ABSTRACT

Periodic phenomena are studied by scientists in many fields. Analysis of complex cycles often involves a representation by Fourier series and tests of significance on each term by means of an analysis of variance. The analysis of variance table provides a useful summary of the regression results and helps the data analyst select an appropriate model by determining how many terms should be included and whether a given model is a good fit to the data. The SPECTRA procedure computes Fourier coefficients but does not provide an ANOVA table. This paper describes a set of SAS macros for ANOVA on results from the SPECTRA procedure and demonstrates its use.

INTRODUCTION

It is quite common in many areas of research to measure a response over a period of time. When the level of response is varied continually, it is said to be cyclic or periodic. More precisely, a function \( f(t) \) is said to be periodic with period \( H \) if, for all \( t \):

\[
f(t) = f(H + t) .
\]  

(1)

The study of periodic phenomena has played an important role in fields as diverse as physics, astronomy, economics, and biology for well over a century. Two different approaches to the analysis of periodic series are available, the frequency domain and the time domain. Time domain methods will not be discussed here. Frequency domain methods derive from the remarkable theorem of Fourier (1807) which states that any function which is continuous over an interval or which has a limited number of finite discontinuities may be expressed in terms of sine and cosine functions as:

\[
f(t) = a_0/2 + \sum a_n \cos n \omega t + b_n \sin n \omega t.
\]  

(3)

The quantity \( \omega = 2\pi f/ H \), represents the angular frequency, which is the number of times the function repeats itself in \( H \) units of time. The quantity \( n \) is the fundamental frequency or the smallest value of \( H \) for which equation (1) holds. The terms \( \cos n \omega t \) and \( \sin n \omega t \) are called the \( k \)-th harmonics.

Harmonic analysis is the application of Fourier’s theorem to the problem of approximating a periodic function by fitting a finite number of terms. The parameters and called the "Fourier coefficients" can be estimated by least squares as:

\[
a_i = (2/n) \int f(t) \cos \omega t dt
\]  

\[
a_n = (2i/n) \int f(t) \sin \omega t dt.
\]  

(2y)

Under certain conditions the SAS procedure SPECTRA can be used to estimate these parameters. The fundamental period must be known. Methods for determining the period are available (Blumberg, 1976; Gottman, 1977) but will not be discussed here. The time intervals must be equally spaced. When this is true, the method used by SPECTRA gives parameter estimates that are orthogonal. It is also necessary that the data form a stationary time series. A time series is stationary if two conditions are met:

(i) \( B(f(t)) \) is constant

(ii) \( \text{Cov}(f(t), f(t+D)) \) depends only on \( D \).

The first condition requires that the mean of a segment of the time series not depend on where in the series that segment comes from. The second condition specifies that the covariance between two random variables does not depend on what part of the time series they come from, but only on the distance between them. If these conditions are not met, the series can be transformed or detrended to make it stationary. For example, a linear trend might be removed by performing a linear regression of the response variable on time. A harmonic analysis could then be done on the residuals of that regression.

Harmonic analysis is really a special case of regression analysis. A statistical approach to regression analysis entails construction of tests to answer some important questions. Those questions usually deal with whether or not the regression coefficients are significantly different from zero and whether a particular model is a good fit to the data. An analysis of variance is normally the vehicle for determination of the answers. When harmonic analysis is approached in this way, it is called periodic regression (Bliss, 1970). When the time series is not based on equally spaced intervals, an alternative to harmonic analysis is possible (Sollberger, 1970). The alternative method also allows for terms that are not harmonics, i.e., not simple fractions of the fundamental period, but
this method will not be computationally as efficient as SPECTBa.

COMPUTING THE ANALYSIS OF VARIANCE TABLE

Periodic regressions can be computed in SAS by using the results from PROC SPECTBA and using several other SAS procedures to complete the analysis of variance table. If the data cover only one fundamental period, they are sent directly to SPECTBA. If the fundamental period is replicated, corresponding points from different cycles are averaged in SPECTBA. The total sum of squares in the response variable is divided into parts due to the individual Fourier terms and error. Each Fourier term is the effect due to the kth harmonics. The sums of squares for the Fourier terms are easily obtained from the periodograms provided by SPECTBA. The periodogram I_ = Cn/2(a_n2 - b_n2), is distributed as chi-square with two degrees of freedom. As the error term is also a chi-square variable, an F-ratio can be used to test the significance of a Fourier term (Hartley 1949)* Because these terms are orthogonal, the error sum of squares can be obtained by subtraction.

In replicated series, the error can be partitioned into sums of squares due to scatter about the curve and pure error. The appropriateness of the model can be assessed by the significance of the scatter term. If the scatter term is not significant, the pure error mean square should be used as the denominator for F tests. Otherwise, the scatter mean square should be used.

The macro ANOVA computes all the necessary statistics. A further set of smaller macros actually produce analysis of variance tables. These macros are to be called to analyse a data set whose time variable t is named XIMB and whose response variable is named DEPEND. To produce a table, specify TABLB n TERM, where n is the number of terms, spelled out.

The following example is given in Bliss, 1958 and Bliss, 1970. The response variable is median effective dose of chorionic gonadotropin causing release of sperm in toads. The data were collected monthly over a two year period. The fundamental period is assumed to be one year. Tables are generated for one Fourier term and for two Fourier terms. The results indicate that a single Fourier term fits the data well.

DATA FIRST:
INPUT DEPEND TIE RK:
TITLE BLISS (1970) P.276, EXExs. 17.2:
LIST:
CARD=:
1.272 1 1
1.258 1 2
1.358 2 1
1.325 2 2
1.494 3 1
1.471 3 2
1.482 4 1
1.420 4 2
1.464 5 1
1.552 5 2
1.502 6 1
1.471 6 2
1.536 7 1
1.520 7 2
1.420 8 1
1.437 8 2
1.405 9 1
1.438 9 2
1.53 10 1
1.388 10 2
1.304 11 1
1.356 11 2
1.322 12 1
1.339 12 2
ANOVA
TABLE ONE TESH
TABLE TWO TBSH

SGFEKBX3


Hartley, H.O. (1949)** Tests of significance In harmonic analysis* Bioaestrllta. 36; 194-201.


Simon, William (1977)* Mathematical ed* Cambridge, Mass*: MIT Press*

APPENDIX 1

SAS MACSOS TO PSODDCE ANOVA TABLE

: MACKO ANOVA
PSOC SORT; & Y XIMB;
* OUTPUT MEANS FOR EACH TIME;
PSOC MEANS N MEAN SUM VAK DSS CSS
NOPRINT DATA=FIJILST :
VAK DEPEND;
By TIME:
OUTPUT OUT=NEW MEAN=CSS MEAN=N:
* SPECTRAL ANAL MEANS;
PSOC SPCTBAL DATA=NOW COEF F S WEIETEST
OnT=TPEC;
VAK H;
FBOC PRINT;
* PUT PAMII ESTIMATES IN NEW DATA SET;
DATA PARA;
SET SPEC;
KEEP COS_01 SIN_01;
IF _N_=1 THEN CONSTANT= COS_01/2;
NnM="N_-l;
FIXE PSINT;
TITLE2 :
TITI.E3 PAAHERGK ESTIMATES;
TITL.E4;
TITLES MODEL IS: Y = CONSTANT +
SnH(A1*COB(WT) — BI*SIN<WT));
IF _N_=1 THEN PUT 10 'CONSTANT =• a 20
CONSTANT /;
ELSE PUT a 10 'A* NUH * =• 9 20 COS_01
a> 35 "B' NUM • => a 45 sin_Ot;
PKOC MBANS DATA=FISST CSS NOPSINT;
VAl DEPEND;
OUTPUT Out=TOTAL CSS=TOTALSS;
PKOC MEANS DATA=NEW SUM MEAN NOPRINT;
VASCSS N:
OUTPUT OnT=EflIOB SnM=efilOESS TN
MEAN=1 SEP;
DATA SS;
SET SPEC;
IF _N_= 1 THEN DELETE;
KEEP P_01;
* PSOC Transpose OUT=TSANS;
DATA MEfGE:
MESGE TOTAL EflIOB TSANS;
TOTALDF=TH-1;
BSSOBSF=TN*BEF *( KEEP >
KKEOJIM=EREOISS/EREOSSJ
DKOF TN _NAME_ LABEL_;
ARASOY COL=COLS;
ASSAY SS SSI-SSS:
AERAY MS MSl-MS9;
AFLAY F F1-F5:
DO OVBK COL;
DO OVBK SS;
DO OVBK MS;
DO OVER F;
SS=SEP*COL; MS=SS/2;
F=US/EROSMS;
END; END; END: END;
MACKO TABLE
DATA BESCLTS;
SET MACSC;
FIXE PRINT;
TITLE2 :
TITLES ANALYSIS OF VASIAMCE OF FODSIEE
TEMS;
APPENDIX 2

EESU.TS FJOU EXAMPLE B

BLISS C1970> P.27b, EXEK. 17.2

PARAMETERS ESTIMATES
HODEL. IS: $Y = \text{CONSTANT} + \text{SULI} (\text{AI} \times \cos (\text{WT}) + \text{BI} \times \sin (\text{VT}))$

<table>
<thead>
<tr>
<th>CONSTANT</th>
<th>A1</th>
<th>A2</th>
<th>A3</th>
<th>A4</th>
<th>A5</th>
<th>A6</th>
</tr>
</thead>
<tbody>
<tr>
<td>= 1.40825</td>
<td>-0.089948</td>
<td>-0.00895833</td>
<td>-0.014250</td>
<td>-0.004625</td>
<td>-0.018801</td>
<td>0.01866667</td>
</tr>
<tr>
<td>B1</td>
<td>0.04459588</td>
<td>B2</td>
<td>-0.00368061</td>
<td>B3</td>
<td>0.0008833333</td>
<td>B4</td>
</tr>
</tbody>
</table>

BLISS (1970> P.276, EXEK. 17.2

ANALYSIS OF VARIANCE OF PODSIES TBKIIIS

<table>
<thead>
<tr>
<th>TBKII</th>
<th>DF</th>
<th>SUM SQ</th>
<th>MEAN SQ</th>
<th>F VALUE</th>
<th>PK &gt; F</th>
</tr>
</thead>
<tbody>
<tr>
<td>A1 + B1</td>
<td>3</td>
<td>0.1209556</td>
<td>0.06047778</td>
<td>67.56665</td>
<td>0.00000</td>
</tr>
<tr>
<td>SCATTEB</td>
<td>9</td>
<td>0.01331394</td>
<td>0.0001479327</td>
<td>1.652725</td>
<td>0.20554</td>
</tr>
<tr>
<td>EBBOB</td>
<td>12</td>
<td>0.010741</td>
<td>0.0008950833</td>
<td></td>
<td></td>
</tr>
<tr>
<td>TOTAL</td>
<td>23</td>
<td>0.1450105</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

NOTE: WHEN THE SCATTEB F-RATIO IS SIGNIFICANTLY GREATER THAN ONE, USE THE F-BATIOS BELOW TO TEST FOUIEB TEEUS

A1 + B1 | 40.88185 |

BLISS C1970> P.276B EXEB. 17.2

ANALYSIS OF VARIANCE OF FODBIBB TBKMS

<table>
<thead>
<tr>
<th>TEEM</th>
<th>BF</th>
<th>SDM SQ</th>
<th>MEAN SQ</th>
<th>F VALUE</th>
<th>PE &gt; F</th>
</tr>
</thead>
<tbody>
<tr>
<td>A1 + B1</td>
<td>2</td>
<td>0.1209586</td>
<td>0.06047778</td>
<td>67.56665</td>
<td>0.00000</td>
</tr>
<tr>
<td>A2 + B2</td>
<td>2</td>
<td>0.001125584</td>
<td>0.0005627918</td>
<td>0.6287581</td>
<td>0.541494</td>
</tr>
<tr>
<td>SCATTEB</td>
<td>7</td>
<td>0.01213836</td>
<td>0.001741184</td>
<td>1.845287</td>
<td>0.14888</td>
</tr>
<tr>
<td>EBBOB</td>
<td>12</td>
<td>0.010741</td>
<td>0.0008950833</td>
<td></td>
<td></td>
</tr>
<tr>
<td>TOTAL</td>
<td>23</td>
<td>0.1450105</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

NOTE: WHEN THE SCATTEB F-RATIO IS SIGNIFICANTLY GREATER THAN ONE, USE THE F-BATIOS BELOW TO TEST FOUGIEB TEEUS

A1 + B1 | 34.73351 |
A2 + B2 | 0.3232217 |