JP Journal of Biostatistics
Volume 4, Number 3, 2010, Pages 253-277
Published Online: February 1, 2011
This paper is available online at http://pphmj.com/journals/jpb.htm

ASYMPTOTIC INFERENCES ABOUT A LINEAR COMBINATION OF TWO PROPORTIONS

A. MARTÍN ANDRÉS* **and I. HERRANZ TEJEDOR**

Facultad de Medicina
Universidad de Granada
18071 Granada, Spain
e-mail: amartina@ugr.es

Facultad de Medicina
Universidad Complutense de Madrid
28040 Madrid, Spain
e-mail: iherranz@med.ucm.es

Abstract

The asymptotic inferences about the difference $(d = p_2 - p_1)$ or ratio $R = (p_2/p_1)$ of two proportions $(p_1 \text{ and } p_2)$ are very common in medicine and in applied statistics, in general. Both the cases may be included within the general case of inferences about the parameter $L = p_2 - \alpha - \beta p_1$ (recently interest has been shown about this parameter from the perspective of clinical trials). In this article, the authors review the 12 principal statistics proposed in the relevant literature, propose 15 new ones, group them into families, correct existing errors in the definitions of some and, finally, define and analyze the most desirable

2010 Mathematics Subject Classification: 62F03, 62F05.

Keywords and phrases: confidence intervals, difference of proportions, risk ratio, *z*-test, chi-squared test.

This research was supported by Consejería de Innovación, Ciencia y Empresa, Junta de Andalucía, Spain, grant number FQM-01459, and by Ministerio de Ciencia e Innovación, Spain, grant number MTM-2009-08886 (with cofinancing of the FEDER).

*Corresponding author

Received September 14, 2010

properties they should have. In addition, they offer a simple formula that allows all the cases to be analyzed in a similar fashion, together with a computer program (http://www.ugr.es/local/bioest/Z_LINEAR.EXE) which permits asymptotic tests related to *L* to be carried out.

1. Introduction

Many clinical trials are aimed at contrasting the proportions of successes of a new treatment (p_2) with a standard treatment or a placebo (p_1). When the difference of proportion $(d = p_2 - p_1)$ is the focus of comparison, it is useful to contrast the hypotheses $H_d : d \leq \delta$ vs. $K_d : d > \delta$, where $-1 < \delta < +1$. In particular, when $\delta = 0$, the classic test of superiority is obtained, the aim of which is to prove that the new treatment is superior to the standard one; when $\delta < 0$, the test of non-inferiority is obtained, the aim of which is to prove that the new treatment is not substantially inferior to the standard one [1], where δ is a prespecified small quantity; finally, when $\delta > 0$, this can be used when we wish to prove that the new treatment is substantially superior to the standard one [2], where δ refers here to the minimum difference *d* that is biologically significant. From another point of view, the test for H_d is also useful for obtaining confidence intervals for the parameter *d* when it is inverted [3] (see Section 4). When the ratio of proportions $(R = p_2/p_1)$ is the focus of comparison (as occurs in vaccine efficacy studies), then the test to be used will be $H_R : R \leq \rho$ vs. $K_R : R > \rho$, where $0 < \rho < \infty$ [4]. In this case, values $\rho < 1$, $\rho = 1$ and $\rho > 1$ exercise a similar role to that of the previous values $\delta < 0$, $\delta = 0$ and $\delta > 0$. In the case of *d*, the value of δ is usually set by the regulatory agencies (FDA in USA, CPMP in Europe); but often δ is not a single value, but varies depending on the value of p_1 (for example, $\delta = -0.15$ for $p_1 = 0.85$ and $\delta = -0.10$ for $p_1 = 0.95$). In order to adapt to these circumstances, Phillips [5] proposed assuming that δ is a linear function of p_1, and this led him to suggest the test $H_L : p_2 \leq \alpha + \beta p_1$ vs. $K_L : p_2 > \alpha + \beta p_1$ (in which case, $\alpha = -0.575$ and $\beta = 1.5$ in the previous example). The same occurs in the case of the model *R* if it is assumed that ρ is a linear function of $1/p_1$.

Cases *d* and *R* have been the subject of abundant publications - see the reviews [3, 6, 7] - and there are hundreds of articles in which the various statistics that allow the asymptotical test to be carried out are proposed and/or analyzed. This huge amount of data is the reason that, all too frequently, researchers "rediscover" old methods that were already known. The prime aim of this paper is to offer a panorama of existing statistics while also proposing new ones. Little attention has been paid to case *L*. Nevertheless, it has the added advantage of containing the particular case of the classic cases *d* and *R*, given that their null hypotheses can also be given as $H_d : p_2 \le p_1 + \delta$ and $H_R : p_2 \le \rho p_1$, respectively. As a result, any property of case *L* is immediately transferred to cases *d* and *R*. Thus, the second aim of this article is the detailed analysis of case *L*.

In all of the cases, the border of the parametric space defining the null hypothesis is a function $p_2 = f(p_1) = \alpha + \beta p_1$ of the unknown parameter p_1, which acts as a nuisance parameter which must be estimated by some means. As it is customary to state the above tests under a null hypothesis formed solely by the said border (see Section 5), in the following it is assumed that the null hypotheses to be contrasted are, for each model:

$$\text{Model } d : H_{d(B)} : p_2 = p_1 + \delta, \text{ where } p_1 = \max\{0; -\delta\} \le p_1 \le \min\{1; 1-\delta\} = p_s, \tag{1}$$

$$\text{Model } R : H_{R(B)} : p_2 = \rho p_1, \text{ where } p_1 = 0 \le p_1 \le \min\{1; 1/\rho\} = p_s, \tag{2}$$

$$\text{Model } L : H_{L(B)} : p_2 = \alpha + \beta p_1, \text{ where } p_1 = \max\left\{0; -\frac{\alpha}{\beta}\right\} \le p_1 \le \min\left\{1; \frac{1-\alpha}{\beta}\right\} = p_s, \tag{3}$$

while the alternative hypotheses *K* are those already mentioned $(p_2 > \alpha + \beta p_1,$ in general). In expressions (1) and (2), the possible values for the parameters p_1, $\delta\ (-1 < \delta < +1)$ and $\rho\ (0 < \rho < \infty)$ are the classic ones published [8]. In expression (3), the possible values for the parameters $\beta\ (\beta > 0)$ and $\alpha\ (-\beta < \alpha < 1)$ proceed from the fact that the straight line $p_2 = \alpha + \beta p_1$ must not decrease in p_1 [9] and must cut the rectangle $[0, 1] \times [0, 1]$; moreover, the limits p_1 can be deduced from conditions $0 \le p_i \le 1$. Note that when $\alpha = \delta$ and $\beta = 1$ (or $\alpha = 0$ and $\beta = \rho$) model *L* yields model *d* (or model *R*).

Finally, it must be pointed out that, in order to make the article easier to follow, the main body of this article is devoted to setting out the results, while most of the principal proofs will be found in the Appendices. The Appendices containing the full proofs may be requested from the authors.

2. Estimating the Parameters

During the inference process, it is necessary to estimate the parameters p_1 and p_2. To do so, we require two independent samples of size n_i and their number of successes x_i, where $i = 1, 2$. Let $y_i = n_i - x_i$, $n = n_1 + n_2$, $a_1 = x_1 + x_2$ and $a_2 = n - a_1 = y_1 + y_2$. As $x_i \sim B(n_i;\ p_i)$ are two independent random binomial variables, then the probability of the observed results will be:

$$L = \Pr(x_1,\ x_2 \mid p_1,\ p_2) = C(n_1;\ x_1)C(n_2;\ x_2)\, p_1^{x_1} q_1^{y_1} p_2^{x_2} q_2^{y_2}, \text{ where } q_i = 1 - p_i. \quad (4)$$

If we ignore the null hypothesis, then the estimators of maximum likelihood for p_i are the classic $\overline{p}_i = x_i / n_i$. It is more appropriate to bear the model in mind and carry out the estimation under the condition that $p_2 = \alpha + \beta p_1$, which implies that only the parameter p_1 need to be estimated. From the conditional point of view [1], the only possible values of p_1 are the ones where the total number of successes a_1 remains constant. In this case, $a_1 = n_1 p_1 + n_2 p_2 = n_1 p_1 + n_2(\alpha + \beta p_1)$ and so the conditioned estimators for the p_i values will be $\tilde{p}_1 = (a_1 - n_2\alpha)/(n_1 + n_2\beta)$ and $\tilde{p}_2 = \alpha + \beta\tilde{p}_1$. The estimators $\tilde{p}_1 = (a_1 - n_2\delta)/n$ in the case d and $\tilde{p}_1 = a_1/(n_1 + n_2\rho)$ in the case R were proposed by Dunnett and Gent [1] and Farrington and Manning [10], respectively. In all the cases, we should proceed with logical caution: when $\tilde{p}_1$ is not an allowed value - because it does not verify expressions (1) to (3) - then it must be made equal to $p_I\,(p_s)$ if $\tilde{p}_1 < p_I$ $(\tilde{p}_1 > p_s)$. Note that $\tilde{p}_1$ is the solution to the equation:

$$n_1(\overline{p}_1 - p_1) + n_2(\overline{p}_2 - p_2) = 0, \quad \text{where} \quad p_2 = \alpha + \beta p_1. \quad (5)$$

From the unconditional viewpoint - Mee [11] in the case d; Koopman [12] in the case R - the estimators $\hat{p}_i$ are obtained, which are more complicated but more efficient in the inference. Now, $\hat{p}_1$ is the value of p_1 which verifies the equality:

$$d \log L/dp_1 = h(p_1) = n_1(\overline{p}_1 - p_1)/p_1 q_1 + \beta n_2(\overline{p}_2 - p_2)/p_2 q_2 = 0,$$

$$\text{where } p_2 = \alpha + \beta p_1, \quad (6)$$

although, when $h(p_1) \neq 0$ in $p_I \leq p_1 \leq p_s$, then $\hat{p}_1 = p_I$ if $h < 0$ or $\hat{p}_1 = p_s$ if

$h > 0$. In Appendix A, it is proven that $\hat{p}_1 = \{-c_2 + 2B^{0.5}\cos\varphi\}/3c_3$, where $c_0 = \alpha(1-\alpha)x_1$, $c_1 = \beta a_1 - n_1\alpha(1-\alpha) - \alpha\beta(n_2 + 2x_1)$, $c_2 = \beta[(n+n_1)\alpha - (n_1 + x_2) - \beta(n_2 + x_1)]$, $c_3 = n\beta^2$, $B = c_2^2 - 3c_1c_3$, $\varphi = [\pi + \cos^{-1}(-A/B^{3/2})]/3$ and $A = 4.5c_3(c_1c_2 - 3c_0c_3) - c_2^3$. In particular, when $\beta = 1$ (model *d*) the value $\hat{p}_1$ given by Miettinen and Nurminen [8] is obtained. The expressions are simpler when $\alpha = 0$ (model *R*); now, $\hat{p}_1 = [b - \{b^2 - 4na_1\rho\}^{0.5}]/2n\rho$, where $b = (n - y_2) + (n - y_1)\rho$.

One consequence is that, according to expression (6), the quantities $(\bar{p}_1 - \hat{p}_1)$ and $(\bar{p}_2 - \hat{p}_2)$ should have opposing signs; this is also true for $\tilde{p}_i$ because of expression (5). Hence, if $\bar{p}_1 \le \hat{p}_1$, then $\bar{p}_2 \ge \hat{p}_2 = \alpha + \beta\hat{p}_1$, that is $\hat{p}_1 \le f^{-1}(\bar{p}_2)$. Thus, $\bar{p}_1 \le \hat{p}_1 \le f^{-1}(\bar{p}_2)$ and $\bar{f}_1 = f(\bar{p}_1) \le \hat{p}_2 \le \bar{p}_2$. As this argument also works the other way round, then:

$$\bar{p}_2 \ge \bar{f}_1 = \alpha + \beta\bar{p}_1 \Leftrightarrow \bar{p}_1 \le \hat{p}_2 \le f^{-1}(\bar{p}_2) = -(\alpha/\beta) + (1/\beta)\,\bar{p}_2 \Leftrightarrow \bar{f}_1 \le \hat{p}_2 \le \bar{p}_2 \quad (7)$$

which, for models *d* and *R* becomes $\bar{p}_2 \ge \bar{p}_1 + \delta \Leftrightarrow \bar{p}_1 \le \hat{p}_1 \le \bar{p}_2 - \delta \Leftrightarrow \bar{p}_1 + \delta \le \hat{p}_2 \le \bar{p}_2$ and $\bar{p}_2 \ge \rho\bar{p}_1 \Leftrightarrow \bar{p}_1 \le \hat{p}_1 \le \bar{p}_2/\rho \Leftrightarrow \rho\bar{p}_1 \le \hat{p}_2 \le \bar{p}_2$, respectively.

Appendix A contains the proof also that $\hat{p}_1$ increases with $\bar{p}_i$, decreases with α and decreases with β (if $\alpha \le 0$); it also decreases (increases) with n_1 (n_2) when $\bar{p}_2 \ge \bar{f}_1$. In particular, $\hat{p}_1$ will decrease with δ (model *d*) and ρ (model *R*). The properties of $\hat{p}_1$ (model *d*) are all well-known from the published literature [2, 13]. In all the cases, it is to be understood that neither the increase nor the decrease is strict.

3. Type *z* Statistics

3.1. Classic statistics

Because contrasting $H_{L(B)}: p_2 = \alpha + \beta p_1$ is equivalent to contrasting $p_2 - \alpha - \beta p_1 = 0$, the contrast statistic will be $\bar{p}_2 - \alpha - \beta\bar{p}_1$. As its asymptotic mean is 0 (under $H_{L(B)}$) and its asymptotic variance is $V(\bar{p}_2) + V(\bar{p}_1)\beta^2 = p_2q_2/n_2 +$

$\beta^2 p_1 q_1/n_1$, then the classic statistic z to be compared with value $z_{1-\gamma}$ (the $1-\gamma$ percentage point of a standard normal distribution) is the square root of the statistic [5]:

$$z_L^2 = (\bar{p}_2 - \alpha - \beta\bar{p}_1)^2 / \{p_2(1-p_2)/n_2 + \beta^2 p_1(1-p_1)/n_1\}. \tag{8}$$

In particular, for models d and R, we obtain the statistics z_d^2 and z_R^2 of Dunnett and Gent [1] and Katz et al. [14], respectively.

In order to carry out the test, it is necessary to estimate the parameters p_i of expression (8). When the estimators $\bar{p}_i$ are used, we obtain the statistic [5]:

$$\bar{z}_L^2 = (\bar{p}_2 - \alpha - \beta\bar{p}_1)^2 / \{\bar{p}_2\bar{q}_2/n_2 + \beta^2 \bar{p}_1\bar{q}_1/n_1\}, \tag{9}$$

whose particular cases $\bar{z}_d^2$ and $\bar{z}_R^2$ were first given by Dunnett and Gent [1] and Katz et al. [14], respectively, although they were rediscovered by Laster et al. [15]. It is better to substitute p_i with $\tilde{p}_i$ in expression (8), in which case the new statistic $\tilde{z}_L^2$ is obtained. Its particular case $\tilde{z}_d^2$ was given by Dunnett and Gent [1] and rediscovered by Wallenstein [16] and Parmet and Schechtman [17]. For model R, the statistic $\tilde{z}_R^2$ [10] is obtained which, when $\tilde{p}_i$ is allowed, is simplified to:

$$\tilde{z}_R^2 = n_1 n_2 (\bar{p}_2 - \rho\bar{p}_1)^2 / [\rho a_1\{1 - \rho n a_1/(n_1 + n_2\rho)^2\}]. \tag{10}$$

It is even better to substitute p_i with $\hat{p}_i$ in expression (8), so obtaining the new statistic $\hat{z}_L^2$. Its particular cases are the statistics $\hat{z}_d^2$ of Mee [11] - rediscovered by Parmet and Schechtman [17] - and $\hat{z}_R^2$ of Miettinen and Nurminen [8]. Note that statistic $\bar{z}_L^2$ is known, but the statistics $\tilde{z}_L^2$ and $\hat{z}_L^2$ are new, with all three having the advantage of containing as particular cases the statistics for the cases d and R (that are already known). Something similar occurs with the other statistics indicated below.

Moreover, the classic Pearson chi-squared statistic (which is obtained by using the score method) can be written as (see Appendix B):

$$\chi_L^2 = \frac{(x_1 - n_1 p_1)^2}{n_1 p_1(1-p_1)} + \frac{(x_2 - n_2 p_2)^2}{n_2 p_2(1-p_2)}, \tag{11}$$

a quantity to be compared with the value $\chi^2_{1-2\gamma}$ (the $1 - 2\gamma$ percentage point of a chi-squared distribution with one degree of freedom). This new statistic contains the particular case of the statistics χ^2_d given by Dunnett and Gent [1] and χ^2_R given by Koopman [12]. As in case z^2_L, in order to carry out this test, it is necessary to estimate p_i, which gives rise to the statistics $\tilde{\chi}^2_L$ and $\hat{\chi}^2_L$ by substituting p_i with $\tilde{p}_i$ or $\hat{p}_i$, respectively ($\overline{\chi}^2_L$ is excluded because $\overline{\chi}^2_L = 0$). In particular, through expression (5):

$$\tilde{\chi}^2_L = (x_1 - n_1\tilde{p}_1)^2\left\{\frac{1}{n_1\tilde{p}_1\tilde{q}_1} + \frac{1}{n_2\tilde{p}_2\tilde{q}_2}\right\}$$

$$= (x_1 - n_1\tilde{p}_1)^2\left\{\frac{1}{n_1\tilde{p}_1} + \frac{1}{n_1\tilde{q}_1} + \frac{1}{n_2\tilde{p}_2} + \frac{1}{n_2\tilde{q}_2}\right\}. \tag{12}$$

Case $\tilde{\chi}^2_d$, in its longest format, was first given by Dunnett and Gent [1]. Cases $\hat{\chi}^2_d$ and $\hat{\chi}^2_R$ were the work of Nam [18] and Koopman [12], respectively.

A point of interest is the possible relation between the type z^2 and type χ^2 statistics. Dunnett and Gent [1] found experimentally that $\tilde{\chi}^2_d \neq \tilde{z}^2_d$; in general, it also happens that $\tilde{\chi}^2_L \neq \tilde{z}^2_L$. However (see Appendix B), the following equality does occur $\hat{\chi}^2_L = \hat{z}^2_L$. This equality was first proved by Nam [18] and rediscovered by Andrés and Tejedor [2] in case *d* and by Gart and Nam [19] in case *R*. In Appendix B, it is also proved that $\hat{\chi}^2_L$ can be written in the following alternative formats:

$$\hat{\chi}^2_L = (\hat{p}_1 - \overline{p}_1)^2/\sigma^2_1 + (\overline{p}_2 - \hat{p}_2)^2/\sigma^2_2 = (\overline{p}_2 - \bar{f}_1)^2/(\sigma^2_2 + \beta^2\sigma^2_1) = \hat{z}^2_L \tag{13}$$

$$= (\hat{p}_1 - \overline{p}_1)(\overline{p}_2 - \bar{f}_1)/\beta\sigma^2_1 = (\overline{p}_2 - \hat{p}_2)(\overline{p}_2 - \bar{f}_1)/\sigma^2_2$$

$$= (\hat{p}_1 - \overline{p}_1)(\overline{p}_2 - \hat{p}_2)\left\{\frac{\beta}{\sigma^2_2} + \frac{1}{\beta\sigma^2_1}\right\} \tag{14}$$

$$= (\hat{p}_1 - \overline{p}_1)^2(1 + \sigma^2_2/\beta^2\sigma^2_1)/\sigma^2_1 = (\overline{p}_2 - \hat{p}_2)^2(1 + \beta^2\sigma^2_1/\sigma^2_2)/\sigma^2_2, \tag{15}$$

where $\sigma^2_1 = \hat{p}_1(1 - \hat{p}_1)/n_1$, $\sigma^2_2 = \hat{p}_2(1 - \hat{p}_2)/n_2$, $\hat{p}_2 = \alpha + \beta\hat{p}_1$ and $\bar{f}_1 = \alpha + \beta\overline{p}_1$.

For cases d and R both, expressions (13) have already been referred to above. For case d, the first two expressions (14) are the work of Andrés and Tejedor [2]. For case R, the first expression (15) is owed to Nam [18] and the second to Gart and Nam [19].

When expressions (13) to (15) are particularized to models d and R, some simplifications are obtained. In case d, the relevant simplification is that $\bar{p}_2 - \bar{f}_1 = \bar{d} - \delta$, where $\bar{d} = \bar{p}_2 - \bar{p}_1$. In particular, the first two expressions of (14) are the cause of:

$$\hat{\chi}_d^2 = \hat{z}_d^2 = \frac{n_1(\hat{p}_1 - \bar{p}_1)(\bar{d} - \delta)}{\hat{p}_1(1 - \hat{p}_1)} = \frac{n_2(\bar{p}_2 - \hat{p}_1 - \delta)(\bar{d} - \delta)}{(\hat{p}_1 + \delta)(1 - \hat{p}_1 - \delta)}. \quad (16)$$

In case R, the second expression in (13) and expressions (15) give rise to:

$$\hat{\chi}_R^2 = \hat{z}_R^2 = n_1 n_2 (\bar{p}_2 - \rho\bar{p}_1)^2 / \rho\hat{p}_1(n_1 + n_2\rho - n\rho\hat{p}_1) \quad (17)$$

$$= \frac{(x_1 - n_1\hat{p}_1)^2}{n_1\hat{p}_1(1 - \hat{p}_1)}\left\{1 + \frac{n_1(1 - \rho\hat{p}_1)}{n_2\rho(1 - \hat{p}_1)}\right\} = \frac{(x_2 - n_2\hat{p}_1)^2}{n_2\rho\hat{p}_1(1 - \rho\hat{p}_1)}\left\{1 + \frac{n_2\rho(1 - \hat{p}_1)}{n_1(1 - \rho\hat{p}_1)}\right\}. \quad (18)$$

Note that the classic test for comparing two proportions $(H : p_2 = p_1)$ is obtained from model L making $\alpha = 0$ and $\beta = 1$. Now, $\hat{p}_1 = \tilde{p}_1 = a_1/n$ and $\tilde{z}_L^2 = \hat{z}_L^2 = \tilde{\chi}_L^2 = \hat{\chi}_L^2 = n(x_1 y_2 - x_2 y_1)^2/(a_1 a_2 n_1 n_2)$, with which the classic test is obtained.

3.2. Statistics based on the logarithmic transformation

Inferences about R are frequently made using the statistic $\log \bar{R}$. If we do the same for model L, then we have to think of the null hypothesis $\log p_2 - \log(\alpha + \beta p_1) = 0$. The contrast statistic will now be $\log \bar{p}_2 - \log(\alpha + \beta\bar{p}_1)$, whose asymptotic mean is 0 (under $H_{L(B)}$) and whose asymptotic variance is $(q_2/n_2 p_2) + [\beta^2 p_1 q_1 / n_1(\alpha + \beta p_1)^2]$. This yields the following statistic:

$$Lz_L^2 = [\log\{\bar{p}_2/(\alpha + \beta\bar{p}_1)\}]^2 / \{q_2/n_2 p_2 + \beta^2 p_1 q_1 / n_1(\alpha + \beta p_1)^2\}. \quad (19)$$

On substituting p_i with $\bar{p}_i$, $\tilde{p}_i$ or $\hat{p}_i$, the statistics $L\bar{z}_L^2$, $L\tilde{z}_L^2$ and $L\hat{z}_L^2$ are obtained, respectively. The case of most interest is that of model R, in which:

$$Lz_R^2 = \{\log(\overline{R}/\rho)\}^2 \Big/ \left(\frac{1}{n_1 p_1} + \frac{1}{n_2 p_2} - \frac{n}{n_1 n_2}\right) = \{\log(\overline{R}/\rho)\}^2 \Big/ \left(\frac{1-p_2}{n_2 p_2} + \frac{1-p_1}{n_1 p_1}\right), \quad (20)$$

which yields the statistics $L\tilde{z}_R^2$, $L\hat{z}_R^2$ and the classic statistic of Woolf [20]:

$$L\bar{z}_R^2 = \{\log(\overline{R}/\rho)\}^2 / \{(1/x_1) + (1/x_2) - (n/n_1 n_2)\}. \quad (21)$$

3.3. Statistics based on the Sterne method

Another way of making the inference is by using the method of Sterne [21]. So, Peskun [22] indicated that the test for $H_{d(B)}: p_2 = p_1 + \delta$ will be significant to the error γ when it is so for the whole value of p_1. This means that $z_d^2 \geq z_{1-\gamma}^2$ in all possible values of p_1, which means working with the statistic $Mz_d^2 = \min_{p_1} z_d^2$. This argument may be extended to any statistic, in particular, to the statistic z_L^2. In Appendix B, it is proven that if $p_0 = (n_1 + n_2\beta - 2n_1\alpha)/(2n\beta)$:

$$Mz_L^2 = \begin{cases} z_L^2 \text{ for } p_1 = p_I \text{ (or } p_1 = p_s), & \text{if } p_0 < p_1 \text{ (or } p_0 > p_s), \\ 4nn_1n_2(\bar{p}_2 - \alpha - \beta\bar{p}_1)^2 / \{(n_1 + n_2\beta)^2 + 4n_1n_2\alpha(1-\alpha-\beta)\}, & \text{otherwise,} \end{cases} \quad (22)$$

$$Mz_d^2 = \begin{cases} (\min_i n_i)(\bar{d} - \delta)^2 / |\delta|\{1 - |\delta|\}, & \text{if } |\delta| > n/2(\max_i n_i), \\ 4nn_1n_2(\bar{d} - \delta)^2/(n^2 - 4n_1n_2\delta^2), & \text{if } |\delta| \leq n/2(\max_i n_i), \end{cases} \quad (23)$$

$$Mz_R^2 = \begin{cases} n_2(\bar{p}_2 - \rho\bar{p}_1)^2 / \{\rho(1-\rho)\}, & \text{if } \rho < n_1/(n + n_1), \\ n_1(\bar{p}_2 - \rho\bar{p}_1)^2/(\rho - 1), & \text{if } \rho > (n + n_2)/n_2, \\ 4nn_1n_2\{(\bar{p}_2 - \rho\bar{p}_1)/(n_1 + n_2\rho)\}^2, & \text{otherwise.} \end{cases} \quad (24)$$

For Peskun [22], the value Mz_d^2 is that of the second expression in (23): but this is incorrect, because the value $p_0 = (n - 2n_1\delta)/(2n)$ in which this minimum is calculated is not always an allowed value of p_1. Feigin and Lumelskii [23] made the same error when they rediscovered the procedure.

3.4. Continuity correction and Agresti type statistics

One interesting point is whether it is right [24] to carry out a continuity correction when the distribution of a discrete random variable (such as the x_i) is approached through that of a continuous random variable (such as the normal one). From the viewpoint of conditional inference, it is usual to carry out Yates' classic correction. This implies [25] substituting the term $(x_1 - n_1\tilde{p}_1)^2$ in expression (12) with the term $\{|x_1 - n_1\tilde{p}_1| - 0.5\}^2$. From the unconditional viewpoint, no corrections are usually performed, although Andrés and Tejedor [26] advise using a rather slight correction in case z_d^2. Their argument, applied to models *L*, *d* and *R* (see Appendix C), leads them to propose that the correction should consist of substituting the numerators $\bar{p}_2 - \alpha - \beta\bar{p}_1$, $\bar{d} - \delta$ and $\bar{p}_2 - \rho\bar{p}_1$ of the statistics z^2 or Mz^2 with the quantities $(\bar{p}_2 - \alpha - \beta\bar{p}_1) - (1+\beta)/2N$, $(\bar{d} - \delta) - 1/N$ and $(\bar{p}_2 - \rho\bar{p}_1) - (1+\rho)/2N$, respectively, where $N = n + n_1 n_2$. It can so be seen that the correction is irrelevant when the n_i values are not excessively small.

A different point is the custom of adding a constant *h* to all the outcomes when using a statistic based on $\bar{p}_i$. In these cases, the statistic is obtained by exchanging x_i and n_i for $(x_i + h)$ and $(n_i + 2h)$, respectively. The value $h = 0.5$ is customary [14] in the statistic $L\bar{z}_R^2$, so obtaining the new statistic $L\bar{z}_R^2(+0.5)$. The value $h = 1$ is newer and is applied to the statistics $\bar{z}_d^2$ [27] and $L\bar{z}_R^2$ [7] in order to obtain the statistics $\bar{z}_d^2(+1)$ and $L\bar{z}_R^2(+1)$, respectively.

3.5. Excluded statistics

The above list of statistics is not exhaustive, but it includes the ones that are most relevant. Other options are: (a) the likelihood ratio statistic for models *d* and *R* [1, 28-31], whose general expression for model *L* is $-2[x_1 \log(\hat{p}_1/\bar{p}_1) + y_1 \log(\hat{q}_1/\bar{q}_1) + x_2 \log(\hat{p}_2/\bar{p}_2) + y_2 \log\{\hat{q}_2/\bar{q}_2\}]$; (b) the statistic based on the arc sine transformation for model *d* [31]; (c) the statistic z^2 with a correction for skewness [19, 32, 33]; (d) the statistic z^2 based on an unbiased estimator of $p_1 q_1$, rather than the one $(\hat{z}_L^2)$ based on an unbiased estimator of p_1; and (e) the statistic based on the individual confidence intervals for p_i [34].

4. Tests and Confidence Intervals

The statistics in the previous section were defined on the basis of the null hypothesis $H_{L(B)} : p_2 = \alpha + \beta p_1$. In order to determine the *p*-value of the observed value, it is necessary to decide if the alternative hypothesis is $K_L : p_2 > \alpha + \beta p_1$ (that of Section 1), $K'_L : p_2 < \alpha + \beta p_1$ or $K''_L : p_2 \neq \alpha + \beta p_1$. If S_L^2 refers to any of the statistics in this paper and $S_L = [\mathrm{sign}(\bar{p}_2 - \alpha - \beta\bar{p}_1)]\sqrt{\mathrm{S}_\mathrm{L}^2}$, then the *p*-values are, respectively,

$$P_L = \Pr\{z \geq S_L\}, \quad P'_L = \Pr\{z \leq S_L\} \text{ and } P''_L = 2\Pr\{z \geq |S_L|\}, \tag{25}$$

where *z* refers to the standard normal random variable.

When the aim is to obtain confidence intervals for the difference of proportions $d = p_2 - p_1$, then it is sufficient to invert the opportune test [4]. In the case that concerns us here, this implies obtaining the two allowed solutions (δ_L and δ_U) to the equation $z_{1-\gamma/2}^2 = S_d^2$. Then $d \leq \delta_U$, $d \geq \delta_L$ or $\delta_L \leq d \leq \delta_U$ will be the confidence intervals for one-tail (right), for one-tail (left) or for two-tails, for the errors $\gamma/2$, $\gamma/2$ or γ, respectively. The same applies if the interval is required for $R = p_2/p_1$. Determining the two values for δ (or for ρ) which verify equation $z_{1-\gamma/2}^2 = S_d^2$ (or S_R^2) may be more or less complex depending on the degree of the polynomial to be solved. In elementary textbooks, it is usual to chose the simplest solution (of degree 1) given by the statistics $\bar{z}_d^2$ and $L\bar{z}_R^2$, that is $d \in \bar{p}_2 - \bar{p}_1 \pm z_{1-\gamma/2}\sqrt{\bar{p}_1\bar{q}_1/n_1 + \bar{p}_2\bar{q}_2/n_2}$ and $R \in (\bar{p}_2/\bar{p}_1)\exp\{\pm z_{1-\gamma/2}\sqrt{(1/x_1) + (1/x_2) - (n/n_1n_2)}\}$.

The statistics $\tilde{z}_d^2$ [16, 35], Mz_d^2 [20] and Mz_R^2 yield an explicit, simple solution (of degree 2). In particular, the first statistic yields the solution (*c* is the continuity correction; make $c = 0$ if one does not want this) $\delta \in -(B/2A) \pm \{(B/2A)^2 - (C/A)\}^{0.5}$, where $A = z^2_{1-\gamma/2}[(n_2 - n_1)^2 + n_1 n_2] + nn_1 n_2$, $B = z^2_{1-\gamma/2}(n_2 - n_1)(a_2 - a_1) - 2nn_1 n_2(\bar{d} \pm c)$ and $C = nn_1 n_2(\bar{d} \pm c)^2 - z^2_{1-\gamma/2} a_1 a_2$.

The statistics which behave best (see Section 6) yield more complex solutions: of degree 3 in the case of $\hat{z}^2_R$ [18] and of degree 4 in the case of $\hat{z}^2_d$. In Appendix D, it can be seen that, when $\bar{p}_2 \neq \alpha + \beta\bar{p}_1$, then the outcomes x_i and n_i, the parameters α and β and the statistic $\hat{z}^2_L$ are directly related (without the need for the intermediate calculation of the value $\hat{p}_1$) using the equality:

$$y = \beta R_1 + R_2 - n\beta + (2\alpha + \beta - 1)\lambda = 0 \tag{26}$$

with $\lambda = \beta\hat{z}^2_L/(\bar{p}_2 - \alpha - \beta\bar{p}_1) \neq 0$, $R^2_1 = \lambda^2 + n^2_1 - 2(n_1 - 2x_1)\lambda$ and $R^2_2 = \lambda^2 + n^2_2\beta^2 + 2\beta(n_2 - 2x_2)\lambda$. This means that expression (26) may be used to determine the unique solution $\hat{z}^2_L \neq 0$ for it or (making $\hat{z}^2_L = z^2_{1-\gamma/2}$) the two allowed solutions α_I and α_S (or β_I and β_S) which permit the two-tailed CI for α (or for β) to the error γ to be obtained. This equation may also be used for carrying out the test simply, without requiring the calculation of $\hat{p}_1$ (see Appendix D). All the above is also valid if we wish to work with the statistic $\hat{z}^2_{Lc}$ with any continuity correction *c*; for this, we only need to make $\lambda = \beta\hat{z}^2_{Lc}\{(\bar{p}_2 - \alpha - \beta\bar{p}_1)/(|\,\bar{p}_2 - \alpha - \beta\bar{p}_1\,| - c)\}^2$ in expression (26).

5. Conditions to be Verified by any Statistic

For a statistic S^2_L to be useful in inference, it has to verify certain coherence properties. Thus, in the context of the exact tests for $H_{d(B)}$ [9], it is necessary for the critical regions not to have holes. Because the critical region is constructed by ordering the points of the sample space from the higher to lower value of a statistic S_d in case K_d, the absence of gaps implies that S_d should increase (decrease) in $\bar{p}_2(\bar{p}_1)$. This is what is known as Barnard's two convexity properties (which we shall call *spatial convexity* here). The argument may be extended to the case of the alternative K_L and to the case of asymptotic tests. In Appendix E, it is proven that the statistics $\bar{z}_L, \tilde{z}_L, \hat{z}_L = \hat{\chi}_L, \tilde{\chi}_L$ and Mz_L verify these properties, while the statistic $L\bar{z}_R$ only verifies the first (it increases with $\bar{p}_2$), for which reason it will not yield a coherent inference. The convexity spatial of $\hat{z}_d$ was initially proved heuristically [4] and then more exactly [2, 36], and it is also true when a continuity correction is performed on $\hat{z}_d$ [31].

Moreover, all the statistics S_L^2 have been obtained under the null hypothesis $H_{L(B)} : p_2 = \alpha + \beta p_1$. However, the real null hypothesis is $H_L : p_2 \leq \alpha + \beta p_1$. From Sterne's principle [21], for S_L to be allowable, it is necessary that it reaches its minimum value on the border of H_L, that is, it must decrease with α (property of *parametric convexity*). It is known [2] that $\hat{z}_d$ verifies this property, and it is proven in Appendix E that the statistics $\bar{z}_L, \tilde{z}_L, \tilde{\chi}_L, \hat{z}_L = \hat{\chi}_L, Mz_L$ and $L\bar{z}_R$ also verify it. Parametric convexity is essential if the confidence interval for the parameter of interest is going to be obtained by resolving the equation $z_{1-\gamma/2}^2 = S_L^2$ rather than the equation $z_{1-\gamma/2}^2 \geq S_L^2$. When the alternative is $K_L' : p_2 < \alpha + \beta p_1$, the statistic will be $S_L' = -S_L$ and its convexity properties are the opposite of those in case $K_L : p_2 > \alpha + \beta p_1$.

The above reasoning is strictly valid only when the variance of the contrast statistic $\bar{p}_2 - \alpha - \beta \bar{p}_1$ has not been obtained under $H_{L(B)}$, something which does not occur with $\hat{z}_L^2$. This is because it makes no sense to apply the principle of maximum likelihood when determining $\hat{z}_L^2$ followed by Sterne's principle when validating $\hat{z}_L^2$ for the whole null hypothesis. This means that $\hat{z}_L^2$ should be defined as the value of z_L^2 in expression (8) with p_i estimated by maximum likelihood within the H_L region. At the end of Appendix A, it is proven that the estimator of maximum likelihood of p_i is $\bar{p}_i$ when $\bar{p}_2 \leq \bar{f}_1$ and $\hat{p}_i$ when $\bar{p}_2 > \bar{f}_1$; as a result the required statistic $(\hat{\hat{z}}_L^2)$ takes the value $\bar{z}_L^2$ in the first case and the value $\hat{z}_L^2$ in the second case. However, this has no relevance in practical terms because, when $\hat{p}_2 \le \hat{f}_1$, the *p*-value P_L in expression (25) will be > 0.5.

Let us look at the classic hypothesis $H_{L(B)} : p_2 = \alpha + \beta p_1$ again, which we shall note down for the moment as $H_O : p_2 = \alpha + \beta p_1$ in order to remind ourselves that it is applied to the original outcomes. If the samples, the p_i and q_i or both together are permuted, then the equivalent hypotheses to H_O for the new presentation of outcomes are $H_S : p_1 = -(\alpha/\beta) + (1/\beta)\, p_2$, $H_C : q_2 = (1 - \alpha - \beta) + \beta q_1$ and $H_{SC} : q_1 = (\alpha + \beta - 1)/\beta + (1/\beta)\, q_2$, respectively, all of which verify model *L*. It appears logical to expect that any statistic S_L^2 takes the same value in all four cases (properties of *equivalence*). This is what occurs with all the statistics described so far, except for Lz_L^2. However, Lz_R^2 does take the same value for H_O and H_S.

6. Selection of the Best Statistic Based on Published Results

Many of the statistics contemplated in this article have been proposed in the context of the hypothesis tests, in the context of the confidence intervals or in both contexts. This has resulted in the various authors comparing different groups of statistics from one perspective or the other. When the comparison is made from the perspective of the hypothesis tests, the authors focus on the power for each test and on the difference $\gamma - \gamma^*$ between the objective (γ) and real errors (γ^*). In the case of the confidence intervals, the centre of attention is the length and the difference $\gamma - \gamma^*$ between its real $(1 - \gamma^*)$ and objective coverage $(1 - \gamma)$. But, in essence, both approaches are the same. In the first, because by inverting a hypothesis tests a confidence interval [3] is obtained, and also the other way round [37]. In the second place (and related to the above), because the more powerful a test is, the lesser the length of the confidence interval it induces. As a result, the following conclusions published in the relevant literature are analyzed without differentiating between the origin of the comparisons (tests or intervals).

Case *d* is the one that has been most analyzed in the literature. Different studies confirm that the statistic $\hat{z}_d^2$ is the best of all, because it is better than $\bar{z}_d^2$ [38, 39], than $\tilde{z}_d^2$ [38] and many others [34]. An exception to this is Wallenstein's conclusion [16] that $\tilde{z}_d^2$ is better than $\hat{z}_d^2$ when the expected quantities are higher than 2. Because the calculation of $\hat{z}_d^2$ is complicated, several authors have concentrated on other, simpler statistics. From a conditional point of view, it is known that $\tilde{\chi}_d^2$ is better than $\tilde{z}_d^2$ and $\bar{z}_d^2$ [1], all of which are calculated with Yates' continuity correction. From a more general point of view, it is known that $\tilde{z}_d^2$ is better than Mz_d^2 [16], which in turn, is better than $\bar{z}_d^2$ [22] and $\bar{z}_d^2(+0.5)$ [23]. The final conclusion is to use the statistic $\hat{z}_d^2$ or, if a simpler one is wanted, then the statistic $\tilde{z}_d^2$. If we want an even simpler statistic, $\bar{z}_d^2(+1)$ [27] may be used for error $\gamma = 5\%$, but never must the statistic $\bar{z}_d^2$ be used without some correction [40, 41].

There are fewer results for case *R*. At the moment it is only known that $\hat{z}_R^2$ is better than $\tilde{z}_R^2$ [10] and than $L\bar{z}_R^2$ [12], which in turn, is better than $\bar{z}_R^2$ [14], while $L\bar{z}_R^2(+1)$ acts quite well [7]. It seems clear that the best option is $\hat{z}_R^2$, although if we are looking for a simpler option, we can use $L\bar{z}_R^2(+1)$ (which, as was pointed out in the previous section, has the disadvantage of not verifying all the properties of convexity and equivalence).

7. Conclusions

The asymptotic inferences about the difference $d = p_2 - p_1$ or the ratio $R = p_2/p_1$ of two proportions p_i are very common in applied statistics, in general. In recent years [5], the aim has been generalized to effecting inference on the difference $L = p_2 - \alpha - \beta p_1$. In order to realize these inferences, it is necessary to define a statistic of test and estimate the nuisance parameter p_1. In this context, the relevant literature has offered 12 different statistics for the cases *d*, *R* or L ($\bar{z}_L^2$, $\bar{z}_d^2$, $\bar{z}_R^2$, $\tilde{z}_d^2$, $\tilde{z}_R^2$, $\hat{z}_d^2$, $\hat{z}_R^2$, $\tilde{\chi}_d^2$, $\hat{\chi}_d^2$, $\hat{\chi}_R^2$, $L\bar{z}_R^2$ and Mz_d^2), complementing these occasionally with different corrections. In this paper, 15 new statistics are defined ($\tilde{z}_L^2$, $\hat{z}_L^2$,

$\tilde{\chi}_L^2$, $\tilde{\chi}_R^2$, $\hat{\chi}_L^2$, $L\bar{z}_L^2$, $L\bar{z}_d^2$, $L\tilde{z}_L^2$, $L\tilde{z}_d^2$, $L\tilde{z}_R^2$, $L\hat{z}_L^2$, $L\hat{z}_d^2$, $L\hat{z}_R^2$, Mz_L^2 and Mz_R^2), errors of definition for some of them are corrected, they are grouped into families, their desirable properties are defined and analyzed and, finally, the optimum statistics for the cases d and R are selected (the latter on the basis of published results). In particular, it is proven that almost all the statistics verify the obligatory properties of convexity in $\bar{p}_2$, in $\bar{p}_1$ and in the parameter which is the object of the inference, with the notable exception of the statistic based on the logarithmic transformation (which is not convex in $\bar{p}_1$). Similarly, it is proven that almost all the statistics verify the desirable properties of equivalence (when the samples and/or the feature being studied are permuted, the absolute value of the statistic does not vary), again with the exception of the case of logarithmic transformation mentioned above.

To give a single conclusion for the three cases (d, R and L), it may be said that the best statistic for effecting the asymptotic inference is $\hat{z}^2$ (classic z based on the estimator of maximum likelihood for the nuisance parameter p_1). As this usually produces complex expressions, the researcher has three options: to use a computer program (like the one given at http://www.ugr.es/local/bioest/Z_LINEAR.EXE) or to apply the simplest methods recommended in this article.

Acknowledgement

The authors would like to thank María Álvarez Hernández for suggestions that led to the improvement of this paper.

Appendices

A. Estimator of maximum likelihood of nuisance parameter p_1

Function $h(p_1)$ of expression (6) always decreases because $\partial h/\partial p_1 = -\Sigma n_i \{\bar{p}_i\bar{q}_i + (p_i - \bar{p}_i)^2\}/p_i^2 q_i^2 \le 0$. Moreover, $h(p_1)$ has a horizontal asymptote in $h = 0$ and, if $0 < \bar{p}_i < n_i$, then $h(p_1)$ also has four vertical asymptotes in $p_1 = 0$, 1, $-\alpha/\beta$ and $(1-\alpha)/\beta$, respectively. Consequently, $h(p_1) = 0$ on three occasions and $\hat{p}_1$ is the second of them due to expression (3). By multiplying expression (6) by $p_1(1 - p_1)\, p_2(1 - p_2)$ the cubic equation in Section 2 is obtained, the second solution of which is the value $\hat{p}_1$ given in Section 2. When $\alpha = 0$, one of the vertical asymptotes disappears. The effect is that the equation is reduced to a second degree equation of type $n\beta p_1^2 - \{(n - y_2) + (n - y_1)\beta\} p_1 + a_1 = 0$, the solution for which is in Section 2.

With the aim of studying the properties of $\hat{p}_L$, we must calculate $\partial h/\partial \overline{p}_1 = n_1/p_1q_1 \ (\geq 0)$, $\partial h/\partial \overline{p}_2 = n_2\beta/p_2q_2 \ (\geq 0)$, $\partial h/\partial\alpha = -n_2\beta\{\overline{p}_2\overline{q}_2 + (\overline{p}_2 - p_2)^2\}/p_2q_2$ (≤ 0) and $\partial h/\partial\beta = -n_2[p_2^2\overline{q}_2 - \alpha\{\overline{p}_2\overline{q}_2 + (\overline{p}_2 - p_2)^2\}]/(p_2q_2)^2$ $(\leq 0$ if $\alpha \leq 0)$. Because with expression (6), $dh/d\theta = 0 = (\partial h/\partial\theta) + (\partial h/\partial\hat{p}_1) \times (d\hat{p}_1/d\theta)$, where $\theta = \overline{p}_1, \overline{p}_2, \alpha$ or β, then $d\hat{p}_1/d\theta = -[(\partial h/\partial\theta)/(\partial h/\partial p_1)]_{p_1=\hat{p}_1}$. Because $[\partial h/\partial p_1]_{p_1=\hat{p}_1}$ ≤ 0, then the sign of $d\hat{p}_1/d\theta$ is the same as that of $[\partial h/\partial\theta]_{p_1=\hat{p}_1}$. Hence the conclusions at the end of Section 2 are obtained.

Up until now, it has been seen that the estimator of maximum likelihood for the pair (p_1, p_2) is $(\hat{p}_1, \hat{p}_2 = \alpha + \beta\hat{p}_1)$ under $H_{L(B)}: p_2 = \alpha + \beta p_1$. The aim now is to effect the estimation of (p_1, p_2) under the true null hypothesis $H_L: p_2 \leq \alpha + \beta p_1$. If the sample values $\overline{p}_i$ verify that $\overline{p}_2 \leq \alpha + \beta\overline{p}_1$, then the estimator of maximum likelihood for (p_1, p_2) is $(\overline{p}_1, \overline{p}_2)$, as these values yield an absolute maximum likelihood in expression (4). When $\overline{p}_2 > \alpha + \beta\overline{p}_1$, for each fixed value α, the estimator of maximum likelihood will be the pair $(\hat{p}_1, \hat{p}_2)$ from before, where $\hat{p}_1$ is a function of α. By substituting these values in expression (4), we obtain $[d \log L/d\alpha]_{\hat{p}_1} = [\partial \log L/\partial\alpha]_{\hat{p}_1} + [\partial \log L/\partial p_1]_{\hat{p}} \cdot (d\hat{p}_1/d\alpha) = [\partial \log L/\partial\alpha]_{\hat{p}_1} = n_2\beta(\overline{p}_2 - \hat{p}_2)/\hat{p}_2(1 - \hat{p}_2) \geq 0$, since $[\partial \log L/\partial p_1]_{\hat{p}_1} = 0$ through expression (6) and $\overline{p}_2 - \hat{p}_2 \geq 0$ through expression (7). This means that $L(\hat{p}_1, \hat{p}_2)$ increases with α and that the maximum likelihood is reached under the border of H_L.

B. Type z^2 or χ^2 statistics

Pearson's classic chi-square statistic is in the form $\chi^2 = \sum (O_i - E_i)^2/E_i$, where O_i and E_i are the observed and expected quantities, respectively. For model L:

$$\chi_L^2 = \sum \left\{ \frac{(x_i - n_i p_i)^2}{n_i p_i} + \frac{(y_i - n_i q_i)^2}{n_i q_i} \right\} = \sum (x_i - n_i p_i)^2 \left\{ \frac{1}{n_i p_i} + \frac{1}{n_i q_i} \right\}$$

$$= \sum \frac{n_i(\overline{p}_i - p_i)^2}{p_i q_i}, \tag{A1}$$

which yields expression (11). In order to obtain alternative expressions for statistic $\hat{\chi}_L^2$, the following equality must first be deduced (Andrés and Tejedor [2] for case *d*):

$$\beta(\overline{p}_2 - \hat{p}_2)/\sigma_2^2 = (\hat{p}_1 - \overline{p}_1)/\sigma_1^2 = \beta(\overline{p}_2 - \bar{f}_1)/(\sigma_2^2 + \beta^2\sigma_1^2). \tag{A2}$$

The first part of the equality is deduced directly from expression (6). In order to see the second part, the term $\pm \bar{f}_1$ must be added to the second numerator in expression (6) and, then, we must remember that $\hat{p}_2 - \bar{f}_1 = \beta(\hat{p}_1 - \overline{p}_1)$. In this way, $0 = (\overline{p}_1 - \hat{p}_1)/\sigma_1^2 + \beta(\overline{p}_2 - \bar{f}_1)/\sigma_2^2 + \beta(\bar{f}_1 - \hat{p}_2)/\sigma_2^2 = \{(\overline{p}_2 - \bar{f}_1)\beta - (\hat{p}_1 - \overline{p}_1)(\sigma_2^2 + \beta^2\sigma_1^2)/\sigma_1^2\}/\sigma_2^2$, and so $(\overline{p}_2 - \bar{f}_1)\beta/(\sigma_2^2 + \beta^2\sigma_1^2) = (\hat{p} - \overline{p}_1)/\sigma_1^2$, which is the second equality in expression (A2). This expression allows any of the terms $(\overline{p}_2 - \hat{f})$, $(\hat{p} - \overline{p}_1)$ and $(\overline{p}_2 - \bar{f}_1)$ to be expressed as a function of the other two. Because expression (A1):

$$\hat{\chi}_L^2 = (\hat{p}_1 - \overline{p}_1)^2/\sigma_1^2 + (\overline{p}_2 - \hat{p}_2)^2/\sigma_2^2, \tag{A3}$$

then by substituting $(\hat{p}_1 - \overline{p}_1)^2$ and $(\overline{p}_2 - \hat{p}_2)^2$ as a function of $(\overline{p}_2 - \bar{f}_1)^2$, we obtain $\hat{\chi}_L^2 = (\overline{p}_2 - \bar{f}_1)^2/\{\sigma_2^2 + \beta^2\sigma_1^2\} = \hat{z}_L^2$. Alternatively, in expression (A3), the two numerators may be written as a function of $(\hat{p}_1 - \overline{p}_1)^2$, $(\overline{p}_2 - \hat{p}_2)^2$, $(\hat{p}_1 - \overline{p}_1)(\hat{p}_2 - \bar{f}_1)$, $(\overline{p}_2 - \hat{p}_2)(\overline{p}_2 - \bar{f}_1)$ or $(\hat{p}_1 - \overline{p}_1)(\overline{p}_2 - \hat{p}_2)$, which leads to expressions (14) and (15).

With respect to statistic Mz_L^2, it must be remembered that expression (8) may be written as $z_L^2 = n_1 n_2(\overline{p}_2 - \alpha - \beta\overline{p}_1)^2/g(p_1)$, where $g(p_1) = n_2\beta^2 p_1 q_1 +$

$n_1(\alpha+\beta p_1)(1-\alpha-\beta p_1)$. Because $dg/dp_1 = 0$ in $p_0 = (n_1 + n_2\beta - 2n_1\alpha)/(2n\beta)$ and $d^2g/dp_1^2 = -2n\beta^2 < 0$, then $g(p_1)$ reaches maximum in $p = p_0$ which produces the minimum needed for z_L^2. As $p_0 = 0.5 + n_1(1-2\alpha-\beta)/(2n\beta)$ then $\alpha + \beta p_0$ $= 0.5 - n_2(1-2\alpha-\beta)/(2n)$, $g(p_0) = (n_2\beta^2 + n_1)/4 - n_1n_2(1-2\alpha-\beta)^2/(4n)$ and Mz_L^2 is the third expression of (22). When $p_0 \notin (p_I;\, p_S)$, then $p_0 < p_I$ (or $p_0 > p_S$), $g(p_1)$ decreases (or increases) in the allowed interval $(p_I;\, p_S)$ and its maximum is reached in $p_1 = p_I$ (or $p_1 = p_S$). Hence the first two expressions of (22).

C. Continuity correction

Haber [42] proposed that a continuity correction should consist of adding or subtracting to the random variable the half its average jump. For model L, the random variable is the contrast statistic $\bar{L} = \bar{p}_2 - \alpha - \beta\bar{p}_1$. Its minimum and maximum values are $-(\alpha+\beta)$ and $(1-\alpha)$, reached at $(\bar{p}_1;\, \bar{p}_2) = (1;\, 0)$ and $(0;\, 1)$, respectively. The total jump of $\bar{L}$ is therefore $(1+\beta)$ and half its average jump will be $c = (1+\beta)/N$, where $N = (n_1+1)(n_2+1) - 1 = n + n_1n_2$ because $(n_1+1)(n_2+1)$ is the total number of possible sample points. Consequently, in the numerator of z_L^2, the statistic $\bar{L}$ must be changed for $\bar{L} - c$, $\bar{L} + c$ or $|\,\bar{L}\,| - c$ depending on whether the alternative hypothesis is K_L, K_L' or K_L'', respectively (see the beginning of Section 4). Cases d and R are deduced from this case.

D. Basic equality for inferences based on model L

Let λ be the value for the three fractions of expression (A2). Due to the third one $\lambda = \beta z_L^2/(\bar{p}_2 - \bar{f}_1)$, due to the first $\lambda\hat{p}_1^2 - (\lambda - n_1)\,\hat{p}_1 - n_1\bar{p}_1 = 0$ and due to the second $\lambda\hat{p}_2^2 - (n_2\beta + \lambda)\,\hat{p}_2 + n_2\beta\bar{p}_2 = 0$. The solutions to the last two equations are $\hat{p}_1 = \{(\lambda - n_1) \pm R_1\}/2\lambda$ and $\hat{p}_2 = \{(\lambda + n_2\beta) \pm R_2\}/2\lambda$, where $R_1^2 = (\lambda - n_1)^2 +$ $4n_1\bar{p}_1\lambda$ and $R_2^2 = (\lambda - n_2\beta)^2 + 4n_2\beta\bar{q}_2\lambda$. If $\bar{p}_2 > \bar{f}_1$ (in which case, $\lambda > 0$), then $R_1 \geq |\,\lambda - n_1\,|$ if $\bar{p}_1 \geq 0$ and $R_2 \geq |\,\lambda - n_2\beta\,|$ if $\bar{q}_2 \geq 0$; this implies that $\hat{p}_1(-) < 0$ and $\hat{p}_2(+) > 1$, and so the true solutions of the above equations are $\hat{p}_1(+)$ and $\hat{p}_2(-)$.

By substitution of these values into $\hat{p}_2 = \alpha + \beta\hat{p}_1$, and simplified, expression (26) is obtained. When $\overline{p}_2 = \bar{f}_1$ it is not necessary to solve any equation because this occurs iff $z_L^2 = 0$ (so that $\lambda = 0$).

It can easily be seen that the function $y(\lambda)$ of expression (26) has the following features: (a) $\lim_{\lambda\to\pm\infty} y = +\infty$; (b) has two oblique asymptotes in $y = 2(\alpha + \beta)\lambda - 2\beta(n_1\overline{q}_1 + n_2\overline{p}_2)$ and $y = -2(1-\alpha)\lambda - 2\beta(n_1\overline{q}_1 + n_2\overline{p}_2)$; (c) its second derivative $d^2y/d\lambda^2 = 4n_1^2\beta\overline{p}_1\overline{q}_1/R_1^3 + 4n_2^2\beta^2\overline{p}_2\overline{q}_2/R_2^3 \geq 0$; and (d) $y = 0$ in $\lambda = 0$. From this, it can be gathered that the function has only one minimum and two cuts $\lambda_1 = 0$ and $\lambda_2 \neq 0$ with the axis λ, where λ_2 is the value sought. Finally, when $\lambda_2 > 0$ (that is, when $\overline{p}_2 > \bar{f}_1$) then the function y will be negative (positive) in values $0 < \lambda < \lambda_2$ $(\lambda_2 < \lambda < \infty)$, so that if its value (y_γ) is calculated in $\hat{z}_L^2 = z_{1-\gamma}^2$ the fact that $y_\gamma \leq 0$ $(y_\gamma > 0)$ is indicative of the test for the alternative $K_L : p_2 - \alpha - \beta p_1 > 0$ is significant (not significant).

E. Properties of the various statistics

Let us consider the hypotheses H_O and H_{SC} in Section 5. In the terminology of Section 2, $H_I : p_2 = \alpha + \beta p_1$ is contrasted based on the data $n_1, n_2, \overline{p}_1, \overline{p}_2$, α and β, while $H_{SC} : p_2' = \alpha' + \beta' p_1'$ is contrasted based on the data $n_1' = n_2$, $n_2' = n_1$, $\overline{p}_1' = 1 - \overline{p}_2$, $\overline{p}_2' = 1 - \overline{p}_1$, $\alpha' = (\alpha + \beta - 1)/\beta$ and $\beta' = 1/\beta$. Similar arguments can be made for $\tilde{p}_i$ and $\hat{p}_i$ because of expressions (4) and (5). In a general manner, it may be affirmed that $p_1' = 1 - p_2$ and $p_2' = 1 - p_1$ with any estimator. Similarly, when H_O is compared with H_{SC} (or H_C) it can be deduced that $p_1' = p_2$ and $p_2' = p_1$ (or $p_1' = 1 - p_1$ and $p_2' = 1 - p_2$). With these data, it can at once be proven that any statistic S_L^2 takes the same value in all four cases. For example, $\chi^2_{L(SC)} = n_1'(\overline{p}_1' - p_1')^2/p_1'(1-p_1') + n_2'(\overline{p}_2' - p_2')^2/p_2'(1-p_2') = n_2(1 - \overline{p}_2 - 1 + p_2)^2/(1-p_2)\,p_2 + n_1(1 - \overline{p}_1 - 1 + p_1)^2/(1-p_1)\,p_1 = \chi^2_{L(O)}$. (The exception is the case of $L\bar{z}_L^2$ in which equivalence only occurs between H_O and

H_S in the case of model R). In reality, the four equivalences are not independent. It is easy to see that if $S^2_{L(O)} = S^2_{L(SC)} = S^2_{L(S)}$, then they are also equal to $S^2_{L(C)}$; also, $S^2_{L(O)} = S^2_{L(SC)}$ iff $S^2_{L(S)} = S^2_{L(C)}$.

Let us assume that the statistic S^2_L verifies the equivalence between H_O and H_{SC}. Because under H_{SC} is $\bar{p}'_2 = 1 - \bar{p}_1$, then $d\,\bar{p}'_2/d\,\bar{p}_1 = -1$ and so $dS_{L(O)}/d\,\bar{p}_1 = -dS_{L(SC)}/d\,\bar{p}'_2$. As H_O and H_{SC} have the same type of alternative (K_L), the previous equality implies that if S_L is convex in $\bar{p}_2$ then it is also convex in $\bar{p}_1$. As all the present statistics (with the exception of $L\bar{z}^2_L$) verify this property of equivalence, the conclusion is that for these, we need only to prove convexity in $\bar{p}_2$ (for example), because that of $\bar{p}_1$ is obtained as a result. Moreover, as $d\beta'/d\alpha = (1-\alpha')/(1-\alpha)^2 \geq 0$, using the same reasoning as before, it can be deduced that the convexity in α implies convexity in β. In addition, as $dS^2_L/d\psi = 2S_L(dS_L/d\psi) \propto [\text{sign}(\bar{p}_2 - \bar{f}_1)](dS_L/d\psi)$, where $\psi = \bar{p}_2$, $\bar{p}_1$ or α, then in order to prove the convexity of S_L in $\bar{p}_2$ and α, we need only to prove that $dS^2_L/d\bar{p}_2 \geq 0$ and $dS^2_L/d\alpha \leq 0$ when $\bar{p}_2 > \bar{f}_1$. In the following, are the principal cases promised in Section 5.

By deriving expression (9), we get $d\bar{z}^2_L/d\alpha \propto -2(\bar{p}_2 - \bar{f}_1) \leq 0$ and $d\bar{z}^2_L/d\bar{p}_2 \propto (\bar{p}_2 - \bar{f}_1)[2\beta^2\bar{p}_1\bar{q}_1/n_1 + \{\bar{f}_1(1-\bar{p}_2) + \bar{p}_2(1-\bar{f}_1)\}/n_2] \geq 0$. By deriving expression (20), we get $dL\bar{z}^2_R/d\rho \propto -2[\log(\bar{p}_2/\rho\bar{p}_1)] \leq 0$ (now the convexity has to be in ρ because $\alpha = 0$) and $dL\bar{z}^2_R/d\bar{p}_2 \propto 2\bar{q}_1/n_1\bar{p}_1 + 2\bar{q}_2/n_2\bar{p}_2 + [\log(\bar{p}_2/\rho\bar{p}_1)]/n_2\bar{p}_2 \geq 0$. However, $L\bar{z}^2_R$ is not convex in $\bar{p}_1$ as the following counter-example shows. For values $\rho = 1$ and $n_1 = n_2 = x_2 = 20$, we obtain $L\bar{z}^2_R = 9.45$ when $x_1 = 1$ and $L\bar{z}^2_R = 11.78$ when $x_1 = 2$, that is, $L\bar{z}^2_R$ increases with $\bar{p}_1$ when $\bar{p}_2 > \bar{f}_1$. Lastly, because $(d\hat{z}^2_L/d\psi) = (\partial\hat{z}^2_L/\partial\psi) + (\partial\hat{z}^2_L/\partial\hat{p}_1)(d\hat{p}_1/d\psi)$, where $\psi = \bar{p}_2$ or α, then, from the first expression of (14), $d\hat{z}^2_L/d\bar{p}_2 = (n_1/\beta)[(\hat{p}_1 - \bar{p}_1)/\hat{p}_1\hat{q}_1 + (d\hat{p}_1/d\bar{p}_2)$

$(\overline{p}_2 - \bar{f}_1)\{\overline{p}_1\overline{q}_1 + (\hat{p}_1 - \overline{p}_1)^2\}/\hat{p}_1^2\hat{q}_1^2] \geq 0,$ $d\hat{z}_L^2/d\alpha = (n_1/\beta)[-(\hat{p}_1 - \overline{p}_1)/\hat{p}_1\hat{q}_1 + (d\hat{p}_1/d\alpha)(\overline{p}_2 - \bar{f}_1)\{\overline{p}_1\overline{q}_1 + (\hat{p}_1 - \overline{p}_1)^2\}/\hat{p}_1^2\hat{q}_1^2] \leq 0,$ where the final notations are due to the fact that when $\overline{p}_2 \geq \bar{f}_1$ (as is being assumed) then $\hat{p}_1 \geq \overline{p}_1$ from expression (7), $d\hat{p}_1/d\overline{p}_2 \geq 0$ because $\hat{p}_1$ increases with $\overline{p}_2$, and $d\hat{p}_1/d\alpha \leq 0$ because $\hat{p}_1$ decreases with α (these two last statements are due to the results at the end of Section 2). Therefore, from what has been pointed out in the second paragraph of this section, it will also occur that $d\hat{z}_L^2/d\overline{p}_1 \leq 0$ and $d\hat{z}_L^2/d\beta \leq 0.$ Strictly speaking, the previous proof is valid when $\hat{p}_1 \neq 0, 1;$ otherwise, we need only to repeat the reasoning based on the second expression of (14).

References

[1] C. W. Dunnett and M. Gent, Significance testing to establish equivalence between treatments, with special reference to data in the form of 2×2 tables, Biometrics 33 (1977), 593-602.

[2] A. Martín Andrés and I. Herranz Tejedor, Exact unconditional non-classical tests on the difference of two proportions, Comput. Stat. Data Anal. 45 (2004), 373-388.

[3] A. Agresti, Dealing with discreteness: making 'exact' confidence intervals for proportions, differences of proportions, and odds ratios more exact, Stat. Meth. Med. Res. 12 (2003), 3-21.

[4] I. S. F. Chan, Exact tests of equivalence and efficacy with a non-zero lower bound for comparative studies, Stat. Med. 17 (1998), 1403-1413.

[5] K. F. Phillips, A new test of non-inferiority for anti-infective trials, Stat. Med. 22 (2003), 201-212.

[6] I. S. F. Chan, Proving non-inferiority or equivalence of two treatments with dichotomous endpoints using exact methods, Stat. Meth. Med. Res. 12 (2003), 37-58.

[7] R. S. Dann and G. G. Koch, Review and evaluation of methods for computing confidence intervals for the ratio of two proportions and considerations for non-inferiority clinical trials, J. Biopharmaceutical Stat. 15 (2005), 85-107.

[8] O. Miettinen and M. Nurminen, Comparative analysis of two rates, Stat. Med. 4 (1985), 213-226.

[9] J. Röhmel and U. Mansmann, Unconditional non-asymptotic one-sided tests for independent binomial proportions when the interest lies in showing non-inferiority and/or superiority, Biometrical J. 41 (1999), 149-170.

[10] C. P. Farrington and G. Manning, Test statistics and sample size formulae for comparative binomial trials with null hypothesis of non-zero risk difference or non-unity relative risk, Stat. Med. 9 (1990), 1447-1454.

[11] R. W. Mee, Confidence bounds for the difference between two probabilities, Biometrics 40 (1984), 1175-1176.

[12] P. A. R. Koopman, Confidence intervals for the ratio of two binomial proportions, Biometrics 40 (1984), 513-517.

[13] A. Martín Andrés, J. M. Tapia Garcia and M. J. del Moral Ávila, Unconditional inferences on the difference of two proportions in cross-sectional studies, Biometrical J. 47 (2005), 177-187.

[14] D. Katz, J. Baptista, S. P. Azen and M. C. Pike, Obtaining confidence intervals for the risk ratio in cohort studies, Biometrics 34 (1978), 469-474.

[15] L. L. Laster, M. F. Johnson and M. L. Kotler, Non-inferiority trials: the 'at least as good as' criterion with dichotomous data, Stat. Med. 25 (2006), 1115-1130.

[16] S. Wallenstein, A non-iterative accurate asymptotic confidence interval for the difference between two proportions, Stat. Med. 16 (1997), 1329-1336.

[17] Y. Parmet and E. Schechtman, On a test of the difference between two binomial proportions, Comm. Statist. Theory Methods 36 (2007), 887-895.

[18] J. M. Nam, Confidence limits for the ratio of two binomial proportions based on likelihood scores: non-iterative method, Biometrical J. 37 (1995), 375-379.

[19] J. J. Gart and J. Nam, Approximate interval estimation of the ratio of binomial parameters: a review and corrections for skewness, Biometrics 44 (1988), 323-338.

[20] B. Woolf, On estimating the relation between blood group and disease, Ann. Human Genetics 19 (1955), 251-352.

[21] T. E. Sterne, Some remarks on confidence of fiducial limits, Biometrika 41 (1954), 275-278.

[22] P. H. Peskun, A new confidence interval method based on the normal approximation for the difference of two binomial probabilities, J. Amer. Statist. Assoc. 88 (1993), 656-661.

[23] P. D. Feigin and Y. P. Lumelskii, On confidence limits for the difference of two binomial parameters, Comm. Statist. Theory Methods 29 (2000), 131-141.

[24] D. R. Cox, The continuity correction, Biometrika 57 (1970), 217-219.

[25] J. Cornfield, A statistical problem arising from retrospective studies, Proceedings of Third Berkeley Symposium on Mathematical Statistics and Probability, University of California Press, Berkeley and Los Angeles, 1956, pp. 135-148.

[26] A. Martín Andrés and I. Herranz Tejedor, Asymptotical tests on the equivalence, substantial difference and non-inferiority problems with two proportions, Biometrical J. 46 (2004), 305-319.

[27] A. Agresti and B. Caffo, Simple and effective confidence intervals for proportions and difference of proportions result from adding two successes and two failures, Amer. Statist. 54 (2000), 280-288.

[28] A. Agresti and Y. Min, On small-sample confidence intervals for parameters in discrete distributions, Biometrics 57 (2001), 963-971.

[29] A. Martín Andrés and I. Herranz Tejedor, Equivalence testing for binomial random variables: which test to use? (Letter to the editor), Amer. Statist. 56 (2002), 253-254.

[30] A. Munk, G. Skipka and B. Stratmann, Testing general hypotheses under binomial sampling: the two sample case - asymptotic theory and exact procedures, Comput. Stat. Data Anal. 49 (2005), 723-739.

[31] I. Herranz Tejedor and A. Martín Andrés, A numerical comparison of several unconditional exact tests in problems of equivalence based on the difference of proportions, J. Stat. Comput. Simul. 78 (2008), 969-981.

[32] X.-H. Zhou, M. Tsao and G. Qin, New intervals for the difference between two independent binomial proportions, J. Statist. Plann. Inference 123 (2004), 97-115.

[33] B. Lecoutre and S. Faure, A note on new confidence intervals for the difference between two proportions based on an Edgeworth expansion, J. Statist. Plann. Inference 137 (2007), 355-356.

[34] R. G. Newcombe, Interval estimation for the difference between independent proportions: comparison of eleven methods, Stat. Med. 17 (1998), 873-890.

[35] A. Martín Andrés and I. Herranz Tejedor, Unconditional confidence interval for the difference between two proportions, Biometrical J. 45 (2003), 426-436.

[36] J. Röhmel, Problems with existing procedures to calculate exact unconditional *P*-values for non-inferiority/superiority and confidence intervals for two binomials and how to resolve them, Biometrical J. 47 (2005), 37-47.

[37] A. Munk and R. Pflüger, $1-\alpha$ equivariant confidence rules for convex alternatives are $\alpha/2$-level tests - with applications to the multivariate assessment of bioequivalence, J. Amer. Statist. Assoc. 94 (1999), 1311-1319.

[38] P. Roebruck and A. Kühn, Comparison of tests and sample size formulae for proving therapeutic equivalence based on the difference of binomial probabilities, Stat. Med. 14 (1995), 1583-1594.

[39] L. Brown and X. Li, Confidence intervals for two sample binomial distribution, J. Statist. Plann. Inference 130 (2005), 359-375.

[40] W. W. Hauck and S. Anderson, A comparison of large-sample confidence interval methods for the difference of two binomial probabilities, Amer. Statist. 40 (1986), 318-322.

[41] L. Xu and J. E. Kolassa, Testing the difference of two binomial proportions: comparison of continuity corrections for saddlepoint approximation, Comm. Statist. Theory Methods 37 (2008), 2213-2218.

[42] M. Haber, A comparison of some continuity corrections for the chi-squared test on 2×2 tables, J. Amer. Statist. Assoc. 75 (1980), 510-515.