = M_2$ for a given week for trace and ultratrace elements (excluding outliers).

	Trace Elements				Ultratrace Elements	
	Mn	Co	Mo	Ni	Pb	Sn
wk 1	4.178565	-0.8536101	inf	0.774101	ind	-0.5549638
wk 2	1.468837	-1.071429	0.4371401	-1.237437	0	0
wk 3	2.353408	0.2265278	0.7173927	0.1121936	2.142857	0
wk 4	2.267787	ind	2.45385	0.3958195	1.071429	0
wk 5	4.160251	-0.2773501	1.158132	-0.1309861	1.196747	0.282076
wk 6	3.663292	-inf	0.8320503	-1.152377	1.38675	0.04952246
wk 7	1.456986	-inf	0.8320503	0	inf	-0.4062173
wk 8	1.994578	inf	inf	-1.507557	1.38675	-0.6581452
wk 12	2.412091	0	inf	-1.01083	inf	-0.9438798

3.1.3 A Permutation Test

The well known permutation test consists of forming $n^* = n_1 \times n_2$ differences $d_j = y_{j_1} - y_{j_2}$ for $j = 1, \dots, n^*$ with $j_1 = 1, \dots, n_1$ and $j_2 = 1, \dots, n_2$, and where y_{j_1} is the j_{1th} observation in the full-term group and y_{j_2} is the j_{2th} observation in the pre-term group. If the median of the n^* differences is significantly different from zero then the difference between the two gestation periods will be significant. When using this approach, the dispersion of the n^* differences are estimated by using the median absolute deviation (MAD) and also by using a Kernel-density based dispersion estimate.

Dispersion Estimate by MAD

This method uses the same procedure as in the ad-hoc test for calculating the dispersion of the differences between two medians.

Let $d = \{d_1, \dots, d_{n^*}\}$ be a set of n^* differences and m_d the median of d_j , $j = 1, \dots, n^*$. Then the MAD as an estimate of the dispersion is given by

$$s^* = \frac{\text{Med}|d_j - m_d|}{0.6745} ; j = 1, 2, \dots, n^*. \quad (3.8)$$

Now, we define the normal test statistic as:

$$z^* = \frac{m_d - M_d}{\frac{s^*}{\sqrt{n^*}}}, \quad (3.9)$$

for testing the $H_0 : M_d = 0$, where M_d is the population median corresponding to m_d .

We apply this procedure to all trace and ultratrace elements and summarize the results in Tables 3.3(a) and 3.3(b). Table 3.3(a) values contain all data values while Table 3.3(b) have the outliers removed. It is clear from Table 3.3(a) that generally, for Mn and Mo, significant differences occur between the full-term group and pre-term group at each time period. For Co,

the difference between the gestation groups appear to be significant only for week 1. For Ni, significant differences exist during the earlier weeks but the difference appears to be insignificant for the the latter weeks. For Pb, there is no difference between full-term and pre-term groups during the early 2 weeks, but differences appear to be significant for the rest of the weeks. For Sn, however, the differences appear to be significant at the beginning as well as at the end, there being no differences in the middle.

Table 3.3(a): The values of the z^* test statistic (3.9) for testing $H_0 : M_d = 0$ for a given week for trace and ultratrace elements (based on complete data).

	Trace Elements				Ultratrace Elements	
	Mn	Co	Mo	Ni	Pb	Sn
wk 1	12.11043	-13.62423	9.082821	4.007127	0	-3.892638
wk 2	8.242526	0	6.181895	0	0	-2.392991
wk 3	7.272817	0	6.181895	2.472758	12.36379	0
wk 4	5.224655	0	5.224655	0	5.224655	-0.6332915
wk 5	6.530818	0	0	0	3.483103	-1.89987
wk 6	3.483103	0	5.224655	-0.2430072	3.483103	-1.099927
wk 7	1.046542	0	4.70944	0.8562618	9.41888	-3.853178
wk 8	4.70944	0	0	-2.018331	4.70944	-2.173588
wk 12	4.285059	0	8.094	0	8.094	-4.047

When possible outliers are removed, the median differences appear to be similar to the case, except for Sn, with all data points in the analysis. For

Sn, the differences appear to be significant for almost all weeks.

Table 3.3(b): The values of the z^* test statistic (3.9) for testing $H_0 : M_d = 0$ for a given week for trace and ultratrace elements (outliers excluded).

	Trace Elements				Ultratrace Elements	
	Mn	Co	Mo	Ni	Pb	Sn
wk 1	9.396211	-12.93915	8.811631	5.060831	ind	-5.286979
wk 2	4.336975	0	10.6006	0	0	-2.177835
wk 3	4.489195	0	5.500381	2.165661	10.44931	2.081136
wk 4	4.326425	0	5.224655	0	9.774434	2.018331
wk 5	5.383996	0	0	0	4.70944	-1.096062
wk 6	7.706356	0	4.7215	-1.5975	1.954887	-2.288889
wk 7	1.748504	0	9.099502	0	9.41888	-3.906152
wk 8	4.032923	0	0	-2.214119	4.37126	-6.937714
wk 12	4.285059	0	11.62412	0	7.074216	-4.047

Kernal-density Based Dispersion Estimate

It is well-known that the sample median is asymptotically normally distributed with mean M_d and dispersion

$$V(m_d) = \frac{1}{4n \cdot \{f(m_d)\}^2} \quad (3.10)$$

where M_d is the median corresponding to m_d , and $f(m_d)$ is the density of the median m_d (cf. Venables and Ripley (1996)). Now, to estimate $V(m_d)$, one needs to estimate $f(m_d)$. This density $f(m_d)$ may be estimated by using

a Kernel-density smoother of the form

$$\hat{f}(m_d) = \frac{1}{b} \sum_{j=1}^{n^*} K\left(\frac{m_d - d_j}{b}\right), \quad (3.11)$$

where $K(\cdot)$ is a fixed Kernel and b is a suitable a bandwidth. The bandwidth b is to be chosen approximately as to smooth the data but not to smooth real peaks. By default S-Plus uses a width b equal to

$$\frac{d_{(n^*)} - d_{(1)}}{2(1 + \log_2 n^*)},$$

Where $d_{(1)}$ and $d_{(n^*)}$ are the minima and maxima of the sample differences d_1, \dots, d_{n^*} respectively. In (3.11), the Kernel K is normally chosen to be a probability density function such as normal, rectangular, and cosine. Note that we have used the normal density function for K as it is chosen by default in S-plus (cf. p136, Venables and Ripley (1996)). Now, similar to (3.9), we use the normal test statistic

$$z^* = \frac{m_d - M_d}{\sqrt{\hat{V}(m)}} \quad (3.12)$$

for testing the $H_0 : M_d = 0$.

We apply this procedure to all trace and ultratrace elements and summarize the results in Tables 3.4(a) and 3.4(b). The results in Table 3.4(a) contain all data values while the results in Table 3.4(b) have been computed after the outliers are removed. From Table 3.4(a) it is clear that the results are the same as the previous method, where dispersion is estimated by MAD.

Table 3.4(a): The values of the z^* test statistic (3.12) for testing $H_0 : M_d = 0$ for a given week for trace and ultratrace elements (based on complete data).

	Trace Elements				Ultratrace Elements	
	Mn	Co	Mo	Ni	Pb	Sn
wk 1	8.278198	-5.538833	7.745976	3.843176	0	-3.882906
wk 2	6.549037	0	3.153925	0	0	-2.824778
wk 3	6.000034	0	3.690489	2.171564	12.7715	0
wk 4	3.877602	0	4.64771	0	4.96724	-0.6799627
wk 5	4.930224	0	0	0	2.572339	-1.465165
wk 6	2.128531	0	3.59976	-0.3322144	2.674456	-1.545367
wk 7	0.9389233	0	3.69814	0.8318205	7.692131	-3.282236
wk 8	3.633175	0	0	-2.026274	1.92837	-2.035035
wk 12	1.648813	0	8.04388	0	8.802648	-3.409725

When possible outliers are removed, the median differences appear to be similar to the case with all data points in the analysis.

Table 3.4(b): The values of the z^* test statistic (3.12) for testing $H_0 : M_d = 0$ for a given week for trace and ultratrace elements (outliers excluded).

	Trace Elements				Ultratrace Elements	
	Mn	Co	Mo	Ni	Pb	Sn
wk 1	6.515692	-10.99779	8.013562	2.040107	0	-4.86795
wk 2	2.626773	0	26.114	0	0	-2.672203
wk 3	2.904784	0	9.367578	2.083484	23.5484	1.610829
wk 4	3.346704	0	4.64771	0	9.621017	2.233526
wk 5	2.728461	0	0	0	5.5883	-1.355917
wk 6	8.260567	0	10.93883	-1.539641	1.640676	-3.453277
wk 7	1.287393	0	6.053874	0	7.692131	-2.920783
wk 8	2.991312	0	0	-2.223285	3.766999	-8.478726
wk 12	1.648813	0	44.8092	0	22.94916	-3.409725

3.2 Tests for Gestation Effects for all Weeks

Combined: A “Working” Independence Approach

In the previous section, we have tested the difference in minerals between the full-term and pre-term groups for a given week. Most of the tests indicated that for Mn and Pb, there were significant differences at any given week. In this section, we are interested to see whether these differences remain when the data are combined together for all weeks. In the same token, we are also interested to see what happens to the other minerals. For the purpose,

for each of the minerals, we combine the data for all 8 weeks together and then develop suitable tests to examine whether there is any overall significant difference in minerals between the full-term and pre-term groups. We use 3 different tests for testing such hypotheses, that is, there is an overall difference between the gestation groups. These tests are: (1) an ad-hoc median test, (2) weighted ad-hoc median test based on 8 strata due to eight time periods, and (3) the ad-hoc median test when dispersion is computed based on delete-1 strata jackknife. Note that in all three tests, it will be assumed that the data for different weeks are independent even though they are actually correlated, the correlation structure being unknown. This approach is referred to as the “working” independence approach for inferences about correlated data.

3.2.1 Ad-hoc Test

This procedure is equivalent to the ad-hoc test discussed earlier in Section 3.1.1. The difference between the earlier test and the present test is that here observations for all weeks are combined together under both the full-term and pre-term groups, whereas in the earlier test, the observations for a given week are used to test the median difference between two groups, namely full-term and pre-term groups.

For a particular *mineral*, let y_{ijt} be the mineral concentration in the milk of the j_{th} ($j = 1, 2, \dots, n_{it}$) mother of the i_{th} ($i = 1, 2$) group at time t . Here group 1 will refer to the full-term mothers and group 2 will refer to the pre-term mothers as before. Also let y_i be a set of observations containing the mineral concentrations for the entire i_{th} group, that is,

$$y_i = \{y_{i11}, \dots, y_{in_{i1}1}, \dots, y_{ijt}, \dots, y_{i1T}, \dots, y_{in_{iT}T}\}.$$

Next suppose that s_i measures the dispersion of the entire $\sum_{t=1}^T n_{it}$ observations in y_i and m_i is the median of these observations. That is,

$$s_i = \frac{\text{Med}(|y_{ijt} - m_i|)}{0.6745}, \quad t = 1, \dots, T. \quad (3.13)$$

The dispersion in (3.13) is now used to construct a ad-hoc normal test statistic given by

$$z^* = \frac{m_1 - m_2}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}, \quad (3.14)$$

where $n_1 = \sum_{t=1}^T n_{1t}$ and $n_2 = \sum_{t=1}^T n_{2t}$ which may be used to test the null hypothesis whether the population median, M_1 , of the full-term group is

equal to the population median, M_2 , of the pre-term group.

We apply this procedure to all trace and ultratrace elements and summarize the results in Table 3.5.

Table 3.5: The values of the z^* test statistic (3.14) for testing $H_0 : M_1 = M_2$ for trace and ultratrace elements.

Trace Elements				Ultratrace Elements	
Mn	Co	Mo	Ni	Pb	Sn
5.845948	0	-2.734795	-0.3319479	7.510914	0

This test procedure indicates that a significant difference exists between the full-term group and pre-term group for Mn, Mo, and Pb while no difference exist between the full-term group and pre-term group for Co, Ni, and Sn. This conclusion is almost in agreement (except for Mo) with the conclusions made in Section 3.1 for individual weeks.

3.2.2 Weighted Ad-hoc Test

The ad-hoc median in the last section does not take into consideration the variability of the observations in individual weeks, which rather took into account the dispersion of all the observations for all weeks combined. In this section, we consider that the observations in a given week form a strata and use the strata dispersion as its weight and then compute the stratified

medians.

Assuming time periods are independent, for a particular *mineral* we let y_{ijt} be the j_{th} observation in the t_{th} week for the i_{th} gestation group and, m_{it} the median of t_{th} week for the i_{th} gestation group. Here also, group 1 will refer to the full-term mothers and group 2 will refer to the pre-term mothers. Now for given $i(i = 1, 2)$ and $t(t = 1, \dots, 8)$, we define the median absolute deviation of n_{it} observations from their median m_{it} as

$$s_{it} = \frac{\text{Med}(|y_{ijt} - m_{it}|)}{0.6745}, \quad j = 1, \dots, n_{it}. \quad (3.15)$$

Next, we define the weighted median as

$$m_i^* = \frac{\sum_{t=1}^8 w_{it} m_{it}}{\sum_{t=1}^8 w_{it}} \quad (3.16)$$

where

$$w_{it} \propto \frac{1}{s_{it}^2} \quad \text{and} \quad \sum_{t=1}^8 w_{it} = 1,$$

with

$$w_{it} = \frac{k_i}{s_{it}^2} \quad \text{and} \quad \sum_{t=1}^8 w_{it} = k_i \sum_{t=1}^8 \frac{1}{s_{it}^2} = 1$$

and

$$k_i = \frac{1}{\sum_{t=1}^8 \frac{1}{s_{it}^2}} \quad \text{and} \quad w_{it} = \frac{\frac{1}{s_{it}^2}}{\sum_{t=1}^8 \frac{1}{s_{it}^2}} = \frac{s_{it}^{-2}}{\sum_{t=1}^8 s_{it}^{-2}}.$$

Suppose that we now estimate the variance of m_i^* as

$$\begin{aligned} \hat{V}(m_i^*) &= \frac{1}{(\sum_{t=1}^8 w_{it})^2} \sum_{t=1}^8 w_{it}^2 \hat{V}(m_{it}) \\ &= \sum_{t=1}^8 w_{it}^2 \hat{V}(m_{it}) \\ &= \sum_{t=1}^8 w_{it}^2 \sum_{t=1}^8 w_{it} (m_{it} - m_i^*)^2 \\ &= \sum_{t=1}^8 w_{it}^2 \sigma_{\text{Med}}^2(i). \end{aligned} \tag{3.17}$$

The variance in (3.17) is now used in a natural way to construct a weighted ad-hoc normal teststatistic given by

$$z^* = \frac{m_1^* - m_2^*}{\sqrt{\hat{V}(m_1^*) + \hat{V}(m_2^*)}}, \tag{3.18}$$

which may be used to test the null hypothesis whether the population weighted median, M_1^* , of the full-term group is equal to the population weighted median, M_2^* , of the pre-term group.

We apply this procedure to all trace and ultratrace elements and summarize the results in Table 3.6.

Table 3.6: The values of the z^* test statistic (3.18) for testing $H_0 : M_1^* = M_2^*$ for trace and ultratrace elements.

Trace Elements				Ultratrace Elements	
Mn	Co	Mo	Ni	Pb	Sn
8.438713	ind	ind	ind	ind	0.7022064

The weighted ad-hoc test indicates that a significant difference exists between the full-term group and pre-term group for Mn but not for any other minerals. The values of the test statistic for most of the other minerals are written as ‘ind’ indicating that at least one of the s_{it} ’s equal zero creating a weight which is undefined. The above problem of obtaining $z^*=ind$ may be solved by simply ensuring none of the s_{it} ’s equal zero.

A simplistic solution to this problem is to give all medians equal weight of $\frac{1}{8}$. We apply this idea of equal weights to all trace and ultratrace elements and summarize the results in Table 3.7.

Table 3.7: The values of the z^* test statistic (3.18) for testing $H_0 : M_1^* = M_2^*$, using a constant weight of $\frac{1}{8}$, for trace and ultratrace elements.

Trace Elements				Ultratrace Elements	
Mn	Co	Mo	Ni	Pb	Sn
8.699439	-1.632993	-1.4399	0.7009566	3.906078	-0.6726248

When the weighted ad-hoc test statistic is computed using equal weights of $\frac{1}{8}$, a difference exists for both Mn and Pb. The results in Table 3.7 are in agreement with those in 3.5.

3.2.3 Comparison of Median When Dispersion is Computed by Delete-1 Group Jackknife Technique.

In Section 3.1.2, we have already studied a jackknife technique based on deleting individual data values in order to estimate the dispersion of a set of data. The jackknife procedure may be extended to estimate the dispersion by deleting one stratum at a time, where the stratum correspond to $T = 1, 2, \dots, 8$ weeks. More specifically, within each gestation period we delete one group at a time and calculate the median of the remaining $T - 1 = 7$ strata.

Therefore we define m_i as the median of the full-term group and $m_{i(t)}$ as the recalculated median of full-term group when the t_{th} stratum is deleted.

The averages of the T recalculated medians for each gestation period may be denoted as:

$$m_{i(.)} = \frac{\sum_{t=1}^T m_{i(t)}}{T} \quad (3.19)$$

where as before $i = 1$ represents the full-term group and $i = 2$ represents the pre-term group.

Along the lines with the Jackknife procedure described in Section 3.1.2 the dispersion for the i_{th} gestation group is now given by

$$\hat{V}(m_i) = \frac{T-1}{T} \sum_{t=1}^T (m_{i(t)} - m_{i(.)})^2, \quad (3.20)$$

yielding the test statistic

$$z^* = \frac{m_1 - m_2}{\sqrt{\hat{V}(m_1) + \hat{V}(m_2)}}. \quad (3.21)$$

This test statistic is then used to determine whether differences exist between the population median, M_1 , of the full-term group and the population median, M_2 , of the pre-term group. We apply this procedure to all trace and

ultratrace elements and summarize the results in Table 3.8.

Table 3.8: The values of the z^* test statistic (3.21) for testing $H_0 : M_1 = M_2$ for trace and ultratrace elements.

Trace Elements				Ultratrace Elements	
Mn	Co	Mo	Ni	Pb	Sn
6.110101	0	-0.693688	-0.1553178	2.285714	0

This test indicates that a significant difference exists between gestation periods for Mn and Pb but there is no difference for Co, Mo, Ni, and Sn. Thus, all the techniques used in this section provide similar results about the difference in gestation effects.

Chapter 4

Gestational Effects: TIME AS A NON-SPECIFIC FACTOR

In the previous chapter, we tested the difference between mineral concentrations in the milk of full-term mothers and pre-term mothers for individual weeks as well as for all the weeks combined. In the former case, as the gestation effects were compared at a specific time point, time was naturally a specific factor. In the latter case, the mothers at a given week were assumed to constitute a stratum, and the overall gestation effects were compared based on such stratified data. There, time was considered as a specific factor.

In this chapter we also study the difference between the two groups

but here the minerals of a mother recorded over the weeks are assumed to form a cluster under the selected mothers. Let there be J_i^* mothers in the i_{th} group ($i = 1, 2$) and suppose that the milk was collected from the j_{th} ($j = 1, 2, \dots, J_i^*$) mother for a period of n_{ij}^* weeks. Consequently, these n_{ij}^* observations for the j_{th} mother will be correlated. Here the correlations among the observations collected over the period will be taken into account in testing the difference between the gestational effects of the two groups. Thus, time will not play any role as a specific factor. In other words, we will refer to "time" as a non-specific factor. Note that as in Section 3.2 of the previous chapter, y_{ijt} will refer to the mineral concentration collected at the t_{th} period from the j_{th} mother and the i_{th} group. Thus, in terms of the notation of Chapter 3, we have $J_i^* = \sum_{t=1}^T n_{it}$, but we now use $t = 1, 2, \dots, n_{ij}^*$, to refer n_{ij}^* minerals recorded under the j_{th} ($j = 1, \dots, J_i^*$) mother of the i_{th} ($i = 1, 2$) group. Further note that in this chapter, we include all observations that were available for analysis, that is, we do not put extra effort to find and discard any outliers.

4.1 Delete-1 Cluster/Mother Jackknifing

In this section we develop a normal test statistic based on the Jackknife procedure discussed in Section 3.2.3. The difference here is that mothers constitute the strata. For a particular *mineral*, let m_i be the median of the i_{th} group and $m_{i(j)}$ be the recalculated median of the i_{th} group when the j_{th} cluster is deleted. Then the average of the recalculated medians for the i_{th} group is given by

$$m_{i(.)} = \frac{\sum_{j=1}^{J_i^*} m_{i(j)}}{J_i^*} \quad (4.1)$$

where its jackknife dispersion is calculated as

$$\hat{V}(m_i) = \frac{J_i^* - 1}{J_i^*} \sum_{j=1}^{J_i^*} (m_{i(j)} - m_{i(.)})^2. \quad (4.2)$$

As in the last chapter, we now use the test statistic

$$z = \frac{m_1 - m_2}{\sqrt{\hat{V}(m_1) + \hat{V}(m_2)}}. \quad (4.3)$$

to test the $H_0 : M_1 = M_2$, where M_1 and M_2 are the unknown population medians for the full-term and pre-term groups respectively. Once again we

treat the test statistic as a normal statistic. We note, however, that although m_1 and m_2 are the same sample medians as in the last chapter, their estimate of their variances are not the same. In the previous chapter, the variances were estimated by delete-1 stratum technique or by other means, but here the variances are computed by delete-1 mother/cluster technique, which are generally different.

When the test (4.3) was applied to the data, we found the z values as recorded in Table 4.1.

Table 4.1: The values of the z test statistic (4.3) with all mothers included in the analysis.

Group	Trace Elements				Ultratrace Elements	
	Mn	Co	Mo	Ni	Pb	Sn
z	2.603343	0	-0.5781516	-0.1670626	1.070529	-0.1601622

The results in Table 4.1 indicates a significant difference between the pre-term and full-term groups for Mn but not for any other minerals.

Note that the n_{ij}^* observations under the j_{th} mother may or may not be consecutive. In order to compare the results of this section with the results of the next section, we also consider the case, where the observations were available for at least three consecutive weeks. The application of (4.3) to such data yields the results in Table 4.2.

Table 4.2: The values of the z test statistic (4.3) when only mothers with at least three consecutive weeks are included in the analysis.

Group	Trace Elements				Ultratrace Elements	
	Mn	Co	Mo	Ni	Pb	Sn
z	3.461248	0	-0.3768427	-0.08583691	1.792843	0

When using the minerals of the mothers only for the consecutive weeks, one obtains the z values as in Table 4.2 which only indicates that a significant difference exists for Mn but not for the other minerals. This is the same conclusion as when we considered all observations in the analysis.

4.2 Ad-hoc Median and Mean Tests: TIME AS A NON-SPECIFIC FACTOR

4.2.1 Modeling Gestation Effects

As mentioned earlier, in this section we consider only consecutive observations recorded for the j_{th} mother. Also suppose that, irrespective of the distributional pattern of the data, the data collected over the consecutive periods follow the $AR(1)$ (autoregressive function of order 1) type relationship

given by

$$y_{ijt} = \theta_{i0} + \phi_i y_{ij,t-1} + \epsilon_{ijt} \quad (4.4)$$

where i and j are fixed and $t = 1, 2, \dots, n_{ij}^*$. In (4.4) we assume that for a given i , ϵ_{ijt} 's are identically and independently distributed with median (Med) 0 and dispersion (Disp) σ_i^2 , for all $j = 1, \dots, J_i^*$ and $t = 1, \dots, n_{ij}^*$. Here σ_i^2 is not necessarily the variance of the errors.

By using the recurrence relation (4.4), one may express y_{ijt} as

$$y_{ijt} = \theta_{i0} \sum_{\ell=0}^{n_{ij}^*-1} \phi_i^\ell + \phi_i^{n_{ij}^*} y_{ij,t-n_{ij}^*} + \sum_{u=0}^{n_{ij}^*} \phi_i^u \epsilon_{ij,t-u}. \quad (4.5)$$

As $|\phi_i| < 1$, for large cluster size ($n_{ij}^* \rightarrow \infty$), it now follows that

$$\begin{aligned} \text{Med}\{y_{ijt}; t = 1, \dots, n_{ij}^*, j = 1, \dots, J_i^*\} &= \theta_{i0}(1 + \phi_i + \dots + \phi_i^{n_{ij}^*} + \dots) \\ &= \theta_{i0}(1 - \phi_i)^{-1} \end{aligned} \quad (4.6)$$

and

$$\text{Disp}(y_{ijt}) = \text{Disp}(\epsilon_{ijt} + \phi_i \epsilon_{ij,t-1} + \phi_i^2 \epsilon_{ij,t-2} + \dots)$$

$$\begin{aligned}
&= \sigma_i^2 + \phi_i^2 \sigma_i^2 + (\phi_i^2)^2 \sigma_i^2 + (\phi_i^2)^3 \sigma_i^2 + \dots \\
&= \sigma_i^2 (1 - \phi_i^2)^{-1},
\end{aligned} \tag{4.7}$$

as, for example,

$$\text{Disp}(\phi_i \epsilon_{ij,t-1}) = \left[\frac{\text{Med}[\phi_i \epsilon_{ij,t-1} - 0]}{0.6745} \right]^2 = \phi_i^2 \sigma_i^2. \tag{4.8}$$

Now, y_{ijt} is the mineral content in the milk that was collected in the t_{th} week from the j_{th} mother of the i_{th} group and ϵ_{ijt} is the t_{th} element of the $n_{ij}^* \times 1$ vector of disturbances $\epsilon_{ij} = (\epsilon_{ij1}, \dots, \epsilon_{ijt}, \dots, \epsilon_{ij, n_{ij}^*})'$ where $\text{Med}(\epsilon_{ij}) = 0$, $\text{Disp}(\epsilon_{ij}) = \sigma_i^2 I_{n_{ij}^*}$. Now, after some matrix algebras, one may show that $y_{ij} = (y_{ij1}, \dots, y_{ijt}, \dots, y_{ij, n_{ij}^*})$ has the dispersion matrix $\sigma_i^2 \Lambda_{ij}$, where

$$\Lambda_{ij} = \frac{1}{1 - \phi_i^2} \begin{pmatrix} 1 & \phi_i & \phi_i^2 & \dots & \phi_i^{n_{ij}^*-1} \\ \phi_i & 1 & \phi_i & \dots & \phi_i^{n_{ij}^*-2} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \phi_i^{n_{ij}^*-1} & \phi_i^{n_{ij}^*-2} & \phi_i^{n_{ij}^*-3} & \dots & 1 \end{pmatrix}$$

is the $n_{ij}^* \times n_{ij}^*$ non-singular matrix and ϕ_i is the parameter of the AR(1) type process as indicated before.

Ad-hoc Median Test

As under the present modeling, the median of the data for the i_{th} group is given by $\theta_{i0}(1 - \phi_i)^{-1}$ (4.6), we now test the equality of the gestation effects by testing the $H_0 : \theta_{10} = \theta_{20}$. For the purpose, we now require to estimate θ_{i0} , which in turn requires the estimate of ϕ_i . As the data is asymmetric, we now estimate this ϕ_i as follows which will be denoted by ϕ_i^* . In order to compute ϕ_i^* , we first compute ϕ_{ij}^* for the j_{th} cluster/mother under the i_{th} group. The formula ϕ_{ij}^* is given by

$$\phi_{ij}^* = \frac{[\text{Med}[(y_{ijt} - m_{ij})(y_{ij,t-1} - m_{ij}), t = 2, \dots, n_{ij}^*]]}{[\text{Med}[(y_{ijt} - m_{ij})^2, t = 1, 2, \dots, n_{ij}^*]]}, \quad (4.9)$$

where m_{ij} is the median of the j^{th} cluster/mother in the i^{th} group.

Now by pooling these estimates of the J_i^* mothers/clusters, we obtain ϕ_i^* as given by

$$\phi_i^* = \frac{\sum_{j=1}^{J_i^*} [\text{Med}[(y_{ijt} - m_{ij})(y_{ij,t-1} - m_{ij}), t = 2, \dots, n_{ij}^*]]}{\sum_{j=1}^{J_i^*} [\text{Med}[(y_{ijt} - m_{ij})^2, t = 1, 2, \dots, n_{ij}^*]]}. \quad (4.10)$$

The results of ϕ_i^* , $i = 1, 2$ are given in Table 4.3 and show that the correlations are generally different for full-term and pre-term groups, the correlations for

Mn being far apart from the others.

Table 4.3: Lag 1 Correlation ϕ_i^*

Group	Trace Elements				Ultratrace Elements	
	Mn	Co	Mo	Ni	Pb	Sn
Full-term (ϕ_1^*)	-0.050	0.214	0.194	-0.303	-0.210	-0.204
Pre-term (ϕ_2^*)	0.758	-0.239	-0.141	-0.040	-0.133	-0.230

Now, turning back to the estimation of θ_{i0} , it follows from (4.6) that one may estimate this parameter as

$$\theta_{i0}^* = (1 - \phi_i^*) \frac{\text{Med}[w_{ijt}y_{ijt}, j = 1, \dots, J_i^*; t = 1, \dots, n_{ij}^*]}{\text{Med}[w_{ijt}, j = 1, \dots, J_i^*; t = 1, \dots, n_{ij}^*]}, \quad (4.11)$$

where w_{ijt} is referred to as the weight for y_{ijt} , which is actually the t_{ih} element in w_{ij} , where $w_{ij} = I'_{ij}\Lambda_{ij}^{-1} = (w_{ij1}, \dots, w_{ijt}, \dots, w_{ijn_{ij}^*})$. In (4.11), ϕ_i^* is the estimate of ϕ_i given by (4.10). We have computed the value of $\hat{\theta}_{i0}$ by (4.11) for both groups and for all minerals. The results are shown in Table 4.4.

Table 4.4: θ_{i0}^* (using ϕ_i^*)

Group	Trace Elements				Ultratrace Elements	
	Mn	Co	Mo	Ni	Pb	Sn
full-term (θ_{10m}^*)	13.650	0.786	1.612	10.424	2.42	14.448
pre-term (θ_{20m}^*)	2.299	1.239	3	9.180	0	15.990

Next, in order to construct a suitable test statistic for testing $H_0 : \theta_{10} = \theta_{20}$, we now require to estimate the dispersion of θ_{i0}^* for $(i = 1, 2)$. Assuming weights w_{ijt} are known, these estimates may be computed as

$$\text{Disp}(\theta_{i0}^*) = (1 - \phi_i^*)^2 \frac{\text{Disp}(\text{Med}[w_{ijt}y_{ijt}; j = 1, \dots, J_i^*; t = 1, \dots, n_{ij}^*])}{(\text{Med}[w_{ijt}; j = 1, \dots, J_i^*; t = 1, \dots, n_{ij}^*])^2} \quad (4.12)$$

where $\text{Disp}(\text{Med}[w_{ijt}y_{ijt}; j = 1, \dots, J_i^*; t = 1, \dots, n_{ij}^*])$ is calculated using the delete-1 cluster/mother Jackknifing technique. Finally by using (4.11) and (4.12) we construct an ad-hoc normal test statistic given by

$$z^* = \frac{\theta_{10}^* - \theta_{20}^*}{\sqrt{\text{Disp}(\theta_{10}^*) + \text{Disp}(\theta_{20}^*)}}. \quad (4.13)$$

The values of z^* computed by (4.13) are shown in Table 4.5.

Table 4.5: z^* (using ϕ_i^*)

	Trace Elements				Ultratrace Elements	
	Mn	Co	Mo	Ni	Pb	Sn
z^*	5.126	-Inf	-0.636	0.227	2.541	-0.264

It is clear for the above table that the gestation effects are significantly different between the pre-term and full-term groups for Mn and Pb. These

results are in agreement with the results of the weighted ad-hoc median test in Section 3.2.2 as well as with the results of the median test in Section 3.2.3, where the variances of the medians were calculated by delete-1 group jackknife technique.

Weighted Average Method: A Large Cluster Size Approach

In this section, we pretend that our data is symmetric and the mean and variance are stationary and follows the AR(1) model (4.4). Thus

$$E(y_{ijt}) = \theta_{i0}(1 - \phi_i)^{-1}, \quad (4.14)$$

and

$$V(y_{ijt}) = \sigma_i^2(1 - \phi_i^2)^{-1}, \quad (4.15)$$

which appear to be similar to those given in (4.6) and (4.7) respectively. Consequently, we estimate ϕ_{ij} by using the traditional time series approach but taking the deviations of the observations from their median. The formula for $\hat{\phi}_{ij}$, an estimate of ϕ_{ij} , is given by

$$\hat{\phi}_{ij} = \frac{\sum_{t=2}^{n_{ij}} (y_{ijt} - m_{ij})(y_{ij,t-1} - m_{ij})}{\sum_{t=1}^{n_{ij}} (y_{ijt} - m_{ij})^2}. \quad (4.16)$$

Next by pooling the J_i^* mother (cf. Quenouille (1958)) we obtain the overall lag 1 correlation estimate given by

$$\hat{\phi}_i = \frac{\sum_{j=1}^{J_i^*} \sum_{t=2}^{n_{ij}} (y_{ijt} - m_{ij})(y_{ij,t-1} - m_{ij})}{\sum_{j=1}^{J_i^*} \sum_{t=1}^{n_{ij}} (y_{ijt} - m_{ij})^2}. \quad (4.17)$$

The results of $\hat{\phi}_i$ ($i = 1, 2$) are given in Table 4.6.

Table 4.6: Lag 1 Correlation $\hat{\phi}_i$

Group	Trace Elements				Ultratrace Elements	
	Mn	Co	Mo	Ni	Pb	Sn
Full-term ($\hat{\phi}_1$)	0.003	0.040	0.099	-0.010	0.240	0.020
Pre-term ($\hat{\phi}_2$)	0.054	-0.026	-0.014	0.260	-0.025	0.137

The results in Table 4.6 shows that the correlations for all minerals are insignificant except for Ni for the pre-term and Pb for the full-term.

As mentioned earlier, the gestation effect of the j_{th} mother may be measured by θ_{ij0} . If we now assume that the indices of deterministic functions

for all mothers are equal, say θ_{i0} , then this constant index may be estimated by using the weighted mean given by

$$\hat{\theta}_{i0} = (1 - \hat{\phi}_i) \left(\sum_{j=1}^{J_i'} I'_{ij} \Lambda_{ij}^{-1}(\hat{\phi}_i) I_{ij} \right)^{-1} \left(\sum_{j=1}^{J_i'} I'_{ij} \Lambda_{ij}^{-1}(\hat{\phi}_i) y_{ij} \right) \quad (4.18)$$

where $y_{ij} = [y_{ij1}, \dots, y_{ijt}, \dots, y_{ij, n_{ij}^*}]'$ and I_{ij} is the $n_{ij}^* \times 1$ vector of unity, yielding the values of $\hat{\theta}_{i0}$ as in Table 4.7.

Table 4.7: $\hat{\theta}_{i0}$ (using $\hat{\phi}$)

Group	Trace Elements				Ultratrace Elements	
	Mn	Co	Mo	Ni	Pb	Sn
Full-term ($\hat{\theta}_{10}$)	15.059	1.293	2.778	22.436	2.689	23.406
Pre-term ($\hat{\theta}_{20}$)	9.304	2.108	3.388	25.850	1.341	22.053

Next, the variance of $\hat{\theta}_{i0}$ may be derived as

$$\begin{aligned} V(\hat{\theta}_{i0}) &= (1 - \hat{\phi}_i)^2 \left(\sum_{j=1}^{J_i'} I'_{ij} \Lambda_{ij}^{-1}(\hat{\phi}_i) I_{ij} \right)^{-2} \sum_{j=1}^{J_i'} (I'_{ij} \Lambda_{ij}^{-1}(\hat{\phi}_i) V(y_{ij}) \Lambda_{ij}^{-1}(\hat{\phi}_i) I_{ij}) \\ &= (1 - \hat{\phi}_i)^2 \left(\sum_{j=1}^{J_i'} I'_{ij} \Lambda_{ij}^{-1}(\hat{\phi}_i) I_{ij} \right)^{-1} \sigma_i^2. \end{aligned} \quad (4.19)$$

where, following (4.12), σ_i^2 may be estimated as $\hat{\sigma}_i^2 = \hat{V}(y_{ijt})(1 - \hat{\phi}_i^2)$. Since

$\hat{V}(y_{ijt})$ may be computed by the formula given by

$$\hat{V}(y_{ijt}) = \frac{\sum_{j=1}^{J_i^*} \sum_{t=1}^{n_{ij}} (y_{ijt} - \bar{y}_{i.})^2}{(\sum_{j=1}^{J_i^*} n_{ij}) - 1}, \quad (4.20)$$

we may use this to obtain the final estimate σ_i^2 as

$$\hat{\sigma}_i^2 = \hat{V}(y_{ijt})(1 - \phi_i^2), \quad (4.21)$$

where $\hat{\phi}_i$ is given by (4.17).

The above formulas are then used to compute an ad-hoc normal test statistic given by

$$z^{**} = \frac{\hat{\theta}_{10} - \hat{\theta}_{20}}{\sqrt{V(\hat{\theta}_{10}) + V(\hat{\theta}_{20})}}. \quad (4.22)$$

Now by applying z^{**} test to the data for all minerals, we obtain the results as in Table 4.8.

Table 4.8: z^{**} (using $\hat{\phi}$)

	Trace Elements				Ultratrace Elements	
	Mn	Co	Mo	Ni	Pb	Sn
z^{**}	3.357	-2.457	-1.338	-0.530	2.618	0.326

The weighted mean method indicates that a significant difference exist

between pre-term and full-term groups for elements Mn, Co and Pb. It is interesting to observe that the results of Mn and Pb are the same as in the previous section, but this test, unlike the test of the previous section, shows a significant difference for Co.

Normal Type Variances

We further continue to compute the normal test statistic, z^{***} , but we now use different formulas for the computation of the variance. We use the following three different formulas:

$$\bar{\sigma}_i^2(1) = \frac{1}{J_i^*} \sum_{j=1}^{J_i^*} \left(\frac{\sum_{t=2}^{n_{ij}} (y_{ijt} - \hat{\theta}_{i0} - \hat{\phi}_i y_{ij,t-1})^2}{n_{ij}^* - 1} \right) \quad (4.23)$$

$$\bar{\sigma}_i^2(2) = \frac{\sum_{j=1}^{J_i^*} \sum_{t=2}^{n_{ij}} (y_{ijt} - \hat{\theta}_{i0} - \hat{\phi}_i y_{ij,t-1})^2}{\sum_{j=1}^{J_i^*} (n_{ij}^* - 1)} \quad (4.24)$$

$$\bar{\sigma}_i^2(3) = \frac{\sum_{j=1}^{J_i^*} \sum_{t=2}^{n_{ij}} \left(y_{ijt} - \frac{\hat{\theta}_{i0}}{1 - \hat{\phi}_i} \right)^2}{\sum_{j=1}^{J_i^*} (n_{ij}^* - 1)} \quad (4.25)$$

and compute the ad-hoc median test statistic given by

$$z^{***}(l) = \frac{\hat{\theta}_{10} - \hat{\theta}_{20}}{\sqrt{\hat{V}_l(\hat{\theta}_{10}) + \hat{V}_l(\hat{\theta}_{20})}}, \quad l=1,2,3, \quad (4.26)$$

where $\tilde{V}(\hat{\theta}_{i0})$ is given by (4.19) with σ_i^2 estimated by (4.23), (4.24) and (4.25).

The values of z^{***} computed by (4.26) are shown in Table 4.9.

Table 4.9: z^{***} (using $\hat{\phi}_l$)

	Trace Elements				Ultratrace Elements	
	Mn	Co	Mo	Ni	Pb	Sn
$z^{***}(1)$	3.759	-2.662	-1.349	-0.518	2.465	0.311
$z^{***}(2)$	4.001	-2.966	-1.543	-0.617	2.800	0.337
$z^{***}(3)$	3.981	-2.973	-1.468	-0.559	2.649	0.328

Table 4.9 indicates that a significant difference exists between pre-term and full-term groups for Mn, Co and Pb. These results are in agreement with the results of the ad-hoc normal test discussed in the previous section.

Chapter 5

Gestation Effects of Macrominerals: Time as a Non-specific Factor

As mentioned earlier there are four important macrominerals in mothers milk and they are: Ca, Mg, Rb and Sr. The longitudinal as well as gestational effects of these minerals were studied by Sutradhar *et. al.* (1998). As far as the distributional pattern is concerned, these minerals were found to be normally distributed. Log transformation of the minerals was considered as it is a special situation under Box-Cox transformation which Sutradhar *et.*

al. (1998) did to determine the status of the distributions of the minerals. Among these four minerals, Sutradhar *et. al.* (1998) concluded that there were significant longitudinal changes in the concentrations of Ca, Mg and Rb, but the gestation effect was significant only for Ca. In order to study the gestational effects, these authors have, however, examined the significance of the mean levels of the minerals at a given week. Thus, 'time' was considered as a specific factor.

Note that in the practicum, we have examined the gestation effects of all the trace and ultratrace elements considering time as a specific as well as non-specific factors. In the same token, in this chapter, we verify the gestational difference of the four macrominerals between the two groups by considering time as a non-specific factor.

5.1 AR(1) Model and Estimates of the Nuisance Parameters

Recall from section 4.2.1 that the mineral concentration recorded over n_{ij}^* time points for the j_{th} mother in the i_{th} group follow the AR(1) model

$$y_{ijt} = \theta_{i0} + \phi_i y_{ij,t-1} + \epsilon_{ijt}. \quad (5.1)$$

But, here ϵ_{ijt} 's are independently and identically normally distributed with mean 0 and variance σ_i^2 . As the distribution of the errors are known and normal, in the next section we construct a partial score test due to Neyman (1959). This test is developed for testing the mean level, that is $H_0 : \theta_{i0} = \theta_{20}$. Thus ϕ_i and σ_i^2 are considered to be nuisance parameters. This test is quite simple to construct, as it does not require the maximum likelihood estimates of the nuisance parameters, rather it requires consistent estimates for them. In this section, we now provide the estimates for ϕ_i and σ_i^2 .

Following Quenouille (1958), by pooling the information from all possible

mothers, we now estimate ϕ_i as

$$\hat{\phi}_i = \frac{\sum_{j=1}^{J_i^*} \sum_{t=2}^{n_{ij}} (y_{ijt} - \bar{y}_{ij})(y_{ijt-1} - \bar{y}_{ij})}{\sum_{j=1}^{J_i^*} \sum_{t=1}^{n_{ij}} (y_{ijt} - \bar{y}_{ij})^2} \quad (5.2)$$

where

$$\bar{y}_{ij} = \frac{\sum_{t=1}^{n_{ij}} y_{ijt}}{n_{ij}}. \quad (5.3)$$

The values of $\hat{\phi}_i$, an estimate of ϕ_i , are recorded in Table 5.1.

Table 5.1: Lag 1 Correlation $\hat{\phi}_i$

Group	Mg	Ca	Rb	Sr ^{1/2}
Full-term ($\hat{\phi}_1$)	0.303	0.170	0.442	0.111
Pre-term ($\hat{\phi}_2$)	0.134	0.124	0.159	-0.021

Note that Sutradhar *et. al.* (1998) calculated this lag 1 parameter in a loose fashion. More specifically, they computed the lag 1 correlation of the mean concentrations instead of the actual observations. These values were found to be 0.654, 0.562, 0.971 and 0.008 for the full-term group and 0.517, 0.024, 0.890 and 0.171 for the pre-term group, for Mg, Ca, Rb and Sr^{1/2}, respectively. Note, however, that the order of the magnitude of the correlations are the same both here and in Sutradhar *et. al.* (1998).

Now we calculate σ_i^2 by using all three formulas given in (4.23), (4.24) and (4.25). Further note that to test $H_0 : \theta_1 = \theta_2 = \theta$ (say), one may also estimate θ or put $\theta = 0$. We, however, choose to estimate θ considering this as a further nuisance parameter. The estimate is given by

$$\hat{\theta} = \frac{\sum_{i=1}^2 \sum_{j=1}^{J_i'} y'_{ij} \Lambda_{ij}^{-1}(\hat{\phi}_i) 1_{ij}}{\sum_{i=1}^2 \sum_{j=1}^{J_i'} 1'_{ij} \Lambda_{ij}^{-1}(\hat{\phi}_i) 1_{ij}} \quad (5.4)$$

and the values of $\hat{\theta}$ are recorded in Table 5.2.

Table 5.2: $\hat{\theta}$ (using $\hat{\phi}_i$)

Group	Mg	Ca	Rb	Sr $^{\frac{1}{2}}$
Full-term ($\hat{\phi}_1$)	20.443	236.759	0.372	4.735
Pre-term ($\hat{\phi}_2$)	25.900	218.168	0.581	5.202

5.2 Score Test for Mean Gestation Levels

In order to derive the score test, one requires to construct the score function

$\frac{\partial l}{\partial \theta}$, where l is the log-likelihood function of the data. Here, following model

(5.1) the log-likelihood is defined as

$$\begin{aligned} \log L &= l \\ &= \text{const} - \frac{1}{2} \sum_{i=1}^2 \frac{\sum_{j=1}^{J_i} [(y_{ij} - \theta 1_{ij})' \Lambda_{ij}^{-1}(\phi_i)(y_{ij} - \theta 1_{ij})]}{\sigma_i^2}. \end{aligned} \quad (5.5)$$

Thus the score function is given by

$$\frac{\partial l}{\partial \theta} = \sum_{i=1}^2 \frac{1}{\sigma_i^2} \sum_{j=1}^{J_i} [y'_{ij} \Lambda_{ij}^{-1} 1_{ij} - \theta 1'_{ij} \Lambda_{ij}^{-1} 1_{ij}]. \quad (5.6)$$

We further require, the second partial derivative of l with respect to θ , which may be simplified as

$$\frac{\partial^2 l}{\partial \theta^2} = - \sum_{i=1}^2 \frac{1}{\sigma_i^2} \sum_{j=1}^{J_i'} 1'_{ij} \Lambda_{ij}^{-1} 1_{ij}. \quad (5.7)$$

Now treat the estimates of the nuisance parameters computed in Section 5.1 as though they have been computed by maximizing the likelihood function under the null hypothesis. Consequently Neyman's (1959) partial score test statistic reduces to Rao's (1948) efficient score test statistic which is given by

$$T = \frac{\left(\frac{\partial l}{\partial \theta}\right)^2}{-\left(\frac{\partial^2 l}{\partial \theta^2}\right)}, \quad (5.8)$$

where $\frac{\partial l}{\partial \theta}$ and $\frac{\partial^2 l}{\partial \theta^2}$ are given by (5.6) and (5.7) respectively and they are evaluated at $\theta = \hat{\theta}$, $\sigma_i^2 = \hat{\sigma}_i^2$ and $\phi_i = \hat{\phi}_i$. The test statistic in (5.8) has a χ^2 distribution with 1 degree of freedom under the null hypothesis. The values of the test statistic are given in Table 5.3.

Table 5.3: T (using $\hat{\phi}_i$)

	Mg	Ca	Rb	Sr $^{\frac{1}{2}}$
$T(1)$	2.287	738.951	0.313	2.351
$T(2)$	3.739	630.743	0.020	0.793
$T(3)$	6.126	385.329	0.628	0.272

In Table 5.3, $T(1)$, $T(2)$ and $T(3)$ indicate that the test statistic T has been computed by using $\hat{\sigma}_i^2$ from (4.23), (4.24) and (4.25) respectively. Now, since T is approximately a χ^2 distribution with one degree of freedom we compare it with critical value $\chi_{0.05,1}^2 = 3.84$. Therefore the test statistic T , recorded in Table 4.3, from the score test, indicates that a significant difference exists between full-term and pre-term groups for Ca but generally no difference exist between groups for Mg, Rb, and Sr $^{\frac{1}{2}}$. These results are in general agreement with the results given in Sutradhar *et. al.* (1998) when time was treated as a specific factor.

Chapter 6

Concluding Remarks

Longitudinal patterns of mineral concentrations in human milk are of practical importance. It is also of practical importance to know whether there is any differences in mineral concentrations in the milk of full-term mothers and pre-term mothers, over time. This is because, as the amount of minerals in the milk of full-term mothers may be considered as the standard measure, one may determine the nutrient amounts that should be used for the premature born infant. As the study of the longitudinal patterns require the usage of statistical tests in determining differences between the two groups, the distributional patterns have to be taken into account when deciding upon the appropriate tests. The trace and ultratrace minerals in this practicum were

found to have asymmetric distributions which required special considerations when deciding upon proper statistical procedures.

When studying asymmetric distributions, the median is usually the parameter of interest and therefore appropriate statistical tests were used for testing the equality of the medians of the two groups. The traditional median tests, for example, those mentioned in Section 3.1.1 based on signs and/or ranks of the observations and ignore the dispersion of the data. As a remedy, in this practicum, we have discussed various statistical tests which take the dispersion of the data into account.

The tests were developed to study either weekly differences between the mineral concentrations of the two groups or overall differences.

When mineral concentrations were studied for individual weeks we found a significant difference between the pre-term and full-term groups in most time periods for Mn and Pb. For Co, we found a significant difference only for the first week. It was clear that in general, there were no significant differences between the full-term and pre-term groups for Mo, Ni and Sn. When the weeks were combined as a group it was again found that an overall significant difference existed between the pre-term and full-term groups for Mn and Pb but not for Co, Mo, Ni and Sn. The results of the tests were, fur-

ther in agreement with the graphical analysis indicating a good performance of the test procedures.

The macrominerals (Mg, Ca, Rb and Sr) were studied in detail by Sutradhar *et. al.* (1998) and therefore they were only discussed in brief. The minerals Mg, Ca and Rb along with the transformation $Sr^{\frac{1}{2}}$ were found to be symmetrically distributed. Therefore we performed a score test to test the equality of the means of the mineral concentrations of the pre-term and the full-term groups. Here an overall significant difference was found between the pre-term and full-term groups for Mg but no significant differences were found for minerals Ca, Rb, and $Sr^{\frac{1}{2}}$. These results are in agreement with the graphical analysis and test procedures performed by Sutradhar *et. al.* (1998).

The results of the practicum should be useful for the Pediatric societies in order to prepare better recommendations for the use of additional nutrients for the premature infants, on top of the milk they consume. The validity of the nonparametric test procedures for more smaller samples could be tested by conducting suitable simulation studies, which is beyond the scope of the present practicum.

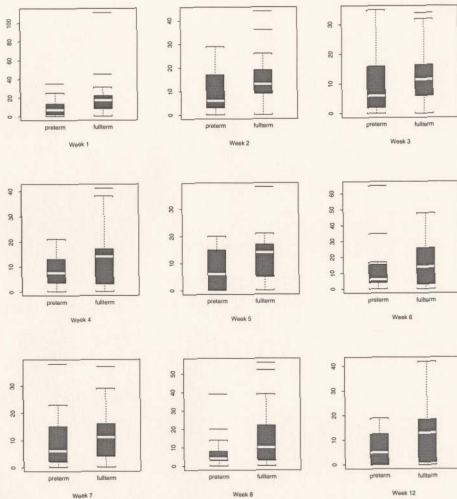


Figure 2.1.1: Boxplots for Mn

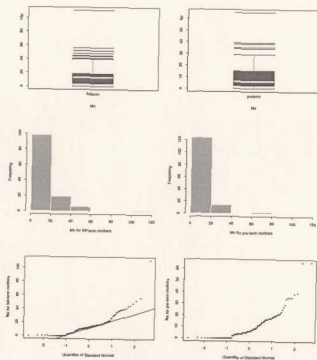


Figure 2.1.2: Boxplot, Histogram and q-q Plot for Mn

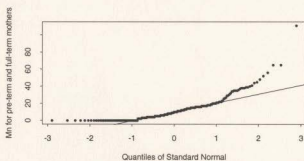
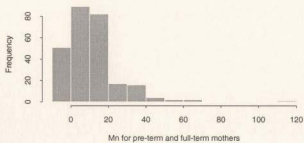


Figure 2.1.3

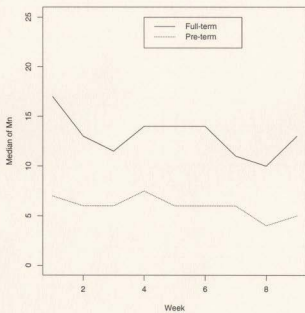


Figure 2.1.4

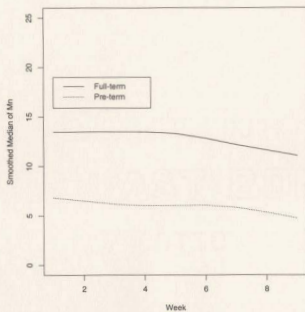


Figure 2.1.5

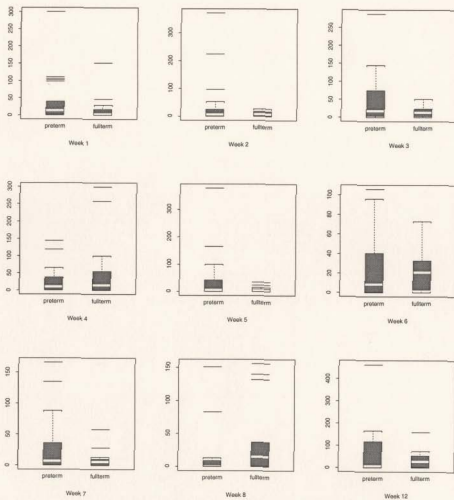


Figure 2.2.1: Boxplots for Ni

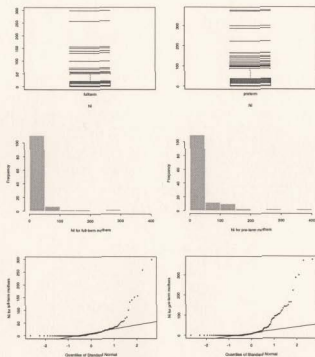


Figure 2.2.2: Boxplot, Histogram and q-q Plots for Ni

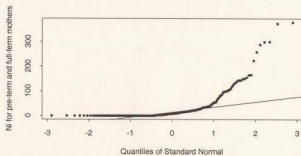
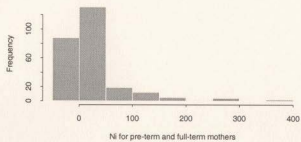


Figure 2.2.3

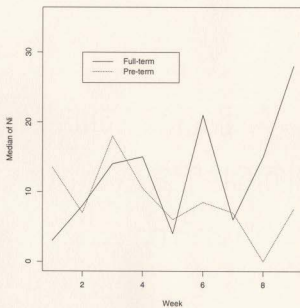


Figure 2.2.4

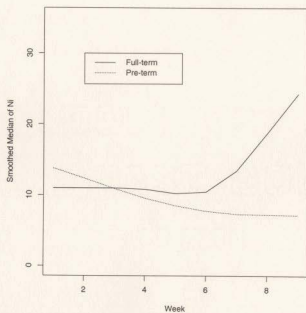


Figure 2.2.5

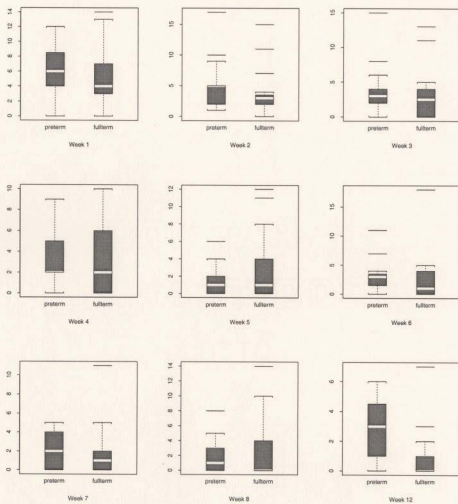


Figure 2.3.1: Boxplots for Mo

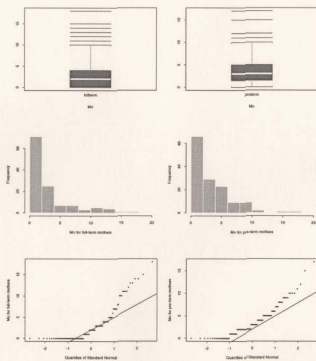


Figure 2.3.2: Boxplot, Histogram and q-q Plots for Mo

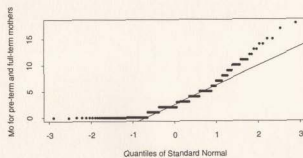
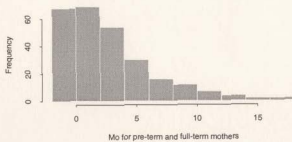


Figure 2.3.3

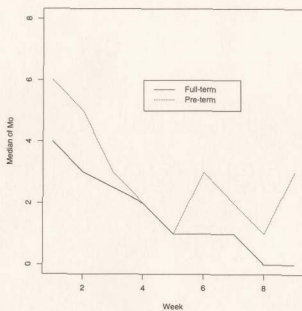


Figure 2.3.4

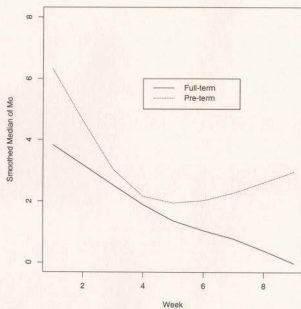


Figure 2.3.5

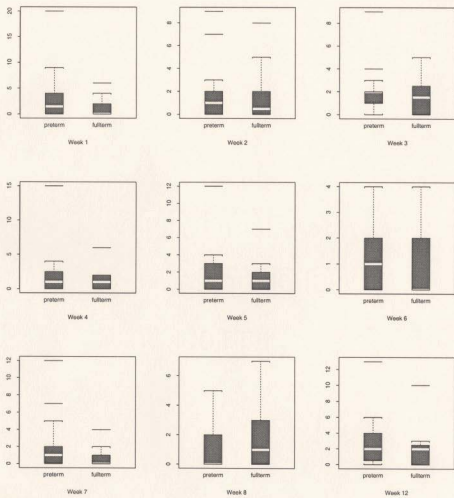


Figure 2.4.1: Boxplots for Co

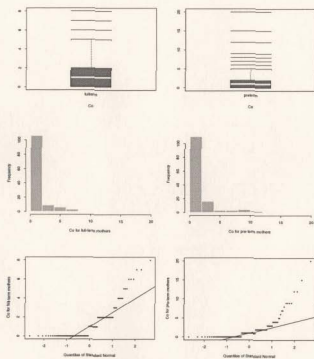


Figure 2.4.2: Boxplot, Histogram and q-q Plots for Co

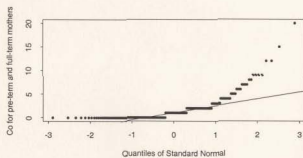
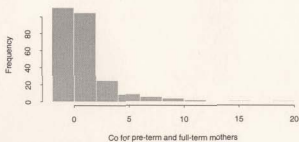


Figure 2.4.3

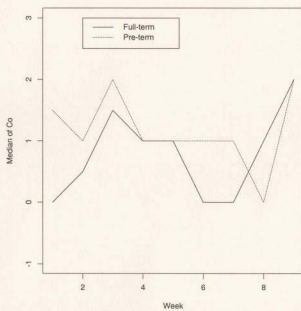


Figure 2.4.4

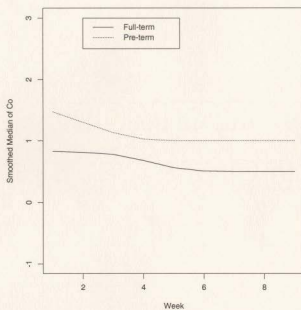


Figure 2.4.5

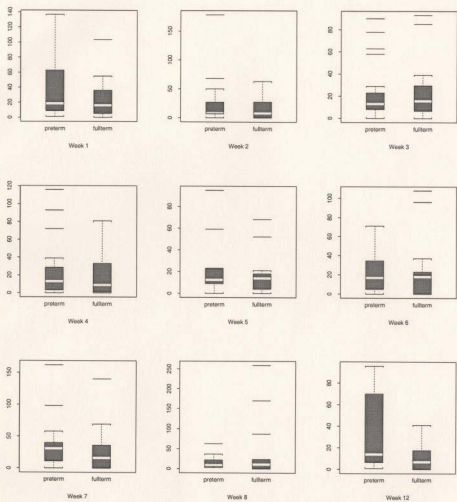


Figure 2.5.1: Boxplots for Sn

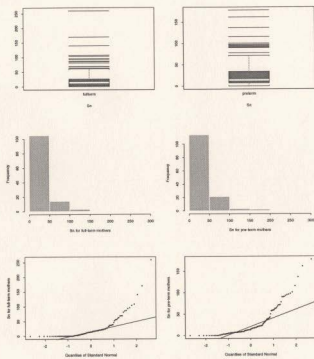


Figure 2.5.2: Boxplot, Histogram and q-q Plots for S_n

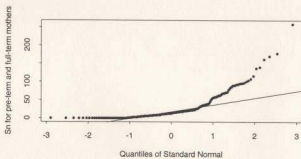
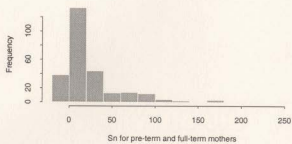


Figure 2.5.3

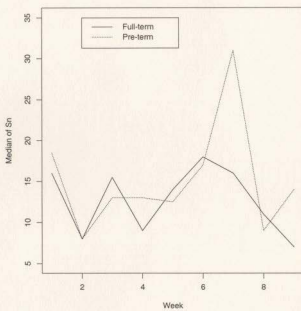


Figure 2.5.4

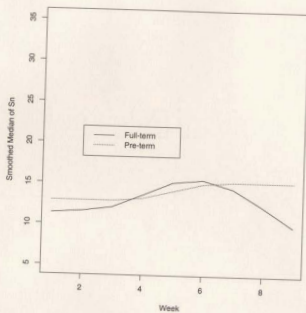


Figure 2.5.5

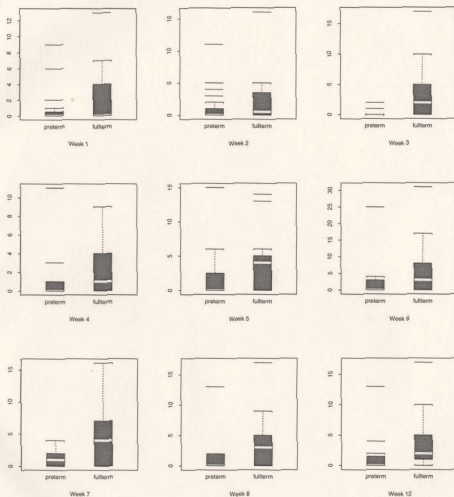


Figure 2.6.1: Boxplots for Pb

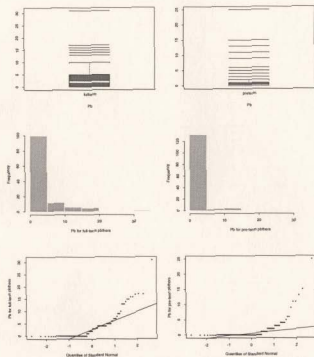


Figure 2.6.2: Boxplot, Histogram and q-q Plots for Pb

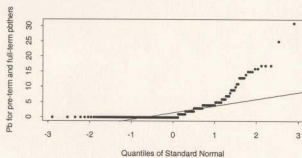
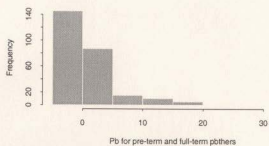


Figure 2.6.3

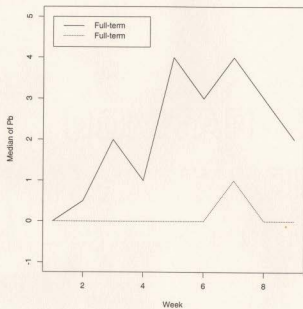


Figure 2.6.4

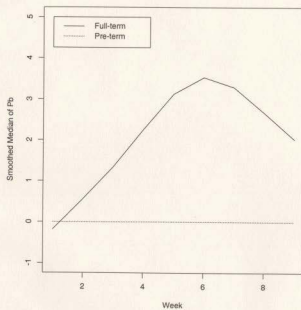


Figure 2.6.5

References

- Atkinson, S. A., Brunton, J., Payes, B., Fraser, D., and Whyte, R. (1990). Calcium and phosphorus fortification of mother's milk and formula for premature infants: metabolic balance of calcium, phosphorus and zinc at two postnatal ages, *FASEB Journal*, **4**, 1393-1398.
- Efron, B. and Gong, G. (1983). A Leisurely Look at the Bootstrap, the Jackknife, and Cross-Validation, *The American Statistician*, **37**, 36-48.
- Fomon, S.J., Ziegler, E. E., and Vazquez, H. D. (1977). Human milk and small premature infant, *American Journal of the Dis Child*, **131**, 463-467.
- Hill, R. W. and Holland, P.W. (1977). Two robust alternatives to Least-Squares regression, *Journal of the American Statistical Association*, **72**, 1041-1067.
- Kashyap, S., Schulze, K., Forsyth, M. and Dell, R. B. (1990). Growth, nutrient retention, and metabolic responses of low-birth-weight infants fed supplemented and unsupplemented preterm human milk, *Journal of Pediatric Gastroenterol Nutrition*, **52**, 254-262.
- Mondanlou, H. D., Lim, M. O., Hansen, J. W. and Sickles, V. (1986). Growth, biochemical status, and mineral metabolism in very-low-birth-weight infants receiving fortified preterm human milk, *Journal of Pediatric Gastroenterol Nutrition*, **5**, 762-767.
- Newman, J. (1959). Optimal asymptotic tests for composite hypothesis. In *Probability and Statistics*, (U. Grenander, Ed), 213-234, Wiley, New York.
- Quenouille, M. H. (1958). The comparison of correlations in Time Series, *Journal of Royal Statistical Society, Series B*, **20**, 158-168.
- Rao, C. R. (1948). Large Sample Tests of Statistical Hypothesis Concerning Several Parameters with Applications to Problems of Estimates, *Proc. Cambridge Philos. Soc.*, **44**, 50-57.

- Sutradhar, B. C., Dawson, B. and Friel, J. (1998). On the Quality of Preterm Infants Formula and the Longitudinal Change in Mineral Contents in Human Milk. *In Quality Improvement Through Statistical Methods*, ed. B. Abraham, Birkhauser, Boston, USA, 423–437.
- Venables and Ripley (1996). *Modern Applied Statistics with S-plus*, Springer, New York.

