An Application of Fisher's Exact Test
in Linguistic Research

One-Soon Her   Rwei-ju Chuang   Kuang-Chao Chang

National Cheng-Chi University   *Fu Jen Catholic University*

ABSTRACT   Fisher's Exact Test (FET), proposed 90 years ago by the famous pioneering statistician Ronald A. Fisher, is one of the most useful and long-lasting statistical methods for finding the exact permutation significance levels of contingency tables. In particular, for testing whether two classification criteria in 2×2 contingency tables are independent in the situation of small cell frequencies, the much more rapid normal approximation and chi-square calculation are liable to be inaccurate; therefore, the FET can be used instead. In this article, we introduce the FET to the broad audience of researchers/teachers/students in theoretical linguistics by applying the FET to a set of categorical data obtained from research on the mathematical interpretation of plural markers such as the English /-s/ suffix in the world's languages. We hope the contents of this article can be of meaningful use for the broad audience of readers mentioned above.

Keywords   Categorical data; Chi-square test; Contingency table; Significance level; Degrees of freedom; Fisher's Exact Test (FET); *P*-value.

1. Introduction

   In social science research, researchers often encounter numerical data that is of discrete type and can be classified into several categories.   Such kind of data is called *categorical data* by statisticians.   One of the most commonly used statistical methods for analyzing categorical data is using *chi-square test* to test the null hypothesis of independence between two response variables for integer type of data arranged in a *contingency table*.   In a contingency table with r rows and c columns, which is often referred to as a $r \times c$ table, there is an integer