

Automated Analysis of Linear-Plateau Models

Damon Disch, Pitman-Moore, Inc.

Introduction:

In 1975, Anderson and Nelson introduced a series of regression models to relate the amount of fertilizer used and the yield response observed. These models are called linear-plateau models and are special cases of the piecewise linear regressions found in some applied statistics texts (for example, Neter and Wasserman, 1974, pp. 313-316). Of particular importance to Pitman-Moore, linear-plateau models have been used by FDA statisticians to relate the dose level of a growth promoting drug to the response observed in such parameters as average daily gain (ADG) and feed per gain (FPG) (BVM, 1975, pp. 9-10). This report concentrates on the four simplest versions of these models, models II, III, IV, and V. The models are described and methods of statistical analysis are discussed. This report continues with a description of an automated analysis of these models using VMS SAS[®] Release 5.18 code to implement the methods discussed. The data are assumed to come from one or more trials possibly with blocking and other factors present. The number of replications may vary from trial to trial, but each dose should have the same amount of data within any single trial. Anderson and Nelson (1975) concentrate mainly on the case of uniformly spaced dose levels although the 1987 reference has some rules for non-uniform spacing. The code documented in this report does not require uniform spacing or integer dose levels. The description of the models follows.

Models:

Each of the linear-plateau models can be thought of as a regression through the average responses at various dose levels. A model II is a single straight line model often known as simple linear regression. The equation for this model is

$$\text{mean response} = B_0 + B_1 * \text{dose}$$

A model III consists of a sloping line from the zero dose to one of the observational doses with a horizontal line continuing on. The horizontal part is called the plateau. If there are 4 non-zero doses, the break point at which the sloping line intersects the plateau could be at any one of the first three non-zero doses. The model III-1 would be a model III with a break point at the first non-zero observation dose, III-2 would be a model III with a break point at the second non-zero dose, etc. See Figure 1 for an example of a III-1 model. The equation for a model III-n where D_n is the nth non-zero dose is

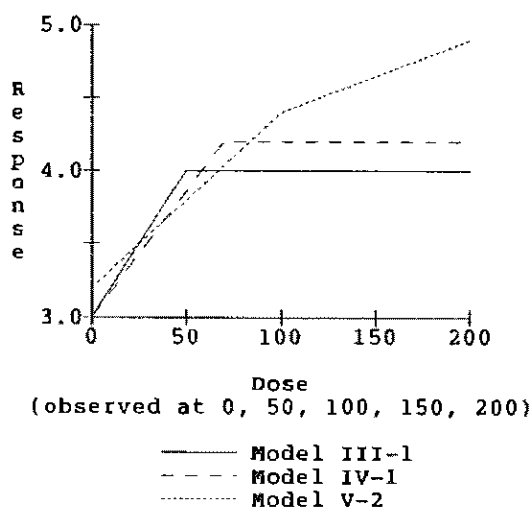
$$\text{mean response} = \begin{cases} B_0 + B_1 * \text{dose} & \text{for dose} \leq D_n \\ B_0 + B_1 * D_n & \text{for dose} > D_n \end{cases}$$

B_0 is the y-intercept of the sloping line, and B_1 is the slope of the sloping line. The height of the plateau is $B_0 + B_1 * D_n$. For there to be

a useful model, the estimate of B_1 must differ significantly from zero. Generally, this test should be one-sided as B_1 should be greater than zero for ADG and less than zero for FPG.

Figure 1

Examples of Linear-Plateau Models



A model IV consists of a sloping line from the zero dose to a point between two observational doses followed by a horizontal line continuing on. If the sloping line intersects the horizontal plateau between non-zero doses 2 and 3, the model is designated a IV-2 model (see Figure 1). In general, a IV-n model has the break point between the nth and (n+1)st non-zero doses. The equation of a IV-n model where D_n is the nth non-zero dose is

$$\text{mean response} = \begin{cases} B_0 + B_1 * \text{dose} & \text{for dose} \leq H \\ B_0 + B_1 * H & \text{for dose} > H \end{cases}$$

where $H = D_n + (D_{n+1} - D_n) * B_2 / B_1$ is the break point. B_0 is the y-intercept of the sloping line, B_1 is the slope of the sloping line, and the height of the plateau is $B_0 + B_1 * D_n + (D_{n+1} - D_n) * B_2$. Note that the height of regression at D_n is $(B_0 + B_1 * D_n)$ so that B_2 is proportional to the added response in going from D_n to the break point dose, H . For ADG a valid model has $B_1 > B_2 > 0$, and for FPG a valid model has $B_1 < B_2 < 0$.

A model V consists of a sloping line from the zero dose to an observational dose followed by a second sloping line. The second line has a shallower slope than that of the first line, although both slopes have the same sign. Essentially, a model V is a model III with a sloping plateau. A model V-2 has a break point at the second non-zero observational dose (see Figure 1). In general, a V-n model has the break point

at the n th non-zero observational dose. The equation of a V-n model where D_n is the n th non-zero dose is

$$\text{mean response} = \begin{cases} B_0 + B_1 * \text{dose} & \text{for dose} \leq D_n \\ B_0 + B_1 * D_n + B_2 * (\text{dose} - D_n) & \text{for dose} > D_n \end{cases}$$

B_0 is the y-intercept of the first sloping line, B_1 is the first slope, and B_2 the second slope. For ADG a valid model has $B_1 > B_2 > 0$, and for FPG a valid model has $B_1 < B_2 < 0$.

Statistical Analysis:

The fitting of linear-plateau models can be based on the use of contrasts. The idea of a contrast links the statistical methods of analysis of variance (ANOVA) and regression together. The major difference between the methods is in how the dose level is treated. The scientific assumption is that there is a functional relation between response (e.g. ADG) and dose level. Regression takes this relation into account directly, while ANOVA does so indirectly through calculation and testing of contrasts.

In ANOVA the dose levels are considered as separate categories without regard to the relation among the categories. Other factors in the ANOVA are regarded as nuisance variables which are brought into the analysis in order to reduce the unexplained variation and so increase the power of statistical tests comparing dose effects. Often block or trial are such nuisance variables. For example, a common model used in an ANOVA is:

$$\text{response}_{i,j,k} = \mu + \text{trial}_i + \text{block}_j(\text{trial}_i) + \text{dose}_k + \text{trial}_i * \text{dose}_k + \text{error}_{i,j,k}$$

The trial and block effects can be averaged over to get a relation of the form:

$$\text{response}_k = \mu + \text{dose}_k + \text{error}_k$$

This equation ignores the functional relation between dose and response and treats each dose level as a separate category. The ANOVA style of analysis yields two important results, however. First, means adjusted for the nuisance variables are calculated, the so called least square means or lsmeans. Second, an error estimate is calculated free of any particular regression model assumption, the pure error. This error can be used in statistical tests of validity of models and tests of whether a parameter in a model is nonzero. Such tests will have no bias caused by a misspecified or invalid model.

Regression uses the relation between dose and response by a model of the form:

$$\text{mean response} = f(\text{dose})$$

where f is the functional form of the model. For a model II, $f(\text{dose}) = B_0 + B_1 * \text{dose}$. The functional forms for the various linear-plateau models have already been given. A common feature of the different forms of f is that they all involve some parameters B_0, B_1, B_2 . While theoretically possible to include nuisance variables in a regression calculation, the set

up of the computations is difficult. It is easier in estimation of parameters to calculate the least square means to account for the nuisance variables and then regress through the means. The least squares method used by proc REG (SAS Institute, 1985) to estimate these parameters reduces mathematically to an estimate which is a linear combination of the adjusted dose means, an estimate of the form:

$$\text{estimate} = A_0 * \text{dose}_0 + A_1 * \text{dose}_1 + \dots + A_n * \text{dose}_n$$

where dose_0 is the lsmean at dose 0 and there are n non-zero doses with lsmeans $\text{dose}_k, k=1, 2, \dots, n$. The constants A_0, A_1, \dots, A_n can be calculated in advance of any data collection and sum to zero for estimates of B_1 and B_2 in the linear-plateau models, so these estimates are contrasts. The estimate of B_0 is not a contrast in the linear-plateau models. For example, in a III-1 model, the constants for estimating B_0 are 1 for A_0 and 0 for all other coefficients. The testing or estimation of a contrast can be carried out by the CONTRAST and ESTIMATE statements in proc GLM (SAS Institute, 1985). Hence, use of proc REG to regress through the lsmeans and use of the ESTIMATE statement in proc GLM can yield the same estimates of B_1 and B_2 . The estimate of B_0 can be gotten from proc REG or by calculating the linear combination by hand once the lsmeans have been calculated using proc GLM.

The standard errors produced by procs GLM and REG will differ since the error in proc REG will be essentially a model error combined with a random error affecting the dose means. If all trials have the same number of blocks, multiplying the standard error from proc REG by the proc REG root mean square and then dividing by the product of the root mean error from proc GLM (the "pure" error) times the root of the number of replications per dose will make the REG standard error equal the GLM standard error. If the trials have different numbers of blocks, the lsmeans will not be independent as assumed by the proc REG analysis. Generally the correlation will be very small and the standard error adjusted by the pure error will be a satisfactory approximation to the true standard error as calculated by the CONTRAST or ESTIMATE statement of GLM. Note that the correct pure error is either the residual error if the dose by location interaction is not significant at the .25 level or the dose by location interaction if it is significant at the .25 level, according to current FDA policy.

Assuming that the number of replicates per dose is p , the error term derived from regressing through the lsmeans should be multiplied by p to express the error as fluctuation of raw data. Once this adjustment is made, if there is no model error, the adjusted error from proc REG estimates only random error. Hence, the ratio of the adjusted error from proc REG, sometimes called the mean square for lack of fit, divided by the error from proc GLM provides an F statistic to test lack of fit (see Anderson and Nelson, 1987, p. 28). A significantly high F indicates lack of fit, and a significantly low F indicates overfitting.

Model selection may be based on several criteria. One traditional method is to pick the model with the highest R^2 . A better method is to pick the model with the lowest mean square for error from the regression or, equivalently, the with the highest adjusted R^2 (SAS Institute, 1985, p. 690). This second method adjusts for the automatic rise in R^2 whenever any variable is added to a least squares regression, an important consideration since there are usually a limited number of dose levels tested and a model IV or V could be "better" than a model III simply by having an extra parameter. Since the mean square for lack of fit is proportional to the mean square for error from the regression, the second criterion is equivalent to selecting the model with the lowest mean square for lack of fit.

The use of the minimum MSE criterion is subjective in that small differences in MSE may be statistically meaningless. Hence, more objective statistical criteria should be used to aid in model selection as well. An acceptable model should not show significant lack of fit by the test for lack of fit described above. Generally for a model to be acceptable, all parameter estimates should be statistically significant. The tests of the parameters should be one-sided since there is an *a priori* assumption on the sign of the parameters. Finally, models may be compared by a statistical test in certain special cases. For example, a model IV-1 is better than a III-2 if the parameter B_2 in the IV-1 model is significantly less than B_1 in magnitude. This comparison can be made easily by testing $(B_1 - B_2)$. If the IV-1 model has a significant B_2 parameter, then this model is preferred to the III-1 model. A significant B_2 parameter in a model V-n shows this model to be preferred to the III-n model. A significant $(B_1 - B_2)$ in a model V shows evidence that the model V is superior to a model II. In general, such statistical tests should replace reliance on model ranking by the R^2 value.

Implementing the proc REG analysis:

The proc REG analysis described above is implemented in SAS through code called LPMODELS.SAS available from the author. Including this code in a driver program provides for the analysis of Models II, III, IV, and V all in one pass.

The driver program must define a macro variable and provide two data sets before including LPMODELS.SAS. The macro variable is BREAKPT and gives the non-zero dose number of the model. If there are 3 non-zero doses, the only valid values of BREAKPT are 1 and 2. In general, if there are n non-zero doses, BREAKPT must be an integer from 1 to (n-1) inclusive.

The first data set to be supplied by the driver program must be called "MEANS" and must contain several variables. The variable DOSE contains dose levels, and MEANS must be sorted by this variable. The variable FREQ contains the number of replications per dose level, and this value should be the same for each dose level. Finally, MEANS must have the dependent variables (e.g. ADG and FPG) which contain the least square means through which to do the regression. In a completely balanced situation

where each trial has the same number of blocks, the MEANS data set can be generated by proc SUMMARY.

The second data set supplied by the driver program must be called "PURE ERR" and have the variables VAR, DF, and MSE. Variable VAR is a character variable holding the name of each dependent variable. Only dependent variables in the MEANS data set are allowed, and PURE ERR must be sorted by VAR. Variable MSE is the mean square for error to test dose effects, and variable DF is the degrees of freedom for MSE. MSE is the pure error and is usually either the residual error from a GLM run or else the MS for dose by location interaction.

Once the driver program has set up the necessary preliminaries, LPMODELS.SAS must be included using a XINCLUDE statement. To illustrate the use of LPMODELS.SAS, a data set concerning the effect of lysocellin on ADG (Darden et al., 1988) is analyzed. Part of the output is in the Appendix. The whole output has four parts. First there is a listing of the data set MODELS generated by the included code along with a reminder in the title of the number of replications per dose on which the analysis is based. The MODELS dataset shows the vectors of the design matrix, X, generated by a data set to do the regression. These are similar to the example vectors in the Anderson and Nelson papers.

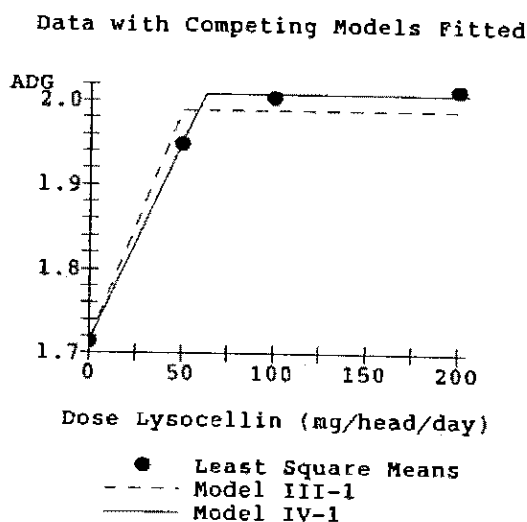
The second section is the regression output from fitting models II, III, IV, and V using a single call for proc REG. The break point for models III, IV, and V was specified in the driver program. This part of the output can be suppressed by setting the macro variable PRINTREG to the value "NOPRINT" (omitting the quote marks!). The OUTEST= and COVOUT options are used on the proc REG step.

A data step is used to break up the output data set from proc REG into a data set containing the MSE terms from each model and a data set containing the estimates and standard errors. After sorting these data sets by VAR, a variable containing the names of the dependent variables, a data set named FIT is created. The third section of the output is a listing of the data set FIT done by each dependent variable. The mean squares for pure error and lack of fit are presented along with their degrees of freedom and the F ratio for lack of fit. If the p-value is too low (say less than .05 or .1), then there is evidence of lack of fit. The minimum dose for maximal effect is the highest dose for the models II and V, the break point dose for model III, and the dose level for the start of the plateau for model IV.

A data step is used to create the data set ESTIMATE which contains parameter estimates, standard errors of the estimates, and t-tests by each dependent variable. A call to proc PRINT lists these values out. The pure error used and the degrees of freedom for the t-tests is displayed next to the dependent variable name. Once the included code is finished, all working data sets generated by the code are deleted except MODELS, FIT, and ESTIMATE. These are retained as temporary data sets so that the driver program may further use them in subsequent operations.

The selection of the best model can be illustrated using the analysis in the Appendix. The lack of fit test eliminates the model II since the p-value for lack of fit is .0106. By minimum mean square for lack of fit, the ordering of the remaining models is IV-1, V-1, and III-1 with the mean squares .00037, .00574, and .00971, respectively. The B_2 parameter has a one-sided p-value of .1358 for model IV-1 and .1746 for model V-1. Hence, they both might be rejected in favor of the simpler III-1 model. However, a look at Figure 2 shows that the model IV-1 provides a better fit than the III-1 model, and the adjusted R^2 value is .9972 for the IV-1 model as opposed to .9383 for the III-1 model. The test of $(B_1 - B_2)$ in model IV-1 is barely significant at the .05 level. Rather than draw a final conclusion on the model and so on the correct breakpoint, more experimentation might be in order.

Figure 2



Implementing an Exact Analysis:

Exact tests of the model parameters can be carried out using contrasts. SAS code, available from the author, called LPCONTRAST.SAS is designed to be included in a driver program to carry out the computation of the coefficients for the contrasts. This code relies on SAS/IML^(r) and leaves proc IML active after the code has been executed by the driver program. Thus, the user may use the matrices generated for further computations or for saving as data sets. The coefficients can be used in a call to proc GLM in ESTIMATE and CONTRAST statements.

Much as LPMODELS.SAS required certain inputs from the driver program, LPCONTRAST.SAS needs the macro variable BREAKPT set and a data set generated. The value of BREAKPT is the non-zero dose number of the break point for the models to be estimated. The data set must be called DOSES and must have the variable DOSE which contains the dose levels in the experiment in increasing order. The data set MEANS used by LPMODELS.SAS may be renamed "DOSES" and used for this requirement.

Once the driver program has set up the necessary preliminaries, the LPCONTRAST.SAS code can be included using a ZINCLUDE statement. An example of part of the output is in the appendix. If all that is desired is a test of a parameter, the coefficients listed next to the parameter name may be used in a CONTRAST statement in proc GLM. If an estimate is to be calculated, the ESTIMATE statement in proc GLM should be used along with the DIVISOR option with the divisor listed next to the coefficients.

For the straight line Model (II), a transformation of the doses is presented which makes the estimate of the slope for the transformed doses statistically independent of the estimate of the intercept. For this model the coefficients listed under "X1" may be used to estimate the transformed slope. No DIVISOR option is required in the ESTIMATE statement. The transformed intercept is the average of the lsmeans. Using the original dose levels, the coefficients listed next to "B1" can be used in a CONTRAST or ESTIMATE statement to test the non-transformed slope. For the original doses the estimate of B_0 is the average of the lsmeans minus the product of the average of the dose levels times the estimate of B_1 . (Note that the transformed dose uses the average of the dose levels in the numerator of the transformation.)

For the other models, the vectors of the design matrix are presented, and the coefficients for using CONTRAST statements are listed. To get appropriate estimates, the divisor is listed for the DIVISOR option in the ESTIMATE statement. Models IV and V have the added feature of coefficients for testing the difference between B_1 and B_2 . For a model IV-n, if B_2 is not significantly less than B_1 , then the model does not differ significantly from a simpler III-(n+1) model. For a model V-n, if B_2 is not significantly less than B_1 , then the model could be replaced by a simpler Model II.

Since the original data consisted of two trials both with four blocks, the proc REG analysis and the proc GLM analysis are equivalent. For example, in the model III-1, the estimate of B_1 is .00549309 and the standard error of the estimate is .0009933 using either method. As the difference in the number of blocks per trial increases, the proc REG and the proc GLM statistical tests will differ more, although some experience has shown that the proc REG analysis is an excellent approximation to the exact analysis. In any case, the estimates will be the same.

SAS Coding Techniques Employed:

The code called LPHODELS.SAS uses data steps extensively to generate the independent variables for the proc REG step, to break apart the output data set from proc REG into an error data set and into an estimate data set, and finally to combine the error terms properly. Macro variables are used to transfer information between data steps. Proc PRINT is used to output results calculated in the data steps.

The code called LPCONTRAST.SAS uses standard proc IML techniques. The only unusual technique is to zero out values in the contrasts below a

"fuzz" value of 1.E-12. An index vector pointing to these low values is generated and then used to point to the elements in the original matrix to be zeroed out. The use of the J function at this point is required because of a bug in the IML software on the VAX^(tm).

Both programs are designed for easy inclusion in a SAS/AF^(r) application with only minor modification. Also, both programs can be adapted for use with PC DOS SAS^(r) Release 6.03, the chief problem being that the output data set from proc REG uses different names than proc REG in VMS SAS^(r) uses.

Trademark Notices:

SAS^(r), SAS/AF^(r), and SAS/IML^(r) are registered trademarks of SAS Institute, Inc., Cary, NC, USA.

VAX^(tm) is a trademark of Digital Equipment Corp., Maynard, MA, USA.

References:

Anderson, R. L. and Nelson, Larry A. (1975). A family of models involving intersecting straight lines and concomitant experimental designs useful in evaluating response to fertilizer nutrients. *Biometrics*, 31: 303-318.

Anderson, R. L. and Nelson, Larry A. (1987). Linear-Plateau and Plateau-Linear-Plateau Models Useful in Evaluating Nutrient Responses. Technical Bulletin 283, North Carolina Research Service, North Carolina State University, Raleigh, NC 27695.

BVM (1975). Preclearance Guidelines for Production Drugs, 3rd Edition, October, 1975. Prepared by the Bureau of Veterinary Medicine of the Food and Drug Administration.

Darden, D. E., W. E. Kunkle and D. B. Bates (1988). Effect of lysocellin on performance of cattle grazing pasture. (Submitted for publication.)

Neter, J. and Wasserman, W. (1974). Applied Linear Statistical Models. Richard D. Irwin, Inc., Homewood, IL.

SAS Institute, Inc. (1985). SAS User's Guide: Statistics (Version 5 Edition). SAS Institute, Inc., Cary, NC.

Author's Address:

Pitman-Moore, Inc., P.O. Box 207, Terre Haute, IN 47808

Appendix -- Sample Output from Proc REG Analysis
Two Pasture Trials -- Data from D. Darden
Test for Lack of Fit

----- Dependent Variable=ADG DF Pure Error=18 MS Pure Error=0.01480108 -----

MODEL	Regression MSE	R-Square	Adjusted R-Square	DF Lack of Fit
II	0.0109382	0.6292	0.4439	2
III_1	0.0012135	0.9589	0.9383	2
IV_1	0.0000465	0.9992	0.9976	1
V_1	0.0007174	0.9878	0.9635	1

MODEL	MS Lack of Fit	F Lack of Fit	P value Lack of Fit	Minimum Dose for Maximum Effect
II	0.0875055	5.91210	0.0106	200.000
III_1	0.0097081	0.65590	0.5309	50.000
IV_1	0.0003722	0.02515	0.8758	62.724
V_1	0.0057392	0.38776	0.5413	200.000

Estimates and Tests of Model Parameters
P-Values are for One-Sided Tests

----- VAR=ADG DF=18 MSE=0.01480108 -----

MODEL	Parameter	Estimate	Standard Error	t	P-Value
II	B0	1.80587	0.0333179	54.2012	0.0000
II	B1	0.00130278	0.0002908	4.4796	0.0001
III_1	B0	1.713872	0.0430132	39.8453	0.0000
III_1	B1	0.005493094	0.0009933	5.5299	0.0000
IV_1	B0	1.713872	0.0430132	39.8453	0.0000
IV_1	B1	0.004696352	0.0012166	3.8602	0.0006
IV_1	B2	0.001195112	0.0010536	1.1343	0.1358
IV_1	B1-B2	0.00350124	0.0020175	1.7354	0.0499
V_1	B0	1.713872	0.0430132	39.8453	0.0000
V_1	B1	0.00498269	0.0011263	4.4237	0.0002
V_1	B2	0.000382803	0.0003982	0.9613	0.1746
V_1	B1-B2	0.004599887	0.0013602	3.3818	0.0017

Sample Output Showing the Computation of Contrast Coefficients

Two Pasture Trials -- Data from D. Darden

Contrasts for Estimating and Testing
Linear-Plateau Model Parameters

Model II -- Straight Line Model

Transformation: $x = (\text{dose} - 87.5000000) / 147.90199458$

	Dose	X0	X1		
	0	1.000000	-0.591608		
	50.000000	1.000000	-0.253546		
	100.000000	1.000000	0.0845154		
	200.000000	1.000000	0.760639		
	0	50	100	200	Divisor
B0	1.000000	1.000000	1.000000	1.000000	4.000000
B1	-7	-3	1.000000	9.000000	1750.000

Model III_1

	Dose	X0	X1		
	0	1.000000	0		
	50.000000	1.000000	50.000000		
	100.000000	1.000000	50.000000		
	200.000000	1.000000	50.000000		
	0	50	100	200	Divisor
B0	1.000000	0	0	0	1.000000
B1	-3	1.000000	1.000000	1.000000	150.000000

Model IV_1

	Dose	X0	X1	X2		
	0	1.000000	0	0		
	50.000000	1.000000	50.000000	0		
	100.000000	1.000000	50.000000	50.000000		
	200.000000	1.000000	50.000000	50.000000		
	0	50	100	200		Divisor
B0	1.000000	0	0	0	1.000000	
B1	-1	1.000000	0	0	50.000000	
B2	0	-2	1.000000	1.000000	100.000000	
B1 - B2	-2	4.000000	-1	-1	100.000000	

Model V_1

	Dose	X0	X1	X2		
	0	1.000000	0	0		
	50.000000	1.000000	50.000000	0		
	100.000000	1.000000	50.000000	50.000000		
	200.000000	1.000000	50.000000	150.000000		
	0	50	100	200		Divisor
B0	1.000000	0	0	0	1.000000	
B1	-7	5.000000	3.000000	-1	350.000000	
B2	0	-4	-1	5.000000	700.000000	
B1 - B2	-2	2.000000	1.000000	-1	100.000000	