ORİJİNAL ARAŞTIRMA ORIGINAL RESEARCH

A Test Procedure for Ordered Hypothesis of Population Proportions Against a Control

Bir Kontrole Karşı Anakütle Oranlarının Sıralı Hipotezi İçin Bir Test Prosedürü

Waheed Babatunde YAHYA,[a]
Oyebayo Ridwan OLANIRAN,[a]
Mohammad Kabir GARBA,[a]
Isiaka OLOYEDE,[a]
Alabi Waheed BANJOKO,[a]
Kazeem Adeshina DAUDA,[b]
Kabir Opeyemi OLOREDE[b]

[a]Department of Statistics,
University of Ilorin,
[b]Department of Statistics and
Mathematical Sciences,
Kwara State University,
Ilorin, Nigeria

Geliş Tarihi/*Received:* 13.01.2016
Kabul Tarihi/*Accepted:* 17.02.2016

Yazışma Adresi/*Correspondence:*
Waheed Babatunde YAHYA
Department of Statistics,
University of Ilorin,
P.M.B. 1515, Ilorin,
NIGERIA/NİJERYA
wbyahya@unilorin.edu.ng

ABSTRACT Objective: This paper aims to present a novel procedure for testing a set of population proportions against an ordered alternative with a control. **Material and Methods:** The distribution of the test statistic for the proposed test was determined theoretically and through Monte-Carlo experiments. The efficiency of the proposed test method was compared with the classical Chi-square test of homogeneity of population proportions using their empirical Type I error rates and powers at various sample sizes. **Results:** The new test statistic that was developed for testing a set of population proportions against an ordered alternative with a control was found to have a Chi-square distribution with non-integer values degrees of freedom v that depend on the number of population groups k being compared. Table of values of v for comparing up to 26 population groups was constructed while an expression was developed to determine v for cases where $k > 26$. Further results showed that the new test method is capable of detecting the superiority of a treatment, for instance a new drug type, over some of the existing ones in situations where only the qualitative data on users' preferences of all the available treatments (drug types) are available. The new test method was found to be relatively more powerful and consistent at estimating the nominal Type I error rates (α), especially at smaller sample sizes than the classical Chi-square test of homogeneity of population proportions. **Conclusion:** The new test method proposed here could find applications in pharmacology where a newly developed drug might be expected to be more preferred by users than some of the existing ones. This kind of test problem can equally exist in medicine, engineering and humanities in situations where only the qualitative data on users' preferences of a set of treatments or systems are available.

Key Words: Ordered hypothesis, test of population proportions;
Chi-square test of homogeneity of population proportions; Gaussian density;
Chi-square distribution; non-inter values degrees of freedom

ÖZET Amaç: Bu çalışmanın amacı bir kontrolle sıralı bir alternatife karşılık anakütle oranlarının bir kümesinin test edilmesi için yeni bir yöntem sunmaktır. **Gereç ve Yöntemler:** Önerilen test için test istatistiğinin dağılımı Monte-Carlo deneyleriyle ve teorik olarak belirlenmiştir. Önerilen test yönteminin etkinliği, farklı örneklem büyüklüklerinde güç ve deneysel Tip I hataları kullanılarak anakütle oranlarının homojenliği için klasik ki-kare testi ile karşılaştırılmıştır. **Bulgular:** Bir kontrolle sıralı bir alternatife karşılık anakütle oranlarının bir kümesinin test edilmesi için geliştirilen yeni test istatistiğinin, karşılaştırılan anakütle grup sayısı k'ya bağlı tam sayı değer olmayan v serbestlik dereceli ki-kare dağılımına sahip olduğu bulunmuştur. Birimlerin $k > 26$ olduğu durumda v'yü belirlemek için bir ifade geliştirilmişken anakütle grubu 26'ya kadar olanları karşılaştırmak için v değerleri tablosu geliştirilmiştir. Ayrıca sonuçlar, yeni test yönteminin örneğin, yeni bir ilacın var olan bir ilaca göre var olan tüm tedavilerin (ilaç tiplerinin) kullanıcı tercihlerinin sadece kalitatif verisinin mevcut olduğu durumlarda üstün olduğu gibi bir tedavinin üstünlüğünü tespit edebildiğini göstermiştir. Yeni test yönteminin, özellikle küçük örneklem büyüklüklerinde nominal Tip I hata oranlarını (α) tahmin etmede anakütle oranlarının homojenliği için klasik ki-kare testinden daha güçlü ve tutarlı olduğu bulunmuştur. **Sonuç:** Bu makalede sunulan yeni test yöntemi kullanıcılar tarafından yeni geliştirilen bir ilacın var olan bazı ilaçlara göre daha fazla tercih edildiği farmakolojide uygulama alanı bulabilir. Bu tür test problemi, bir tedavi ya da sistem kümesinde sadece kullanıcı tercihlerine dayalı kalitatif verinin mevcut olduğu tıpta, mühendislikte, beşeri bilimlerde aynı şekilde ortaya çıkabilir.

Anahtar Kelimeler: Sıralı hipotezler; anakütle oranlarının testi;
anakütle oranlarının homojenliği için ki-kare testi; Gauss yoğunluğu;
ki-kare dağılımı; tam sayı olmayan değerler serbestlik derecesi

doi: 10.5336/biostatic.2016-50196

Turkiye Klinikleri J Biostat 2016;8(1):1-12

The study of test of hypothesis involving ordered alternatives started by the earlier works of Jonckheere and Bartholomew which was later revisited about a decade later by R. E. Barlow and in his team in their work in 1971.[1-3] However, the significance of this study area further gave it prominence in the last few years. The classical case of the test hypothesis problems often addressed takes the form of

$$H_0: \mu_1 = \mu_2 = \cdots = \mu_k \quad \text{against } H_a: \mu_1 < \mu_2 < \cdots < \mu_k \qquad (1)$$

for a k comparison of population means where the k populations are in no particular order of preference.[4-7]

Here, the underlying distribution of the random variable $X_i, i = 1,2, \dots, k$ that governed the test procedure for hypothesis (1) was assumed to have a Gaussian distribution with mean μ_i and (possibly) constant variance $\sigma^2$for all the k populations.

In a more controlled study, the situation that requires a comparison of all the $k-1$ populations with the k^{th} population being the control might be desirable. Such situations are quite common in Pharmacological studies (in drug discovery) where it is often of interest to determine the efficacy of a new drug relative to some of the existing ones. For instance, suppose the superiority of a new analgesic drug type d_k over some of the existing ones d_i, $i = 1,2, \dots, k-1$, is to be determined by the number of minutes it takes the users to get relieved of pain after its intake. The hypothesis of interest under this scenario would be of the form:

$$H_0: \mu_k = \mu_i \;\; i = 1, 2, \dots, k-1 \text{ against } H_a: \mu_k < \mu_i \forall i, \;\; i = 1, 2, \dots, k-1 \qquad (2)$$

The procedures to handle the ordered hypothesis set in (2) have been provided elsewhere when the underlying random variables X_i (the number of minutes it takes users to get relieved of pain after the intake of drug type d_i, $i = 1, 2, \dots, k$) are from Gaussian densities as earlier remarked.[8,9]

The test hypothesis problem (2) is also suitable to establish, for example, the superiority of a new antiretroviral drug over some of the existing ones in terms of its relative contribution at enhancing the CD4 counts in HIV patients over a period of time. Here, the new antiretroviral drug would be treated as a control whose supremacy over some of the existing drugs in terms of efficacy is to be determined.

Without loss of generality, it is obvious that any test problem that requires either of the ordered hypothesis set (1) and (2) pre-supposes that the sample data, apart from emanating from Gaussian populations, are also available and obtainable.

However, in many of the real life situations, the required information on the sample data to prosecute hypothesis (1) or (2) can only be imagined but difficult to obtain. For instance, it is quite easy to imagine obtaining information on the number of minutes it takes a set of patients to get relieved of pain after taken a certain analgesic drug. Due to low literacy level of some patients, especially in developing countries like Nigeria, it might be difficult, if not impossible to get this kind of information accurately from such people. In a controlled study where this information can be obtained accurately on the patients, another challenge is that the observed sample measurements might not truly satisfy the desired distributional assumption of normality to guarantee efficient test results.

In the two examples illustrated above and generally in several other real life situations, the quick and easy (qualitative) information that can be truly obtained could be on whether a person get cured or not after receiving a particular chemotherapy treatment (or a drug type) or whether a particular antiretroviral drug boosted the CD4 counts of an HIV patient or not and so on. In other words, information on the number and proportion of people that are positive to the use of a particular drug type among those that were exposed could be readily and easily obtainable.

Therefore, in a situation where it is only this kind of information that is available to evaluate the superiority of new therapy treatment or drug over a number of existing ones, the test of hypothesis involving population proportions with ordered alternative set may be desirable. This kind of test problem motivated the development

of the test procedure presented in this work. A test method to handle similar hypothesis problem involving two population proportions has been presented elsewhere.[10] The Bayesian approach of non-Inferiority test between two independent binomial population proportions has also been developed by Yahya et al.[11]

DEVELOPMENT OF THE PROPOSED TEST PROCEDURE

Suppose $d_1, d_2, \dots, d_k$, in no particular order, is a set of k chemotherapy treatments (treatment groups) or drug types that are meant to cure a particular disease with the c^{th} drug, d_c, $d_c \in \{d_1, d_2, \dots, d_k\}$ being the newly developed (best) drug while the remaining $k-1$ drugs, excluding d_c are the existing drug types or chemotherapy treatments. For convenience, let $d_k = d_c$ be the control drug type (control group), $c \in \{1, 2, \dots, k\}$.

Also, let $X_1, X_2, \dots, X_k$ be the number of people that are positive to the use of drug types or treatments $d_1, d_2, \dots, d_k$ among the $n_1, n_2, \dots, n_k$ that were treated by the drug types respectively. It then follows that X_i is the number of people that are positive to drug d_i among the n_i patients that are exposed to it while $\hat{p}_i = \frac{X_i}{n_i}$ represents the corresponding proportion. Thus, for $i = 1,2, \dots k$ we have that

$$\hat{p}_1 = \frac{X_1}{n_1}, \hat{p}_2 = \frac{X_2}{n_2}, \dots, \hat{p}_k = \frac{X_k}{n_k}.$$

Necessarily, each of the X_i, $i = 1, 2, \dots, k$ is an independent binomial random variable with parameters n_i and p_i. That is, $X_i \sim Bin(n_i, p_i)$. The unbiased estimate of the population proportion p_i is $\hat{p}_i = \frac{X_i}{n_i}$.

Since drug type d_c is regarded as the control (best) drug among all the k drug types, it is expected to be most efficacious and most preferred among all the drug types. Hence, the proportion of sample units that would be positive to drug d_c, $c \in \{1, 2, \dots, k\}$, is expected to be higher than those of the remaining $k-1$ existing drugs. As a result, the hypothesis of interest to be tested would be of the form:

$$H_0: p_i = p_c \text{ against } H_a: p_i < p_c \quad \forall_i, \ i = 1, 2, \dots, k-1 \tag{3}$$

where p_c is the proportion of the users of the control drug type d_c, $c \in \{1, 2, \dots, k\}$. The unbiased estimator of p_c is $\hat{p}_c = \frac{X_c}{n_c}$.

The inequality statement in the ordered alternative hypothesis set (3) can be reconstructed in two ways as presented below.

CASE I

In the ordered hypothesis (3), the inequality statement in the ordered alternative set H_a is $p_i < p_c \ \forall_i, \ i = 1, 2, \dots, k-1$.

$$\rightarrow \qquad p_i - p_c < 0 \ \forall_i, \ i = 1, 2, \dots, k-1.$$

This yields the following $k-1$ differences of population proportions:

$$p_1 - p_c < 0, p_2 - p_c < 0, \dots, p_{k-1} - p_c < 0 \tag{4}$$

If we consider the representation $\delta_i = p_i - p_c$, then all the proportion differences in (4) can be represented by vector

$$\delta = \{\delta_1, \delta_2, \dots, \delta_{k-1}\} \tag{5}$$

However, for ordered alternative set H_a in (3) to be sustained under the significance hypothesis testing, the maximum of the proportion differences in vector δ must be less than zero. That is,

$$max(\delta_i) < 0 \ \forall_i, i = 1,2, \dots, k-1 \tag{6}$$

and if this is defined over the δ_i 's we have

$$max_{1 \le i \le k-1} \delta_i < 0, \forall_i, i = 1,2, \dots, k-1 \tag{7}$$

Hence, the hypothesis set (3) can be re-parameterized as follows:

$$H_0: \delta_i = 0 \text{ against } H_a: max_{1 \le i \le k-1} \delta_i < 0 \quad \forall_i, \ i = 1, 2, \dots, k-1 \tag{8}$$

The unbiased estimator of vector δ in (5) is

$$\hat{\delta} = \{\hat{\delta}_1, \hat{\delta}_2, \dots, \hat{\delta}_{k-1}\} \tag{9}$$

where

$$\hat{\delta}_i = \hat{p}_i - \hat{p}_c \tag{10}$$

and $\hat{p}_i = \frac{X_i}{n_i}$. However, $\hat{\delta}_i$ can be reasonably assumed to have a Gaussian density with mean $\delta_i = p_i - p_c$ and variance $Var(\hat{\delta}_i)$. That is, $\hat{\delta}_i \sim N(\delta_i, Var(\hat{\delta}_i))$. Both the $E(\hat{\delta}_i)$ and $Var(\hat{\delta}_i)$ are determined as follows;

$$E(\hat{\delta}_i) = E(\hat{p}_i - \hat{p}_c)$$

$$\rightarrow \quad E(\hat{\delta}_i) = E\left[\frac{X_i}{n_i} - \frac{X_c}{n_c}\right]$$

$$= \frac{1}{n_i}E(X_i) - \frac{1}{n_c}E(X_c)$$

$$= \frac{1}{n_i}n_i p_i - \frac{1}{n_c}n_c p_c$$

$$\rightarrow \quad E(\hat{\delta}_i) = p_i - p_c \tag{11}$$

Also,

$$Var(\hat{\delta}_i) = \quad Var(\hat{p}_i - \hat{p}_c)$$

$$\rightarrow \quad Var(\hat{\delta}_i) = Var(\hat{p}_i) + Var(\hat{p}_c)$$

$$= Var\left(\frac{X_i}{n_i}\right) + Var\left(\frac{X_c}{n_c}\right)$$

$$= \frac{1}{n_i^2}Var(X_i) + \frac{1}{n_c^2}Var(X_c)$$

$$= \frac{n_i p_i(1-p_i)}{n_i^2} + \frac{n_c p_c(1-p_c)}{n_c^2}$$

$$\rightarrow \qquad Var(\hat{\delta}_i) = \quad \frac{p_i(1-p_i)}{n_i} + \frac{p_c(1-p_c)}{n_c} \tag{12}$$

However, in the null hypothesis set H_0 in (3), $p_i = p_c = p$ (say). Therefore, (12) becomes

$$Var(\hat{\delta}_i) = \quad \left(\frac{n_i+n_c}{n_i n_c}\right)p(1-p) \tag{13}$$

where parameter p in (12) is estimated by

$$\hat{p} = \frac{X_i+X_c}{n_i+n_c} \tag{14}$$

Now, since $\hat{\delta}_i \sim N\left(\delta_i, Var(\hat{\delta}_i)\right)$, let statistic $\hat{Z}_i$ be defined as follows:

$$\hat{Z}_i = \frac{\hat{\delta}_i}{\sqrt{Var(\hat{\delta}_i)}} \quad \forall_i, i = 1, 2, \dots, k-1 \tag{15}$$

and when (5) is defined on vector (9) it yielded a vector of $\hat{Z}_i$ of the form

$$\hat{Z} = \left(\hat{Z}_1, \hat{Z}_2, \dots, \hat{Z}_{k-1}\right) \tag{16}$$

Obviously, each of the $\hat{Z}_i$ in $\hat{Z}$ has a unit Gaussian density. That is,

$$\hat{Z}_i = \frac{\hat{\delta}_i}{\sqrt{Var(\hat{\delta}_i)}} \sim N(0,1) \quad \forall_i, i = 1, 2, \dots, k-1 \tag{17}$$

Since each of the $\hat{Z}_i$ in vector $\hat{Z}$ ($i.e.\ \hat{Z}_i \in \hat{Z}$) has a unit Gaussian, it then follows that statistic $\hat{Z}_{max} = max\,\hat{Z}_i \in \hat{Z}$ can be assumed to have a unit Gaussian. Thus, the test statistic for testing the ordered hypothesis set (8) is

$$\hat{Z}_{max} = \underset{1\le i\le k-1}{max}\left(\frac{n_i n_c}{n_i+n_c}\right)^{\frac{1}{2}}\frac{\hat{p}_i - \hat{p}_c}{\sqrt{p(1-p)}} \sim N(0,1) \tag{18}$$

If the weight w_i is defined by

$$w_i = \left(\frac{n_i n_c}{n_i+n_c}\right)^{\frac{1}{2}} \tag{19}$$

Therefore, the test statistic in (18) becomes

$$\hat{Z}_{max} = \underset{1\le i\le k-1}{max} w_i \frac{\hat{p}_i - \hat{p}_c}{\sqrt{p(1-p)}} \sim N(0,1) \tag{20}$$

Hence, for testing the ordered hypothesis (8), the test statistic is given by (20).

However, it is important to remark that a unit Gaussian density was only assumed for the distribution of the test statistic $\hat{Z}_{max}$ as indicated in (18) and (20). If the set of the $k-1$ statistics $\hat{Z}_i$ in vector $\hat{Z} = \left(\hat{Z}_1, \hat{Z}_2, \dots, \hat{Z}_{k-1}\right)$ of (16) are regarded as ordered statistics for which the statistic $\hat{Z}_{max}$ is their maximum, i.e. $\hat{Z}_{max} = max\,\hat{Z}_i \in \hat{Z}$, then, it may be necessary to determine the true density function of $\hat{Z}_{max}$ from the distribution of ordered statistics.

Now, since $\hat{Z}_i \sim N(0,1)\ \forall_i, i = 1, 2, \dots, k-1$, it then follows that the density function of $\hat{Z}_{max}$ in vector $\hat{Z}$ is of the form;

$$f_{\hat{Z}_{max}}(\hat{\zeta}) = (k-1)\left[\Phi(\hat{\zeta})\right]^{k-2} f(\hat{\zeta}) \tag{21}$$

where $f(\hat{\zeta}) = \frac{1}{\sqrt{2\pi}}\,exp\left(\frac{\hat{\zeta}^2}{2}\right)$, $\Phi(\hat{\zeta}) = F(\hat{\zeta})$ and $\hat{\zeta} \in \hat{Z}$ is the random variable defined on $\hat{Z}_i$, $\forall_i$ since all the $\hat{Z}_i$ are identically distributed.

If we let $(k-1)\left[\Phi(\hat{\zeta})\right]^{k-2} = \eta$, then, the pdf in (21) reduces to

$$f_{\hat{Z}_{max}}(\hat{\zeta}) = \eta f(\hat{\zeta}) \tag{22}$$

Therefore, for the density of $\hat{Z}_{max}$ in (22) to be a unit Gaussian as earlier assumed, the condition that $\eta = 1$ is necessary. It then follows that the necessary condition for the normal approximation to hold as used in (18) and (20) for the distribution of the test statistic $\hat{Z}_{max}$ is that

$$(k-1)\left[\Phi(\hat{\zeta})\right]^{k-2} = 1 \tag{23}$$

However, the equality statement in (23) is only true if and only if $k = 2$, a condition that is only suitable for comparing two population proportions.

If the assumption of a unit Gaussian density is sustained for the test statistic $\hat{Z}_{max}$ in (20), it then follows that the statistic

$$X^2_{max} = \left(\hat{Z}_{max}\right)^2 \sim \chi^2_v \quad (24)$$

where v in (24) is the degree of freedom of the Chi-Square distribution to be determined. Therefore, for $i = 1,2,\dots,k-1$, the alternative test statistic for testing the ordered hypothesis (8) is

$$X^2_{max} = \underset{1\le i\le k-1}{max}\left(\frac{n_i n_c}{n_i+n_c}\right)\left(\frac{\hat{p}_i - \hat{p}_c}{\sqrt{p(1-p)}}\right)^2 \sim \chi^2_v \quad (25)$$

which, when the weight w_i as defined in (19) is substituted, becomes

$$X^2_{max} = \underset{1\le i\le k-1}{max} w_i^2\left(\frac{\hat{p}_i - \hat{p}_c}{\sqrt{p(1-p)}}\right)^2 \sim \chi^2_v \quad (26)$$

Obviously, if $k = 2$, the statistic $\hat{Z}_{max}$ in (20) has a unit Gaussian density, and by extension the test statistic X^2_{max} in (26) is distributed Chi-square with one degree of freedom (with $v = 1$). That is $X^2_{max} \sim \chi^2_1$. For the cases where $k > 2$ as conjectured in this work, the true distribution of $\hat{Z}_{max}$ cannot be perfectly approximated by a unit Gaussian. By extension, the distribution of the test statistic X^2_{max}, although coming from the Chi-square family of distributions with v degrees of freedom, cannot be said to possess a Chi-square distribution with $v = 1$ for all values of $k > 2$. Therefore, the true density of X^2_{max}, more specifically the values of v at various values of $k > 2$ has to be determined. These results, as obtained from Monte-Carlo studies are provided in a later section.

When the group sample sizes n_i in (25) are equal for all i, that is $n_1 = n_2 = \cdots = n_k = n$, the test statistic (18) and (25) become

$$\hat{Z}_{max} = \underset{1\le i\le k-1}{max}\left(\frac{n}{2}\right)^{\frac{1}{2}}\frac{\hat{p}_i - \hat{p}_c}{\sqrt{p(1-p)}} \sim N(0,1) \quad (27)$$

and

$$X^2_{max} = \underset{1\le i\le k-1}{max}\left(\frac{n}{2}\right)\left(\frac{\hat{p}_i - \hat{p}_c}{\sqrt{p(1-p)}}\right)^2 \sim \chi^2_v \quad (28)$$

respectively.

CASE II

Similarly, given the inequality statement $p_i < p_c \ \forall_i,\ i = 1,2,\dots,k-1$, it then follows that $p_i < p_c$ iff $p_c - p_i > 0 \ \ \forall_i, i = 1,2,\dots,k-1$. This consequently yielded the following $k-1$ difference of proportions

$$p_c - p_1 > 0, p_c - p_2 > 0, \dots, P_c - P_{k-1} > 0 \quad (29)$$

If $\delta_i^* = p_c - p_i$, we have the vector representation of (29) in the form

$$\delta^* = \{\delta_1^*, \delta_2^*, \dots, \delta_{k-1}^*\} \quad (30)$$

Therefore, for ordered alternative set H_a in (3) to be sustained under the significance hypothesis testing, the minimum of the proportion differences in vector δ^* must be greater than zero. That is,

$$min_{1\le i\le k-1}\,\delta_i^* > 0, \forall_i, i = 1,2,\dots,k-1 \quad (31)$$

Hence, the hypothesis set (3) can be re-parameterized as:

$$H_0: \delta_i^* = 0 \ \text{against}\ H_a: min_{1\le i\le k-1}\,\delta_i^* > 0, \forall_i, i = 1,2,\dots,k-1 \quad (32)$$

Similarly, the unbiased estimator of δ_i^* is $\hat{\delta}_i^* = \hat{p}_c - \hat{p}_i$ which shows that $E\left(\hat{\delta}_i^*\right) = p_c - p_i$ and $Var\left(\hat{\delta}_i^*\right) = \left(\frac{n_i+n_c}{n_i n_c}\right)p(1-p)$ as earlier determined.

The following statistic $\hat{Z}_i^* = \frac{\hat{\delta}_i^*}{\sqrt{Var(\hat{\delta}_i^*)}}\ \forall_i, i = 1,2,\dots,k-1$ was also considered for the development of the test statistic for the ordered hypothesis set (32). As a result, we have a vector of statistics $\hat{Z}^* = \left(\hat{Z}_1^*, \hat{Z}_2^*, \dots, \hat{Z}_{k-1}^*\right)$. Also, each of the elements in $\hat{Z}^*$ has a unit Gaussian density. That is $\hat{Z}_i^* \sim N(0,1)\ \forall_i, i = 1,2,\dots,k-1$ which therefore presupposes that the probability density of the minimum element in $\hat{Z}^*\left(i.e.\hat{Z}_{min} = min\left(\hat{Z}_i^*\right)\right)$ can be assumed to have a unit Gaussian density. That is, $\hat{Z}_{min} = min\left(\hat{Z}_i^*\right) \sim N(0,1)$, $min\left(\hat{Z}_i^*\right) \in \hat{Z}$.

Therefore, the test statistic for testing the ordered hypothesis set (32) becomes

$$\hat{Z}_{min} = \underset{1\le i\le k-1}{min}\left(\frac{n_i n_c}{n_i+n_c}\right)^{\frac{1}{2}}\frac{\hat{p}_c - \hat{p}_i}{\sqrt{p(1-p)}} \sim N(0,1) \quad (33)$$

Similarly, a square transform of (33) yielded alternative test statistic

$$X^2_{min} = \underset{1\le i\le k-1}{min}\left(\frac{n_i n_c}{n_i+n_c}\right)\left(\frac{\hat{p}_c - \hat{p}_i}{\sqrt{p(1-p)}}\right)^2 \sim \chi^2_v \quad (34)$$

Following similar line of arguments as in Case I regarding the true distribution of the test statistic for the hypothesis set (8), the development of the true distribution of the test statistics

(33) or (34) shall be pursued by examining the distribution of ordered statistics based on vector of statistics $\hat{Z}^* = (\hat{Z}_1^*, \hat{Z}_2^*, \dots, \hat{Z}_{k-1}^*)$. By this, the following probability distribution for the test statistic $\hat{Z}_{min}$ in (33)

$$f_{\hat{Z}_{min}}(\hat{\zeta}^*) = (k-1)[1-\Phi(\hat{\zeta}^*)]^{k-2} f(\hat{\zeta}^*) \quad (35)$$

was derived where $\hat{\zeta}^* \in \hat{Z}^*$ is a random variable defined on $\hat{Z}_i^*$.

Again, the distribution of the $\hat{Z}_{min}$ in (35) can only be a unit Gaussian if and only if $k = 2$, which is only desirable while comparing two-population proportions. For the cases where $k > 2$ as conjectured in this work, the true distribution of $\hat{Z}_{min}$ cannot be perfectly approximated by unit Gaussian and by extension, the test statistic $\hat{Z}_{min}^2 = X_{min}^2$ cannot possess a Chi-square distribution with one degree of freedom for all values of $k > 2$. So, the true density of X_{min}^2 has to be determined for various values of $k > 2$.

It should be noted that either of the test statistics developed under Case I or Case II can be used for testing equality of population proportions against ordered alternative with a control depending on whether hypothesis statement (8) or (32) is to be employed respectively. The two tests would essentially provide the same conclusion. Hence, further discussions on the development of the proposed test procedure in this work were presented for Case I only due to space.

SIMULATION STUDIES

The scheme followed in the Monte-Carlo studies of the proposed test statistics for testing hypothesis of equality of population proportions with ordered alternative involving a control group is presented in this section.

To determine the power and empirical Type I error rate of the proposed test statistics, five independent population groups ($k = 5$) were conjectured. Therefore, five independent random samples $X_1, \dots, X_5$ of numbers of people that are positive to drug types $d_1, \dots, d_5$ (say) were simulated from binomial distributions with varying sample sizes n_i set between 20 and 100 for different combinations of population proportions p_i, $i = 1, \dots, 5$. All the tests were performed at 5% significance level.

Under this setting, the fifth drug d_5 was taken to be the best and the control drug. Hence, the p_i were set such that $p_5 > p_i$, $i = 1, \dots, 4$. Specifically, one of the set of values of the five population proportions employed in the simulations were $p_1 = 0.1$; $p_2 = 0.15$; $p_3 = 0.2$; $p_4 = 0.25$; and $p_5 = 0.45$.

In order to determine the degrees of freedom of the proposed test statistics, data were simulated for up to 100 population groups. The values of the test statistics (26) and (34) for the hypothesis sets (8) and (32) were determined respectively at various samples sizes and at different population proportions. At a particular sample size for instance, the average of the computed test statistics was determined over 1000 replications in order to obtain its degree of freedom.

All the simulations and analysis were performed within the environment of R statistical package.[12]

RESULTS

Results from the Monte-Carlo experiments for the development and applications of the proposed test statistics are presented in this section. However, only the simulation results for theoretical development of the proposed test statistic and its applications under Case I were presented due to space. Results for Case II were similar to those obtained under Case I and the two results essentially yielded similar conclusions.

The simple histogram and the density graphs of the computed values of the test statistic (26) were plotted at various sample sizes for comparing population proportions in five groups. The histogram and the density of the test statistic values for sample sizes 20 and 100 are presented by Figures 1a-d. respectively due to space. In all the four graphs, the closeness of the distribution of the test statistic to the family of Chi-square distribution is apparent as shown in Figures 1a-d.

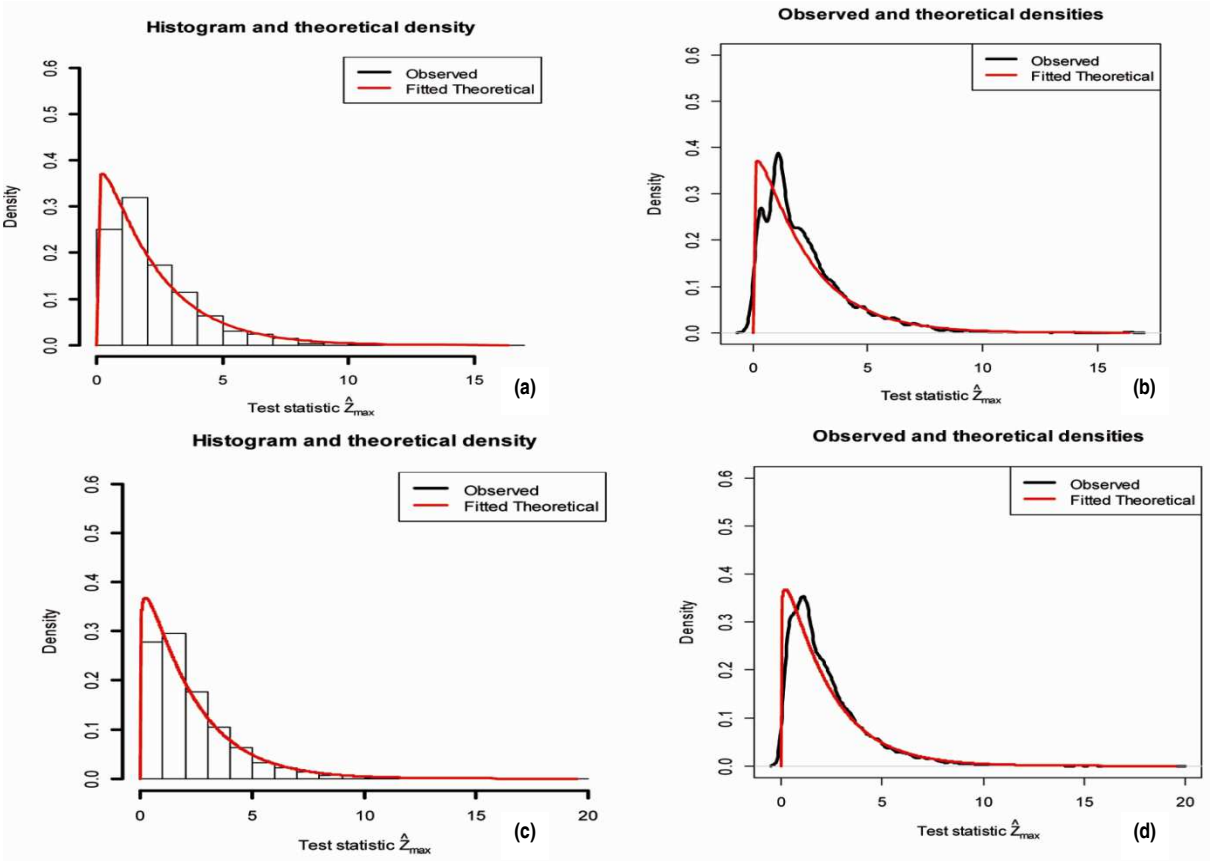

FIGURE 1: The histogram of the proposed test statistic values $\hat{X}^2_{max}$ overlaid by its theoretical density curve and the density plot of the observed $\hat{X}^2_{max}$ values (black curve) overlaid by its theoretical density plot (red curve) at sample sizes 20 (graphs (a.) & (b.)) and 100 (graphs (b.) & (c.)).

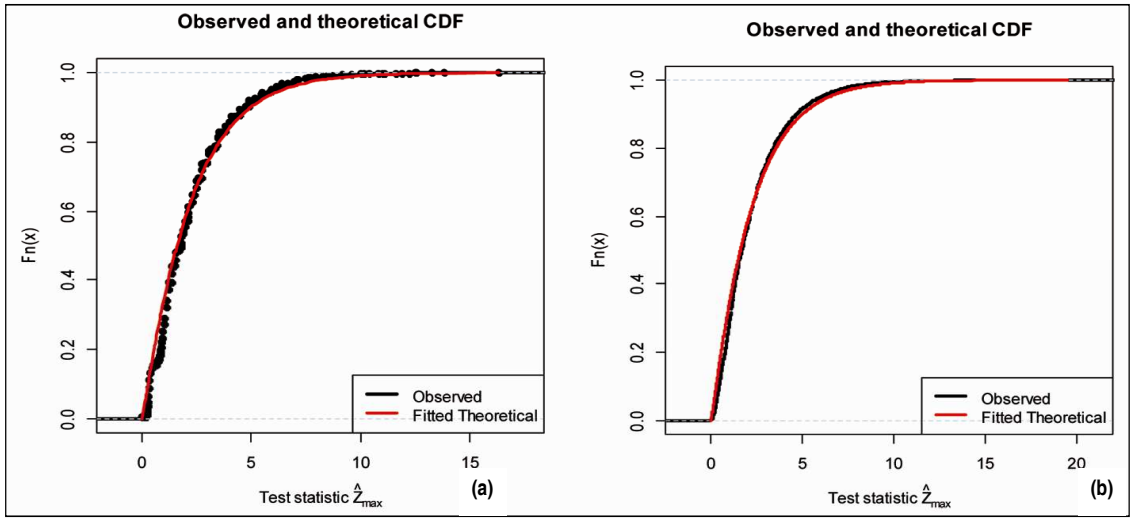

FIGURE 2: Plot of observed and theoretical cumulative distribution function of the proposed test statistic $\hat{X}^2_{max}$ at sample sizes 20 (left graph) and 100 (right graph).

As a further confirmation of the distribution of the test statistic, the graphs of the observed (empirical) and theoretical cumulative distributions of the computed values are plotted for sample sizes 20 and 100 as shown in Figure 2. Again, the two graphs showed that the distribution of the test statistic X^2_{max} is from Chi-square family of distributions.

TABLE 1: Table of degrees of freedom for the proposed test statistic for ordered hypothesis test of population proportions with a control. The table provided the degrees of freedom (v) for number of groups (k) being compared up to k = 26 populations (groups). For more than 26 groups (k > 26), the following equation $v = 1.6478 \times k^{0.2765}$ should be used to determine v.

Number of Group(k)	Degree of freedom (v)	Number of Group (k)	Degree of freedom (v)	Number of Group (k)	Degree of freedom (v)
3	1.619051	11	3.064805	19	3.661092
4	1.981779	12	3.138456	20	3.712839
5	2.204617	13	3.205245	21	3.767818
6	2.432621	14	3.264881	22	3.811624
7	2.604624	15	3.346759	23	3.867405
8	2.770404	16	3.424155	24	3.944560
9	2.857594	17	3.475543	25	4.012377
10	2.971645	18	3.575597	26	4.052725

Now that the distribution of the proposed test statistic X^2_{max} has been determined to be Chi-square, it is therefore important to further determine the degrees of freedom for all possible $\hat{X}^2_{max}$ values for different population groups k as they would emanate from various test problems as conjectured in this work. This would enable easy inference to be drawn as in the conventional chi-square test of independence.

However, various results from simulation studies showed that the degrees of freedom of the proposed test statistics are only influenced by the number of groups (k) being compared but not by the group sample sizes n_i whether they are equal, in which case $n_1 = n_2 = \cdots = n_k = n$ or not. Thus, only the results for the case of homogenous group sample sizes are reported in this work.

However, it can be recalled that the degree of freedom of a typical Chi-square random variable X can be estimated by its arithmetic mean $\bar{X}$. To this end, the degrees of freedom v of the proposed test statistic X^2_{max} were simply determined in the Monte-Carlo experiment by computing the means of the estimated test statistic values $\hat{X}^2_{max}$ for number of groups $k > 2$ up to k = 100 with group sample size $n_i = 100$ for all i at 1000 replications. These degrees of freedom for values of $k > 2$ up to $k = 26$ (i.e. k = 3, 4, 5, …, 26) are provided in Table 1 due to space. The graph of the plot of the degrees of freedom of this test statistic against the number of population groups being compared is provided by Figure 3.

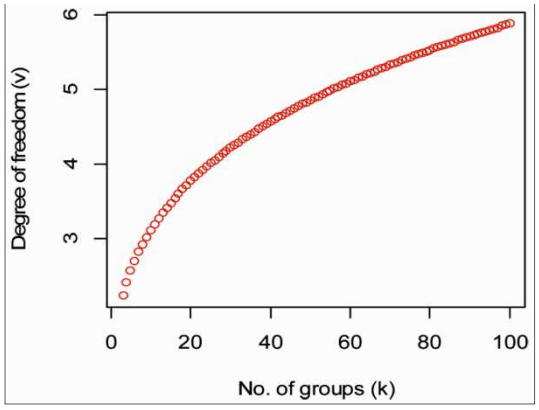

FIGURE 3: Graph of the plot of number of group (k) against empirical degree of freedom (v) for the proposed chi-squared distributed test statistic X^2_{max} for comparing ordered population proportions with a control.

Intuitively, comparisons among population groups with k = 26 might turn out to be a needless exercise in real life situations. Nevertheless, for test hypothesis cases for which k > 26, their degrees of freedom can be determined using the equation

$$v = 1.6478 \times k^{0.2765} \quad (36)$$

where v is the desired degree of freedom and k is the number of groups being compared as specified in the hypothesis test problem. The above equation that connects v and k was constructed using the simulated v of the proposed test statistic X^2_{max} at various number of population groups k (for $k > 26$) over 1000 replications.

It is worthy to remark that various results from Monte Carlo experiments performed showed that the degrees of freedom for the proposed test statistics X^2_{max} or X^2_{min} are not necessarily integers as shown in Table 1 unless if truncated at one decimal place. This truncation if performed, will affect the corresponding p-values and possibly the inferences based on such test statistics values.

To compare the goodness of the proposed test procedure with the conventional Chi-square test of homogeneity of population proportions, we considered hypothesis case to compare the proportions of five populations with the fifth group being the control. The values of the five population proportions are $p_1 = 0.1$, $p_2 = 0.15$, $p_3 = 0.2$, $p_4 = 0.25$ and $p_5 = 0.45$ as specified in Section 3 under the simulation scheme. The observed values $X_1, X_2, X_3, X_4$and X_5were simulated independently through Bernoulli distributions with the above respective specified population proportions at small sample sizes 20, 40, 60; moderate sample sizes 80, 100; and large sample sizes 120, 140, 160, 180, 200.

Given the above data structure as simulated, the desired ordered hypothesis of interest, as conjectured here, based on the proposed test statistic χ^2_{max} using (8) is of the form

$$H_0: \delta_i = 0 \text{ against } H_a: max_{1\le i\le 4}\,\delta_i < 0 \quad \forall_i,\ i = 1,2,3,4 \qquad (37)$$

where $\delta_i = p_i - p_5$ for $i = 1, 2, 3, 4$. On the other hand, the hypothesis of interest for the conventional Chi-square test of homogeneity of the five population proportions is of the form $H_0: p_i = p_{i'} \forall_{i,i'}$ $i, i' = 1,2,3,4,5$ against $H_a: p_i \neq p_{i'}$, (38) for at least a pair (i, i').

The two hypotheses (37) and (38) were tested at 5% Type I error rate using the proposed tests statistic X^2_{max} and the classical Chi-square statistic for testing homogeneity of population proportions. The empirical Type I error rates as provided by each test method and their corresponding powers at various sample sizes are presented in Table 2.

The well-known general Chi-square statistic for testing homogeneity of k population proportions is of the form

$$X^2 = \sum_{i=1}^{k} \frac{(O_i - e_i)^2}{e_i} \sim \chi^2_{k-1} \qquad (39)$$

where O_i and e_i are the observed count (X_i) value and the expected count (under the null hypothesis H_0) from the i^{th} group respectively. The e_i is estimated as $e_i = np_i$ where $n = \sum_{i=1}^{k} n_i$ and $p_i = 1/k$ under H_0.

For ease of assessment of the performance of the two tests methods, the plot of the estimated empirical Type I error rates and powers (in %) of these tests at various sample sizes are provided as shown by Figure 4.

ILLUSTRATIVE EXAMPLE

To demonstrate the application of the proposed test method for testing ordered hypothesis of

TABLE 2: Table of Empirical Type -1 error rates (ALPHA-α) and Powers of the proposed test procedure and the conventional chi-square test of homogeneity of population proportions at various (homogeneous) group sample sizes $n_i = n$.

Sample sizes (n)	Empirical ALPHA (α) at 5%		Power	
	Proposed X^2_{max} Test	X^2- Test	Proposed X^2_{max} Test	X^2- Test
20	4.57	7.55	**48.63**	28.96
40	4.77	4.84	**66.78**	48.04
60	4.76	4.75	**80.25**	69.65
80	4.73	4.73	**87.69**	82.77
100	4.95	5.09	**92.99**	91.29
120	4.88	4.95	95.68	96.49
140	4.98	4.75	97.37	98.15
160	4.84	4.90	98.69	99.38
180	4.43	4.69	99.29	99.60
200	4.98	5.20	99.67	99.94

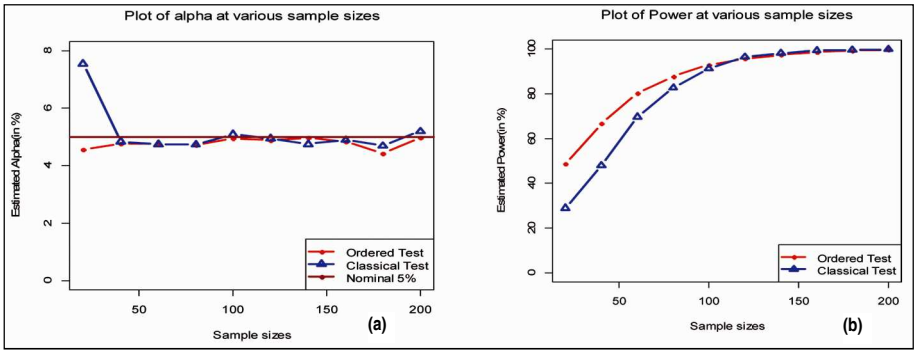

FIGURE 4: Graph of the estimated empirical Type I error rates (Figure 4 (a)) by the proposed test for ordered hypothesis set (37) and the classical Chi-square test of homogeneity of population proportions (38) at different sample sizes. The solid horizontal line in the plot indicated the nominal 5% Type I error rate used for the two tests. Figure 4 (b) is the plots of the estimated powers of the proposed and the classical Chi-square tests at various sample sizes.

TABLE 3: Table of results of the proposed test for testing ordered hypothesis of population proportions with a control of hypothesis statement (37) and the conventional chi-square test of homogeneity of population proportions of hypothesis statement (38) at sample size 20 for comparing five populations with the population proportions $p_1 = 0.1, p_2 = 0.15, p_3 = 0.2, p_4 = 0.25,$ and $p_5 = 0.45$.

Test Procedure	Test Statistic Value	Degree of freedom	P-Value	Decision
Proposed Ordered Test	11.4342	2.204617	0.0042	Reject H_0
Conventional Chi-square Test	9.3144	4	0.0537	Do not reject H_0

population proportions with a control, the results of the test hypothesis problem (37) at group sample size 20 are reported. The corresponding results of the conventional chi-square test of homogeneity of population proportions for hypothesis problem (38) are equally reported as shown in Table 3 for easy comparison.

The values of the proposed test statistics X^2_{max} and that of the conventional chi-square test of homogeneity of population proportions, their corresponding degrees of freedom and their p-values are all provided in Table 3. The decisions based on the p-values of the two tests are equally provided in which the proposed test rejected the null hypothesis H_0 in hypothesis set (37) while the conventional chi-square test did not reject the null hypothesis set H_0 in hypothesis set (38) at 5% significance level.

The p-value, $\hat{p}_v$ of a Chi-square statistic value $\hat{X}^2$ at v degree of freedom is obtained by a straightforward general procedure by computing

$$\hat{p}_v = p\left(X^2 > \hat{X}^2\right)_v = 1 - p\left(X^2 \leq \hat{X}^2\right)_v \qquad (40)$$

The $\hat{p}_v$ can be easily obtained from any general Chi-square table if v is an integer. However, if a non-integer valued degree of freedom v is obtained from Table 1 or determined from equation (36), the associated p-value can be easily obtained using any of the flexible statistical packages. For instance, the p-value $\hat{p}_v$ of the computed test statistic $\hat{X}^2_{max}$ at v degree of freedom is determined by $\hat{p}_v = 1 - p\left(X^2 \leq \hat{X}^2_{max}\right)_v$. Therefore, using R statistical package, with $\hat{X}^2_{max} = 11.4342$ and $v = 2.204617$, the p-value is obtained by invoking the R function **p.value = 1 - pchisq(q = 11.4342, df = 2.204617)**which yielded the needed p-value of 0.0042 as reported in Table 3.

The argument **q** in the above R code is the quantile value of the Chi-square distribution which corresponds to the computed value of the test statistic X^2_{max} while the **df** is its associated (non-integer value) degree of freedom. What any interested users should therefore provide to determine the p-value of the proposed test statistic X^2_{max} are the computed test statistic value $\hat{X}^2_{max}$ and its associated degree of freedom v from Table 1 or determined by equation (36) based on the number of groups k being compared.

DISCUSSION

The theoretical development of the proposed test statistic X^2_{max} or X^2_{min} for testing equality of k population proportions against ordered alternative with a control suggested that its true density cannot be Gaussian as it might otherwise be quickly conjectured from the test statistic for comparing two population groups. Indeed, results from Monte-Carlo experiments showed that the true density of X^2_{max} or X^2_{min} is Chi-square with non-integer values degrees of freedom v.

Although, the distribution of the new proposed test statistic has been found to be chi-square, a major attribute of this statistic is that its degree of freedom is not constrained to be necessarily integers as shown in Table 1, the table of degrees of freedom constructed for this new test statistic. However, if any of the constructed degrees of freedom is rounded up to integer, the value of the p-value that is associated with it would be distorted which might result in misleading conclusion. For instance, in the illustrative example in Section 4.1, the degree of freedom of the test statistic for five group comparison is 2.204617 which yielded a p-value of 0.0042 for the computed value 11.4342 of the proposed test statistic indicating strong evidence in support of the rejection of the null hypothesis H_0 in favour of the ordered alternative set H_a in the hypothesis set (37).

If the degree of freedom 2.204617 is rounded up to integer of 2, the corresponding p-value of the test statistic value would be 0.0033 which is relatively smaller than 0.0042 earlier obtained for the test, a difference of 0.0009. Although, these results essentially yielded the same conclusion, but such a difference might at times yield contradictory conclusion in some other data structure.

A thorough comparison of the proposed test statistic with the conventional chi-square test of homogeneity of population proportions was performed for which the test hypotheses set (37) for the proposed test, and hypothesis set (38) for the conventional chi-square test were constructed and tested at 5% Type I error rate.

The goodness of these two tests were compared based on the closeness of their estimated Type I error rates to the nominal Type I error rate of 5% set for each of the tests. The plots of the empirical Type I error rates of the two tests against the various sample sizes as presented by Figure 4(a) showed that the newly proposed test is relatively more efficient than the conventional test. The proposed test statistic yielded efficient estimates of Type I error rates that are closer to the nominal 5% used at all the sample sizes. On the other hand, the conventional Chi-square test statistic yielded poor estimates of the Type I error rate at smaller sample size of 20 before it got stabilized at other higher sample sizes.

In terms of power, the new test statistic proved to be more powerful at smaller sample sizes than the classical Chi-square test statistic as shown in Figure 2 (right graph). The powers of the new test statistic are higher than that of the Chi-square test statistic beginning from sample size 20 up to 100 as shown in Table 2 while they both yielded appreciable powers at relatively higher sample sizes as expected.

More importantly, the goodness of the new proposed test method over the classical Chi-square test is also apparent in its ability to establish the superiority of a new drug type (say) over some of the existing ones with respect to its relative preference by users over others. This point was made very clear by the illustrative hypothetical example given in Section 4.1 in which drug type five was more preferred by 45% of users over other four exiting drugs with $p_5 = 0.45$.

From the results provided in Table 3, the classical Chi-square test of homogeneity of population proportions using the test statistic (39) simply concluded that the preferences for the five drug types are the same across the various users thereby failed to reject the null hypothesis H_0 in (37). Whereas, the proposed test method for ordered hypothesis test of population proportions actually established the superiority of the fifth drug type being the most preferred by users over others by rejecting the null hypothesis in favour of the alternative hypothesis H_a in hypothesis set (37).

CONCLUSION

A novel test statistic for testing hypothesis of equality of population proportions against ordered alternatives with a control is proposed in this work. A suitable test statistics for this hypothesis test problem was developed and its statistical distribution established.

The distribution of the new proposed test statistic was found to be Chi-square but with non-integer degrees of freedom. A table that contained degrees of freedom for k groups' comparisons was provided with $k \leq 26$ groups. For comparing population groups more than 26, an equation was developed to determine the required degrees of freedom for the test statistic.

The proposed test method was compared with the classical Chi-square test of homogeneity of population proportions. Results showed that the classical Chi-square test, unlike the proposed test, suffers from its inability to establish superiority of a population group over others even if such superiority actually exists in a set of population groups.

Not only this, the powers of the proposed test method are quite higher than those provided by the classical chi-square test of homogeneity of population proportions, especially at lower group sample sizes.

Finally, it can be concluded from the results in this work that if determination of superiority of a quantity in terms of its acceptability by people or otherwise is of prime interest as often the case in pharmacology and drug discovery, the test method proposed here could be a viable tool to determine such.

It should be reiterated at this point that, the application of the new test method proposed in this work is not limited to medicine or pharmacology. The method could equally find its use in all other disciplines and scenarios where it is desirable to establish the superiority of one method/system over others in which only qualitative data sets are available.

Acknowledgment

The authors thanked Prof. B. A. Oyejola of Department of Statistics, University of Ilorin for taken time to read through the first draft of this paper and for his useful comments.

REFERENCES

1. Jonckheere AR. A distribution-free k-sample test against ordered alternatives. Biometrika 1954;41(1-2):133-45.
2. Bartholomew DJ. A test of homogeneity for ordered alternatives. II. Biometrika 1959; 46(3-4):328-35.
3. Barlow RE, Bartholomew DJ, Bremmer JM, Brunk HD. Statistical Inference under Order Restrictions: The Theory and Application of Isotonic Regressions. Wiley series in probability and mathematical statistics, no. 8. New York: John Wiley; 1972. p.388.
4. Chacko, VJ. Testing homogeneity against ordered alternatives. Ann Math Statist 1963;34(3):945-56.
5. Adegboye OS, Gupta AK. On testing against restricted alternatives about the means of gaussian models with common unknown variance. Sankhyā: The Indian Journal of Statistics Series B 1986;48(3);331-41.
6. Kwong KS, Cheung SH, Chan WS. Multiple testing to establish superiority/equivalence of a new treatment compared with k standard treatments for unbalanced designs. Biometrics 2004;60(2):491-8.
7. Yahya WB, Jolayemi ET. A note on the test of hypothesis with ordered alternatives. Scientific Annals of "Alexandru Ioan Cuza" 2008;54(1):197-208.
8. Yahya WB, Jolayemi ET. Testing ordered means against a control. Journal of Nigerian Statistical Association 2003;16:40-52.
9. Yahya WB, Jolayemi ET, Ige SO. Superiority test in design and analysis of experiment. International Journal of Pure and Applied Sciences 2006;5(2):1-7.
10. Yahya WB, Ezenweke, CP, Olaniran, OR and Adeniyi, IA. Non-Inferiority Tests Between Two Independent Binomial Proportions: A Comparison of Bayesian and Frequentist Techniques. Proceedings of 39th Annual Conference of the Nigerian Statistical Association, Nigeria, 2015a. http://www. nsang.org/.
11. Yahya WB, Olaniran, OR, Ezenweke, CP, Adeniyi, IA and Garba, MK. Bayesian Non-Inferiority Test Between Two Independent Binomial Population Proportions. Proceed-ings of 34th Annual Conference of the National Mathematical Society (NMS), Nigeria, 2015b. http://nigerianmathematical-society.org/
12. R Core Team. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria, 2014. URL http://www.r-project.org/