Defining the normal range of α-tocopherol for pregnant women

Essay in Statistics, Spring 2007

Department of Economics and Society, Dalarna University

Author **Tong Zhou**

Supervisor **Johan Bring**

Date **June, 2007**

Abstract

Pregnant women are under the risk of developing pre-eclampsia. Some studies indicate that women with pre-eclampsia have higher values of isoprostanes, which includes α-and γ-tocopherol, compared to women without pre-eclampsia. To be able to predict whether the pregnant women have the danger of pre-eclampsia or not, it is important to know the normal values of α-and γ-tocopherol. In my paper, I concentrate on α-tocopherol and based on a sample of 52 women, estimate the normal range for pregnant women without pre-eclampsia. With the normal range, we can identify the women with high values. In order to estimate the values of each week, we derived a model that describes the percentiles as a function of the time (weeks) during the gestational period was estimated.

Through the parametric and non-parametric methods, the percentiles for each observed week were computed. The percentile curves were derived basing on the percentile estimates. Also use mean and sd method to estimate percentiles as another method. We can estimate the α-tocopherol of specific weeks through the function.

Key words: normal range, α-tocopherol, pregnant women, percentiles

1. Introduction

Along with the developments in medical sciences, the births of the babies are safer for the mothers and infants nowadays than before. But pregnant women are under the risk of developing pre-eclampsia. Oxidative stress is associated with normal human pregnancy. Several studies on animals and humans have been published on the etiology of pre-eclampsia focusing on the role of oxidative stress in recent years (Palm). However, we know little about the involvement of oxidative stress in uncomplicated normal human pregnancy (Palm). There are studies indicating that women with pre-eclampsia have higher values of isoprostanes (α -and γ-tocopherol) as compared to the women without pre-eclampsia. It occurs in 5–10 percent of all pregnancies and is still one of the leading causes of maternal and fetal mortality in developed and developing countries (Plasma Concentrations of Carotenoids,Retinol, and Tocopherols in Preeclamptic and Normoyensive Pregnant Women). We can see that pre-eclampsia is quite dangerous for pregnant women and it is threatening them.

To be able to predict who will develop pre-eclampsia, it's important to know what are the normal values of α-and γ-tocopherol so that women with high values can be identified. In my paper, α-tocopherol is concentrated on, which is traditionally recognized as the most active form of vitamin E in humans, and is a powerful biological antioxidant. Tocopherol, or vitamin E, is a fat-soluble vitamin in eight forms that is an important antioxidant. Natural vitamin E exists in eight different forms or isomers, four tocopherols and four tocotrienols.
(http://en.wikipedia.org/wiki/Tocopherol, 1 May 2007).

2. Data

The data set of 52 pregnant women is taken from an outpatient antenatal clinic in Uppsala City, which is consecutively recruited into this study during 2003-2004.

(1) Variables

The variable which studied is α-tocopherol, which was labeled atocopherol when the data was analyzed and its unit is mg/mmol.

The time was measured as per day and can be calculated per week, 2-week and month according to the days. Therefore, the time variable (the length of pregnancy) can be measured as per day, per week, per 2-weeks or per month.

The observations of α-tocopherol disperse on 168 different days between the 53rd and the 300th pregnant day when the time variable measured by days. The observations are dispersed all over the time points and each time point does not have enough values. Further, we can see the box plot of α-tocopherol changes per days do not show a clear picture. There is a single value on various time points.

When the time variable is measured in weeks, the observations disperse on 33 different weeks i.e. between the 8th and 43rd pregnancy week. The situation is clearer when we use weeks as time unit.

The measurement of α-tocopherol in every two weeks is taken from the weekly data i.e. the odd weeks are classified to the nearest even weeks. Therefore, the observations disperse on 18 different weeks i.e. between the 8th and the 42nd pregnant week. Obviously, this time the data set is better than the above mentioned data sets. On each time point, we have more observations and it will not affect the accuracy of the data information. Thus the 2-week time is taken as a time variable for the data set. In the Method and Result parts, the data which I used is 2-week data if there is no special mentioned.

(2) Deal with part of the raw data

When the data is transformed from week to 2-week, sometimes it causes a problem. That is the same woman can have two observations for the same twoweek-period.

All the data with this problem is shown in Table 1. According to Table 1, the number 110 pregnant woman has two α-tocopherol values in the 32nd week. These two values were tested on the 218th and 231st day of pregnancy separately. And 218 days make 31 weeks and 1 day which are approximately 32 weeks, whereas 231days make 33 weeks which is not the approximation of 32 weeks. So we choose the value of the first row and delete the second row. Similarly, about the third and the fourth row 272 days is nearer to 38 weeks than 273 days. Of course, there is only one day difference so it seems there no significant difference. Through the same method, the values of the number 1, 3,6,7,9,11,13,16 and 17 rows were chosen and omit the other rows.

So when the variables which are measured in 2-week will be analyzed and the same woman have two observations for the same twoweek-period, the values which I have chosen will be used and the omitted ones will be ignored.

Here one of the two values was used, because there will be less variation than using mean of them.

Table 1 The different α-tocopherol values in the same week of the same pregnant woman

Number	patientnr	α-tocopherol	week	Days
1	110	1.81	32	218
2	110	1.95	32	231
3	110	2.08	38	272
4	110	1.84	38	273
5	112	2.08	40	276
6	112	2.28	40	283
7	113	1.87	38	261
8	113	1.80	38	273
9	133	1.64	40	278
10	133	1.69	40	284
11	138	2.22	40	277
12	138	2.18	40	285
13	141	1.85	36	251
14	141	1.94	36	257
15	144	1.53	40	275
16	144	1.70	40	283
17	145	1.54	40	277
18	145	1.65	40	284

The study aim is to estimate the normal range of α-tocopherol for the women during their gestational period. Therefore, only data during their pregnancy period was used.

As mentioned above, 2-week was chosen as the time variable. In the 8^{th}, 14^{th}, 16^{th} and 18^{th} week, the number of observations is less than 10 which is not good for analyzing because of the small sample size. But deleting them can cause a loss of some information. Hence those observations will be kept while analyzing. But for some part of the analysis, these values will be deleted as some analysis require of the sample size.

3. Method

3.1 Definition of normal range

In this article, I want to define the normal range of the data set which comes from the pregnant women. Sometimes normal range also called reference interval or reference range. In the health related issues, the doctors or other related professionals often use the normal range to interpret some results of a particular case. There are two kinds of definition about normal range: one is 2.5^{th} to 97.5^{th} percentile; the other is biologically motivated limits.

(1) 2.5^{th} to 97.5^{th} percentile

The normal range is set to cover a ninety-five percent of values from a general population. Five percent of the results fall outside the normal range. Generally, the normal range is an interval between two symmetrical percentiles of a size variable.

(2) Biological motivated limits

This definition follows the standards of risk measurement. Usually, the professionals of that special field define the standards.

Take Normal blood pressure as an example.

There is a standard of blood pressure in medical science.

For most adults, a normal blood pressure is less than 120/80 mm Hg. If the blood pressure is at or above 140/90mm Hg, people needs drug treatment. And drug treatment is needed for a blood pressure level of 130/80 mm Hg or higher if people with diabetes or chronic kidney disease.

Table 2 Ranges for the four blood pressure categories

Category	Systolic (mm Hg)		Diastolic (mm Hg)	Blood Pressure Reading (mm Hg)
Normal**	below 120	and	below 80	below 120/80
Pre-hypertension	120-139	or	80-89	120/80-139/89
Stage 1 High Blood Pressure (hypertension)	140-159	or	90-99	140/90 - 159/99
Stage 2 High Blood Pressure (hypertension)	160+	or	100 +	160/100 +

Note: Blood pressure that is too low can be dangerous.

(http://www.diovan.com/info/answers/normal_blood_pressure.jsp, 30 April 2007)

3.2 Estimate the percentiles from a data set

Estimate the percentiles for each week (meaning twoweek). This could be done in two different ways.

(1) Parametric

Under this method, first of all we should know the distribution of the data set. In this article, to study the percentiles for each week was the aim. So the α-tocopherol distribution of each week was interested in. If it follows a certain distribution for example, normal or gamma distribution, the distribution will be chosen to estimate. Here whether they have normal distribution or not were checked. If they are normal, 2.5^{th} to 97.5^{th} percentile will be used to estimate the percentiles.To know whether the data per each week is normal or not, Shapiro-Wilk normality Test (Shapiro and Wilk 1965) is used. The null hypothesis is that the data set is normally distributed. This test requires the sample size must between 3 and 5000.

If the data do not follow a normal distribution, the data set can be transformed. Logarithmic Transformation was used in this article. Logarithmic Transformation transforming Y based on a logarithmic function $\ln(Y-\tau)$. (Royston 1991)

If the values of the week have a normal distribution, the mean and standard deviation for this week will be calculated and the 2.5th and 97.5th percentiles will be computed. According to the formula,

$$centile_{week} = mean_{week} + K \times SD_{week} ,$$

where $mean_{week}$ and SD_{week} are the mean and standard deviation at the required week respectively. K is the desired normal equivalent deviate (NED).

I want to estimate 95% normal range (i.e. 2.5th and 97.5th percentiles), which requires $K = \pm 1.96$.

(2) Non-parametric

Quantile returns estimates of underlying distribution quantiles which based on one or two order statistics from the supplied elements in x at probabilities in R. (The help of R) There are nine quantile algorithms in R. Type 7 is the default one. Its formula is p(k) = (k - 1) / (n - 1),where p(k) = mode[F(x[k])]. And Type 9’s formula is p (k) = (k - 3/8) / (n + 1/4). If the data follows a normal distribution, the resulting quantile estimates are approximately unbiased for the expected order statistics when using type 9. (The help of R) Since the samples for most of the weeks have a normal distribution, the Type 9 was used to estimate, and for the sample which is not normal, Type 7 was chosen.

Since these estimates will be very unstable due to the small sample size we need to smooth the estimates over time. This could be done in many different ways. The samples for most of the weeks have a normal distribution, so only one smoothing technique was used.

3.3 Centile-regression

One approach is to estimate a regression model with the percentile as the dependent variable and time as the explanatory variables. This could be done to both the parametric and non-parametric estimates above. Here this was done to the parametric estimates. Since the variability of individual order statistics, the sample quantiles are inefficient.

Fitting a model which is a percentile of α-tocopherol (y) against the number of week (t), the formula is $y = a + bt + ct^2 + dt^3$. If c and d are significant then they will be part of the model, if not, there will be no b or c in the final results.

3.4 Regression and residuals (Mean and SD method)

This method is not based on any percentile estimates (like 3.2 (1) and (2) above).

(1) Estimate a mean curve

With this method we first estimate a function describing how the mean value varies over time. I would estimate this function based on the original data.

Fitting a polynomial model of mean uses $y = a + bt + ct^2 + dt^3$, where y is the mean of different weeks from the raw data and t is the number of week. Firstly, a cubic polynomial will be fitted. If coefficient d is not significant, a quadratic one will be kept. And if c is also not significant, a linear model will be gotten.

(2) Estimate SD of residuals

From the regression analysis we get the residuals[1]. If the variation is constant over time we can estimate the standard deviation by just calculate the standard deviation among the residuals. If the variation is not constant we need a more sophisticated method to estimate the standard deviation.

Since the data set per each week is normally distributed, the residuals from the mean model should also be normally distributed. That is the absolute residuals have a half normal distribution. The mean of a half normal distribution is $\sqrt{2/\pi}$, so the estimate of the SD of residuals is the mean of the absolute residuals multiplied by $\sqrt{\pi/2}$.(Silverwood and Cole,2007)

If the absolute residuals show no trend with week, SD is estimated by the mean of the absolute residuals multiplied by $\sqrt{\pi/2}$.

If the variation is not constant, a polynomial regression will be used to estimate a SD curve according to the weeks.

(3) Estimate percentiles based on (1) and (2)

When we have the estimates for the mean (as a function of time) and the standard

[1] Observed value minus predicted value

deviation (also perhaps a function of time), then we can estimate the percentiles as:

$$\hat{\mu}_t \pm 1.96 \cdot \hat{\sigma}_t$$

For my data there seems to be a trend in the mean value while the variation seems to be the same over time.

(4) Centile-regression

Using the estimates of (3) smooth percentile curves.

3.5 Assessing model fit

Z-scores are a useful tool in assessing model fit. Using Z-scores to assess the model fit,

$$Z = \frac{observedatocopherolvalue - mean_{week}}{SD_{week}},$$

where $mean_{week}$ and SD_{week} are the mean and standard deviation given by the models above for each week.

(1) Plotting the Z-scores against week.

(2) Shapiro-Wilk normality test

4. Results

4.1 Overview of all the data

(1) Here a density histogram of α-tocopherol was added which has the smooth density curve of α-tocopherol.

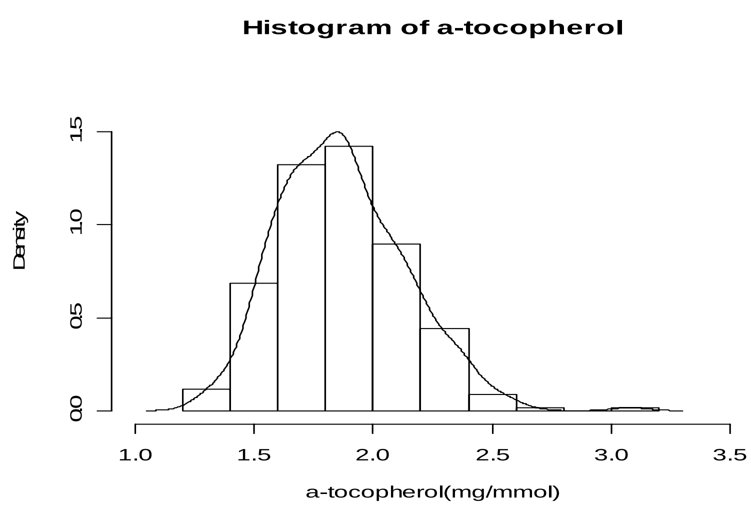


Figure 1 Histogram of the all α-tocopherol data.

For all the α -tocopherol observations which include all weeks of pregnancy during the gestational period. We apply Shapiro-Wilk normality test. The null hypothesis of Shapiro-Wilk normality test is the data follow a normal distribution. And we get W = 0.9812 whereas p-value = 0.0001943.Since p-value= 0.0001943<0.05, we reject H_0 i.e. we do not think α -tocopherol has a normal distribution.

(2) The box plots of the α -tocopherol per week and per 2-week are shown below.

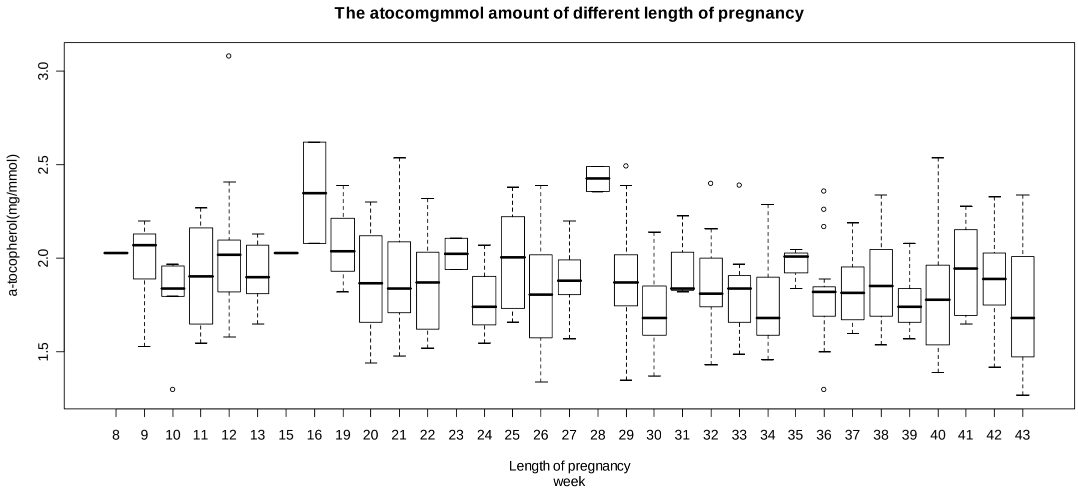


Figure 2 Box plot of the α -tocopherol per week.

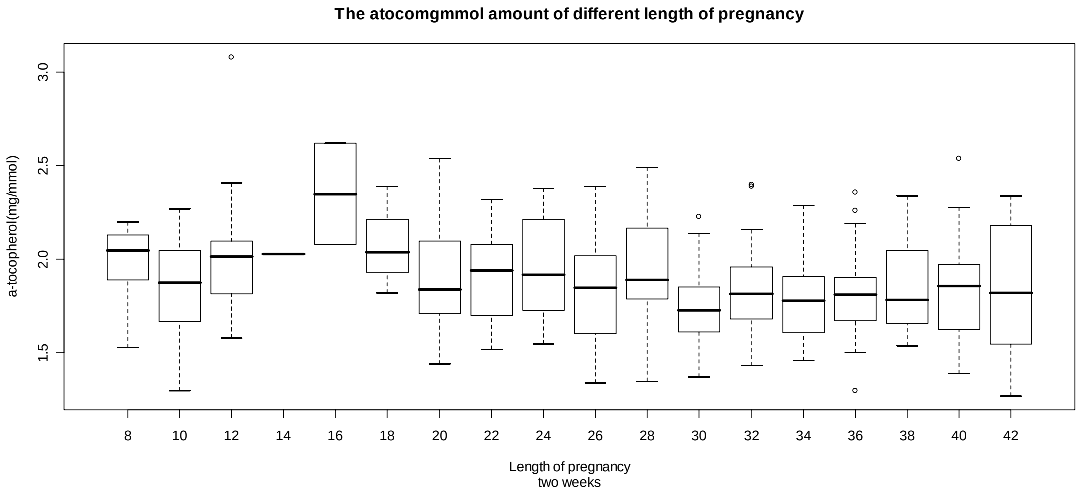
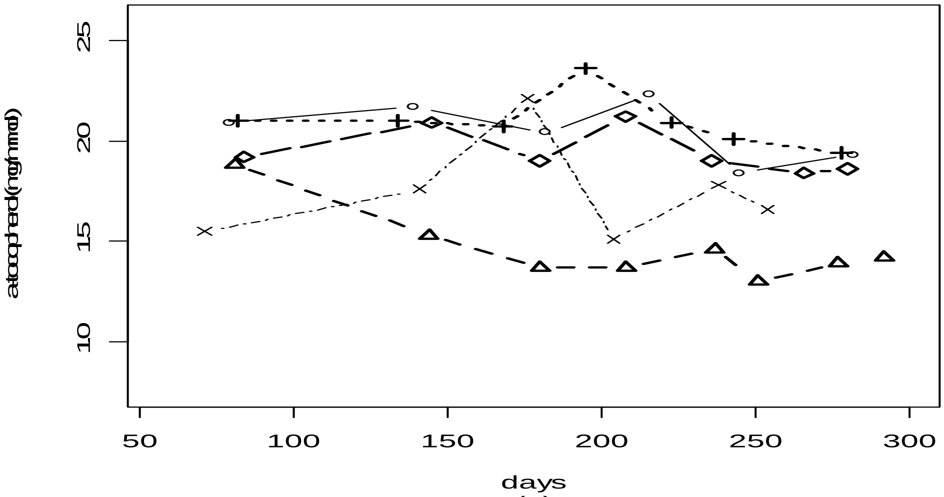
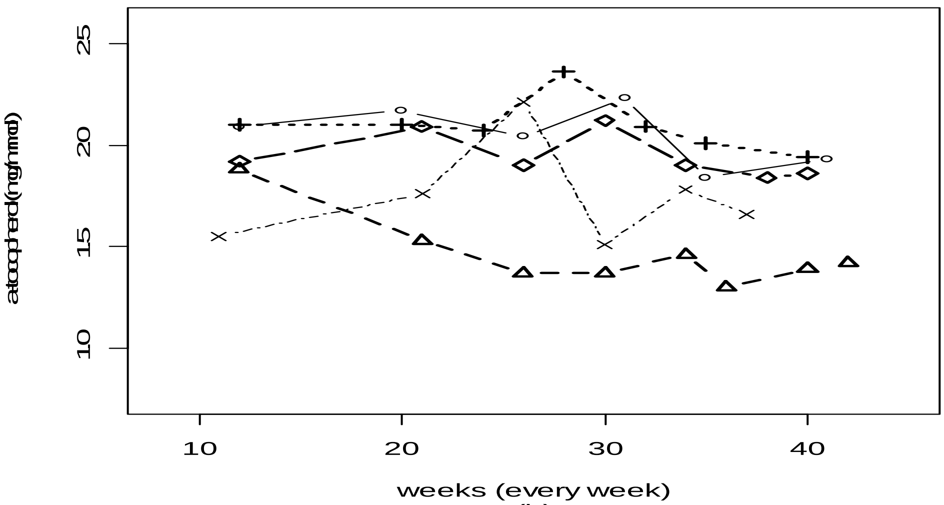


Figure 3 Box plot of the α -tocopherol per 2-week.

(3) The trend of α -tocopherol changing individually for the raw data

For the raw data, I think it is necessary to show the trend of α -tocopherol amount. Here I track No.100,101,104,105,106 pregnant women per day and per week respectively.


25
20
15
10
50
100
150
200
250
300
days
(a)


25
20
15
10
10
20
30
40
weeks (every week)
(b)

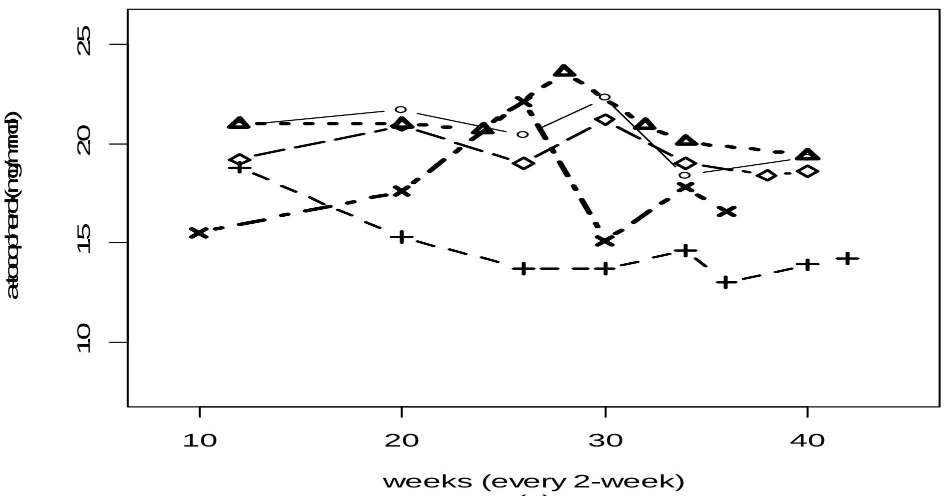

25
20
15
10
10
20
30
40
weeks (every 2-week)
(c)

Figure 4 (a) The amount of α-tocopherol changes in terms of the pregnancy days for 5 pregnant women; (b) The amount of α-tocopherol changes in terms of the pregnancy weeks for 5 pregnant women; (c) The amount of α-tocopherol changes in terms of the pregnant even weeks for 5 pregnant women.

These are graphs to indicate that how the amount of α-tocopherol changes in terms of pregnancy days, weeks and 2-weeks for each pregnant woman.

4.2 Estimate the percentile from a data set

(1) Parametric

(a) Distribution

Figure 5 (a) shows a scatter plot of α-tocopherol amount against 2-week for all the observed women during their pregnancy period. The data can be analyzed to get normal range for each week. Figure 5 (b) follows three pregnant women over time respectively.

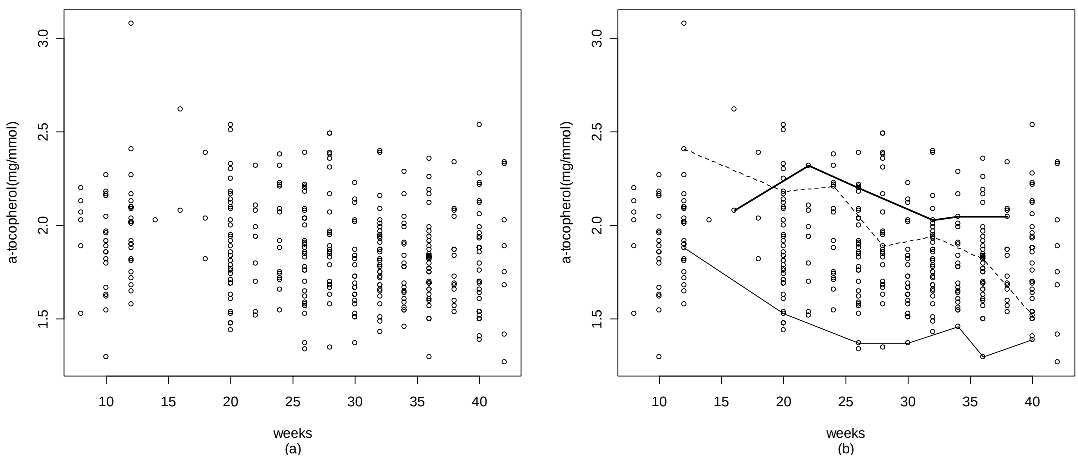


Figure 5 (a) Scatter plot of α-tocopherol against two week; (b) Follow 3 pregnant women over time respectively.

Shapiro-Wilk normality test was used to test whether the observations have a normal distribution or not. The results are shown in Table 3.

We do not reject H_0 for all weeks except the 12th week, i.e. we do not have enough evidence to reject that the α-tocopherol has a normal distribution.

The measurement which we are interested within 12th week is not normally distributed. Therefore, the data of the distribution of 12th week should be transformed, i.e. do some transformations of the data then define a distribution of it, and we try our best to transform it towards normal distribution.

Table 3 Shapiro-Wilk normality test

week	W	p-value	Distr.	week	W	p-value	Distr.
8	0.8621	0.1965	Normal	26	0.9766	0.6449	Normal
10	0.9692	0.7828	Normal	28	0.9469	0.2129	Normal
12	0.8508	0.002261	Non.[2]	30	0.9585	0.4329	Normal
14	NA[3]	NA		32	0.9636	0.3434	Normal
16	NA	NA		34	0.9476	0.3063	Normal
18	0.983	0.75	Normal	36	0.9676	0.3639	Normal
20	0.969	0.3811	Normal	38	0.9247	0.2571	Normal
22	0.9601	0.7874	Normal	40	0.9684	0.4766	Normal
24	0.9278	0.2529	Normal	42	0.9485	0.6962	Normal

The box plots below are α-tocopherol (the left one) and log (α-tocopherol) (the right one) of the 12th week. I use the logarithmic transformation, computing the natural logarithms.

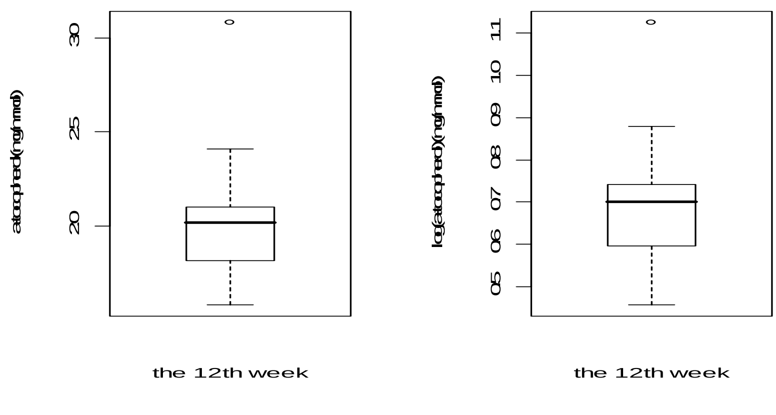


Figure 6 Box plots of α-tocopherol (the left one) and log (α-tocopherol) (the right one) of the 12th week.

[2] Non. means nonnormal.

[3] The sample size must be between 3 and 5000 when using the Shapiro-Wilk normality test. The number of observations of the 14th and 16th week is less than 3, so there are no outputs.

Using Shapiro-Wilk normality test after the logarithmic transfornation, W = 0.9208, p-value = 0.06099. Since 0.06099> 0.05, then it indicate that it follows normal distribution after logarithmic transformation.

(b) Estimates

Under the assumption that in each 2-week α -tocopherol has a normal distribution,

$$centile_{week} = mean_{week} + K \times SD_{week} ,$$

where $mean_{week}$ and SD_{week} are the mean and standard deviation at the required week respectively. K is the desired normal equivalent deviate (NED).

I want to estimate 95% normal range (i.e. 2.5^{th} and 97.5^{th} percentile curves), which requires $K = \pm 1.96$ .The results are shown in Table 4.

Table 4 Percentiles of α -tocopherol for each week (parametric)

week	mean	Standard Deviation	2.5^{th} percentile	97.5^{th} percentile
8	1.98	0.24	1.50	2.45
10	1.87	0.25	1.38	2.36
18	2.08	0.29	1.52	2.65
20	1.90	0.28	1.34	2.45
22	1.90	0.26	1.39	2.39
24	1.96	0.27	1.44	2.49
26	1.83	0.25	1.34	2.33
28	1.96	0.30	1.37	2.56
30	1.76	0.22	1.32	2.20
32	1.84	0.23	1.39	2.28
34	1.79	0.22	1.35	2.22
36	1.81	0.22	1.38	2.24
38	1.83	0.24	1.37	2.30
40	1.83	0.28	1.28	2.38
42	1.84	0.39	1.07	2.60

For the 12th week, the observations after logarithmic transformation, we have a normal distribution. Therefore, the percentiles was calculated basing on the log (α-tocopherol), then transform them with natural exponential and compute the percentiles.

After logarithmic transformation, the mean is 0.6864101, standard deviation is 0.1398193 and the percentiles of log (α-tocopherol) are 0.4123643 and 0.960456. We can get mean and percentiles as 1.986571, 1.510385 and 2.612888 respectively.

(2) Non-parametric method

Table 5 is the non-parametric 2.5th and 97.5th percentile in term of every 2-week. Here non-parametric percentiles of the weeks which had the sample size not less than 10 were calculated.

For 12th week, the type 7 method in R was used to calculate it, for other weeks I use type 9 method because the data of these weeks have normal distributions.

Table 5 Percentiles of α-tocopherol for each week (non-parametric)

week	2.5th percentile	97.5th percentile	week	2.5th percentile	97.5th percentile
10	1.30	2.27	30	1.37	2.23
12	1.62	2.69	32	1.44	2.40
20	1.45	2.53	34	1.46	2.29
22	1.52	2.32	36	1.36	2.33
24	1.55	2.38	38	1.54	2.34
26	1.35	2.35	40	1.39	2.50
28	1.35	2.49	42	1.27	2.34

(c) Comparison

To compare the percentiles of parametric and non-parametric, a plot was given.

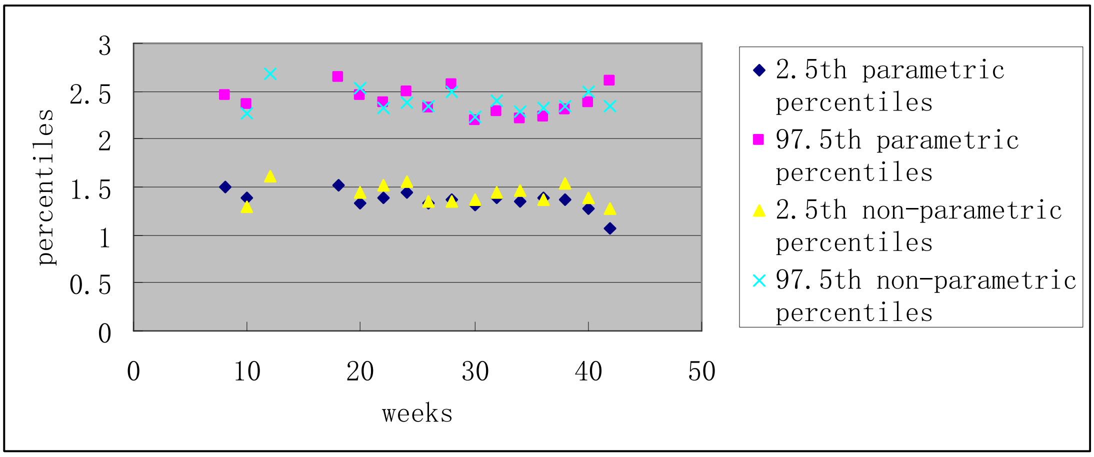


Figure 7 A plot of parametric and non-parametric percentiles.

4.3 Centile-regression

Basing the parametric percentile values above, we get the 2.5th and 97.5th percentile regression curves as y_1=1.55444-0.00696t and y_2=2.534381-0.004824t respectively, where y_1 is the 2.5th percentile, y_2 is the 97.5 percentile and t is the number of week.

The mean curve is y_{mean1}=2.023307-0.005267t.

The graph of three lines is shown in Figure 8.

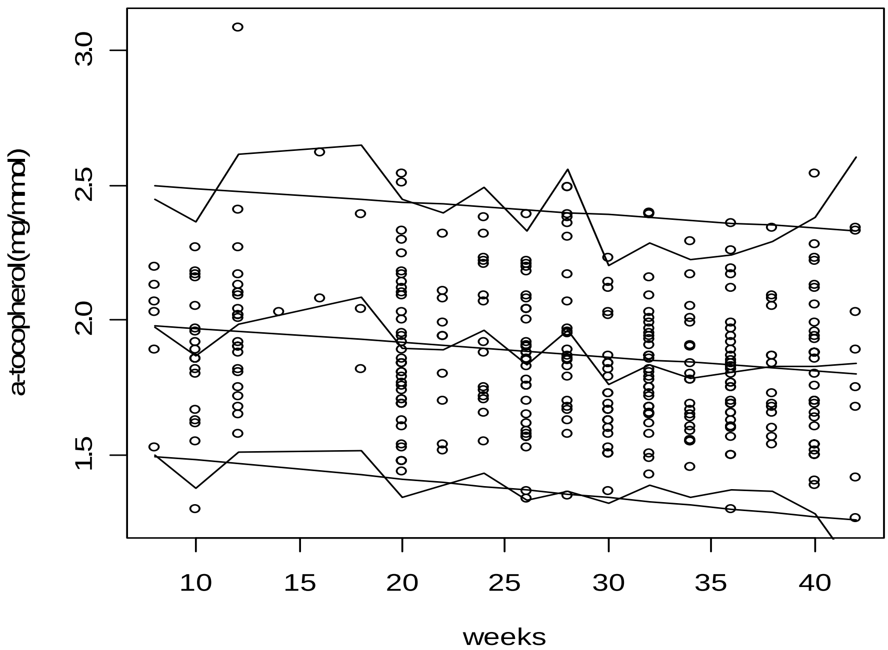


Figure 8 Fitted 2.5th, mean, 97.5th curves of α-tocopherol versus weeks.

4.4 Regression and residuals

(1) Estimate a mean curve

Using the mean per week and the number of week in Table 5, the mean is modeled by fitting a polynomial curve. y_{mean2} =2.014221-0.004998t, where y_{mean2} is the mean and t is the number of the week.

(2) Estimate SD of residuals

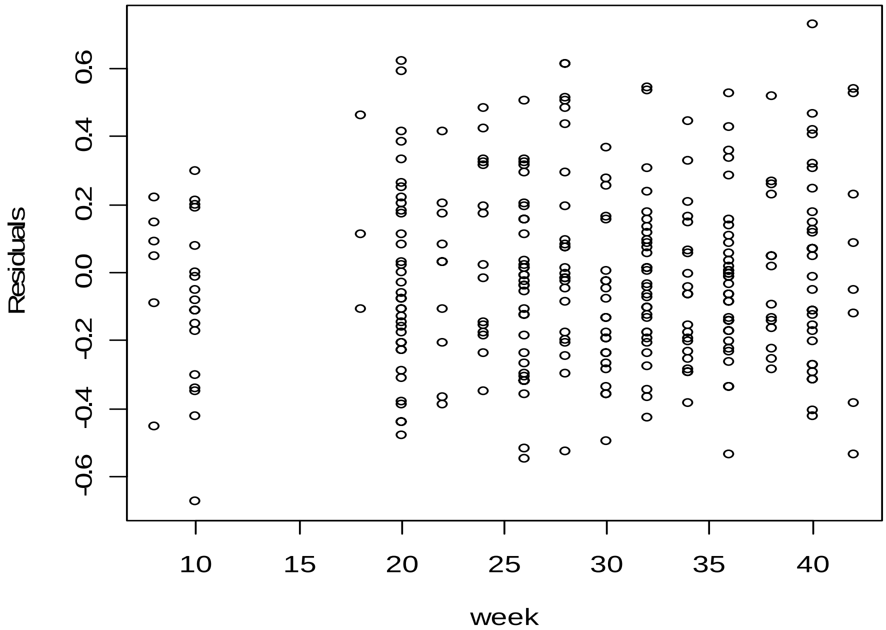


Figure 9 Plot of residuals of the fitted mean model versus weeks.

Since the observations follow normal distributions, the absolute residuals have a half normal distribution. So an estimate of the SD of the residuals is the mean of the absolute residuals multiplied by $\sqrt{\pi/2}$.

If we thought that the variability does not change with week. Therefore, SD is estimated by the mean of the absolute residuals multiplied by $\sqrt{\pi/2}$ .The mean of the absolute residuals is 0.2074050, SD of residuals is 0.2599437.

If we think there seems to be a trend in the mean value while the variation seems to be the same over time. A polynomial regression was used to estimate a SD curve according to the weeks. SD per week is estimated by the mean of absolute residuals per week multiplied by $\sqrt{\pi/2}$. Then I use polynomial regression.

The fit of the SD model is y_{sd} =0.2390821+0.0009878t, where t is the number of the week.

(3) Estimate percentiles based on (1) and (2)

(a) No variability

According to the formula:

$$centile_{week} = mean_{week} + K \times SD_{week} ,$$

where $mean_{week}$ and SD_{week} are the mean and standard deviation given by the models and $K = \pm 1.96$.

Table 6 Percentiles of α -tocopherol for each week (no variability)

week	Estimated mean	Estimated sd	2.5^{th} percentile	97.5^{th} percentile
8	1.97	0.26	1.46	2.48
10	1.96	0.26	1.45	2.47
18	1.92	0.26	1.41	2.43
20	1.91	0.26	1.40	2.42
22	1.90	0.26	1.39	2.41
24	1.89	0.26	1.38	2.40
26	1.88	0.26	1.37	2.39
28	1.87	0.26	1.36	2.38
30	1.86	0.26	1.35	2.37
32	1.85	0.26	1.34	2.36
34	1.84	0.26	1.33	2.35
36	1.83	0.26	1.32	2.34
38	1.82	0.26	1.31	2.33
40	1.81	0.26	1.30	2.32
42	1.80	0.26	1.29	2.31

(b) Have variability

If we think there seems to be a trend in the mean value while the variation seems to be the same over time.

Percentiles are estimated using the estimates of mean and SD by the model.

Table 7 Percentiles of α -tocopherol for each week (have variability)

week	Estimated mean	Estimated sd	2.5^{th} percentile	97.5^{th} percentile
8	1.97	0.25	1.49	2.46
10	1.96	0.25	1.48	2.45
18	1.92	0.26	1.42	2.43
20	1.91	0.26	1.41	2.42
22	1.90	0.26	1.39	2.42
24	1.89	0.26	1.38	2.41
26	1.88	0.26	1.37	2.40
28	1.87	0.27	1.35	2.40
30	1.86	0.27	1.34	2.39
32	1.85	0.27	1.32	2.38
34	1.84	0.27	1.31	2.38
36	1.83	0.27	1.30	2.37
38	1.82	0.28	1.28	2.37
40	1.81	0.29	1.27	2.36
42	1.80	0.28	1.25	2.35

(4) Percentile regression

For the situation (3) (a), basing on the percentile values in Table 6, we get the 2.5^{th} and 97.5^{th} percentile regression curves as y_1 =1.504731-0.004998t and y_2 =2.523711-0.004998t respectively, where y_1 is the 2.5^{th} percentile, y_2 is the 97.5 percentile and t is the number of week.

For the situation (3) (b), basing on the percentile values in Table 7, we get the 2.5^{th}

and 97.5th percentile regression curves as y_1 = 1.545620-0.006934t and y_2 =2.482822-0.003062t respectively, where y_1 is the 2.5th percentile, y_2 is the 97.5 percentile and t is the number of week.

4.5 Assessing model fit

Using Z-scores to assess the model fit,

$Z = \dfrac{observedatocopherolvalue - mean_{week}}{SD_{week}}$, where $mean_{week}$ and SD_{week} are the mean and SD given by the model for each week.

(1) Plot a scatterplot of Z-scores[4] against weeks to check for the existence of any patterns.

The Z-scores should be scattered around zero per all the weeks which are observed. Any deviation from it indicates that the mean curve may require modification. It was shown in Figure 10.

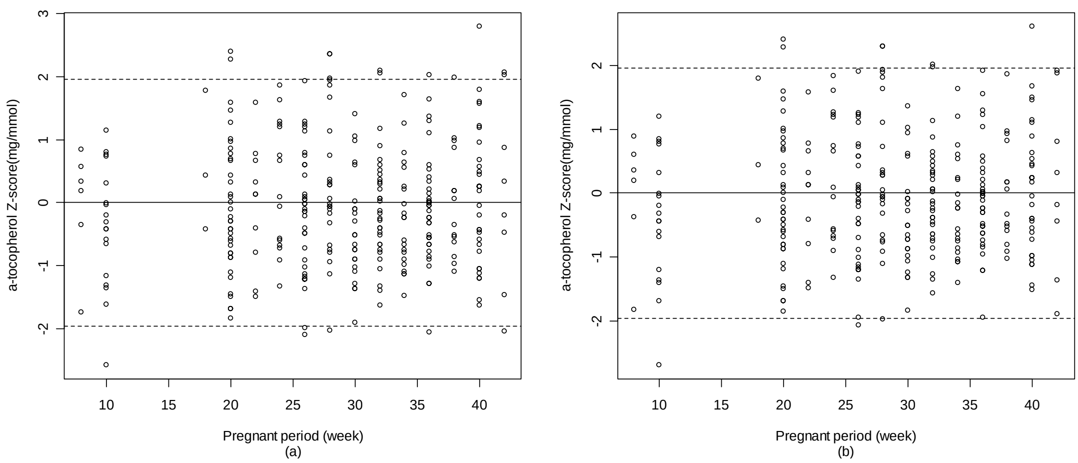


Figure 10 Plot of calculated Z-scores against pregnancy period. In (a) the Z-scores are calculated the under the assumption that absolute residuals show no trend with week. In (b) the assumption is that absolute residuals show trend with week.

[4] The Z-scores do not contain the values from the 12th week.

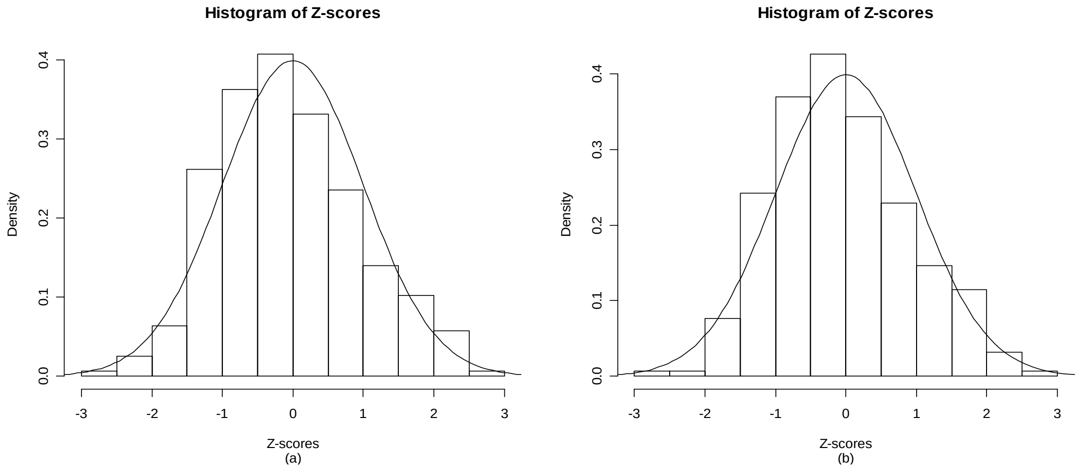


Figure 11 Histogram of calculated Z-scores with overlaid standard normal distribution. In (a) the Z-scores are calculated the under the assumption that absolute residuals show no trend with week. In (b) the assumption is that absolute residuals show trend with week.

(2) Shapiro-Wilk normality test

Under the assumption that absolute residuals show no trend with week, according to the Shapiro-Wilk normality test, the test statistic W = 0.987 with p-value = 0.006456. Since p-value =0.006456<0.05, we do not think Z-scores follow a normal distribution. Whereas under the assumption that there is a trend, according to the Shapiro-Wilk normality test, W = 0.9881 with p-value = 0.01100. Since p-value = 0.01100<0.05, we do not think Z-scores follow a normal distribution.

Since 0.01100>0.006456, the models under the assumption of the absolute residuals show trend with week is better. It seems to be a trend in the mean value while the variation seems to be the same over time for the data set.

5. Discussions

The data set used in this essay has only one such week which is not normally distributed. For that, I used logarithmic transformation. But I do not think that the transformed data can be used when I fit the goodness of the mean curve model. But maybe it can be used after some retransformation or other method.

The Z-scores are not normally distributed. It is necessary to know that how well the curves are fitted and how to modify the curves if they are not good-fit.

References

[1] Armitage, P., [no date] *ENCYCLOPEDIA OF BIOSTATISTICS. John Wiley& Sons* Volume 4 3053-3075

[2] Palm, M., Axelsson, O., Wernroth, L. and Basu, S., [no date] Oxidative stress is associated with normal human pregnancy: A hallmark of free radical mediated lipid peroxidation in physiology. Uppsala University.

[3] Siliverwood, R.J. and Cole, T.J.,[ca. 2007]. Statistical Opinion Statistical methods for constructing gestational age-related reference intervals and centile charts for fetal size. *Ultrasound Obstet Gynecol 2007;* 29:6-13

[4] Wright, E.M. and Royston, P., (1997) A comparison of Statistical Methods for Age-related Reference Intervals. *J.R.Statist. Soc.A*(1997) 160,Part1, pp.47-69

[5] Waltz, M., 2003. *Webster's New World™ Medical Dictionary* [on lline]. (*2nd Edition).* Wiley Publishing, Inc.. Available from:
http://www.medterms.com/script/main/art.asp?articlekey=4582
[Accessed 30 April 2007]

[6]*What Is Normal Blood Pressure?* [on line] Available from:
http://www.diovan.com/info/answers/normal_blood_pressure.jsp?usertrack.filter_applied=true&NovaId=3350119491241679437 [Accessed 30 April 2007]

[7] Wikipedia, the free encyclopedia, *Tocopherol*[on line].Available from:
http://en.wikipedia.org/wiki/Tocopherol [Accessed 1 May 2007]

[8] R help for package stats, R version 2.4.0.