THE EFFECT OF SAMPLE SIZE ON THE TRUE \( \alpha \) LEVEL WITH FUNCAT - A SIMULATION STUDY

John L. Stedl, Abbott Laboratories

Abstract

FUNCAT analyzes categorical data using the Grizzle-STARREY-Koch methodology. This methodology exploits test statistics which have asymptotic \( \chi^2 \) distributions in order to carry out tests of significance. A simulation program was written to investigate the effect of sample size on the true \( \alpha \) levels for various \( 2 \times 2 \times k \) tables when the null hypothesis was true. This simulation program examined two response functions: log linear and linear. Many true \( \alpha \) levels were much greater or much less than the nominal levels of 0.01, 0.05 and 0.10, unless the sample sizes were very large. The results indicate that this methodology may not be appropriate for a 100 patient clinical trial with 2 treatments, 2 responses and 10 investigators. Cochran-Mantel-Haenszel analysis or ordinary \( \chi^2 \) analysis, however, of the same simulated data achieved true \( \alpha \) levels close to the nominal levels of 0.01, 0.05 and 0.10.

The Model and Data

Data from \( k \) \( 2 \times 2 \) contingency tables were generated using the UNIFORM function. Within each table the rows refer to two populations, A and B, and the columns refer to the responses, success (+) and failure (-). Let \( P_A = (P_{a1}, P_{a2}, \ldots, P_{ak}) \) and \( P_B = (P_{b1}, P_{b2}, \ldots, P_{bk}) \) be the success probabilities for the \( k \) tables. And let the sample sizes \( S = (S_1, S_2, \ldots, S_k) \) be equal within table for A and B.

Thus, for example, consider simulating a clinical trial with three investigators (\( K = 3 \)) that have respective sample sizes of 10, 20 and 16 (\( S = (5, 10, 8) \)), and success probabilities \( P_A = (0.5, 0.4, 0.1) \) and \( P_B = (0.7, 0.8, 0.1) \).

This simulation can be described as

\[
\begin{array}{cccccccc}
\text{INVESTIGATOR} & - & - & - & - & 1 & 2 & 3 \\
\text{Response} & + & - & + & - & \text{Response} \\
T & A & 0.5 & 0.5 & 0.4 & 0.6 & 10 & A & 0.1 & 0.9 & 8 \\
& B & 0.7 & 0.3 & 0.8 & 0.2 & 10 & B & 0.1 & 0.9 & 8 \\
\end{array}
\]

where the cell probabilities are shown in the tables and the sample sizes are given in the margins.

The Hypotheses and Test Statistics

Four test statistics were computed from each run of the simulation:

- Cochran-Mantel-Haenszel (CMH)
- Ordinary \( \chi^2 \) (pooling all tables)
- FUNCAT (log-linear response function)
- FUNCAT (linear response function)

The null hypothesis common to each statistic was one of no difference between A and B with respect to the success probabilities, i.e.,

\[ H_0: P_A = P_B \]

In addition, FUNCAT also produces a test of no table differences as well as a test for table by treatment interaction.

The Simulations and Results

The following six cases were simulated,
### K = 10 - TESTS FOR EQUAL TREATMENT EFFECT

- **FUNCAT -**

\[ \log \left( \frac{1}{x} \right) \text{ Linear } \chi^2 \]

<table>
<thead>
<tr>
<th>( \alpha = 0.01 )</th>
<th>5</th>
<th>( 0.003^* )</th>
<th>0.009</th>
<th>0.010</th>
</tr>
</thead>
<tbody>
<tr>
<td>( S = 10 )</td>
<td>0.015*</td>
<td>0.006</td>
<td>0.010</td>
<td>0.010</td>
</tr>
<tr>
<td>( S = 30 )</td>
<td>0.010</td>
<td>0.010</td>
<td>0.010</td>
<td>0.010</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>( \alpha = 0.05 )</th>
<th>5</th>
<th>( 0.02^* )</th>
<th>0.05</th>
<th>0.06</th>
</tr>
</thead>
<tbody>
<tr>
<td>( S = 10 )</td>
<td>0.10*</td>
<td>0.03*</td>
<td>0.05</td>
<td>0.05</td>
</tr>
<tr>
<td>( S = 30 )</td>
<td>0.08*</td>
<td>0.05</td>
<td>0.07*</td>
<td>0.07*</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>( \alpha = 0.10 )</th>
<th>5</th>
<th>( 0.06 )</th>
<th>0.10</th>
<th>0.10</th>
</tr>
</thead>
<tbody>
<tr>
<td>( S = 10 )</td>
<td>0.17*</td>
<td>0.07*</td>
<td>0.10</td>
<td>0.10</td>
</tr>
<tr>
<td>( S = 30 )</td>
<td>0.14*</td>
<td>0.11</td>
<td>0.12</td>
<td>0.12</td>
</tr>
</tbody>
</table>

### K = 10 - TESTS FOR NO TABLE EFFECT

AND NO TABLE X TRI INTERACTION

<table>
<thead>
<tr>
<th>( \alpha = 0.01 )</th>
<th>5</th>
<th>( 0.003^* )</th>
<th>0.004</th>
<th>0.005</th>
</tr>
</thead>
<tbody>
<tr>
<td>( S = 10 )</td>
<td>0.10*</td>
<td>0.001*</td>
<td>0.006</td>
<td>0.006*</td>
</tr>
<tr>
<td>( S = 30 )</td>
<td>0.03*</td>
<td>0.002*</td>
<td>0.03*</td>
<td>0.010</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>( \alpha = 0.05 )</th>
<th>5</th>
<th>( 0.004^* )</th>
<th>0.007*</th>
<th>0.008*</th>
</tr>
</thead>
<tbody>
<tr>
<td>( S = 10 )</td>
<td>0.12*</td>
<td>0.007*</td>
<td>0.16*</td>
<td>0.03*</td>
</tr>
<tr>
<td>( S = 30 )</td>
<td>0.11*</td>
<td>0.03*</td>
<td>0.08*</td>
<td>0.06</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>( \alpha = 0.10 )</th>
<th>5</th>
<th>( 0.007^* )</th>
<th>0.008*</th>
<th>0.009*</th>
</tr>
</thead>
<tbody>
<tr>
<td>( S = 10 )</td>
<td>0.14*</td>
<td>0.03*</td>
<td>0.23*</td>
<td>0.09</td>
</tr>
<tr>
<td>( S = 30 )</td>
<td>0.17*</td>
<td>0.08</td>
<td>0.14*</td>
<td>0.10</td>
</tr>
</tbody>
</table>

@FUNCAT issued warning to the user that the matrix was singular and results are not valid.
@Right of the 1000 runs at \( S = 5 \) resulted in warnings to the user.

#### Discussion of Results

The CMH and \( \chi^2 \) tests for treatment differences performed well for all cases studied with only two of the observed values of \( \alpha \) being outside of "normal" ranges. FUNCAT, on the other hand, produced many estimates of \( \alpha \) more than two standard errors above or below the nominal values. This was particularly true for tests of interaction and tests for table effects. These results for FUNCAT were more pronounced with the linear response function than the log-linear response function. In general larger sample sizes produced observed \( \alpha \) levels closer to the nominal levels than smaller sample sizes.

A final problem was noted with FUNCAT; with small \( n \) (5=5) the procedure sometimes issued a warning to the user that the matrix was singular and results were not valid.

#### A Clinical Trials Example

This example considered simulated data from a 10 investigator, 2 treatment 2 responses (cure, failure) clinical trial. The total sample size was 102 with \( S = (2, 3, 4, 4, 4, 5, 5, 6, 8, 10) \).
DATA INDEX;
N1=11 N2=100 ;
MACRO HATT
DATA AI
PROC MATRIX
FETCH CONST DATA=INDEX;
L=0;
N1=CONST(1,1);
N2=CONST(1,2);
REPEAT=1 1/1 2/2 1/2 2/2;
RESPJ=(4,1,0);
PROBA=5 .5 .5 .5 .5 .5 .5 .5 .5 .5 ;
PROBB=5 .5 .5 .5 .5 .5 .5 .5 .5 .5 ;
N=2 ;
SS=30 30 5 5 5 5 5 5 5 5;
CMH=J(100,1,0);
DO I=N1 TO N2 ;
L=I+1;
G=0;
V=0;
DO K=1 TO H.;
A=UNIFORM((1,SS(I,K)-0));
B=UNIFORM((1,SS(I,K)-0));
RESP(I,1)=ROWSUM(A*PROBA(I,K));
RESP(2,1)=SS(I,K)-RESP(I,1);
RESP(3,1)=ROWSUM(B*PROBB(I,K));
RESP(4,1)=SS(I,K)-RESP(3,1);
F=/(J(4,1,K))**REPEAT**RESP**J(4,1, )
PLUS=RESP(1,1)+RESP(3,1);
MINUS=RESP(2,1)+RESP(4,1);
II=RESP(1,1) -.5*PLUS;
G=G+II;
I2=(PLUS-MINUS)*INV((4*2*SS(1,K)-1));
V=V+I2;
CMH(I,1)=(G**2)*INV(V); END;
OUTPUT F OUT=$ MACRO RE (RENAME=(COL1=INV COL2=TRT COL3=RESP COL4=COUNT COL5=INDEX));X
MACRO JOHNNY OUTPUT CMH OUT= %
MACRO JASON (RENAME=(COL1=P_CMH)) ;
MATT FUN1 RE JOHNNY CMH1 JASON DATA INDEX; N1=101 N2=200 ;
MATT FUN2 RE JOHNNY CMH2 JASON DATA INDEX; N1=201 N2=300 ;
MATT FUN3 RE JOHNNY CMH3 JASON DATA INDEX; N1=301 N2=400 ;
MATT FUN4 RE JOHNNY CMH4 JASON DATA INDEX; N1=401 N2=500 ;
MATT FUN5 RE JOHNNY CMH5 JASON DATA INDEX; N1=501 N2=600 ;
MATT FUN6 RE JOHNNY CMH6 JASON DATA INDEX; N1=601 N2=700 ;
MATT FUN7 RE JOHNNY CMH7 JASON DATA INDEX; N1=701 N2=800 ;
MATT FUN8 RE JOHNNY CMH8 JASON DATA INDEX; N1=801 N2=900 ;
MATT FUN9 RE JOHNNY CMH9 JASON DATA INDEX; N1=901 N2=1000 ;
MATT FUNT RE JOHNNY CMH10 JASON DATA INDEX;
DATA FUN1; SET FUND FUN2 FUN3 FUN4 FUN5 FUN6 FUN7 FUN8 FUN9 FUN10;
DATA CH1;SET CH1 CH2 CH3 CH4 CH5 CH6 CH7 CH8 CH9 CH10;
DATA SAVE1;SET CHH;
DROP ROW;
DATA;
IF EOF THEN PUT CMH01= CMH05= CMH10=;*;
SET CMH END=EDF;f
IF P_CMH<0.01 THEN CMH01+1;
IF P_CMH<0.05 THEN CMH05+1;
IF P_CMH<0.10 THEN CMH10+1;
PROC PRINTTO UNIT=20 NEW;
PROC FUMCAT DATA=FUN ; BY INDEX; WEIGHT COUNT;
MODEL RESP=INV TRT ;
RESPONSE 1 -1;
PROC PRINTTO;
DATA BIFILE PRINT NOPRINT;
INFILE FT20F001#;
INPUT @2 NAME $ @;
IF NAME='INV' THEN INPUT DF CHI P_INV;
IF NAME='TRT' THEN INPUT DF CHI P_TRT;
IF NAME='RESIDUAL' THEN INPUT DF CHI P_RES;
RETAIN P_TRT P_INV;
IF NAME='RESIDUAL';
DROP NAME DF CHI;
DATA SAVE;SET DATA;
DATA Y;
IF EOF THEN PUT TRT01= TRT05= TRT10= / INV01= INV05= INV10= / RES01= RES05= RES10= RES50= /;
SET D END=EOF;
IF P_TRT<.01 THEN TRT01=1;
IF P_TRT<.05 THEN TRT05=1;
IF P_TRT<.10 THEN TRT10=1;
IF P_INV<.01 THEN INV01=1;
IF P_INV<.05 THEN INV05=1;
IF P_INV<.10 THEN INV10=1;
IF P_RES<.01 THEN RES01=1;
IF P_RES<.05 THEN RES05=1;
IF P_RES<.10 THEN RES10=1;
IF P_RES<.50 THEN RES50=1;
PROC PRINTTO UNIT=20 NEW;
PROC FREG DATA=FUN;BY INDEX;WEIGHT COUNT;
MODEL RESP=INV TRT;
PROC PRINTTO;
DATA EFFILE PRINT NOPRINT;
INFILE FT20FOO1;
INPUT @2 NAME & @;
IF NAME='INV' THEN INPUT DF CHI LP_INV;
IF NAME='TRT' THEN INPUT DF CHI LP_TRT;
IF NAME='RESIDUAL' THEN INPUT DF CHI LP_RES;
RETAIN LP_TRT LP_INV;
IF NAME='RESIDUAL';
DROP NAME DF CHI;
DATA SAVE;SET E;
DATA LOG;
IF EOF THEN PUT L_TRT01= L_TRT05= L_TRT10= / L_INV01= L_INV05= L_INV10= / L_RES01= L_RES05= L_RES10= L_RES50= /;
SET E END=EOF;
IF LP_TRT<.01 THEN L_TRT01=1;
IF LP_TRT<.05 THEN L_TRT05=1;
IF LP_TRT<.10 THEN L_TRT10=1;
IF LP_INV<.01 THEN L_INV01=1;
IF LP_INV<.05 THEN L_INV05=1;
IF LP_INV<.10 THEN L_INV10=1;
IF LP_RES<.01 THEN L_RES01=1;
IF LP_RES<.05 THEN L_RES05=1;
IF LP_RES<.10 THEN L_RES10=1;
IF LP_RES<.50 THEN L_RES50=1;
DATA CHI;
PROC PRINTTO UNIT=20 NEW;
PROC FREQ DATA = FUN;BY INDEX;
TABLES RESP*TRT/CHISQ;WEIGHT COUNT;
PROC PRINTTO;
DATA FREQFILE PRINT NOPRINT;
INFILE FT20FOO1;
INPUT @2 NAME & $10. @;
IF NAME='CHI-SQUARE' THEN INPUT +43 PROB;
RETAIN PROB;
IF NAME='CONTINUITY';
DROP NAME;
DATA SAVE;SET FREQ;
DATA FCUM;IF EOF THEN PUT CHI01= CHI05= CHI10= ;
SET FREQ END=EOF;
IF PROB<.01 THEN CHI01=1;
IF PROB<.05 THEN CHI05=1;
IF PROB<.10 THEN CHI10=1;
165