Exact Testing Procedures in SAS® for Categorical Data Analysis

Guogen Shan, University of Nevada Las Vegas, Las Vegas, NV
Darren Liu, University of Nevada Las Vegas, Las Vegas, NV

ABSTRACT

In categorical data analysis, the majority of methods are based on asymptotic approaches. As we all know, asymptotic approaches are very easy to calculate, however, they may have an unsatisfied type I error control. Actual type I error rates of testing procedures could be underestimated or overestimated for the discrete data. In addition, when power of the study is overestimated, sample size is definitely underestimated. The sample size in a study is actually not enough to attain the pre-specified power. In order to avoid such problems, exact approaches can be considered as alternatives in data analysis. SAS provides the option EXACT in several procedures for producing exact p-values as well as exact confidence intervals. We will review the commonly used conditional exact approach (Fisher’s test) by using examples in the SAS help files. We will also compare the EXACT statement in SAS to exact testing procedures in other statistical software packages.

INTRODUCTION

In categorical data analysis, asymptotic approaches are often used for statistical inference. The first topic in statistical inference would be the testing of hypotheses, and the second topic is to construct confidence intervals for the parameter of interest. PROC FREQ provides the EXACT statement to calculate exact p-values based on Fisher’s conditional approach. From SAS 9.3, Barnard test (Barnard, 1947) (an unconditional approach) is available in the EXACT statement for comparing proportions (see, What’s New in the Base SAS 9.3 Statistical Procedures). There are still some efficient testing procedures not included in SAS (Lloyd, 2008; Shan, Ma, Hutson, & Wilding, 2012, 2013; Shan, 2013a).

Asymptotic confidence intervals are often the default option for most statistical software packages. These confidence intervals do not guarantee the nominal level, where infimum coverage probabilities of these asymptotic approaches may be much less than the coverage size. For this reason, PROC FREQ presents exact confidence limits using the Clopper-Pearson method. Recently, Wang (2010) proposed a method for optimal exact confidence limits for the difference between two independent or dependent proportions with the use of an R package, ExactCldiff, which was developed by Shan and Wang (2013) to implement the new optimal confidence limits.
Table 1: Output from PROC FREQ for comparing two independent proportions

<table>
<thead>
<tr>
<th>Statistic</th>
<th>DF</th>
<th>Value</th>
<th>Prob</th>
</tr>
</thead>
<tbody>
<tr>
<td>Chi-Square</td>
<td>1</td>
<td>4.9597</td>
<td>0.0259</td>
</tr>
<tr>
<td>Likelihood Ratio Chi-Square</td>
<td>1</td>
<td>5.0975</td>
<td>0.0240</td>
</tr>
<tr>
<td>Continuity Adj. Chi-Square</td>
<td>1</td>
<td>3.1879</td>
<td>0.0742</td>
</tr>
<tr>
<td>Mantel-Haenszel Chi-Square</td>
<td>1</td>
<td>4.7441</td>
<td>0.0294</td>
</tr>
<tr>
<td>Phi Coefficient</td>
<td></td>
<td>0.4644</td>
<td></td>
</tr>
<tr>
<td>Contingency Coefficient</td>
<td></td>
<td>0.4212</td>
<td></td>
</tr>
<tr>
<td>Cramer’s V</td>
<td></td>
<td>0.4644</td>
<td></td>
</tr>
</tbody>
</table>

ASYMPTOTIC AND EXACT P-VALUES

We use an example from the PROC FREQ procedure in the SAS help file to compare asymptotic and exact p-values. This is the example used in for analyzing a 2 by 2 contingency table (Example 36.5 in SAS 9.3).

```sas
proc freq data=FatComp order=data;
  format Exposure ExpFmt . Response RspFmt .;
  tables Exposure*Response / chisq relrisk ;
  exact riskdiff binomial pdiff ;
  weight Count ;
  title 'Case−Control Study of High Fat/Cholesterol Diet' ;
  ods output chisq=chisq ;
run ;
```

The data is from a case-control study of patients with high cholesterol diet \( n_1 = 15 \) and low cholesterol diet \( n_2 = 8 \). The numbers of response are \( r_1 = 11 \) and \( r_2 = 2 \) in the high and low cholesterol diet, respectively. The response rate in the high cholesterol diet group \( 11/15=0.7333 \) is much higher than that in the low cholesterol diet group \( 2/6=0.25 \). The asymptotic p-value for testing the equality of these two proportions is 0.0259, see Table 1. One has to reject the null hypothesis of equal proportions. The exact test based on Fisher’s approach by fixing both marginal totals, also shows a significant difference between two proportions with p-value 0.0393.

The exact test based on Fisher’s approach may be conservative due to the limited sample space in small to medium sample settings. An alternative to the Fisher’s exact test is one that based on an unconditional framework, see Barnard (1947). The Barnard option is added in the EXACT statement of the PROC FREQ in SAS 9.3. While the Barnard test is a general approach, it is only available for testing risk difference in the current SAS procedures. Barnard test is computational intense as compared to the Fisher’s exact test. The discussion between Barnard test and Fisher’s test is almost as old as statistical history. Another unconditional test developed by Lloyd (2008) is based on the parametric bootstrap and Barnard’s test. This new approach has been shown to be generally more powerful as compared to Barnard’s test in many statistical problems (Shan et al., 2012; Shan, 2013b, 2013a; Shan et al., 2013).
Table 2: Output from PROC FREQ for comparing confidence intervals.

<table>
<thead>
<tr>
<th>Risk</th>
<th>ASE</th>
<th>(Asymptotic) 90% Confidence Limits</th>
<th>(Exact) 90% Confidence Limits</th>
</tr>
</thead>
<tbody>
<tr>
<td>Row 1</td>
<td>0.7333</td>
<td>0.5455</td>
<td>0.9211</td>
</tr>
<tr>
<td>Row 2</td>
<td>0.2500</td>
<td>0.0000</td>
<td>0.5018</td>
</tr>
<tr>
<td>Total</td>
<td>0.5652</td>
<td>0.3952</td>
<td>0.7352</td>
</tr>
<tr>
<td>Difference</td>
<td>0.4833</td>
<td>0.1692</td>
<td>0.7975</td>
</tr>
</tbody>
</table>

Figure 1: Actual coverage probability of upper confidence intervals for $n_1 = n_2 = 15$ using an asymptotic approach at a 0.05 nominal level.

**ASYMPTOTIC AND EXACT CONFIDENCE INTERVALS**

The confidence intervals for the proportion difference are displayed in Table 2. PROC FREQ produces asymptotic (Wald-type) and exact (Clopper-Pearson) confidence limits by default. There are another three asymptotic confidence limits available from PROC FREQ. For instance, the 90% exact confidence limits based on Clopper-Pearson is (0.1074,0.7510). Wang (2010) developed exact smallest (optimal) one-sided confidence limits for the difference of two independent and dependent proportions. These confidence limits can be implemented by an R package ExactCIdiff (Shan & Wang, 2013). The associated 90% confidence limits based on this implementation are (0.09202,0.75105).

Actual coverage probability of upper confidence intervals based on an asymptotic approach and the exact approach are shown in Figure 1 and Figure 2, respectively. The coverage probability of a confidence interval is the proportion that the calculated interval would contain the true value of the parameter of interest. As can be seen, the actual coverage is below the nominal level for majority of the proportion difference $\theta_d$. Similar results are observed for other asymptotic confidence intervals. Exact one-sided confidence limits in Wang (2010)
CONCLUSION

Exact tests and exact confidence limits should be used in practice for small to medium sample settings. It takes some time to compute exact p-values, and needs more efforts to calculate exact confidence limits. Shan and Wang (2013) provided the first available package in R to calculate exact confidence intervals for the difference of proportions. One of our ongoing project will be coming up with a method using SAS macros for calculating exact p-value and exact confident limits.

REFERENCE

trends based on the modified Baumgartner–Weiβ–Schindler statistics. *Computa-
tional Statistics & Data Analysis, 57*(1), 246–261.
*Journal of Statistical Planning and Inference, 136*(12), 4293–4306.
Wang, W. (2010, April). On construction of the smallest one-sided confidence interval for the

**ACKNOWLEDGMENTS**

Shan’s research is partially supported by Faculty Opportunity Awards from UNLV.

**INFORMATION**

Your comments and questions are valued and encouraged. Contact the author at:

Guogen Shan
Department of Environmental and Occupational Health
School of Community Health Sciences
University of Nevada Las Vegas,
Las Vegas, NV 89154
guogen.shan@unlv.edu

SAS and all other SAS Institute Inc. product or service names are registered trademarks
or trademarks of SAS Institute Inc. in the USA and other countries. ® indicates USA registra-
tion.
Other brand and product names are trademarks of their respective companies.