Published in final edited form as:
Commun Stat Simul Comput. 2010 ; 39(3): 612–623. doi:10.1080/03610910903528335.

Exact Statistical Tests for Heterogeneity of Frequencies Based on Extreme Values

Chih-Chieh Wu[1], **Roger C. Grimson**[2], and **Sanjay Shete**[1]

[1]Department of Epidemiology, The University of Texas M. D. Anderson Cancer Center, Houston, Texas 77030

[2]Clinical Statistics Consulting, Stony Brook, New York 11790

Abstract

Sophisticated statistical analyses of incidence frequencies are often required for various epidemiologic and biomedical applications. Among the most commonly applied methods is Pearson's χ^2 test, which is structured to detect non-specific anomalous patterns of frequencies and is useful for testing the significance for incidence heterogeneity. However, the Pearson's χ^2 test is not efficient for assessing the significance of frequency in a particular cell (or class) to be attributed to chance alone. We recently developed statistical tests for detecting temporal anomalies of disease cases based on maximum and minimum frequencies; these tests are actually designed to test of significance for a particular high or low frequency. We show that our proposed methods are more sensitive and powerful for testing extreme cell counts than is the Pearson's χ^2 test. We elucidated and illustrated the differences in sensitivity among our tests and the Pearson's χ^2 test by analyzing a data set of Langerhans cell histiocytosis cases and its hypothetical sets. We also computed and compared the statistical power of these methods using various sets of cell numbers and alternative frequencies. Our study will provide investigators with useful guidelines for selecting the appropriate tests for their studies.

Keywords

classical occupancy model; extreme value; maximum; minimum; temporal anomalies

1. Introduction

Investigators are often required to perform sophisticated statistical analyses of frequencies for various epidemiologic and biomedical applications. Among the most widely applied methods is the Pearson's χ^2 test of goodness-of-fit, which provides a quantitative measure of the overall discrepancies between the observed and expected frequencies in comparisons. The Pearson's χ^2 test is a robust general-purpose test, because any types of deviations between the observed and expected counts would make χ^2 values large when present with sufficiently large force. This generality feature may also make it inefficient, in some applications, at showing in what way the observed and expected numbers deviate. For

Address correspondence to: Chih-Chieh Wu, Ph.D. Department of Epidemiology, Unit 1340, The University of Texas M. D. Anderson Cancer Center, 1155 Pressler Street, Houston, Texas 77030, U.S.A. ccwu@mdanderson.org, Tel: 713-745-3977, Fax: 713-792-8261.

instance, in data analysis of epidemiology or life sciences, we are often particularly interested in an unusually high or low incidence frequencies and must test the significance for this specific frequency rather than for the overall homogeneity of frequencies.

Recently, we proposed, derived, and illustrated exact statistical tests for epidemiologic anomalies in a time series based on the minimum cell counts (Wu et al, 2008). Earlier, we also developed the exact tests for temporal clustering of disease cases on the basis of maximum cell counts (Grimson, 1993; Grimson, 1994; Grimson and Oden, 1996). In addition to our proposed exact tests, Ederer, Myers, and Mantel (EMM) (1964) and Mantel, Kryscio, and Myers (1976) proposed and discussed the related test statistics for detecting disease clustering over a space-time series based on maximum frequency. An informative discussion of the EMM statistic, the Grimson models, and numerous applications is given by J. Krauth (Glaz and Balakrishnan, 1999). These tests are based on the null hypothesis, in which health-related events are randomly assigned to consecutive cells that underlies a symmetric multinomial distribution. They have been shown to be useful for analyzing anomalous health-related incidence patterns over a time series. In fact, these statistical tests are structured to test for heterogeneity of frequencies and have wider applications than those used in temporal incidence analysis. They are based on extreme value random variables and are designed to test of significance for the largest or smallest frequencies. In contrast, the Pearson's χ^2 test is designed to assess overall discrepancies between the expected and observed frequency counts and is not efficient at assessing the deviation of a particular frequency to be attributed to sampling fluctuations.

When investigators are considering the $1\times c$ table of frequencies and have some idea about the type of departure the data are likely to show, if any. They first apply the Pearson's χ^2 test to assess the overall frequency deviations on the cells. If the overall null hypothesis is rejected, some subsequent “subgroup” data explorations are required to test for the specific frequency. Snedecor and Cochran described and illustrated this approach of χ^2 analysis (p. 196, 1989). In many if not most $1\times c$ tables, the null frequency is symmetric (that is, T/c per cell for T objects), as we illustrate in this paper. They pointed out the caveats of developing new hypothesis after examining the data and regarded such analysis as exploratory. These authors also described an direct approach, in which a statistical test is constructed that is sensitive to the expected or sought departure, if any, under the null hypothesis. “Often, the initial χ^2 test is omitted in this situation.” (Snedecor and Cochran, 1989, p. 198) The statistical tests based on maximum and minimum frequencies are of this type, sensitive to specific and not unusual expected departures from random allocations. Furthermore, compared with the approach based on the use of the χ^2 test, our statistical tests are robust for testing for extreme frequency because they do fully take into account the information on each of the cell frequencies simultaneously.

In this report, we show that our methods are more sensitive and efficient for testing extreme cell counts than the Pearson's χ^2 test and could therefore provide important and valuable information in epidemiologic and biomedical studies. We elucidated the differences in sensitivity among our proposed methods and the Pearson's χ^2 test by analyzing a data set of Langerhans cell histiocytosis cases and its hypothetical sets. Langerhans cell histiocytosis is a rare disorder with a heterogeneous clinical spectrum whose cause and pathogenesis remain

poorly understood. We also computed and compared the statistical power over various sets of cell numbers and alternative frequencies. The investigation of statistical sensitivity and power presented in this work will provide investigators with useful guidelines for selecting the appropriate statistical methods for their studies. We also provide a link to a suite of JAVA programs for efficiently computing the exact p-value, mean, and variance of our tests.

2. Statistical Tests for Homogeneity of Frequencies

In this section, we describe statistical tests based on the maximum and minimum cell frequencies and compare them with the Pearson's χ^2 test. We illustrate their differences in sensitivity and utility by analyzing the incidence distribution of Langerhans cell histiocytosis and its hypothetical sets.

2.1 Test Statistics

Consider a mathematical model in which *T* distinct objects (e.g., health-related events such as cancers) are randomly allocated into *c* consecutive cells (or classes) in such a manner that all possible and distinguishable occupancy configurations are equally likely in probability. This model underlies a symmetric multinomial distribution with parameters (*T*; 1/*c*) and is termed as the classical occupancy model or the Maxwell-Boltzmann model (Johnson and Kotz, 1977; Karr, 1993). On the basis of this model, Grimson (1993) developed the exact statistical test based on the random variable of the maximum frequency in a cell, denoted by *Max*. Recently, we proposed and developed the use of the random variable of the minimum frequency in a cell, denoted by *Min*, as a test for disease incidence anomalies in a time series. We derived exact expressions of the p-values and moments for the test using combinatorial mathematics (Wu et al, 2008). These statistical methods are based on relatively simple assumptions and can be used to test for heterogeneity of frequencies in various applications. The expressions of the exact p-value formulae for *Max* and *Min* are presented as follows:

$$Pr(\text{Max} \geq \max) = c^{-T} \sum_{k \geq 1} (-1)^{k-1} \binom{c}{k} \sum_{\substack{\max \leq j_i \leq T \\ i=1,\ldots,k}} (c-k)^{T-j_1-\ldots-j_k} \binom{T}{j_1 \ldots j_k} \quad (1)$$

$$Pr(\text{Min} \leq \min) = c^{-T} \sum_{k \geq 1} (-1)^{k-1} \binom{c}{k} \sum_{\substack{0 \leq j_i \leq \min \\ i=1,\ldots,k}} (c-k)^{T-j_1-\ldots-j_k} \binom{T}{j_1 \ldots j_k} \quad (2)$$

where *max* and *min* are the observed maximum and minimum cell frequencies, respectively.

The Pearson's χ^2 test is a robust general-purpose test of goodness-of-fit that provides an empirical quantitative measure for the overall discrepancies between the observed and expected cell frequencies. Any major deviations between the observed and expected frequencies that originate either from a single cell or jointly from multiple cells would make χ^2 values large. The Pearson's χ^2 test in this model setting is based on the assumption that

$$D_c^2 = \frac{\sum_{i=1}^{c}(n_i - T/c)^2}{T/c}$$

is approximately a χ^2 distribution with $c - 1$ degrees of freedom, where n_i represents the number of objects assigned to the i-th cell and $\sum_{i=1}^{c} n_i = T$. The Pearson's χ^2 test rejects the null hypothesis of equal cell frequencies at the a nominal significance level if

$$D_c^2 \geq \chi^2_{c-1,\alpha} \quad (3)$$

Where $\chi^2_{c-1,\alpha}$ is the critical value of the upper 100α% probability point of the χ^2 distribution with $c - 1$ degrees of freedom.

2.2 Analysis of Hypothetical Cases

To closely investigate the differences in sensitivity and utility among the 3 tests of interest, we considered the following 2 hypothetical cases of the 32 incidence cases occurring into 4 classes.

1. For the observed incidence distribution (12, 3, 14, 3), the p-values were 0.100 for *Min*, 0.063 for *Max*, and 0.005 for the Pearson's χ^2 test.
2. For the observed incidence distribution (9, 11, 2, 10), the p-values were 0.027 for *Min*, 0.857 for *Max*, and 0.100 for the Pearson's χ^2 test.

In the first example, in which there was no discernable maximum or minimum cell frequency, significant incidence heterogeneity was noted using the Pearson's χ^2 test. However, in the second example, a significantly low minimum cell frequency was found using *Min*, but this was not picked up as “heterogeneity” with the Pearson's χ^2 test.

2.3 Analysis of Langerhans Cell Histiocytosis

The data on childhood Langerhans cell histiocytosis (LCH) patients in Taiwan provide an opportunity to illustrate the differences in sensitivity and utility among the statistical tests of interest while testing for homogeneity of frequencies.

LCH is a rare disorder with a heterogeneous clinical spectrum, in which monoclonal proliferation of cells that phenotypically resemble Langerhans cells accumulate in various organs. The cause and pathogenesis of this disease have remained poorly understood since it was first described in 1893 (Hand, 1893). LCH can occur at any age with a peak between ages 1 and 3 years, and ≥ 50% of patients diagnosed are between 1 and 15 years old. The reported annual incidence is approximately 5.4 per 1 million children in the United States (Starling and Fernbach, 1984). In a nationwide study of LCH incidence in children less than 15 years old in Taiwan, 32 LCH incidence cases occurred from 1997 through 1998, when the most severe El Niño of the 20th century caused extreme global weather changes (Chen et al, 2003). In contrast, there were only 23 cases diagnosed in 1995, 1996, and 1999

combined. In this work, we used the observed seasonal incidence distribution of LCH during the El Niño period, which is (1, 15, 8, 8), to test for homogeneity of frequencies.

Assuming that the 32 childhood LCH cases are randomly allocated within 4 cells for the null hypothesis, the value of the Pearson's χ^2 test would be 12.25, and its p-value would be 0.007 with 3 degrees of freedom, suggesting that the deviations between the observed and expected cell frequencies are too large to be attributed to chance alone. When the individual cell deviations and their relative contributions to the total χ^2 value are examined, we can comment that the significant deviation for the observed frequencies (1, 15, 8, 8), compared with the expected count of 8, is only explained by the first and second cell frequencies equally. This example shows that the Pearson's χ^2 test may be less powerful and ideal than *Max* or *Min* for evaluating some specific alternative hypotheses. We used *Max* and *Min* in this study and found exact p-value of 0.024 for *Max* [$pr(Max \geq 15|T = 32, c = 4)$] by equation (1) and exact p-value of 0.005 for *Min* [$pr(Min \leq 1|T = 32, c = 4)$] by equation (2).

The results of the analysis of the observed LCH incidence distribution and its two hypothetical cases demonstrate that the tests based on *Max*, and *Min* are more sensitive to test for extreme cell counts and the Pearson's χ^2 test is more useful for testing overall incidence heterogeneity. In many applications, these three tests are not structured to replace with one another, but may be complementary to one another, because they characterize observed incidence patterns in different ways by making full use of available information about the data set in the analysis.

3. Comparisons of Statistical Power

In addition to elucidating the sensitivity differences, as described in the last section, we computed the empirical power figures and delineated the relative powers of the 3 tests of interest in this section. This information should provide investigators with useful guidelines.

The test statistic D_c^2 for the Pearson's χ^2 test expressed in (1), *Max*, and *Min*, are all discrete, and none of them can exactly reach the given nominal significance level, in general. For these discrete test statistics, it is necessary to closely examine their exact significance levels before computing the power figures. Because these tests must be adjusted to have identical significance levels in comparisons of power (Wu and Amos, 2003). We first calculated the figures of exact significance levels for the Pearson's χ^2 test in section 3.1. The exact p-values for *Max* and *Min* generally do not attain these given exact significance levels. Therefore, we assigned the power of *Max* and *Min* through linear interpolation adjusted to the exact significance levels of the Pearson's χ^2 test and presented these power figures of the tests in Section 3.2.

3.1 Testing for Equal Cell Frequencies

The enumeration of the exact significance levels for the Pearson's χ^2 test enables us to investigate the accuracy of the distributional approximation of the discrete test, D_c^2 to χ^2_{c-1} distributions. Note that if $c = 2$, the χ^2 approximation coincides with the usual normal approximation to the binomial distributions. At nominal significance levels of 0.05 and 0.01, Table 1 shows the exact significance levels of the Pearson's χ^2 test for $c = 3$, 5, and 7 with

sample sizes ranging from 20 to 100 in increments of 10 and from 100 to 160 in increments of 20, denoted by $T = 20(10)100(20)160$.

At 0.05 nominal significance level, the Pearson's χ^2 test for $c = 3$ is consistently a little liberal (the exact significance level is higher than the nominal significance level) and, for $c = 5$, it fluctuates near the nominal significance level of 0.5. The Pearson's χ^2 test slightly overestimates the exact significance level for $c = 3, 5$ at 0.01 nominal significance level and more closely approaches the exact values as the sample size increases. The Pearson's χ^2 test for $c = 7$ does not maintain appropriate true significance levels for nominal significance levels of 0.05 or 0.01 as for $c = 3, 5$. It tends to be conservative and becomes worse as the sample size increases (up to 160).

3.2 Computation of Statistical Power

With respect to various sets of alternatives that contain all but 1 equal cell frequencies, we computed and compared the power of the Pearson's χ^2 test for the null hypothesis of equal cell frequencies with those of the tests based on *Max* and *Min*. The power of the Pearson's χ^2 test is obtained by

$$\sum \frac{T!}{n_1! \dots n_c!} p_1^{n_1} \dots p_c^{n_c} \quad (4)$$

where the summation is over all c-tuples $(n_1,\dots, n_c)$ subject to $\sum_{i=1}^{c} n_i = T$ and satisfying constraint (1) with respect to specific alternative cell probabilities $(p_1,\dots, p_c)$, $\sum_{i=1}^{c} p_i = 1$. The calculation of power for *Max* and *Min* is also based on (2), but the summation is subject to different constraints: *maximum*$\{n_1,\dots, n_c\} \geq max$ for *Max* and *minimum*$\{n_1,\dots, n_c\} \leq min$ for *Min*, where *max* and *min* are the critical values of the rejection regions for *Max* and *Min*, respectively.

Letting α_e represent the exact significance level for the Pearson's χ^2 test at α nominal significance level, there exist 2 adjacent integers, a and $a + 1$, as the critical values of the rejection regions for *Max*, such that the exact p-value based on a as the critical value, denoted by p_a, is the smallest value that is higher than α_e and the exact p-value based on $a + 1$, denoted by p_{a+1}, is the largest value that is lower than α_e (thus, $p_a > \alpha_e > p_{a+1}$). We assign the power of *Max* at α_e through linear interpolation between the power values at p_a and p_{a+1}. For example, the exact significance level of D_c^2 at $\alpha = 0.05$ is 0.0502 for $c = 3$, $T = 50$, as shown in Table 1. The exact p-values of *Max* are 0.0666 for *max* = 24 and 0.0325 for *max* = 25. Assuming the alternative cell probabilities to be (0.2, 0.2, 0.6), the exact power values of *Max* are 0.9686 and 0.9427 for *max* = 24 and 25, respectively. Therefore, the power of *Max* adjusted at $\alpha_e = 0.0502$ is 0.9562 through linear interpolation between 0.9686 and 0.9427. The power calculation for *Min* at α_e can be performed in a similar way.

We assumed that exactly 1 cell frequency differed from the others for the alternative cell probabilities; that is, $p_1 = \dots = p_{c-1} = p^0$, $p_c = 1 - (c - 1)p^0$. For $c = 3$, we let p^0 range from 0.10 to 0.45 in increments of 0.05, denoted by $p^0 = 0.10(0.05)0.45$ (thus, $p_3 = 0.80(-$

0.10)0.10). We let p^0 range from 0.05 to 0.225 in increments of 0.025 for $c = 5$; that is $p^0 = 0.05(0.025)0.225$ and $p_5 = 0.80(-0.10)0.10$. For $c = 7$, we set $p^0 = 0.10$, 0.125, 0.150, and 0.16 (thus, $p_7 = 0.40$, 0.25, 0.10, and 0.04, respectively). The select power values for $c = 3$, 5, and 7 are presented in Tables 2, 3, and 4, respectively, at nominal significance levels of 0.05 and 0.01. The power figures of *Min* were omitted and denoted by "-" for a few very small numbers of T in Tables 3 and 4, because even the least p-value of *Min* ($pr(Min = 0)$) is greater than the value of α_e and the power figures of *Min* adjusted at α_e cannot be approached through linear interpolation.

3.3 Power Comparisons

Our results showed that, when $p_c > p^0$, *Max* is more powerful than the Pearson's χ^2 test for $c = 3$, 5, 7, but *Max* has less power when $p_c < p^0$. When p_c is much larger than p^0 (e.g., (p^0, p_3) = (0.10, 0.80), (0.15, 0.70), (0.20, 0.60) and (p^0, p_5) = (0.05, 0.80), (0.075, 0.70), (0.10, 0.60)), both *Max* and the Pearson's χ^2 test have very high power figures (data not shown) and the relative advantage of *Max* is unclear. The largest excess of power for *Max* over the Pearson's χ^2 test occurred at (p^0, p_3) = (0.25, 0.50), (p^0, p_5) = (0.15, 0.40), and (p^0, p_7) = (0.125, 0.25), as shown in Tables 2, 3, and 4, respectively. For (p^0, p_3) = (0.25, 0.50), *Max* has higher power figures than the Pearson's χ^2 test by ≥ 5.49% for $T \geq 70$ at $\alpha = 0.01$. For (p^0, p_5) = (0.15, 0.40), power excess is as high as 10.05% at $\alpha = 0.05$ and 8.70% at $\alpha = 0.01$. The substantial gains in power by *Max* increase as the number of cells increases. For (p^0, p_7) = (0.125, 0.25) and $T \geq 60$, *Max* has more power by ≥ 12.4% and by ≥ 21.2% at nominal significance levels of 0.05 and 0.01, respectively.

In contrast with the performance of *Max* in power, *Min* functioned in the opposite way. Compared with *Max* and the Pearson's χ^2 test, *Min* has least power when $p_c > p^0$. However, when $p_c < p^0$, the power of *Min* increases dramatically and *Min* becomes much more powerful than the Pearson's χ^2 test and *Max*. For instance, *Min* has more power than the Pearson's χ^2 test by ≥ 14.2% at 0.01 nominal significance level for (p^0, p_3) = (0.40, 0.20), as shown in Table 2. For (p^0, p_5) = (0.225, 0.10) and $T \geq 40$, *Min* has more power by ≥ 27.5% at 0.05 nominal significance level and ≥ 76.5% at 0.01 nominal significance level, as shown in Table 3. The substantial advantage in power for *Min* increases as the number of cells increases. This same phenomenon was seen with *Max*. For (p^0, p_7) = (0.16, 0.04), as shown in Table 4, *Min* has more power than the Pearson's χ^2 test by ≥ 39.8% at 0.05 nominal significance level and ≥ 135.7% at 0.01 nominal significance level.

We conclude that *Max* has higher power than the Pearson's χ^2 test and *Min* when $p_c > p^0$ and the least power when $p_c < p^0$. In contrast, *Min* is the most powerful among the 3 tests when $p_c < p^0$ and the least powerful when $p_c > p^0$. *Max* and *Min* are very sensitive to test of significance for the highest and lowest frequencies, respectively, whereas the Pearson's χ^2 test is more useful to test for overall heterogeneity. As anticipated, *Max* and *Min* are sensitive in opposite directions. The relative sensitivity of the Pearson's χ^2 test is mediocre for either $p_c < p^0$ and $p_c > p^0$, compared with *Max* and *Min*. It is noteworthy that the superiority of *Max* or *Min* to the Pearson's χ^2 test is more substantial when there are larger numbers of cells. Here, we found that *Max* or *Min* are most dominant over the Pearson's χ^2 test for $c = 7$ and least for $c = 3$.

4. Discussion

In this paper, we revisited statistical tests that we originally developed for detecting epidemiologic anomalies in a time series based on the extreme value random variables (Grimson, 1993; Grimson, 1994; Grimson and Oden, 1996; Wu et al, 2008). These statistical tests, *Max* and *Min*, are based on relatively mild assumptions and have wider applications than those used in temporal incidence analysis. In fact, they can be applied to test for frequency heterogeneity in various applications when the models underlie a symmetric multinomial distribution. We performed an analysis to articulate the differences in sensitivity and presented comprehensive comparisons of empirical power figures among these tests and the Pearson's χ^2 test.

In contrast to the extreme value tests, *Max* and *Min*, which are designed to be sensitive to the occurrence of a particular high or low frequency, we showed that the Pearson's χ^2 test was structured to detect non-specific anomalous patterns of incidence frequencies. As anticipated, our results showed that *Max* and *Min* have substantially more power than the Pearson's χ^2 test over 2 different regions, $\{p_c > p^0\}$ and $\{p_c < p^0\}$, respectively, which are exclusive and exhaustive in the general alternatives, composing all but one equal cell frequencies. More important, we showed that the superiority of *Max* or *Min* to the Pearson's χ^2 test is more substantial for cases with larger numbers of cells. In our analysis, *Max* or *Min* are most dominant over the Pearson's χ^2 test for $c = 7$ and least for $c = 3$.

Although we compared statistical power only with respect to the alternative hypothesis in which all but one have equal cell frequencies, our analysis of LCH incidence distribution and the 2 hypothetical examples described in the second section illustrates the sensitivity differences among the 3 tests of interest and shows that their applicability and utility are not restricted to this situation only. *Max* and *Min* provide exact testing results unlike the Pearson's χ^2 test, which gives approximate values. The Pearson's χ^2 *test* has wider applicability because it is useful for probability distributions other than symmetric multinomial distributions. In this study, we show that *Max* and *Min* have higher power and more efficient to test for extreme frequencies than the Pearson's χ^2 *test*. This feature could provide more valuable and important information in epidemiologic or biomedical studies than do the tests based on the detection of overall non-specific frequency anomalies. We provide a suite of JAVA programs to efficiently compute the exact p-value, mean, and variance for *Min* and *Max* on our website (http://www.epigenetic.org/software.php).

Acknowledgments

This work was supported by U.S. National Cancer Institute grants 1R03-CA128103 (Wu CC) and 1R01-CA131324 (Shete S).

Bibliography

Chen RL, Lin KS, Chang WH, et al. Childhood Langerhans cell histiocytosis increased during El Niño 1997-98: a report from Taiwan Pediatric Oncology Group. Acta Paediatrica Taiwanica. 2003; 44:14–20. [PubMed: 12800378]

Ederer F, Myers MH, Mantel N. A statistical problem in space and time: do leukemia cases come in clusters? Biometrics. 1964; 20:626–638.

Glaz, J.; Balakrishnan, N. Scan Statistics and Applications. Vol. Chapter 3. Birkhauser; Boston/Berlin: 1999.

Grimson RC. Disease clusters, exact distributions of maxima, and *P*-values. Statistics in Medicine. 1993; 12:1773–1794. [PubMed: 8272660]

Grimson, RC. 1994 Proceedings of the Epidemiology Section. American Statistical Association; 1994. Tests based on the maximum occupancy frequency; p. 64-69.

Grimson RC, Oden N. Disease clusters in structured environments. Statistics in Medicine. 1996; 15:851–871. [PubMed: 9132911]

Hand A Jr. Polyuria and tuberculosis. Archives of Pediatrics. 1893; 10:673–675.

Johnson, NL.; Kotz, S. Urn models and their applications: an approach to modern discrete probability theory. John Wiley & Sons; New York: 1977.

Karr, AF. Probability. Spring-Verlag; New York: 1993.

Mantel N, Kryscio RJ, Myers MH. Tables and formulas for extended use of the Ederer-Myers-Mantel disease-clustering procedure. American Journal of Epidemiology. 1976; 104:576–584. [PubMed: 984032]

Snedecor, GW.; Cochran, WG. Statistical Methods. 8th. Iowa State University Press; Iowa: 1989.

Starling, KA.; Fernbach, DJ. Histiocytosis. In: Sutow, WW.; Fernbach, DJ.; Vietti, TJ., editors. Clinical pediatric oncology. Mosby; St. Louis: 1984. p. 498-515.

Wu CC, Amos CI. Statistical properties of affected sib-pair linkage tests. Human Heredity. 2003; 55:153–162. [PubMed: 14566093]

Wu CC, Grimson RC, Amos CI, Shete S. Statistical methods for anomalous discrete time series based on minimum cell count. Biometrical Journal. 2008; 50:86–96. [PubMed: 17853406]

Table 1

Exact significance levels of the Pearson's χ^2 test for c – 3, 5, 7 and T = 20(10)100(20)160 at nominal significance levels of 0.05 and 0.01

	c = 3		c = 5		c = 7	
T	α = 0.05	α = 0.01	α = 0.05	α = 0.01	α = 0.05	α = 0.01
20	0.0556	0.0072	0.0505	0.0095	0.0460	0.0095
30	0.0478	0.0099	0.0461	0.0098	0.0455	0.0092
40	0.0514	0.0083	0.0483	0.0087	0.0456	0.0092
50	0.0502	0.0091	0.0465	0.0098	0.0472	0.0095
60	0.0536	0.0093	0.0503	0.0096	0.0452	0.0090
70	0.0513	0.0095	0.0498	0.0099	0.0471	0.0087
80	0.0538	0.0089	0.0502	0.0098	0.0434	0.0086
90	0.0555	0.0098	0.0482	0.0096	0.0432	0.0085
100	0.0533	0.0099	0.0505	0.0099	0.0408	0.0080
120	0.0500	0.0094	0.0491	0.0097	0.0377	0.0072
140	0.0503	0.0093	0.0500	0.0099	0.0345	0.0058
160	0.0509	0.0096	0.0499	0.0097	0.0307	0.0047

Commun Stat Simul Comput. Author manuscript; available in PMC 2015 January 02.

Table 2

Statistical power of the tests for $c = 3$ and $T = 20(10)100$ at nominal significance levels of 0.05 and 0.01 when $p_1 = p_3 = p^0$ and $p_3 = 1 - 2p^0$

			$\alpha = 0.05$			$\alpha = 0.01$		
p^0	p_3	T	*Max*	χ^2	*Min*	*Max*	χ^2	*Min*
0.25	0.50	20	0.291	0.290	0.185	0.097	0.094	0.036
		30	0.396	0.374	0.224	0.201	0.195	0.075
		40	0.518	0.49	0.301	0.270	0.250	0.091
		50	0.613	0.592	0.360	0.372	0.359	0.126
		60	0.707	0.683	0.444	0.467	0.436	0.166
		70	0.769	0.754	0.490	0.556	0.518	0.202
		80	0.831	0.811	0.558	0.625	0.587	0.238
		90	0.878	0.854	0.619	0.704	0.666	0.290
		100	0.906	0.887	0.664	0.762	0.722	0.333
0.30	0.40	20	0.088	0.089	0.078	0.016	0.016	0.012
		30	0.096	0.094	0.077	0.028	0.028	0.019
		40	0.120	0.117	0.093	0.031	0.028	0.019
		50	0.136	0.134	0.101	0.041	0.04	0.023
		60	0.163	0.158	0.118	0.051	0.047	0.027
		70	0.178	0.175	0.124	0.061	0.057	0.031
		80	0.205	0.198	0.139	0.069	0.063	0.033
		90	0.231	0.220	0.153	0.084	0.076	0.039
		100	0.246	0.234	0.160	0.096	0.085	0.043
0.35	0.30	20	0.062	0.063	0.063	0.009	0.009	0.009
		30	0.057	0.058	0.058	0.013	0.013	0.013
		40	0.064	0.067	0.066	0.011	0.012	0.012
		50	0.065	0.069	0.068	0.013	0.014	0.015
		60	0.072	0.077	0.077	0.014	0.016	0.016
		70	0.072	0.078	0.078	0.016	0.018	0.018
		80	0.078	0.086	0.086	0.015	0.018	0.019
		90	0.084	0.093	0.093	0.018	0.021	0.022

			$\alpha = 0.05$			$\alpha = 0.01$		
p^0	p_3	T	*Max*	χ^2	*Min*	*Max*	χ^2	*Min*
		100	0.083	0.093	0.095	0.019	0.023	0.024
		20	0.145	0.190	0.218	0.028	0.039	0.052
		30	0.170	0.257	0.287	0.052	0.088	0.124
		40	0.219	0.365	0.395	0.060	0.129	0.167
		50	0.256	0.429	0.481	0.080	0.173	0.237
0.40	0.20	60	0.310	0.526	0.584	0.100	0.242	0.319
		70	0.344	0.570	0.645	0.121	0.310	0.386
		80	0.397	0.656	0.716	0.136	0.366	0.457
		90	0.449	0.728	0.775	0.167	0.445	0.533
		100	0.479	0.764	0.818	0.191	0.522	0.596
		20	0.315	0.564	0.682	0.082	0.197	0.309
		30	0.415	0.773	0.841	0.166	0.489	0.648
		40	0.556	0.921	0.941	0.214	0.696	0.799
		50	0.665	0.963	0.979	0.302	0.822	0.903
0.45	0.10	60	0.785	0.989	0.995	0.390	0.922	0.964
		70	0.854	0.995	0.998	0.485	0.967	0.985
		80	0.921	0.999	1	0.562	0.986	0.995
		90	0.963	1	1	0.670	0.995	0.998
		100	0.978	1	1	0.751	0.999	0.999

Commun Stat Simul Comput. Author manuscript; available in PMC 2015 January 02.

Table 3

Statistical power of the tests for c – 5 and T – 20(10)100 at nominal significance levels of 0.05 and 0.01 when p_1=p_2 = p_3 = p_4 =p^0 and p_3 = 1 – 4p^0

			α = 0.05			α = 0.01		
p^0	p_5	T	*Max*	χ^2	*Min*	*Max*	χ^2	*Min*
		20	0.750	0.701	-	0.528	0.497	-
		30	0.897	0.864	0.311	0.761	0.741	0.093
		40	0.963	0.948	0.416	0.888	0.864	0.136
		50	0.987	0.980	0.489	0.958	0.942	0.198
0.125	0.50	60	0.996	0.993	0.613	0.985	0.976	0.255
		70	0.999	0.998	0.692	0.995	0.991	0.332
		80	1	0.999	0.758	0.998	0.996	0.412
		90	1	0.999	0.820	0.999	0.998	0.462
		100	1	1	0.866	0.999	0.999	0.544
		20	0.411	0.373	-	0.205	0.189	-
		30	0.566	0.517	0.164	0.351	0.331	0.042
		40	0.701	0.657	0.211	0.481	0.444	0.054
		50	0.794	0.752	0.243	0.625	0.575	0.075
0.15	0.40	60	0.877	0.837	0.308	0.735	0.677	0.092
		70	0.921	0.891	0.349	0.814	0.766	0.116
		80	0.952	0.930	0.395	0.871	0.830	0.140
		90	0.970	0.953	0.435	0.915	0.878	0.157
		100	0.983	0.971	0.485	0.944	0.915	0.187
		20	0.130	0.124	-	0.040	0.038	-
		30	0.168	0.156	0.080	0.063	0.060	0.019
		40	0.219	0.204	0.095	0.085	0.077	0.020
0.175	0.30	50	0.263	0.242	0.101	0.122	0.108	0.025
		60	0.329	0.297	0.121	0.158	0.134	0.028
		70	0.375	0.340	0.131	0.194	0.166	0.033
		80	0.428	0.386	0.143	0.23	0.197	0.036

			$\alpha = 0.05$				$\alpha = 0.01$			
p^0	p_5	T	Max	χ^2	Min	Max	χ^2	Min
		90	0.470	0.420	0.150	0.271	0.226	0.039
		100	0.523	0.469	0.166	0.311	0.261	0.044
		20	0.087	0.108	-	0.019	0.024	-
		30	0.092	0.136	0.179	0.023	0.035	0.056
		40	0.107	0.186	0.253	0.024	0.045	0.089
		50	0.115	0.230	0.312	0.031	0.070	0.135
0.225	0.10	60	0.136	0.297	0.405	0.034	0.094	0.181
		70	0.145	0.352	0.475	0.039	0.125	0.241
		80	0.158	0.411	0.542	0.042	0.158	0.307
		90	0.164	0.456	0.605	0.046	0.194	0.353
		100	0.182	0.521	0.664	0.050	0.238	0.421

Table 4

Statistical power of the tests for $c = 7$ and $T = 20(10)100$ at nominal significance levels of 0.05 and 0.01 when $p_1 = p_2$, $= p_3 = p_4 = p_5 = p_6 = p^0$ and $p_7 = 1 - 6p^0$

			$\alpha = 0.05$			$\alpha = 0.01$		
p^0	p_7	T	*Max*	χ^2	*Min*	*Max*	χ^2	*Min*
		20	0.617	0.586	-	0.428	0.404	-
		30	0.817	0.773	-	0.646	0.611	-
		40	0.911	0.880	0.194	0.808	0.762	-
		50	0.960	0.940	0.253	0.906	0.868	0.070
0.10	0.40	60	0.975	0.964	0.307	0.950	0.921	0.092
		70	0.978	0.972	0.373	0.965	0.950	0.111
		80	0.974	0.971	0.398	0.969	0.961	0.129
		90	0.965	0.964	0.443	0.964	0.96	0.151
		100	0.954	0.954	0.469	0.954	0.953	0.170
		20	0.139	0.136	-	0.052	0.050	-
		30	0.204	0.188	-	0.083	0.076	-
		40	0.264	0.240	0.079	0.123	0.107	-
		50	0.333	0.298	0.091	0.173	0.146	0.021
0.125	0.25	60	0.391	0.348	0.098	0.221	0.180	0.023
		70	0.460	0.406	0.112	0.266	0.218	0.024
		80	0.506	0.446	0.112	0.316	0.260	0.026
		90	0.560	0.493	0.118	0.365	0.300	0.028
		100	0.601	0.526	0.120	0.409	0.337	0.029
		20	0.053	0.056	-	0.012	0.012	-
		30	0.056	0.060	-	0.012	0.013	-
		40	0.058	0.066	0.064	0.013	0.015	-
0.15	0.10	50	0.062	0.074	0.075	0.014	0.017	0.018
		60	0.062	0.077	0.082	0.014	0.018	0.021
		70	0.066	0.086	0.096	0.014	0.019	0.024
		80	0.063	0.087	0.099	0.014	0.021	0.027

			$\alpha = 0.05$			$\alpha = 0.01$		
P^0	p_7	T	*Max*	χ^2	*Min*	*Max*	χ^2	*Min*
		90	0.065	0.094	0.109	0.014	0.023	0.031
		100	0.064	0.097	0.114	0.014	0.024	0.034
		20	0.075	0.100	-	0.018	0.023	-
		30	0.086	0.135	-	0.020	0.033	-
		40	0.094	0.185	0.313	0.023	0.048	-
		50	0.106	0.251	0.457	0.026	0.070	0.198
0.16	0.04	60	0.111	0.313	0.581	0.028	0.094	0.317
		70	0.124	0.399	0.689	0.029	0.127	0.404
		80	0.123	0.461	0.740	0.032	0.169	0.482
		90	0.131	0.536	0.796	0.033	0.221	0.570
		100	0.133	0.594	0.831	0.034	0.273	0.643