Volume 16 Number 3
September 1993

ISSN 0167-9473
CSDA 16(3) (1993) 257-378

COMPUTATIONAL STATISTICS & DATA ANALYSIS

Incorporating Statistical Software Newsletter

The official journal of

The International Association for Statistical Computing
A Section of The International Statistical Institute

North-Holland

Statistical Software Newsletter

Using Common Statistical Program Packages for the Friedman Test when there are Ties Among Treatments and a Multiple Comparison Procedure is Needed

Ralf BENDER

Heinrich-Heine-Universität Düsseldorf, Abteilung Ernährung und Stoffwechsel
Moorenstr. 5, D-40225 Düsseldorf, Germany

Abstract: It is investigated which form of the Friedman test statistic is implemented in the statistical program packages BMDP, SAS, SPSS, and STATGRAPHICS. Special attention is focused on the handling of ties and whether a multiple comparison procedure (MCP) is available. As none of the considered software systems offer both, namely adjustment for ties and an MCP, two SAS/IML modules are proposed, which use the formula given by CONOVER (1980) for computing the Friedman test statistic with adjustment for ties and the sequentially rejective test procedure of HOLM (1979) for performing multiple comparisons.

I. Introduction

One of the well known nonparametric methods is the test of FRIEDMAN (1937), used for comparison of k treatments in a randomized block design when no assumption of normality can be made. BERNHARD et al. (1988) examined the availability and reliability of the Friedman test and other elementary nonparametric methods in statistical program packages. However, two important items had been neglected, namely the possibility of adjustment for ties and the availability of a multiple comparison procedure (MCP). As in practice ties frequently occur and investigators often wish to perform pairwise comparisons of the treatments, this paper focuses on these subjects.

Beside the classical Friedman statistic as given in many textbooks - see for example BÜNING, TRENKLER (1978, p. 219) - some other statistics had been proposed. The classical statistic is approximately χ^2 distributed and assumes that there are no ties at all. A statistic with adjustment for ties, which is easy to compute, is given by CONOVER (1980, p. 300). On the same page CONOVER (1980) recommended to use an asymptotically F distributed statistic, because this approximation is superior to the χ^2 approximation.

In this paper it is investigated, which form of the Friedman test statistic is implemented in the PC software systems BMDP (version BMDP/386 Dynamic), SAS (version 6.04), SPSS (version 4.0.1) and STATGRAPHICS (version 5.5) and whether a multiple comparison procedure is available.

II. Comparison of Computer Packages

To find out which form of the Friedman statistic is implemented in the considered software systems, the corresponding manuals of BMDP (1990), SAS (1987), SPSS (1986), and STATGRAPHICS (1992) were consulted. Additionally, the computer procedures were tested using a data set given by CONOVER (1980, p. 301). BERNHARD et al. (1988) already pointed out that nearly all software houses chose approximations instead of exact p values. In Table 1 it is summarized, which approximation is used and whether a multiple comparison procedure is available for the Friedman test.

In Table 1 it is shown that the only system which offers an MCP for the Friedman test is BMDP (for details see III.), but no adjustment for ties is made. The only system which offers a direct computation of the Friedman test with adjustment for ties is STATGRAPHICS, but here no MCP is available. SPSS offers the classical Friedman test without adjustment for ties and without MCP. No warning was printed neither by SPSS nor by BMDP that there are ties among treatments, which are not considered in the computations of p values. SAS has no possibility to call up the Friedman test in a direct way. However, the statistics with adjustment for ties can be computed by using PROC FREQ with the option SCORES=RANK (χ^2 approximation) and by using PROC RANK together with PROC ANOVA (F approximation). However, it is difficult to find the correct statistics in the output and therefore these ways are not convenient for practice, in particular for non-specialists. More-

over, it is misleading to replace nonparametric methods by parametric methods applied to ranks, because the formal relationships between these approaches are limited (HILGERS, 1985).

In summary, none of the considered software systems offers the possibility to compute the Friedman test with adjustment for ties and to perform multiple comparisons between the treatments.

III. Discussion

The classical Friedman test is conservative if there are ties among treatments. In the following the results of the different statistics are compared by using the data set of CONOVER (1980, p. 301) with n=12 blocks and k=4 treatments and a hypothetical data set with n=50, k=4, which was formed arbitrary with the objective to obtain a large amount of ties (45 of 50 blocks contain ties). The results are summarized in Table 2.

In Table 1 it is shown that the only system which offers an MCP for the Friedman test is BMDP, but no adjustment for ties is made. The only system which offers a direct computation of the Friedman test with adjustment for ties is STATGRAPHICS, but here no MCP is available. SPSS offers the classical Friedman test without adjustment for ties and without MCP. No warning was printed neither by SPSS nor by BMDP that there are ties among treatments, which are not considered in the computations of p values. SAS has no possibility to call up the Friedman test in a direct way. However, the statistics with adjustment for ties can be computed by using PROC FREQ with the option SCORES=RANK (χ^2 approximation) and by using PROC RANK together with PROC ANOVA (F approximation). However, it is difficult to find the correct statistics in the output and therefore these ways are not convenient for practice, in particular for non-specialists. Moreover, it is misleading to replace nonparametric methods by parametric methods applied to ranks, because the formal relationships between these approaches are limited (HILGERS, 1985).

In summary, none of the considered software systems offers the possibility to compute the Friedman test with adjustment for ties and to perform multiple comparisons between the treatments.

Table 1: *Survey of implemented Friedman statistics and MCPs*

	without adjustment for ties			with adjustment for ties		
	approximation		MCP	approximation		MCP
	χ^2	F		χ^2	F	
BMDP	**YES**	NO	**YES**	NO	NO	NO
SAS	NO	NO	NO	(YES)[1]	(YES)[2]	NO
SPSS	**YES**	NO	NO	NO	NO	NO
STATGRAPHICS	NO	NO	NO	**YES**	NO	NO

1: To get the χ^2 approximation with adjustment for ties PROC FREQ can be used.

2: To get the F approximation with adjustment for ties PROC RANK and PROC ANOVA can be used.

Table 2: *Results of the different Friedman statistics*

	Conover's data		hypothetical data	
	F	p	F	p
classical Friedman test (BMDP, SPSS)	7.63	0.054	5.35	0.148
with adjustment for ties:				
χ^2 approximation (STATGRAPHICS, SAS: FREQ)	8.10	0.044	14.29	0.003
F approximation (SAS: RANK & ANOVA)	3.19	0.036	5.16	0.002

From Table 2 it is obvious that the results of the classical Friedman test are not satisfactory, because of its conservative character. Hence, software systems like BMDP and SPSS which offer no adjustment for ties are useless for practice if there is more than a negligible amount of ties. Regarding MCPs for the Friedman test the offer of the software systems is also a disappointment. The only package offering an MCP in this context (BMDP) uses a conservative procedure in a double sense. Firstly, no adjustment for ties are made and secondly the MCP is based on the Bonferroni procedure, which is also known to be a method with low power. CONOVER (1980) proposed an MCP for the Friedman test which is based on Fisher's least significant difference (LSD) method.

This method has better power properties but does not control the experimentwise error rate. WRIGHT (1992) pointed out that modified Bonferroni-based procedures, which control the experimentwise error rate and have more power than the Bonferroni approach deserve wider use.

To overcome the deficiencies of the software systems, in this paper the sequentially rejective procedure of HOLM (1979) is applied and computed by means of a SAS/IML module using matrix language (see Appendix). In Table 3 the results of the MCPs computed with BMDP, the procedure of CONOVER (1980) and the procedure of HOLM (1979) are given for the same data as in Table 2. In particular, the differences of the rank sums used for the MCPs of BMDP and CONOVER (1980) and the individual p values computed by means of the Friedman test (F approximation) for each pairwise comparison are given. These pairwise tests formed the basis for the sequentially rejective test procedure of HOLM (1979). Significant results of the three MCPs were indicated by symbols (α=0.10: ★, α=0.05: ★★).

Table 3 shows that the use of BMDP is insufficient for detecting differences between treatments, because the applied statistics have very low power. Even clear differences, shown by significant outcome of the Holm procedure are overlooked, when there are ties among treatments. The results of Holm's procedure detected less significant differences than that of Conover, due to the fact that the experimentwise error rate is under control. Moreover, the result of a single pairwise comparison is not influenced by treatments not considered.

IV. Concluding Remarks

A few years ago HAUX et al. (1988) gave some recommendations on software for nonparametric statistical methods. Focusing the attention on the Friedman test, even simple demands like adjustment for ties are still not met by commonly used statistical computer packages. Another critical point is the availability of convenient multiple comparison procedures. Until the software systems offer the Friedman test with adjustment for ties and a satisfactory multiple comparison procedure, the SAS/IML modules given in the Appendix can be used.

Table 3: *Results of the MCPs*

	Conover's data					hypothetical data				
	rank sum diff.	significance			individual p value	rank sum diff.	significance			individual p value
		BMDP	Conover	Holm			BMDP	Conover	Holm	
1 vs. 2	14.5	-	★★	-	0.027	3.0	-	-	-	0.674
1 vs. 3	13.5	-	★★	★★	0.002	22.5	-	★★	★	0.023
1 vs. 4	4.0	-	-	-	1.000	22.5	-	★★	★★	0.002
2 vs. 3	-1.0	-	-	-	0.551	19.5	-	★★	★	0.017
2 vs. 4	-10.5	-	★	-	0.266	19.5	-	★★	★★	0.006
3 vs. 4	-9.5	-	-	-	0.266	0.0	-	-	-	0.766

★ significant for α=0.10

★★ significant for α=0.05

References

BERNHARD, G., ALLE, M., HERBOLD, M., MEYERS, W. (1988): Investigation on the reliability of some elementary nonparametric methods in statistical analysis systems. *Statistical Software Newsletter* 14, 19-26.

BMDP (1990): *Statistical Software Manual Vol.1* (edited by W. J. Dixon). University of California Press, Berkeley.

BÜNING, H., TRENKLER, G. (1978): *Nichtparametrische statistische Methoden*. De Gruyter, Berlin.

CONOVER, W.J. (1980): *Practical Nonparametric Statistics*. Wiley, New York.

FRIEDMAN, M. (1937): The use of ranks to avoid the assumption of normality implicit in the analysis of variance. *Journal of the American Statistical Association* 32, 675-701.

HAUX, R., HILGERS, R., HÖRMANN, A., JÖCKEL, K.-H., LEHMACHER, W. (1988): Recommendations on software for nonparametric statistical methods. *Statistical Software Newsletter* 14, 27.

HILGERS, R. (1985): Rangverfahren und Grenzen ihrer Anwendung. In: Jedinsky, H.J., Trampisch, H.J. (Hrsg.): *Prognose- und Entscheidungsfindung in der Medizin; 30. Jahrestagung der GMDS, Düsseldorf, Sept. 1985, Proceedings*. Springer, Berlin, 472-493.

HOLM, S. (1979): A simple sequentially rejective multiple test procedure. *Scandinavian Journal of Statistics* 6, 65-70.

SAS (1987): *SAS/STAT Guide for Personal Computers, Version 6 Edition*. SAS Institute Inc., Cary, NC.

SPSS (1986): *SPSS/PC+ for IBM PC/XT/AT*. SPSS Inc., Chicago.

STATGRAPHICS (1992): *Referenzhandbuch der statistischen Prozeduren, Version 5*. Manugistics Inc., Rockville.

WRIGHT, S.P. (1992): Adjusted p-values for simultaneous inference. *Biometrics* 48, 1005-1013.

APPENDIX

```
*************************************************************************
*                  COMPUTING OF THE FRIEDMAN TEST                       *
*                     WITH ADJUSTMENT FOR TIES                          *
*           AND MULTIPLE COMPARISONS USING THE HOLM PROCEDURE           *
*************************************************************************;

*-----------------------------------------------------------------------*
|                           READING DATA                                |
|                                                                       |
| Be sure that your data set "DATA" has a block*treatment structure,    |
|  i.e. the rows are the blocks and the columns are the treatments      |
|  and that there are no missing values !                               |
*-----------------------------------------------------------------------*;
PROC IML;   use DATA;   read all into X;
  reset center linesize=80 nolog spaces=0;
  n=nrow(X);     k=ncol(X);
  if n<10 then print
    'WARNING: The following results are based on approximations',
    '         and should be interpreted carefully, because the ',
    '         sample size is small !!!                         ',,,,;
*-----------------------------------------------------------------------*
|                           FRIEDMAN TEST                               |
*-----------------------------------------------------------------------*;
START FRIEDMAN;
*----------------------*computing of mean ranks*--------------------*;
  do i=1 to n;     Y  = Y // ranktie(X[i,]);     end;
  do j=1 to k;     RS = RS || sum(Y[,j]);        end;
  SSQX = ssq(Y);  SSRS = ssq(RS);
*----------------------*some different statistics*------------------*;
  F_cla = 12/(n*k*(k+1))*SSRS - 3*n*(k+1);
  F_tie = (k-1)* (SSRS-(n**2*k*(k+1)**2)/4) / (SSQX-(n*k*(k+1)**2)/4);
  F_Con = (n-1)*F_tie / (n*(k-1)-F_tie);
*---------------------------*p values*-----------------------------*;
  p_cla = 1-probchi(F_cla,k-1);
  p_tie = 1-probchi(F_tie,k-1);
  p_Con = 1-probf(F_Con,(k-1),(n-1)*(k-1));
```

```
*-------------------------*printing of results*---------------------*;

  print 'RESULTS OF THE FRIEDMAN TEST',;
  print 'Number of Blocks and Treatments: ' N K,;
  print 'Classical Friedman Test:                    '
        F_cla [format=8.3] '  ' p_cla [format=7.5],;
  print 'Friedman Test with Adjustment for Ties:     '
        F_tie [format=8.3] '  ' p_tie [format=7.5];
  print 'F Approximation of CONOVER (1980, p. 300):  '
        F_Con [format=8.3] '  ' p_Con [format=7.5],,,;
FINISH;

*-------------------------------------------------------------------------*
|                    MULTIPLE COMPARISONS (HOLM, 1979)                    |
*-------------------------------------------------------------------------*;
START HOLM;
  create MC var{v w f p};
*-----------------*Friedman test of all possible pairs*-------------*;
  M = k*(k-1)*0.5;           * M is the Number of pairwise comparisons;
  do v=1 to k-1;
    do w=v+1 to k;
      XP = X[,v] || X[,w];
      free YP;
      do i=1 to n;   YI = ranktie(XP[i,]);   YP = YP // YI;   end;
      free RS;
      do j=1 to 2;   RJ = sum(YP[,j]);       RS = RS || RJ;   end;
      SSQX = ssq(YP);   SSRS = ssq(RS);
      F_tie = (SSRS-(n**2*2*9)/4) / (SSQX-(n*2*9)/4);
      F_Con = (n-1)*F_tie / (n-F_tie);
      p_Con = 1-probf(F_Con,1,n-1);
      MC    = v || w || F_Con || p_Con;
      pairv = v || w;
      append from MC;
    end;
  end;
  close MC;
*-------------------------*sorting of p values*---------------------*;
  sort MC by p;
  use MC;    read all into HOLM;
*-------------------------------*decisions*-------------------------*;
  do i=1 to M;
       if HOLM[i,4] <= 0.10/(M-i+1) then do;
     d10i=1;  d10=d10//d10i;
    end;  else goto D10R;
  end;
  D10R:  d10r=J(M-i+1,1,0);  D10=d10//d10r;
  do i=1 to M;
    if HOLM[i,4] <= 0.05/(M-i+1) then do;
      d05i=1;  d05=d05//d05i;
    end;  else goto D05R;
  end;
  D05R:  d05r=J(M-i+1,1,0);  D05=d05//d05r;
  PAIR = HOLM[,1:2];     P = HOLM[,4];
*-------------------------*printing of results*---------------------*;
  print 'RESULTS OF THE HOLM PROCEDURE';
  print
  'Individual p-Values and Simultaneous Decisions (Alpha=0.05,0.10)',;
  print  PAIR[format=1.0] '   ' P[format=7.5] D05 D10;
FINISH;
  run friedman;
  run holm;
quit;
run
```