

1-2009

Testing for the Supremacy of a Multinomial Cell Probability

Dan Nettleton

Iowa State University, dnett@iastate.edu

Follow this and additional works at: http://lib.dr.iastate.edu/stat_las_preprints

 Part of the [Statistics and Probability Commons](#)

Recommended Citation

Nettleton, Dan, "Testing for the Supremacy of a Multinomial Cell Probability" (2009). *Statistics Preprints*. 92.
http://lib.dr.iastate.edu/stat_las_preprints/92

This Article is brought to you for free and open access by the Statistics at Iowa State University Digital Repository. It has been accepted for inclusion in Statistics Preprints by an authorized administrator of Iowa State University Digital Repository. For more information, please contact digirep@iastate.edu.

Testing for the Supremacy of a Multinomial Cell Probability

Abstract

Tests for the supremacy of a multinomial cell probability are developed. The tested null hypothesis states that a particular cell of interest is not more probable than all others. Rejection of this null leads to the conclusion that the cell of interest has a strictly greater probability than all other cells. The null hypothesis constrains the multinomial probability vector to a non-convex region that is a union of closed convex cones. The likelihood ratio test for this problem is derived and shown to be equivalent to an intersection-union test. The least favorable configuration of the multinomial probability vector in the null parameter space is derived, and the limiting null distribution of the test statistic that is stochastically greatest is shown to be a mixture of point mass at zero and a chi-square distribution with a single degree of freedom. Asymptotic and valid finitesample testing procedures are proposed and examined via a simulation study and the analysis of two data sets. The proposed procedures are extended to test whether the cell with the largest observed frequency is uniquely most probable. An equivalence between a likelihood ratio test for this problem and a union-intersection test is demonstrated.

Keywords

chi-bar-square distribution, cone, discrete-choice experiments, intersection-union test, multiple comparisons with the best, order restricted inference, preference testing, selection, union-intersection test

Disciplines

Statistics and Probability

Comments

This preprint was published as Dan Nettleton, "Testing for the Supremacy of a Multinomial Cell Probability", *Journal of the American Statistical Association* (2009): 1052-1059, doi: [10.1198/jasa.2009.tm08213](https://doi.org/10.1198/jasa.2009.tm08213).

Testing for the Supremacy of a Multinomial Cell Probability

Dan Nettleton

Department of Statistics
Iowa State University
Ames, IA 50011-1210
(*dnett@iastate.edu*)

January 14, 2009

Abstract

Tests for the supremacy of a multinomial cell probability are developed. The tested null hypothesis states that a particular cell of interest is not more probable than all others. Rejection of this null leads to the conclusion that the cell of interest has a strictly greater probability than all other cells. The null hypothesis constrains the multinomial probability vector to a non-convex region that is a union of closed convex cones. The likelihood ratio test for this problem is derived and shown to be equivalent to an intersection-union test. The least favorable configuration of the multinomial probability vector in the null parameter space is derived, and the limiting null distribution of the test statistic that is stochastically greatest is shown to be a mixture of point mass at zero and a chi-square distribution with a single degree of freedom. Asymptotic and valid finite-sample testing procedures are proposed and examined via a simulation study and the analysis of two data sets. The proposed procedures are extended to test whether the cell with the largest observed frequency is uniquely most probable. An equivalence between a likelihood ratio test for this problem and a union-intersection test is demonstrated.

Key words: chi-bar-square distribution, cone, discrete-choice experiments, intersection-union test,

multiple comparisons with the best, order restricted inference, preference testing, selection, union-intersection test.

1. Introduction

In experiments or surveys with a multinomial response, it is often desirable to know whether a specific cell probability is larger than all others. For example, a candidate for elected office will wish to know whether he or she is leading a race given survey data where each respondent chooses one of multiple candidates. A company that has developed a new or improved version of a product will wish to know whether their product is the favorite of consumers based on a preference test in which consumers choose one favorite from multiple competing products. An exam evaluator may wish to know whether the correct answer to a multiple-choice question is selected more often than any other based on a sample of exam results. These and other scenarios involve situations where – prior to data collection – there is interest in knowing whether one particular multinomial cell probability is larger than all others. This paper presents tests that can be used to directly address this question.

Our main results are closely related to the order restricted inference literature on estimation and testing of a multinomial probability vector \mathbf{p} . Chacko (1966) derived the maximum likelihood estimator of \mathbf{p} subject to a simple order restriction (e.g., $p_1 \leq \dots \leq p_k$) and obtained the asymptotic null distribution of a chi-square-type statistic for testing $p_1 = \dots = p_k$ against the simple order. Robertson (1978) generalized the work of Chacko (1966) to include estimation of \mathbf{p} subject to the restriction $\mathbf{p} \in \mathcal{C}$, where \mathcal{C} is a closed convex cone in \mathbb{R}^k . For a specified multinomial probability vector $\mathbf{q} \in \mathcal{C}$, Robertson (1978) also established the likelihood ratio tests of $\mathbf{p} = \mathbf{q}$ versus $\mathbf{p} \in \mathcal{C} \setminus \{\mathbf{q}\}$ and $\mathbf{p} \in \mathcal{C}$ versus $\mathbf{p} \notin \mathcal{C}$. As we will illustrate, the likelihood ratio test for our problem

is similar to the latter test of Robertson (1978) except that our null hypothesis does not constrain p to a closed convex cone but rather to a finite union of closed convex cones that is not convex.

When testing a null hypothesis that constrains the parameter to a union of sets, an Intersection-Union Test (IUT) is a natural approach. The concept of an IUT – named by Gleser (1973) – can be traced back to Lehmann (1952). Berger (1982) proposed the use of IUTs for acceptance sampling problems, and Berger and Hsu (1996) studied IUTs in the context of bioequivalence trials. Useful introductions to IUTs and more extensive references to the relevant literature are included in Berger (1997) and Sections 5.3 and 9.2 of Silvapulle and Sen (2005). We will demonstrate that the likelihood ratio test for our problem is an IUT. Furthermore, we will present other IUTs that could be used as alternatives to the likelihood ratio test.

Our proposed procedures can be extended to test whether the cell with the largest observed frequency is uniquely most probable. For this test, it is not necessary to specify the cell of interest prior to observing the data. Rather, the cell with the largest frequency is automatically tested for superiority over all other cell probabilities. This test can be viewed as a multinomial analog of the “test for the best” procedures that follow from Hsu’s (1981, 1984) pioneering work on simultaneous confidence intervals for multiple comparisons with the “best” of k location parameters estimated from k independent samples. Our work is very similar to Berry (2001) who developed a likelihood ratio test for the existence of a uniquely most probable multinomial cell. We develop a union-intersection test (UIT) that is equivalent to the likelihood ratio test and can be used to determine whether the cell with the largest observed frequency is uniquely most probable.

Our paper is organized as follows. In Section 2, we use ideas from order restricted inference to determine a closed-form expression for the likelihood ratio test statistic. Furthermore, we derive the asymptotic null distribution of the likelihood ratio test statistic and establish an appropriate

reference null distribution by determining the least favorable limiting null distribution, i.e., the distribution that is stochastically largest among all possible limiting null distributions. In Section 3, we show that the likelihood ratio test is an IUT and propose other IUTs for testing our null hypothesis. We evaluate the performance of our proposed tests through simulation in Section 4 and illustrate their use on example data sets in Section 5. In Section 6, we consider the related problem of testing whether the cell with the largest observed frequency is uniquely most probable and discuss the relationships between our proposed approach and existing work. The paper concludes with a brief summary.

2. The Likelihood Ratio Test for the Supremacy of a Multinomial Cell Probability

Suppose $\mathbf{X} = (X_1, \dots, X_k)'$ has a multinomial distribution with n trials and cell probabilities p_1, \dots, p_k . The parameter $\mathbf{p} = (p_1, \dots, p_k)'$ lies in the set \mathcal{P} of vectors of order k whose components are nonnegative and sum to one. We wish to test

$$H_0 : p_k \leq \max\{p_1, \dots, p_{k-1}\} \quad \text{vs.} \quad H_A : p_k > \max\{p_1, \dots, p_{k-1}\}. \quad (1)$$

The null and alternative portions of the parameter space \mathcal{P} are depicted in Figure 1 for the case of $k = 3$. Henceforth, it will be convenient to stretch our notation slightly and use H_0 to denote both the null hypothesis and the set of vectors in \mathcal{P} that satisfy the null hypothesis.

Let $\hat{p}_j = X_j/n$ so that $\hat{\mathbf{p}} \equiv (\hat{p}_1, \dots, \hat{p}_k)'$ denotes the unrestricted maximizer of the multinomial likelihood function $L(\mathbf{p}) = \frac{n!}{X_1! \dots X_k!} \prod_{j=1}^k p_j^{X_j}$. To establish the likelihood ratio test of H_0 vs. H_A , we seek a maximizer of $L(\mathbf{p})$ over $\mathbf{p} \in H_0$ or, equivalently, a maximizer of

$$\log \lambda(\mathbf{p}) \equiv \sum_{j=1}^k X_j \log(p_j / \hat{p}_j)$$

over $\mathbf{p} \in H_0$, where 0 times any logarithm is taken as 0 in the definition of $\log \lambda(\mathbf{p})$ and $\log \lambda(\mathbf{p}) = -\infty$ for $\mathbf{p} \notin \mathcal{P}$.

Note that H_0 may be written as

$$H_0 : \mathbf{p} \in \bigcup_{j=1}^{k-1} \mathcal{C}_j \text{ where } \mathcal{C}_j \equiv \{\mathbf{p} : p_k \leq p_j\} \text{ for } j = 1, \dots, k-1. \quad (2)$$

Each \mathcal{C}_j is a closed convex cone corresponding to what Robertson, Wright, and Dykstra (1988) refer to as a partial order. However, the union of these sets is neither a cone nor convex, and the resulting order restriction on \mathbf{p} is of a different character than those typically considered in order restricted inference. Nonetheless, results from the order restricted inference literature are useful for addressing the problem. It follows from Example 1.5.7 of Robertson *et al.* (1988) that, for each $j = 1, \dots, k-1$, the multinomial likelihood is maximized over \mathcal{C}_j by $\hat{\mathbf{p}}^{(j)} = (\hat{p}_1^{(j)}, \dots, \hat{p}_k^{(j)})'$, where, for $i = 1, \dots, k$,

$$\hat{p}_i^{(j)} = \begin{cases} \hat{p}_i & \text{if } i \notin \{j, k\} \text{ or if } \hat{p}_k \leq \hat{p}_j, \\ (\hat{p}_j + \hat{p}_k)/2 & \text{otherwise.} \end{cases}$$

Thus, $\hat{\mathbf{p}}^{(j)}$ is simply the unrestricted maximum likelihood estimator if $\hat{p}_k \leq \hat{p}_j$, or the unrestricted maximum likelihood estimator with \hat{p}_j and \hat{p}_k each replaced by the average $(\hat{p}_j + \hat{p}_k)/2$ when $\hat{p}_k > \hat{p}_j$. For example, if $X = (2, 4, 1, 3)'$, then $\hat{p}^{(2)} = (.2, .4, .1, .3)'$ and $\hat{p}^{(3)} = (.2, .4, .2, .2)'$. It follows that the multinomial likelihood or, equivalently, $\lambda(\mathbf{p})$ is maximized over $\mathbf{p} \in \bigcup_{j=1}^{k-1} \mathcal{C}_j$ by some vector in the set $\{\hat{\mathbf{p}}^{(1)}, \dots, \hat{\mathbf{p}}^{(k-1)}\}$. Now note that

$$\log \lambda(\hat{\mathbf{p}}^{(j)}) = \begin{cases} 0 & \text{if } \hat{p}_k \leq \hat{p}_j, \\ X_j \log \left(\frac{X_j + X_k}{2X_j} \right) + X_k \log \left(\frac{X_j + X_k}{2X_k} \right) & \text{otherwise.} \end{cases} \quad (3)$$

Straightforward differentiation shows that $x \log \left(\frac{x+y}{2x} \right) + y \log \left(\frac{x+y}{2y} \right)$ is an increasing function of x for $0 \leq x < y$. Thus, $\hat{\mathbf{p}}^{(j)}$ maximizes the multinomial likelihood subject to the restriction imposed

by H_0 whenever j is such that $X_j = M \equiv \max\{X_1, \dots, X_{k-1}\}$. We will use $\tilde{\mathbf{p}}$ to denote any such maximizing vector. Figure 1 depicts an example value for $\hat{\mathbf{p}}$ and the corresponding value of $\tilde{\mathbf{p}}$. Note that, because of the possibility of ties, the maximum likelihood estimator of \mathbf{p} under H_0 is not necessarily unique, but the value of the likelihood ratio test statistic ($-2 \log$ likelihood) is unique and is given by

$$T_n = \begin{cases} 0 & \text{if } X_k \leq M, \\ 2 \left[M \log \left(\frac{2M}{M+X_k} \right) + X_k \log \left(\frac{2X_k}{M+X_k} \right) \right] & \text{if } X_k > M. \end{cases} \quad (4)$$

2.1. The Asymptotic Null Distribution of the Likelihood Ratio Test Statistic

We now derive the asymptotic distribution of T_n for \mathbf{p} satisfying H_0 . First note that if $p_k < p_j$ for some $j = 1, \dots, k-1$, then T_n converges to 0 almost surely as n goes to infinity because of the almost sure convergence of $\hat{\mathbf{p}}$ to \mathbf{p} . The more interesting case occurs when \mathbf{p} is on the boundary of H_0 , i.e., when $p_k = \max\{p_1, \dots, p_{k-1}\}$ which we will denote by $\mathbf{p} \in \partial H_0$. As in Section 2 of Berry (2001) and in the proofs of Theorems 5.2.1 and 5.2.2 in Robertson, Wright, and Dykstra (1988), a second-order Taylor series expansion of the logarithm function yields

$$\begin{aligned} & 2 \left[M \log \left(\frac{2M}{M+X_k} \right) + X_k \log \left(\frac{2X_k}{M+X_k} \right) \right] \\ &= 2M \left[\frac{\left(\frac{M}{n} - \frac{X_k}{n} \right) / 2}{\frac{M}{n}} + \frac{\left\{ \left(\frac{M}{n} - \frac{X_k}{n} \right) / 2 \right\}^2}{2a_n^2} \right] + 2X_k \left[\frac{\left(\frac{X_k}{n} - \frac{M}{n} \right) / 2}{\frac{X_k}{n}} + \frac{\left\{ \left(\frac{X_k}{n} - \frac{M}{n} \right) / 2 \right\}^2}{2b_n^2} \right] \\ &= n \left(\frac{M}{n} - \frac{X_k}{n} \right)^2 \left(\frac{M}{na_n^2} + \frac{X_k}{nb_n^2} \right) / 4, \end{aligned} \quad (5)$$

where a_n is between M/n and $(M+X_k)/(2n)$ and b_n is between $(M+X_k)/(2n)$ and X_k/n .

For $\mathbf{p} \in \partial H_0$, $\hat{\mathbf{p}} \xrightarrow{a.s.} \mathbf{p}$ implies that M/n , X_k/n , a_n , and b_n each converge almost surely to p_k . It follows that (5) has the same limiting distribution as $n(M/n - X_k/n)^2 / (2p_k)$ for $\mathbf{p} \in \partial H_0$. For a

geometric interpretation of this random variable, note that $(M/n - X_k/n)^2 / 2 = \|\hat{\mathbf{p}} - \tilde{\mathbf{p}}\|^2$, which is, for example, the squared length of the dashed line segment in Figure 1.

Now note that

$$\sqrt{n} \left(\frac{M}{n} - \frac{X_k}{n} \right) / \sqrt{2p_k} = \max \left\{ \sqrt{n} (\hat{p}_j - \hat{p}_k) / \sqrt{2p_k} : j = 1, \dots, k-1 \right\}. \quad (6)$$

Because $\hat{\mathbf{p}}$ converges to \mathbf{p} almost surely, (6) has the same limiting distribution as

$$\max \left\{ \sqrt{n} (\hat{p}_j - \hat{p}_k) / \sqrt{2p_k} : j < k, p_j = p_k \right\} \quad (7)$$

for $\mathbf{p} \in \partial H_0$. Because $\sqrt{n}(\hat{\mathbf{p}} - \mathbf{p})$ converges in distribution to $N(\mathbf{0}, \text{diag}(\mathbf{p}) - \mathbf{p}\mathbf{p}')$, (7) converges in distribution to $M_Z \equiv \max\{Z_1, \dots, Z_h\}$, where h is the number of $j < k$ such that $p_j = p_k$ and $\mathbf{Z} \equiv (Z_1, \dots, Z_h)' \sim N_h(\mathbf{0}, \frac{1}{2}I + \frac{1}{2}J)$ with I the $h \times h$ identity matrix and J the $h \times h$ matrix of ones. This result combined with the definition of T_n in (4) yields the following theorem.

Theorem 1. *For $\mathbf{p} \in \partial H_0$, the likelihood ratio test statistic T_n converges in distribution to the random variable $M_Z^2 \mathbb{1}(M_Z < 0)$, where M_Z is the largest component of a $N_h(\mathbf{0}, \frac{1}{2}I + \frac{1}{2}J)$ vector and $\mathbb{1}(\cdot)$ denotes the indicator function.*

2.2. The Least Favorable Asymptotic Null Distribution of the Likelihood Ratio Test Statistic

The asymptotic distribution of T_n under the null hypothesis is not fixed but instead will depend on the true value of $\mathbf{p} \in H_0$. To establish an approximate significance level α likelihood ratio test, we seek a critical value t_α such that

$$\sup_{\mathbf{p} \in H_0} \lim_{n \rightarrow \infty} P_{\mathbf{p}}(T_n \geq t_\alpha) \leq \alpha,$$

where $P_{\mathbf{p}}(\cdot)$ denotes the probability distribution when the true parameter is \mathbf{p} . As noted in the

previous section, T_n converges almost surely to 0 for $\mathbf{p} \in H_0 \setminus \partial H_0$. Thus, it is sufficient to consider the supremum over $\mathbf{p} \in \partial H_0$. By Theorem 1, the asymptotic distribution T_n for $\mathbf{p} \in \partial H_0$ depends only on h , the number of $j < k$ such that $p_j = p_k$. For any $t \geq 0$,

$$P\{M_Z^2 \mathbb{1}(M_Z < 0) \geq t\} = P(M_Z \leq -\sqrt{t}) \leq \Phi(-\sqrt{t}), \quad (8)$$

where $\Phi(\cdot)$ denotes the standard normal cumulative distribution function. For $h = 1$, $M_Z \sim N(0, 1)$ so that the upper bound in (8) is attained. Thus, we have established that the least favorable asymptotic null distribution is that of the random variable $M_Z^2 \mathbb{1}(M_Z < 0)$ when $h = 1$. The distribution of this random variable is easily seen to be an equal mixture of unit mass at zero and a χ^2 distribution with 1 degree of freedom. This distribution is known as a chi-bar-square distribution in order restricted inference literature (see, for example, Robertson *et al.* 1988). We have established the following theorem.

Theorem 2. *For any $t \in \mathbb{R}$,*

$$\sup_{\mathbf{p} \in H_0} \lim_{n \rightarrow \infty} P_{\mathbf{p}}(T_n \geq t) = \frac{1}{2} \mathbb{1}(t \leq 0) + \frac{1}{2} P(\chi_1^2 \geq t),$$

where χ_1^2 is a χ^2 random variable with 1 degree of freedom.

The following corollary is a simple consequence of Theorem 2 and provides a critical value for an approximate significance level α test of H_0 vs. H_A .

Corollary. *For $\alpha \in (0, 1/2)$,*

$$\sup_{\mathbf{p} \in H_0} \lim_{n \rightarrow \infty} P_{\mathbf{p}}\{T_n \geq \chi_1^2(1 - 2\alpha)\} = \alpha,$$

where $\chi_1^2(1 - 2\alpha)$ denotes the $1 - 2\alpha$ quantile of a χ^2 distribution with 1 degree of freedom.

Thus, for sufficiently large n , an approximate significance level α likelihood ratio test of H_0 vs. H_A rejects H_0 if and only if $T_n \geq \chi_1^2(1 - 2\alpha)$. The approximate p -value for the test is given by $P(\chi_1^2 \geq T_n | T_n)/2$, where χ_1^2 denotes a χ^2 random variable with 1 degree of freedom that is independent of T_n .

3. Intersection Union Tests for the Supremacy of a Multinomial Cell Probability

Berger (1997) recommends the IUT method when a null parameter space is naturally expressed as a union, as in our expression (2). To establish an IUT for our problem, we simply test $H_{0j} : \mathbf{p} \in \mathcal{C}_j$ vs. $H_{Aj} : \mathbf{p} \notin \mathcal{C}_j$ for each $j = 1, \dots, k - 1$. We reject H_0 at level α if and only if H_{0j} is rejected at level α for all $j = 1, \dots, k - 1$. This approach provides a level- α test of H_0 by Theorem 1 of Berger (1982), Theorem 15.1.1 in Berger (1997), or Proposition 5.3.1 in Silvapulle and Sen (2005).

As a natural starting point, consider an IUT based on asymptotic size- α likelihood ratio tests of H_{0j} for $j = 1, \dots, k - 1$. By the work in Section 2, the likelihood ratio statistic for testing H_{0j} is given by $L_j \equiv -2 \log \lambda(\hat{\mathbf{p}}^{(j)})$, where $\lambda(\hat{\mathbf{p}}^{(j)})$ is as defined in (3) for $j = 1, \dots, k - 1$. Furthermore, for any $j = 1, \dots, k - 1$; it is straightforward to show that the asymptotic size- α test of H_{0j} rejects H_{0j} if and only if $L_j \geq \chi_1^2(1 - 2\alpha)$. Because $x \log \left(\frac{x+y}{2x} \right) + y \log \left(\frac{x+y}{2y} \right)$ is an increasing function of x for $0 \leq x < y$, the smallest of these statistics is obtained for any j such that $X_j = M \equiv \max\{X_1, \dots, X_{k-1}\}$. It follows that $\min\{L_j : j = 1, \dots, k - 1\} = T_n$ and that the asymptotic size- α IUT for testing H_0 that is based on likelihood ratio tests of H_{0j} for $j = 1, \dots, k - 1$ rejects H_0 if and only if $T_n \geq \chi_1^2(1 - 2\alpha)$. Hence, the IUT is the same as the likelihood ratio test established in Section 2.

Note that for testing problems in which the sample size is fixed, Theorem 15.2.1 of Berger

(1997) provides sufficient conditions for the size- α likelihood ratio test to be the same as the IUT based on individual size- α likelihood ratio tests. Although Berger's theorem is not directly applicable in our case, it is easy to establish an asymptotic version of the theorem that covers our special case by making simple adjustments to the proofs in Berger (1982) and Berger (1997).

Rather than using likelihood ratio tests to form an IUT, we can construct an IUT of H_0 using alternative procedures for testing H_{0j} for $j = 1, \dots, k - 1$. For example, score tests or Wald tests could be used to test each H_{0j} . For each $j = 1, \dots, k - 1$, the score and Wald test statistics are

$$S_j = \frac{n\{(\hat{p}_k - \hat{p}_j)_+\}^2}{\hat{p}_k + \hat{p}_j} \quad \text{and} \quad W_j = \frac{n\{(\hat{p}_k - \hat{p}_j)_+\}^2}{\hat{p}_k + \hat{p}_j - (\hat{p}_k - \hat{p}_j)^2},$$

respectively, where $(x)_+ = x$ if $x > 0$ and 0 otherwise. The asymptotic size- α likelihood ratio, score, and Wald tests each reject H_{0j} if and only if the test statistic (L_j , S_j , or W_j) exceeds $\chi^2(1 - 2\alpha)$. The tests are asymptotically equivalent when $p_j = p_k$ but are not equivalent in general. The question of which test to use is similar to the issue discussed by Robbins (1977) and Eberhardt and Fligner (1977) for comparing the success probabilities of two independent binomial random variables. Note that $S_j \leq W_j$ with strict inequality whenever $\hat{p}_j < \hat{p}_k$. Thus, the Wald test is more powerful than the score test. We investigate the finite-sample performance of these tests through simulations presented in Section 4.

In addition to the asymptotic IUTs for testing H_0 , we consider a finite-sample level- α IUT based on conditional tests of H_{0j} for each $j = 1, \dots, k - 1$. Note that when $\mathbf{p} \in \partial H_{0j} = \{\mathbf{p} : p_j = p_k\}$, the conditional distribution of X_k , given $X_j + X_k = m_j$, is Binomial($m_j, 1/2$). Thus, a p -value for testing H_{0j} that is valid for all n is $P(B_j \geq x_k)$, where $B_j \sim \text{Binomial}(m_j, 1/2)$ and x_k denotes the observed value of X_k . Thus, the corresponding IUT rejects H_0 at level α if and only if $\max\{P(B_j \geq x_k) : j = 1, \dots, k - 1\} \leq \alpha$ or, equivalently, if and only if $P(B \geq x_k) \leq \alpha$

where $B \sim \text{Binomial}(m, 1/2)$ with $m = \max\{x_1, \dots, x_{k-1}\} + x_k$.

4. A Simulation Study

We conducted a simulation study to evaluate the finite-sample performance of our proposed testing procedures. We considered significance level $\alpha = 0.05$ tests with $n = 50, 200,$ and 1000 for various choices of \mathbf{p} in the null (Table 1) and the alternative (Table 2) parameter spaces. For each simulation setting, 100,000 replications were used to estimate the type I error rates and powers in Tables 1 and 2, respectively. Thus, each standard error is bounded above by 0.0016. For symmetry of presentation, we used $k = 5$, but note that in many cases some components of the vector \mathbf{p} are 0, which is equivalent to considering choices of $k < 5$.

The results in Table 1 show that the type I error rates of the proposed testing procedures are at or below nominal levels for most of the scenarios studied. The largest type I error rate relative to the nominal level occurred for the case of $\mathbf{p} = (0, 0, 0, 1/2, 1/2)'$ and $n = 50$, where the estimated type I error rate was 0.059 for the likelihood ratio test and the asymptotic IUTs. Although this is significantly above the nominal level based on 100,000 simulation replications, it is not alarmingly high from a practical standpoint. Note that the actual type I error rates appear to be substantially lower than the nominal levels when more than two cells are tied for the highest probability ($h \geq 2$). This is consistent with the results of Section 2 which show that such vectors \mathbf{p} are not least favorable in H_0 . As predicted by the analytical work in Section 2.2, type I error rates are closer to nominal (especially for larger sample sizes) for least favorable vectors \mathbf{p} with $h = 1$. The type I error rates for the IUT based on conditional tests are guaranteed to be bounded above by 0.05, and in many cases, the simulation results suggest that the actual type I error rates are substantially less than 0.05 and less than the type I error rates of the asymptotic procedures.

The results in Table 2 illustrate the power characteristics of the proposed procedures. Power increases with sample size n and the distance of \boldsymbol{p} from H_0 , which is given by the discrepancy between p_k and M . The asymptotic procedures tend to be more powerful than the IUT based on conditional tests. The IUT based on Wald tests seems to hold a very slight power advantage over the IUT based on likelihood ratio tests or score tests, though the differences among these asymptotic procedures appears practically inconsequential.

5. Example Data Analysis

In this section, we present a brief example analysis for each of two data sets. The first data set served as the motivation for this work. The second is simply readily available data that shows the utility of the method for examining a scientific question regarding voter behavior. As noted in the Introduction, the proposed method has application in many other areas.

5.1. An Investigation of Parent-Child Resemblance

To objectively investigate commonly made claims of striking resemblance between parents and infant children, Nettleton and Froelich (2008) conducted a study to determine whether the resemblance between a mother and her infant daughter was stronger than the resemblance between the mother and each of three unrelated babies. Study participants were presented with a photograph of the mother along with the photographs of four babies including the daughter of the mother. Participants were asked to guess which of the four babies was the daughter.

The multinomial response vector was $\boldsymbol{X} = (12, 14, 43, 54)'$, where the fourth cell represents the number of correct guesses among the total of 123 responses. If participants were guessing randomly in such a way that each baby would be equally likely to be selected, we would ex-

pect approximately 30.75 subjects to select each baby. The observed vector deviates significantly from that expectation. More specifically, the probability of 54 or more correct guesses under the “equally likely” model is about 0.0000038. This provides some evidence of resemblance between mother and child. However, to claim that the mother looks more like her own child than any of the other three babies, we should test H_0 vs. H_A and find in favor of the alternative. Based on $\mathbf{X} = (12, 14, 43, 54)'$, the p -values for our asymptotic tests are each approximately 0.13, and the p -value for the IUT based on conditional tests is approximately 0.15. These results provide only very weak evidence in favor of the alternative hypothesis.

5.2. An Investigation of the Effect of Candidate Ballot Order on Election Results

Brook and Upton (1974) studied the effect of candidate ballot order on the results of local government elections in England and Wales. They considered a sample of 948 combinations of political party and ward in which exactly three candidates within a single party were running for election. Within a particular such party/ward combination, the results of voting were summarized by a permutation of the letters a , b , and c . In their notation, the permutation abc denotes the case where the first candidate listed on the ballot received the most votes, followed by the second candidate on the ballot, followed by the third candidate. Similarly, bca denotes the case where the second candidate on the ballot received the most votes and the first candidate on ballot the least. Other permutations are defined analogously. The counts for permutations abc , acb , bac , bca , cab , and cba were 232, 136, 174, 151, 114, and 141, respectively.

Brook and Upton hypothesized that the permutation abc – in which the final order of votes matches the order of the names on the ballot – may be more likely than any other when many voters have little information to distinguish candidates within their preferred party. We can test for

evidence in favor of this hypothesis by considering the counts as a multinomial response vector and conducting a test of H_0 vs. H_A with cell abc corresponding to $k = 6$. The p -values for all four of our testing procedures were 0.002 when rounded to the third decimal place. This provides strong evidence in favor of the hypothesis that the final order of votes is more likely to match the ballot order than any other.

6. Testing whether the cell with the largest observed frequency is uniquely most probable

This paper has focused on testing whether one particular multinomial cell – chosen prior to observing \mathbf{X} – is more probable than all others. In some cases, special interest in a particular multinomial cell may arise only after observing that it is most frequent in the multinomial vector \mathbf{X} . For example, suppose that there were no *a priori* interest in permutation abc in the voter data discussed in Section 5.2. After seeing the election results for the 948 party/ward combinations, it would be natural to wonder if permutation abc was more probable than all other permutations.

To address the question with an asymptotic testing procedure, we need only consider the case where multiple cells are tied for the largest probability. Without loss of generality, suppose $\mathbf{p} = (p_1, \dots, p_{k-h-1}, p, \dots, p)'$ for some $h > 0$ where $p_j < p$ for all $j = 1, \dots, k - h - 1$. For all $j = 1, \dots, k$, let $T_n^{(j)}$ denote the statistic T_n in (4) obtained by exchanging labels on the j^{th} and k^{th} cells. Define $T_n^* = \max\{T_n^{(j)} : j = 1, \dots, k\}$ so that T_n^* is simply the value of T_n that is obtained

by defining the cell with the largest observed frequency to be the k^{th} . It follows that

$$\begin{aligned}
\lim_{n \rightarrow \infty} P_{\mathbf{p}}(T_n^* \geq t) &= \lim_{n \rightarrow \infty} \sum_{j=k-h}^k P_{\mathbf{p}}(T_n^* \geq t | T_n^* = T_n^{(j)}) P_{\mathbf{p}}(T_n^* = T_n^{(j)}) \\
&= \lim_{n \rightarrow \infty} \sum_{j=k-h}^k P_{\mathbf{p}}(T_n^* \geq t | T_n^* = T_n^{(k)}) \frac{1}{h+1} \\
&= \lim_{n \rightarrow \infty} P_{\mathbf{p}}(T_n^* \geq t | T_n^* = T_n^{(k)}) \\
&= \lim_{n \rightarrow \infty} P_{\mathbf{p}}(T_n \geq t | \hat{p}_k > \hat{p}_j \forall j = k-h, \dots, k-1) \\
&= P(M_Z^2 \geq t | M_Z < 0),
\end{aligned} \tag{9}$$

where the last equality follows from the same argument used to establish Theorem 1.

Now for any $t > 0$,

$$P(M_Z^2 \geq t | M_Z < 0) \leq P(Z^2 \geq t | Z < 0) = P(Z^2 \geq t) = P(\chi_1^2 \geq t), \tag{10}$$

where $Z \sim N(0, 1)$ and χ_1^2 is a χ^2 random variable with 1 degree of freedom. When $h = 1$, $M_Z \sim N(0, 1)$ and the upper bound in (10) is obtained. Combining this result with (9) yields the following theorem.

Theorem 3. *Suppose $h + 1$ elements of the vector \mathbf{p} are equal to one another and strictly greater than all other elements of \mathbf{p} for some $h > 0$. Then for all $\alpha \in (0, 1)$,*

$$\lim_{n \rightarrow \infty} P_{\mathbf{p}}\{T_n^* \geq \chi_1^2(1 - \alpha)\} \leq \alpha,$$

where $\chi_1^2(1 - \alpha)$ denotes the $1 - \alpha$ quantile of a χ^2 distribution with 1 degree of freedom. Furthermore, equality holds if $h = 1$.

It follows from Theorem 3 that an asymptotic α -level test can be obtained by concluding that the cell with the largest observed frequency is uniquely most probable if and only if $T_n^* \geq$

$\chi_1^2(1 - \alpha)$. If we use this procedure to test whether the cell associated with permutation abc is uniquely most probable for the voter data, we obtain a p -value of approximately 0.004, which is simply twice the p -value obtained in Section 5 where our interest in testing cell abc was assumed to exist prior to observing \mathbf{X} .

6.1. Relationship to the likelihood ratio test for the existence of a uniquely most probable cell

As remarked in the Introduction, our results are very closely related to the work of Berry (2001). Berry considered a likelihood ratio test for the existence of a uniquely most probable cell. Berry showed that the asymptotic rejection region of his likelihood ratio test is (using our notation) of the form $(h + 1)P(Z_1 \geq \sqrt{t}, \dots, Z_h \geq \sqrt{t})$. Our test statistic T_n^* is easily seen to be identical to Berry's test statistic given in equation (3) of Berry (2001). Furthermore, our result in equation (9) can be rewritten as

$$\begin{aligned}
\lim_{n \rightarrow \infty} P_{\mathbf{p}}(T_n^* \geq t) &= P(M_Z^2 \geq t | M_Z < 0) \\
&= P(M_Z^2 \geq t, M_Z < 0) / P(M_Z < 0) \\
&= P(Z_1 \leq -\sqrt{t}, \dots, Z_h \leq -\sqrt{t}) / P(Z_1 < 0, \dots, Z_h < 0) \\
&= (h + 1)P(Z_1 \geq \sqrt{t}, \dots, Z_h \geq \sqrt{t}),
\end{aligned}$$

where the last equality can be proved by noting that

$$\begin{aligned}
P(Z_1 < 0, \dots, Z_h < 0) &= P(Y_1 - Y_{h+1} < 0, \dots, Y_h - Y_{h+1} < 0) \\
&= P(Y_1 < Y_{h+1}, \dots, Y_h < Y_{h+1}) \\
&= P(Y_{h+1} = \max\{Y_1, \dots, Y_{h+1}\}) = \frac{1}{h + 1},
\end{aligned}$$

where Y_1, \dots, Y_{h+1} are independent and identically distributed normal random variables with mean

0 and variance 1/2. Thus, the main results of Berry (2001) are equivalent to results that we obtained in Section 6.

Berry claimed that rejection of his null hypothesis implies the existence of a uniquely most probable cell whose identity is unknown. In contrast, we claim that the cell with the highest estimated probability may be declared uniquely most probable when T_n^* exceeds $\chi^2(1 - \alpha)$ and that the asymptotic probability of an incorrect selection will be no larger than α when applying this procedure. To see why it is safe to make this additional claim, simply note that the asymptotic probability of incorrect selection is 0 when a uniquely most probable cell exists.

6.2. Relationship between the Likelihood Ratio Test and a Union-Intersection Test

In this subsection, we show that our proposed testing procedure for identifying a uniquely most probable cell can be viewed as a Union-Intersection Test (UIT). UITs are closely related to IUTs and have a long history that dates back to at least Roy (1953). A contemporary review is provided by Silvapulle and Sen (2005). UITs are potentially useful when the null parameter space can be written as an intersection of component null parameter spaces whose corresponding component null hypotheses are relatively straightforward to test individually. Because the null corresponding to the intersection must be false if any component null hypothesis is false, the rejection region of an UIT is the union of rejection regions for tests of the component null hypotheses.

In our case, note that the null hypothesis of no uniquely most probable cell can be written as

$$H_0 : \mathbf{p} \in \bigcap_{j=1}^k \mathcal{U}_j, \text{ where } \mathcal{U}_j \equiv \bigcup_{j' \neq j} \{\mathbf{p} : p_j \leq p_{j'}\} \text{ for } j = 1, \dots, k. \quad (11)$$

Consider the component null hypotheses $H_0^{(j)} : \mathbf{p} \in \mathcal{U}_j$, defined for $j = 1, \dots, k$. Note that $H_0^{(k)}$ is precisely the null hypothesis (2). Thus, the work in Sections 2 and 3 of this paper provides like-

likelihood ratio tests and IUTs, respectively, that can be used to test each component null hypothesis. In particular, if we choose the likelihood ratio test or, equivalently, the IUT based on likelihood ratio tests to test $H_0^{(j)}$, we reject $H_0^{(j)}$ for large values of $T_n^{(j)}$. Thus, the corresponding UIT rejects H_0 in (11) if any one of the statistics $T_n^{(1)}, \dots, T_n^{(k)}$ is sufficiently large or, equivalently, if $\max\{T_n^{(j)} : j = 1, \dots, k\} = T_n^*$ is sufficiently large. Thus, the test that we proposed at the beginning of Section 6 is a UIT. More specifically, it is a UIT for which each component null hypothesis is tested with an IUT.

6.3. Relationship to the classic and modern selection literature

The problems that we have considered are similar to problems studied in the classic selection literature (see, for example, Gupta and Panchapakesan 1979). Two primary formulations of the selection problem have been investigated extensively. In the *indifference zone* formulation, interest rests on determining the sample size necessary for the most probable multinomial cell to have the largest observed frequency with probability at least $1 - \alpha$, given that the largest multinomial cell probability is larger than the next largest by an additive (or multiplicative) factor δ . This formulation is not so relevant for our problem because we must be concerned with the situation in which multiple cells are tied for the largest cell probability.

The second major formulation of the selection problem is known as the *probability of correct selection* formulation. The goal is to determine a set of cells that contains the most probable cell with probability at least $1 - \alpha$. In principle, this type of selection procedure could be used to declare that the cell with the largest observed frequency is uniquely most probable if the selected set contains only one cell. Unfortunately, if multiple cells are tied for the largest probability, it is typically assumed that one of the top cells is arbitrarily designated the “best” so that selection of

only that “best” cell is considered correct even though other cells may be equally probable (Gupta and Panchapakesan 1979, pp. 10–11). Converting such a selection procedure into a test could result in a testing procedure with a higher than nominal type I error rate when multiple cells are tied for the largest probability.

More recent work on selection by Ng and Panchapakesan (2007) provides a method for testing whether the cell first to reach a specified count in sequential multinomial sampling has a larger success probability than all others. More specifically, Ng and Panchapakesan consider sequentially drawing independent observations, each distributed as $\text{Multinomial}(1; p_1, \dots, p_k)$, until the total count in any cell first reaches a pre-specified value Q . Using the difference between Q and the next largest observed cell count as a statistic for testing whether the first cell to reach Q is uniquely most probable, Ng and Panchapakesan (2007) show how to compute a p -value from a single integral involving the gamma distribution. The problem that we have addressed in Section 6 can be viewed as the fixed-sample-size analog of this sequential analysis problem.

Ng and Panchapakesan (2007) utilized some key results from Gutmann and Maymin (1987) who considered the problem of determining whether the population with the largest sample mean has the uniquely largest population mean in the context of independent samples from k normal populations. Our problem is the natural analog of this problem for the case of multinomial sampling.

Other recent work has considered selection of the “best treatment” in more general settings and with a focus on decision-theoretic properties of selection rules. Some examples include Abughalous and Bansal (1995) on selection of the one-parameter exponential family with the largest mean, Bansal and Gupta (1997) on selection of the best treatment in the normal general linear model, and Bansal and Miescke (2006) on selection in the framework of generalized linear

models. The focus of this research is on determining experimental designs for which selection rules can be obtained that are Bayes rules with respect to given loss functions and priors on model parameters. Although this work is general in scope, it is not directly applicable to our multinomial problem.

7. Conclusion

We have developed the likelihood ratio test and four IUTs for testing whether a specified multinomial cell probability is larger than all others. We have shown that the likelihood ratio test is equivalent to the IUT based on likelihood ratio tests of relatively simple component null hypotheses. We have also demonstrated that Berry's (2001) likelihood ratio test for the existence of a uniquely most probable cell is equivalent to a UIT whose component null hypotheses are tested with IUTs. Although most of our proposed testing procedures are asymptotic, we have presented an IUT based on conditional tests that is valid – regardless of sample size – for testing whether a specified cell probability is larger than all others. Our simulations suggest that the asymptotic procedures are preferred for moderate to large sample sizes, but the the IUT based on conditional tests is a useful option when a small sample size casts doubt on the validity of the asymptotic procedures.

Acknowledgements

The author wishes to thank an anonymous referee who suggested the examination of intersection-union tests for the problems considered in this article.

References

- Abughalous, M. M., and Bansal, N. K. (1995), "On selecting the best natural exponential families with quadratic variance functions," *Statistics and Probability Letters*, 25, 341–349.
- Bansal, N. K., and Gupta, S. (1997), "On the natural selection rule in general linear models," *Metrika*, 46, 59–69.
- Bansal, N. K., and Miescke, K. J. (2006), "On selecting the best treatment in a generalized linear model," *Journal of Statistical Planning and Inference*, 136, 2070–2086.
- Berger, R. L. (1982), "Multiparameter hypothesis testing and acceptance sampling," *Technometrics*, 24, 295–300.
- Berger, R. L. (1997), "Likelihood ratio tests and intersection union tests," in *Advances in Statistical Decision Theory and Applications*, eds. S. Panchapakesan and N. Balakrishnan, Boston: Birkhäuser, pp. 225–237.
- Berger, R. L., and Hsu, J. C. (1996), "Bioequivalence trials, intersection-union tests and equivalence confidence sets," *Statistical Science*, 11, 283–302.
- Berry, J. C. (2001), "On the existence of a unique most probable category," *Journal of Statistical Planning and Inference*, 99, 175–182.
- Brook, D., and Upton, G. J. G. (1974), "Biases in local government elections due to position on the ballot paper," *Applied Statistics*, 23, 414–419.
- Chacko, V. J. (1966), "Modified chi-square test for ordered alternatives," *Sankhya B*, 28, 185–190.
- Eberhardt, K. A., and Fligner, M. A. (1977), "A comparison of two tests for equality of two proportions," *The American Statistician*, 31, 151–155.

- Gleser, L. J. (1973), "On a theory of intersection-union tests (preliminary report)," *IMS Bulletin*, 2, 233.
- Gupta, S. S., and Panchapakesan, S. (1979), *Multiple Decision Procedures: Theory and Methodology of Selecting and Ranking Populations*, New York: Wiley.
- Gutmann, S., and Maymin, Z. (1987), "Is the selected population the best?" *Annals of Statistics*, 15, 456–461.
- Hsu, J. C. (1981), "Simultaneous confidence intervals for all distances from the best," *The Annals of Statistics*, 9, 1026–1034.
- Hsu, J. C. (1984), "Constrained simultaneous confidence intervals for multiple comparisons with the best," *The Annals of Statistics*, 12, 1136–1144.
- Lehmann, E. L. (1952), "Testing multiple hypotheses," *Annals of Mathematical Statistics*, 23, 541–552.
- Nettleton, D., and Froelich, A. G. (2008), "Does my baby really look like me? Testing for resemblance between parent and child," In preparation.
- Ng, H. K. T., and Panchapakesan, S. (2007), "Is the selected multinomial cell the best?" *Sequential Analysis*, 26, 415–423.
- Robbins, H. (1977), "A fundamental question of practical statistics," *The American Statistician*, 31, 97.
- Robertson, T. (1978), "Testing for and against an order restriction on multinomial parameters," *Journal of the American Statistical Association*, 73, 197–202.

Robertson, T., Wright, F. T., and Dykstra, R. L. (1988), *Order Restricted Statistical Inference*,
New York: Wiley.

Roy, S. N. (1953), "On a heuristic method of test construction and its use in multivariate analysis,"
Annals of Mathematical Statistics, 24, 220-238.

Silvapulle, M. J., and Sen, P. K. (2005), *Constrained Statistical Inference: Inequality, Order, and
Shape Restrictions*, Hoboken: Wiley.

Table 1. Type I error rates of the nominal 0.05-level likelihood ratio test (LRT) and intersection union tests based on likelihood ratio tests (LRT), Wald tests (IUT_{Wald}), score tests (IUT_{score}), and conditional tests ($IUT_{\text{cond.}}$) for various sample sizes (n) and null parameter configurations. Values are estimated from 100,000 replications; thus, each standard error is bounded above by 0.0016.

p_1	p_2	p_3	p_4	p_5	n	LRT	IUT_{Wald}	IUT_{score}	$IUT_{\text{cond.}}$
0	0	0	1/2	1/2	50	0.059	0.059	0.059	0.032
0	0	1/3	1/3	1/3		0.009	0.013	0.009	0.007
0	0	1/5	2/5	2/5		0.045	0.050	0.044	0.033
0	1/4	1/4	1/4	1/4		0.004	0.005	0.004	0.002
0	1/7	2/7	2/7	2/7		0.010	0.012	0.010	0.006
0	1/6	1/6	1/3	1/3		0.039	0.044	0.039	0.027
1/5	1/5	1/5	1/5	1/5		0.002	0.002	0.002	0.001
1/9	2/9	2/9	2/9	2/9		0.004	0.005	0.004	0.002
1/8	1/8	1/4	1/4	1/4		0.010	0.011	0.010	0.006
1/7	1/7	1/7	2/7	2/7		0.033	0.038	0.033	0.022
0	0	0	1/2	1/2	200	0.052	0.052	0.052	0.038
0	0	1/3	1/3	1/3		0.013	0.013	0.012	0.010
0	0	1/5	2/5	2/5		0.051	0.052	0.050	0.042
0	1/4	1/4	1/4	1/4		0.005	0.005	0.004	0.003
0	1/7	2/7	2/7	2/7		0.012	0.012	0.011	0.009
0	1/6	1/6	1/3	1/3		0.051	0.052	0.051	0.042
1/5	1/5	1/5	1/5	1/5		0.002	0.002	0.002	0.001
1/9	2/9	2/9	2/9	2/9		0.004	0.004	0.004	0.003
1/8	1/8	1/4	1/4	1/4		0.012	0.012	0.011	0.008
1/7	1/7	1/7	2/7	2/7		0.050	0.050	0.049	0.041
0	0	0	1/2	1/2	1000	0.046	0.053	0.046	0.046
0	0	1/3	1/3	1/3		0.012	0.012	0.012	0.011
0	0	1/5	2/5	2/5		0.050	0.051	0.050	0.047
0	1/4	1/4	1/4	1/4		0.005	0.005	0.005	0.004
0	1/7	2/7	2/7	2/7		0.012	0.012	0.012	0.011
0	1/6	1/6	1/3	1/3		0.050	0.050	0.050	0.046
1/5	1/5	1/5	1/5	1/5		0.002	0.002	0.002	0.002
1/9	2/9	2/9	2/9	2/9		0.005	0.005	0.005	0.004
1/8	1/8	1/4	1/4	1/4		0.012	0.012	0.012	0.010
1/7	1/7	1/7	2/7	2/7		0.050	0.051	0.050	0.046

Table 2. Powers of the 0.05-level likelihood ratio test (LRT) and intersection union tests based on likelihood ratio tests (LRT), Wald tests (IUT_{Wald}), score tests (IUT_{score}), and conditional tests ($IUT_{\text{cond.}}$) for various sample sizes (n) and alternative parameter configurations. Values are estimated from 100,000 replications; thus, each standard error is bounded above by 0.0016.

p_1	p_2	p_3	p_4	p_5	n	LRT	IUT_{Wald}	IUT_{score}	$IUT_{\text{cond.}}$
.00	.00	.00	.45	.55	50	0.199	0.199	0.199	0.128
.00	.00	.00	.35	.65		0.725	0.725	0.725	0.621
.00	.00	.30	.30	.40		0.085	0.102	0.085	0.070
.00	.00	.20	.20	.60		0.919	0.926	0.917	0.893
.00	.10	.10	.35	.45		0.196	0.206	0.194	0.158
.00	.10	.10	.25	.55		0.784	0.795	0.781	0.736
.00	.10	.10	.15	.65		0.996	0.997	0.996	0.994
.10	.10	.10	.30	.40		0.209	0.223	0.209	0.168
.10	.10	.10	.20	.50		0.827	0.841	0.827	0.783
.10	.10	.10	.10	.60		0.997	0.997	0.997	0.995
.00	.00	.00	.45	.55	200	0.418	0.418	0.418	0.364
.00	.00	.00	.35	.65		0.997	0.997	0.997	0.995
.00	.00	.30	.30	.40		0.363	0.363	0.359	0.327
.00	.00	.20	.20	.60		1.000	1.000	1.000	1.000
.00	.10	.10	.35	.45		0.477	0.480	0.472	0.443
.00	.10	.10	.25	.55		0.999	0.999	0.999	0.999
.00	.10	.10	.15	.65		1.000	1.000	1.000	1.000
.10	.10	.10	.30	.40		0.521	0.524	0.519	0.484
.10	.10	.10	.20	.50		1.000	1.000	1.000	1.000
.10	.10	.10	.10	.60		1.000	1.000	1.000	1.000
.00	.00	.00	.45	.55	1000	0.932	0.939	0.932	0.932
.00	.00	.00	.35	.65		1.000	1.000	1.000	1.000
.00	.00	.30	.30	.40		0.972	0.972	0.972	0.969
.00	.00	.20	.20	.60		1.000	1.000	1.000	1.000
.00	.10	.10	.35	.45		0.971	0.972	0.971	0.969
.00	.10	.10	.25	.55		1.000	1.000	1.000	1.000
.00	.10	.10	.15	.65		1.000	1.000	1.000	1.000
.10	.10	.10	.30	.40		0.984	0.984	0.984	0.983
.10	.10	.10	.20	.50		1.000	1.000	1.000	1.000
.10	.10	.10	.10	.60		1.000	1.000	1.000	1.000

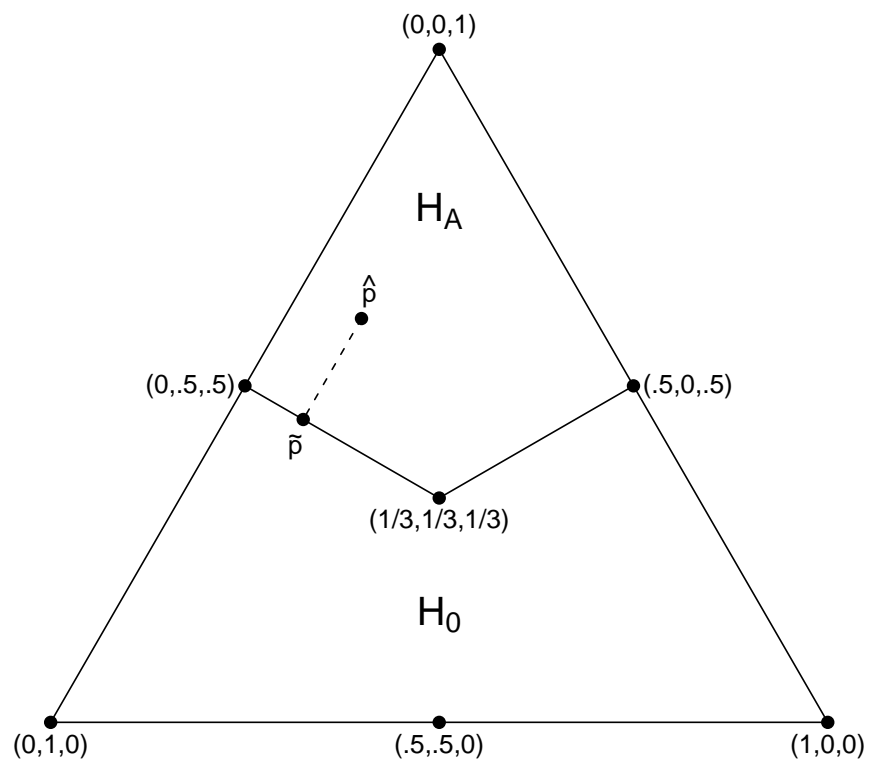


Figure 1: Barycentric plot of the parameter space \mathcal{P} partitioned into the null (H_0) and alternative (H_A) subsets for the case of $k = 3$. Those unfamiliar with the barycentric representation may wish to interpret it as a view of the parameter space looking from the origin in \mathbb{R}^3 toward the point $(1/3, 1/3, 1/3)$. An example maximum likelihood estimator of \boldsymbol{p} (denoted $\hat{\boldsymbol{p}}$) is plotted along with the corresponding maximum likelihood estimator of \boldsymbol{p} subject to the restriction imposed by H_0 (denoted $\tilde{\boldsymbol{p}}$).