DISCUSSION PAPER / gary m. grandon

Using SPSS MULT RESPONSE to Generate Filtered Marginals

Gary M. Grandon
Social Science Data Center
The University of Connecticut, Storrs

With the release of SPSS's Version 7.0 this year, several additions have been added to its already extensive battery of analysis programs. MULT RESPONSE is one of these programs. It generates frequency counts and bivariate tables for "dummy" coded multiple response questions. These multiple response variables are a real nuisance to the analyst and the program at face value provides a means for their interpretation. Considerations for the analysis of multiple response data are not the subject of this paper though, but rather the use of this program for quite another purpose; the generation of "filtered marginals."

Filtered marginals are typically basic frequency counts and percentages for specified variables for each of a number of subpopulations within a study. SPSS has in the past provided *SELECT IF statements to facilitate the processing of subpopulations. Each such statement followed by FREQUENCIES procedure statements will generate filtered marginals. The inherent problem with this type of coding
is that each *SELECT IF statement requires a complete reading of the data set. For large studies this procedure can be quite inefficient! Another alternative to the *SELECT IF procedure would be to create a new variable with codes indicating each filter group (subpopulation). Then the very fast CROSSTABS procedure can be used with the new filter group variable as one variable in the bivariate table and the variable upon which frequencies are to be generated as the other. A large number of frequencies can be generated in this manner with only one read of the data. The problem with this CROSSTABS procedure is that it requires that each filter group or subpopulation be discrete. That is, any one subject can only be in one filter group. This restriction is removed by using the MULT RESPONSE technique suggest in this paper. The following setup using the SPSS - OSIRIS interface to shorten the presentation illustrates the use of the MULT RESPONSE procedures to perform filtered marginals:

```
1
RUN NAME YANKELOVICH--COMBINED STUDIES-1976
OSIRIS VARS V7,V8,V11,V13,V18,V28,V64,V75
INPUT MEDIUM TAPE
N OF CASES UNKNOWN
WEIGHT V75
IF (((V7 EQ 1 OR 2 OR 3) AND (V8 EQ 6) AND (V11 EQ 1) AND (V18 EQ 1)F1=1
IF (((V7 EQ 5 or 6) AND (V8 EQ 1 OR 2) AND (V11 EQ )3) AND (V18 EQ 1)F2=1
MULT RESPONSE GROUPS=FILTERS (F1 F2(1))/
VARIABLES=V13(1,4),V28(1,3),V64(1,3)/
TABLES=FILTERS BY V13 TO V64/
STATISTICS 1
READ INPUT DATA
FINISH
```

Not only does the technique illustrated above produce the same kind of results as the *SELECT IF process but it formats the output in such a way as to allow the user to scan differences across subpopulations within a single tabular display. A single procedure statement like the one above can process 20 filter groups and up to 100 analysis variables. Each frequency variable will have its own table.

A resource expenditure comparison between the *SELECT IF procedure and the MULT RESPONSE procedure was performed on the 1976 Yankelovich Combined Study archived at The Social Science Data Center of the University of Connecticut. The study has a N of 7,977 subjects. Nineteen filter groups were identified and frequencies for 45 analysis variables were obtained using both techniques. Table 1 provides relevant resource comparisons for the techniques. The CROSSTABS technique was not included due to its unique group membership constraint. Clearly costs are much lower for the MULT RESPONSE procedure.
Table 1. A resource comparison* of the *SELECT IF and MULT RESPONSE techniques for filtered marginals.

<table>
<thead>
<tr>
<th>RESOURCE</th>
<th>*SELECT IF</th>
<th>MULT RESPONSE</th>
</tr>
</thead>
<tbody>
<tr>
<td>NUMBER OF JOBS</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>KBYT-SEC</td>
<td>21,016</td>
<td>24,957</td>
</tr>
<tr>
<td>ELAP-KBS</td>
<td>180,493</td>
<td>73,752</td>
</tr>
<tr>
<td>READER EXCP**</td>
<td>93</td>
<td>31</td>
</tr>
<tr>
<td>PRINTER EXCP</td>
<td>14,630</td>
<td>3,844</td>
</tr>
<tr>
<td>PUBLIC DISK EXCP</td>
<td>10,476</td>
<td>360</td>
</tr>
<tr>
<td>PRIVATE DISK EXCP</td>
<td>10</td>
<td>5</td>
</tr>
<tr>
<td>TAPE EXCP</td>
<td>72</td>
<td>36</td>
</tr>
<tr>
<td>TAPE MOUNTS</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>COST</td>
<td>$31.70</td>
<td>$13.39</td>
</tr>
<tr>
<td>ELAPSED TIME</td>
<td>13.48 min.</td>
<td>6.23 min.</td>
</tr>
</tbody>
</table>

*Comparisons were performed on the University of Connecticut's Research Computer Center IBM 360/65 - IBM 370/155 systems running under OS and shared spool HASP. Local costs may vary from installation to installation.

**Each EXCP represents the "Execution of a channel program" and indicates the movement of one block of data.

The MULT RESPONSE procedure can further be extended to filtered bivariate tables by including a second BY statement on the TABLES card. This technique is not illustrated due to space limitations. Readers are referred to their local installations for further documentation of SPSS Version 7.0.

DISCUSSION PAPER/ richard c. roistacher

The following article describes the work of Richard C. Roistacher and Barbara Noble at the Center for Advanced Computation at the University of Illinois. They are involved in the development of guidelines for the descriptive materials which accompany a data file.

A Source Documentation Style Manual

by
Richard Roistacher
Center for Advanced Computation
University of Illinois
Urbana, Illinois

Barbara Noble and Richard Roistacher of the University of Illinois' Center for Advanced Computation are currently developing a style manual for the documentation of machine readable data. The manual, which is being developed as part of a project funded by the U.S. Department of Justice's Law Enforcement Assistance Administration, is presently available in draft form. The manual, conforming to the