Increasing sample sizes do not always increase the power of UMPU-tests for 2×2 tables [1]

By H. Finner and K. Strassburger

Deutsches Diabetes-Forschungsinstitut an der
Heinrich-Heine-Universität Düsseldorf
Abteilung Biometrie und Epidemiologie,
Düsseldorf, Germany

May 2, 2000

Abstract. We consider the uniformly most powerful unbiased (UMPU) one-sided test for the comparison of two proportions based on sample sizes m and n, i.e., the randomized version of Fisher's exact one-sided test. It will be shown that the power function of the one-side UMPU-test based on sample sizes m and n can coincide on the entire parameter space with the power function of the UMPU test based on sample sizes $m+1$ and n for certain levels. A characterization of all such cases with identical power functions is derived. Finally, this characterization is closely related to number theoretical problems concerning Fermat-like binomial equations. Some consequences for Fisher's original exact test will be discussed, too.

1. Introduction

One of the oldest and apparent most basic problems in statistics is the evaluation of a 2×2-table. Originally, our aim was to develope certain optimal multiple decision procedures as for example optimal selection and partitioning procedures in k-sample situations with underlying binomial distributions. We expected that everything should be clarified concerning the evaluation of 2×2-tables. After inspection of uncountable papers on 2×2-tables we learnt that there are more unsolved than solved problems and that some of the problems we had are not even mentioned in the literature. Some of these issues concerning structural properties of Fisher's exact test (Fisher (1934/35), Yates (1934), Irwin (1935)) as well as the corresponding UMPU-tests for 2×2-tables (Tocher (1950)) are discussed and partially solved in Finner & Strassburger (2000). Some of the results for the one-sided UMPU-test derived there will be applied here, too. In the present paper we investigate a surprising phenomenon arising in connection with power considerations for one-sided UMPU-tests for comparing two proportions. It will be shown that increasing sample sizes do not always increase the (unconditional) power of the one-sided UMPU-test.

To set up notation, suppose we have two sets of Bernoulli random variables $X_i,\ i = 1, \ldots, m$ and $Y_i,\ i = 1, \ldots, n$, respectively, with success probabilities p and q, respectively, where $X_1, \ldots, X_m$, $Y_1, \ldots, Y_n$ are independently distributed, m, n are fixed and known and p, q are unknown. We are faced with testing the one-sided hypothesis $H : p \leq q$ versus the one-sided alternative $K : p > q$.

[1] MSC 1991 classification numbers. Primary 62F03, 62C99. Secondary 11D41.

Key words and phrases. diophantine equation, Fermat-like binomial equation, Fisher's exact test, one-sided hypothesis, 2×2 table, uniformly most powerful unbiased test.

Then $X = \sum_{i=1}^{m} X_i$ and $Y = \sum_{i=1}^{n} Y_i$ are independent random variables having a binomial distribution with parameters $m \in I\!N$, $p \in [0,1]$ and $n \in I\!N$, $q \in [0,1]$, respectively. The UMPU-test at level $\alpha \in (0,1)$ for testing H versus K is based on conditioning on $S = X + Y = s$ and is given by

$$\psi(x|s,m,n,\alpha) = \begin{cases} 0, & x < c_{s,m,n,\alpha}, \\ \gamma_{s,m,n,\alpha}, & x = c_{s,m,n,\alpha}, \\ 1, & x > c_{s,m,n,\alpha}, \end{cases} \tag{1.1}$$

where $c = c_{s,m,n,\alpha} \in \{0,\ldots,m\}$ and $\gamma = \gamma_{s,m,n,\alpha} \in [0,1)$ are determined such that

$$F(c|s,m,n) - \gamma f(c|s,m,n) = 1 - \alpha.$$

Here

$$f(x|s,m,n) = \begin{cases} \binom{m}{x}\binom{n}{s-x}/\binom{m+n}{s}, & \text{for } \max(0,s-n) \le x \le \min(s,m), s,x \in I\!N_0, \\ 0, & \text{otherwise }, \end{cases} \qquad m,\ n \in I\!N,$$

denotes the probability mass function (pmf) of the hypergeometric distribution with parameters s,m,n and F denotes the corresponding cumulative mass function (cmf). Let $g(x|m,p) = \binom{m}{x}p^x(1-p)^{m-x}$, $x = 0,\ldots,m$, denote the pmf of the binomial distribution with parameters p and m. Setting

$$\beta_1(p|y,m,n,\alpha) = \sum_x \psi(x|x+y,m,n,\alpha)g(x|m,p),$$

the (unconditional) two-dimensional power function of ψ can be calculated by

$$\beta(p,q|m,n,\alpha) = \sum_y \beta_1(p|y,m,n,\alpha)g(y|n,q),\ \ p,q \in [0,1]. \tag{1.2}$$

It is well known, that the power of the UMPU-test is non-decreasing on the alternative K in the sample sizes m and n, respectively. This is based on the fact, that the UMPU-test, given m and n, is based on the sufficient statistic $(\sum_{i=1}^{m} X_i\ ,\ \sum_{i=1}^{n} Y_i)$. One might expect, that the power function is not only non-decreasing in m and n on K but even strictly increasing for at least some (if not all) parameter points in K. It will be shown, that this expectation turns out to fail, that is, there exist $\alpha \in (0,1)$ and m,n, such that

$$\beta(p,q|m,n,\alpha) = \beta(p,q|m+1,n,\alpha) \qquad \forall p,q \in [0,1]. \tag{1.3}$$

A full characterization of situations where (1.3) is satisfied is given in Section 2. The technical and rather lengthy part of the proof of our main result is deferred to the Appendix. Consequences for the classical exact test of Fisher are discussed in Section 3. Finally, in Section 4 we show that the whole problem results in interesting for the most part unsolved number theoretical problems concerning Fermat-like binomial equations.

2. The main result

Noting that the class of binomial distributions with pmf $g(\cdot|m,p)$, $p \in [0,1]$, is complete, we get that (1.3) is equivalent to

$$\beta_1(p|y,m,n,\alpha) = \beta_1(p|y,m+1,n,\alpha) \qquad \forall p \in [0,1],\ y = 0,\ldots,n. \tag{2.1}$$

Since $\beta_1(p|y,m,n,\alpha) - \beta_1(p|y,m+1,n,\alpha)$ is a continuous function in p and $\beta(p,p|m,n,\alpha) = \beta(p,p|m+1,n,\alpha)$ for all $p \in [0,1]$, (1.2) yields that (2.1) is equivalent to

$$\beta_1(p|y,m,n,\alpha) = \beta_1(p|y,m+1,n,\alpha) \qquad \forall p \in (0,1),\ y = 0,\dots,n-1. \tag{2.2}$$

To further characterize the situations where (1.3) holds, we make use of the following Lemma the proof of which is given in the Appendix.

Lemma 1. Let $\alpha \in (0,1)$.

a) $\beta_1(p|0,m,n,\alpha) = \beta_1(p|0,m+1,n,\alpha)$ holds for all $p \in (0,1)$ if and only if there exits an $u \in \{0,\dots,m-1\}$ such that $F(u|u+1,m,n) = 1-\alpha$.

b) Let $y \in \{1,\dots,n-1\}$ and suppose there exists an integer $u \in \{0,\dots,m-1\}$ with $F(u|u+y,m,n) = 1-\alpha$. Then $\beta_1(p|y,m,n,\alpha) = \beta_1(p|y,m+1,n,\alpha)$ holds for all $p \in (0,1)$ if and only if there exits a $v \in \{u+1,\dots,m-1\}$ such that $F(v|v+y+1,m,n) = 1-\alpha$.

Lemma 1 immediately yields the following characterization of (1.3).

Theorem 2. For $\alpha \in (0,1)$ equation (1.3) holds, if and only if there exists a sequence of integers $0 \le u_1 < \dots < u_n < m$ such that

$$F(u_i|u_i+i,m,n) = 1-\alpha \ \text{ for all } i = 1,\dots,n. \tag{2.3}$$

It remains the question whether there exist m,n and a sequence of integers $0 \le u_1 < \dots < u_n < m$ such that (2.3) is satisfied. In case of $n \in \{1,2\}$ and in case of $m = n,\ n \in I\!N$, we obtain the following answer.

Theorem 3. Let $\alpha \in (0,1)$, $m \in I\!N$.
(a) For $n = 1$, (1.3) holds if and only if $\alpha = \ell/(m+1)$, $\ell \in \{1,\dots,m\}$.
(b) For $n = 2$, (1.3) holds if and only if there exist $u,v \in \{0,\dots,m-1\}$ with $\alpha = f(u+1|u+1,m,2) = 1 - f(v|v+2,m,2)$. or, equivalently,

$$m = \frac{(v+1)(v+2)}{2(u+1)} + \frac{u}{2} - 1 \quad \text{and} \quad \alpha = 1 - \frac{(v+1)(v+2)}{(m+1)(m+2)}. \tag{2.4}$$

(c) For $m = n$, (1.3) holds if $\alpha = 1/2$.

Proof. Parts (a) and (b) follow immediately from Theorem 2. In case of $m = n$ and $\alpha = 1/2$ the UMPU-test is given by

$$\psi(x|x+y,m,m,1/2) = \begin{cases} 0, & x < y \\ 1/2, & x = y \\ 1, & x > y \end{cases}. \tag{2.5}$$

This can easily be seen by noting the symmetry of $f(\cdot|s,m,m)$. Hence we also get part (c) from Theorem 2. $\square$

Remark. (a) The set of all solutions (m,u,v) of (2.4) is infinite, since for all $u \in I\!N \cup \{0\}$ the triples $(5u+5, u, 3u+2)$ and $(5u+7, 2u+2, 4u+5)$ belong to this set.

(b) Setting $a = m - u - 1$, $b = v$ and $c = m$, it can easily be seen that the equation $f(u+1|u+1, m, 2) = 1 - f(v|v+2, m, 2)$ is equivalent to

$$(a+1)(a+2) + (b+1)(b+2) = (c+1)(c+2) \quad \text{or} \quad \binom{a+2}{2} + \binom{b+2}{2} = \binom{c+2}{2}.$$

The set of all positive integer solutions for these equations can be found e. g. in Harborth (1988), related references are Khatri (1955), Sierpiǹski (1962) and Fraenkel (1971). At this place we refer to Section 4 of this paper for a more detailed discussion concderning the relationship to number theoretic problems.

We display two examples where the power functions of two different tests coincide. The UMPU tests at level $\alpha = 1/2$ for $(m, n) = (5, 5)$ and $(m, n) = (6, 5)$ are given by

$\psi(x|x+y, 5, 5, 1/2)$

x \ y	0	1	2	3	4	5
5	1	1	1	1	1	$\frac{1}{2}$
4	1	1	1	1	$\frac{1}{2}$	0
3	1	1	1	$\frac{1}{2}$	0	0
2	1	1	$\frac{1}{2}$	0	0	0
1	1	$\frac{1}{2}$	0	0	0	0
0	$\frac{1}{2}$	0	0	0	0	0

$\psi(x|x+y, 6, 5, 1/2)$

x \ y	0	1	2	3	4	5
6	1	1	1	1	1	$\frac{1}{2}$
5	1	1	1	1	$\frac{7}{12}$	$\frac{1}{12}$
4	1	1	1	$\frac{2}{3}$	$\frac{1}{6}$	0
3	1	1	$\frac{3}{4}$	$\frac{1}{4}$	0	0
2	1	$\frac{5}{6}$	$\frac{1}{3}$	0	0	0
1	$\frac{11}{12}$	$\frac{5}{12}$	0	0	0	0
0	$\frac{1}{2}$	0	0	0	0	0

The UMPU tests at level $\alpha = 2/7$ for $(m, n) = (5, 2)$ and $(m, n) = (6, 2)$ are given by

$\psi(x|x+y, 5, 2, 2/7)$

x \ y	0	1	2
5	1	1	$\frac{2}{7}$
4	1	$\frac{1}{2}$	0
3	1	$\frac{1}{4}$	0
2	$\frac{3}{5}$	0	0
1	$\frac{2}{5}$	0	0
0	$\frac{2}{7}$	0	0

$\psi(x|x+y, 6, 2, 2/7)$

x \ y	0	1	2
6	1	1	$\frac{2}{7}$
5	1	$\frac{7}{12}$	$\frac{1}{21}$
4	1	$\frac{1}{3}$	0
3	$\frac{3}{4}$	$\frac{1}{8}$	0
2	$\frac{8}{15}$	0	0
1	$\frac{8}{21}$	0	0
0	$\frac{2}{7}$	0	0

All cases for $3 \leq m \leq 62$, where (2.4) is fulfilled with corresponding values $\alpha = 1 - \frac{(v+1)(v+2)}{(m+1)(m+2)} = \frac{(m-u)(m-u+1)}{(m+1)(m+2)}$ with $\alpha \leq 1/2$ are given in Table 3.

m	α	m	α	m	α	m	α
2	$\frac{1}{2}$	20	$\frac{1}{11}, \frac{26}{77}$	39	$\frac{19}{82}$	52	$\frac{35}{477}, \frac{176}{477}$
5	$\frac{2}{7}$	22	$\frac{15}{92}, \frac{11}{46}, \frac{35}{92}$	40	$\frac{40}{287}, \frac{11}{41}, \frac{100}{287}$	53	$\frac{14}{99}$
7	$\frac{5}{12}$	25	$\frac{40}{117}$	42	$\frac{351}{946}$	54	$\frac{1}{28}$
9	$\frac{2}{11}$	26	$\frac{13}{63}$	43	$\frac{7}{33}$	55	$\frac{187}{532}$
10	$\frac{7}{22}$	27	$\frac{2}{29}, \frac{55}{406}, \frac{153}{406}$	44	$\frac{1}{23}$	56	$\frac{126}{551}$
12	$\frac{36}{91}$	30	$\frac{171}{496}$	45	$\frac{91}{1081}, \frac{378}{1081}$	57	$\frac{171}{1711}, \frac{630}{1711}$
14	$\frac{1}{8}$	32	$\frac{70}{187}$	47	$\frac{145}{392}$	58	$\frac{26}{59}$
15	$\frac{45}{136}$	35	$\frac{2}{37}, \frac{35}{222}, \frac{77}{222}$	48	$\frac{38}{245}, \frac{17}{35}$	60	$\frac{351}{1891}, \frac{406}{1891}, \frac{666}{1891}$
17	$\frac{22}{57}$	36	$\frac{325}{703}$	50	$\frac{155}{442}$	61	$\frac{100}{651}$
19	$\frac{1}{2}$	37	$\frac{92}{247}$	51	$\frac{153}{1378}$	62	$\frac{247}{672}$

Table 3. All values for m, $3 \le m \le 62$ and corresponding values of $\alpha \le 1/2$ for which the assumptions of Theorem 3 (b) are satisfied.

3. Consequences for Fisher's exact test

Fisher's (nonrandomized) exakt Test $\tilde{\psi}$ is given by

$$\tilde{\psi}(x|x+y,m,m,\alpha) = \begin{cases} 1, & \psi(x|x+y,m,m,\alpha) = 1, \\ 0, & \text{otherwise.} \end{cases} \tag{3.1}$$

The (unconditional) two-dimensional power function of $\tilde{\psi}$ can be calculated by

$$\tilde{\beta}(p,q|m,n,\alpha) = \sum_y \sum_x \tilde{\psi}(x|x+y,m,n,\alpha) g(x|m,p) g(y|n,q), \;\; p,q \in [0,1].$$

The next Theorem yields that there are situations where the power function of Fisher's exact test for sample sizes $(m+1,n)$ is strictly greater than for sample sizes (m,n). Some of these situations are covered by Theorem 2.

Theorem 4. If (1.3) is valid, then

$$\tilde{\psi}(x|x+y,m,n,\alpha) = \tilde{\psi}(x+1|x+y,m+1,n,\alpha), \quad x = 0,\dots,m, \quad y = 0,\dots,n-1. \tag{3.2}$$

Moreover, (3.2) implies

$$\tilde{\beta}(p,q|m,n,\alpha) > \tilde{\beta}(p,q|m+1,n,\alpha) \quad \forall p,q \in (0,1).$$

Proof. Using similar arguments as in the beginning of the proof of Lemma 1a,b, Theorem 2 yields that (1.3) implies the existence of integers $0 \le u_1 < \cdots < u_n < m$, such that

$$\tilde{\psi}(x|x+y,m+1,n,\alpha) = \begin{cases} 0, & \text{for } 0 \le x \le u_{y+1}+1, \\ 1, & \text{otherwise,} \end{cases} \tag{3.3}$$

and

$$\tilde{\psi}(x|x+y,m,n,\alpha) = \begin{cases} 0, & \text{for } 0 \le x \le u_{y+1}, \\ 1, & \text{otherwise.} \end{cases} \tag{3.4}$$

holds for all $y = 0, \dots, n-1$. This implies (3.2).

On the other hand, (3.2) implies the existence of integers $0 \le u_1 < \cdots < u_n < m$, such that (3.3) and (3.4) holds. By noting that $\tilde{\psi}(x+1|x+n, m+1, n, \alpha) = \tilde{\psi}(x|x+n, m, n, \alpha) = 0$, $x = 0, \dots, m$ and $\tilde{\psi}(0|y, m+1, n, \alpha) = 0$, $y = 0, \dots, n$, we conclude that

$$\tilde{\beta}(p, q|m, n, \alpha) - \tilde{\beta}(p, q|m+1, n, \alpha) = \sum_{y=0}^{n-1} g(y|n, q)[H(u_{y+1}+1|m, p) - H(u_{y+1}+2|m+1, p)].$$

Using (A.5) we get

$$\begin{aligned} & H(u_{y+1}+1|m,p) - H(u_{y+1}+2|m+1,p) \\ &\qquad = g(u_{y+1}+1|m,p) + H(u_{y+1}+2|m,p) - H(u_{y+1}+2|m+1,p) \\ &\qquad = g(u_{y+1}+1|m,p) - pg(u_{y+1}+1|m,p) \\ &\qquad = (1-p)g(u_{y+1}+1|m,p). \end{aligned}$$

This completes the proof of Theorem 4. □

There exit situations where (1.3) is not valid but (3.2) holds. Numereous examples for this phenomenon can be constructed using very small values of α, m and n. But there exist some practical relevant cases, too. For instance, (3.2) is fullfilled for $(m, n) = (12, 12)$ and $\alpha = 0.02$. In this case Fisher's exact test is given by

$\tilde{\psi}(x, x+y, 12, 12, 0.02)$

12	1	1	1	1	1	1	1	1	0	0	0	0	0
11	1	1	1	1	1	1	0	0	0	0	0	0	0
10	1	1	1	1	1	0	0	0	0	0	0	0	0
9	1	1	1	1	0	0	0	0	0	0	0	0	0
8	1	1	1	0	0	0	0	0	0	0	0	0	0
7	1	1	0	0	0	0	0	0	0	0	0	0	0
6	1	0	0	0	0	0	0	0	0	0	0	0	0
5	1	0	0	0	0	0	0	0	0	0	0	0	0
4	0	0	0	0	0	0	0	0	0	0	0	0	0
3	0	0	0	0	0	0	0	0	0	0	0	0	0
2	0	0	0	0	0	0	0	0	0	0	0	0	0
1	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0
	0	1	2	3	4	5	6	7	8	9	10	11	12

and condition (3.2) is satisfied.

The reason for the observed phenomenon is the discreteness of the underlying distributions as well as fixing the level α in advance. It should be noted that the unconditional size of Fisher's exact test for $(m, n) = (13, 12)$ is approximately 0.0077574 which is strictly less than the corresponding unconditional size 0.0114223 for $(m, n) = (12, 12)$, i.e., the loss in power yields a gain with respect to (unconditional) type I errors. The maximum difference between the unconditional power functions over the alternative K is approximately 0.1370881 and is attained in $(p, q) = (0.3846154, 0.0)$. The maximum conditional size (0.0195629) for $(m,n) = (12,12)$ is attained for $s = 12$, the maximum conditional size (0.0149068) for $(m,n) = (13,12)$ is attained at $s = 20$.

4. Open problems

The question arises whether there exist further parameter configurations $(m,n) \neq (m',n')$ with identical unconditional power functions than given in Theorem 3. The existence of a solution of (2.3) may be described via n urn experiments. Suppose we have an urn with m red and n black balls. The i'th experiment consists in drawing exactly $u_i + i$ balls. Let $A(i,u_i)$ denote the event of drawing at most u_i red balls in the i'th experiment. Then the question is, whether there exist integers $0 \leq u_1 < \cdots < u_n < m$ such that all the events $A(i,u_i)$, $i = 1,\ldots,n$ have the same probability.
The attempt to find a set $(\alpha, u_1, \ldots, u_n)$ of solutions of (2.3) for $3 \leq m,n \leq 500$ for $\alpha \in (0,1)$, ($\alpha \neq 1/2$ if $m = n$) by using an algebraic computer package failed, moreover, for $3 \leq m,n \leq 100$ we found that if $F(u|u+i,m,n) = 1-\alpha$ for some i,u, then there exists at most one tuple $(j,v) \neq (i,u)$ with $F(v|v+j,m,n) = 1-\alpha$, except for $(m,n,\alpha) = (85,35,1/2)$ where $F(1|2,85,35) = F(42|60,85,35) = F(83|118,85,35) = 1/2$. For $n = 3$ we found that (2.3) is not solvable for $3 \leq m \leq 66500$ whenever $\alpha \neq 1/2$. These observations lead us to the following conjecture.

Conjecture. Let $\alpha \in (0,1)$ and $m,n \geq 3$. Then (1.3) holds, if and only if $m = n$ and $\alpha = 1/2$.

Finally the whole bag of tricks results in number theoretic problems which also show that the most critical point of our conjecture is the case $n = 3$. For example, the equation $F(u|u+1,m,n) = F(v|v+n,m,n)$ (which is a necessary but not sufficient condition for the validity of (1.3), consider $i = 1,n$, $u_1 = u$, $u_n = v$ in (2.3)) is equivalent to $1 - f(u+1|u+1,m,n) = f(v|v+n,m,n)$. A straight forward calculation shows that the last equation is equivalent to

$$\prod_{i=1}^{n}(m-u-1+i) + \prod_{i=1}^{n}(v+i) = \prod_{i=1}^{n}(m+i).$$

In other words, we are looking for integer solutions a,b,c satisfying the diophantine equation

$$\prod_{i=1}^{n}(a+i) + \prod_{i=1}^{n}(b+i) = \prod_{i=1}^{n}(c+i). \tag{3.5}$$

Harborth (1988) called this type of equation a Fermat-like binomial equation. More precisely, he as well as Wunderlich (1962) for $n = 3,4$, considered the equation

$$\binom{a+n}{n} + \binom{b+n}{n} = \binom{c+n}{n}.$$

Fraenkel (1971) considered a more general class of diophantine equations by replacing i in (3.5) by $i\delta$, δ being a positive integer.

However, (3.5) has at least one solution for each $n \geq 2$, namely $a = b = n-1, c = n$. This has already been mentioned in Fraenkel (1971). By the way, this corresponds to the case $\alpha = 1/2$ and $m = n$. Now the question is, whether there exist further solutions for $n \geq 3$. For $n = 3$ it is mentioned in Wunderlich (1962) that S. Chowla (in a paper which seems to be unpublished) proved that the number of solutions of (3.5) is infinite. A proof can be found in Sierpiǹski (1962), confer also the remarks in Fraenkel (1971) on this topic. Setting $a = u - v - 1$, $c = u + v - 1$ and $b = 4v - 1$, for $n = 3$ (3.5) is equivalent to $31u^2 - 3v^2 = 1$, or, setting $x = 31u$, we get

$$x^2 - 93v^2 = 31, \tag{3.6}$$

which is a diophantine equation (or Pell equation) of the type $x^2 - Cy^2 = R$. It is well known, that a diophantine equation of this type has either infnitely many solutions or no solution. Since $x = 14 \times 31$, $v = 45$ solves (3.6), we obtain in fact that (3.5) has infinitely many positive integer solutions for $n = 3$.
For $n \geq 4$, the number of further solutions seems to be rather small although Harborth (1988) proved that there exist infinitely many n's with further solutions. Fraenkel (1971) conjectures that the number of solutions is finite for any fixed $n > 3$. Numerical calculations have shown, that there exist only seven solutions (n, a, b, c) for $4 \leq n \leq 500$, $a \leq b$, $1 \leq c \leq 500$ namely $(4, 128, 186, 196)$, $(6, 8, 9, 10)$, $(6, 13, 13, 15)$, $(35, 83, 83, 85)$, $(40, 63, 64, 65)$, $(204, 491, 491, 493)$ and $(273, 440, 441, 442)$. We found no further solution for $n = 4$ for $c \leq 126900$, for $n = 5, 6$ for $c \leq 10000$. As far as we know it is not known whether the set of solutions of (3.5) for fixed $n \geq 4$ is finite.
Some readers may be interested in visiting the WWW-address of Sloane (2000) where sequences of solutuions (which are partially slightly shifted) for $n = 2, 3, 4$ of (3.5) can be found, cf. the sequences A012132, A002311, A020329.

A further question is whether we can conclude that the power functions of the UMPU-tests based on sample sizes (m, n) and $(m + 1, n)$ satisfy $\beta(p, q|m, n, \alpha) \neq \beta(p, q|m + 1, n, \alpha)$ for all $p \neq q$, $p, q \in (0, 1)$, provided that they do not coincide on the entire parameter space. If the power functions coincide on an open subset of the parameter space, then it is immediate that they coincide on the entire parameter space.

It remains the question whether there exit further practically relevant models with testing situations, where the power of optimal tests considered as a function of the sample sizes is not uniformly increasing.

Appendix

Proof of Lemma 1. To prove part a) and b) we make use of the following facts, which are valid for all $m, n \in I\!N$, $x \in \{0, \dots, m\}$, $y \in \{0, \dots, n\}$.

$$\psi(x|x + y, m, n, \alpha) \in (0, 1) \Rightarrow \psi(x|x + y, m, n, \alpha) = \frac{\alpha - 1 + F(x|x + y, m, n)}{f(x|x + y, m, n)} \tag{A.1}$$

$$\left.\begin{array}{l}\psi(x - 1|x + y, m, n, \alpha) = 0 \\ \psi(x + 1|x + y, m, n, \alpha) = 1\end{array}\right\} \Rightarrow \psi(x|x + y, m, n, \alpha) = \frac{\alpha - 1 + F(x|x + y, m, n)}{f(x|x + y, m, n)} \tag{A.2}$$

(A.3) $$\psi(x|x+y,m+1,n,\alpha) \le \psi(x|x+y,m,n,\alpha) \le \psi(x+1|x+1+y,m+1,n,\alpha)$$

(A.4) $$\psi(x|x+y,m,n,\alpha) \le \psi(x+1|x+1+y,m,n,\alpha) \le \psi(x+1|x+y,m,n,\alpha)$$

(A.5) $$H(x|m,p) - H(x|m-1,p) = pg(x-1|m-1,p),$$

where $H(x,m,p) = \sum_{y=x}^{m} g(y|m,p)$ for $x \in \{0,\dots,m\}$ and $H(x,m,p)=0$ for $x > m$.

(A.6) $$\frac{g(x|m,p)}{f(x|x+y,m,n,p)} = \frac{g(x+y|m+n,p)}{g(y|n,p)}$$

(A.7) $$\frac{m}{m-x}(1-p)g(x|m-1,p) = g(x|m,p) = \frac{m}{x}pg(x-1|m-1,p) \qquad x \notin \{0,m\}$$

(A.8) $$F(x|x+y,m,n) = (1-\frac{x+y}{m+n})F(x|x+y,m-1,n) + \frac{x+y}{m+n}F(x-1|x-1+y,m-1,n)$$

The monotonicity properties (A.3) and (A.4) can be found in Finner & Strassburger (2000). All remaining properties are easily verified.

Proof of Lemma 1 Part a). Since $\psi(0|0,m,n,\alpha) = \alpha$ for all $m,n \in I\!N$ there exists an integer $u \in \{1,\dots,m\}$, such that

$$\psi(x|x,m+1,n,\alpha) = \begin{cases} \dfrac{\alpha}{f(x|x,m+1,n)}, & \text{for } 0 \le x \le u+1, \\ 1, & \text{otherwise.} \end{cases}$$

Combining (A.1) with (A.3) we conclude that there exist $r \in [0,1]$, so that

$$\psi(x|x,m,n,\alpha) = \begin{cases} \dfrac{\alpha}{f(x|x,m,n)}, & \text{for } 0 \le x \le u, \\ r, & \text{if } x = u+1, \\ 1, & \text{otherwise.} \end{cases}$$

For $p \in (0,1)$ we get with (A.6)

$$\begin{aligned} \beta_1(p|0,m+1,n,\alpha) &= H(u+2|m+1,p) + \alpha \sum_{x=0}^{u+1} g(x|m+1,p)/f(x|x,m+1,n) \\ &= H(u+2|m+1,p) + \alpha \sum_{x=0}^{u+1} (1-p)^{-n} g(x|m+n+1,p) \\ &= H(u+2|m+1,p) + \alpha(1-p)^{-n}[1 - H(u+2|m+n+1,p)]. \end{aligned}$$

The identity (A.5) yields

(A.9) $$\begin{aligned} \beta_1(p|0,m+1,n,\alpha) &= H(u+2|m,p) + pg(u+1|m,p) \\ &\quad +\alpha(1-p)^{-n}[1 - H(u+2|m+n,p) - pg(u+1|m+n,p)]. \end{aligned}$$

Similarly we get

$$\beta_1(p|0,m,n,\alpha) = H(u+2|m,p) + rg(u+1|m,p) + \alpha(1-p)^{-n}[1 - H(u+1|m+n,p)].$$

Combining this equation with (A.9) we obtain

(A.10) $$\begin{aligned} &\beta_1(p|0,m+1,n,\alpha) - \beta_1(p|0,m,n,\alpha) \\ &\quad = (p-r)g(u+1|m,p) + \alpha(1-p)^{-n}[g(u+1|m+n,p) - pg(u+1|m+n,p)]. \end{aligned}$$

If $u = m$, the r.h.s. of (A.10) is equal to $\alpha(1-p)^{1-n}g(m+1|m+n,p)$, which is positive for all $p \in (0,1)$ so we can reduce our attention to the case $u \leq m-1$. Now the r.h.s. of (A.10) equals $p^{u+1}(1-p)^{m-u-1}[(p-r)\binom{m}{u+1} + \alpha\binom{m+n}{u+1}(1-p)]$. This term equals zero for all $p \in (0,1)$ if and only if

$$p\left[\binom{m}{u+1} - \alpha\binom{m+n}{u+1}\right] = r\binom{m}{u+1} - \alpha\binom{m+n}{u+1} \qquad \forall p \in (0,1).$$

But this can happen if and only if $\binom{m}{u+1}/\binom{m+n}{u+1} = \alpha$ and $r = 1$, or, equivalently $F(u|u+1,m,n) = 1-\alpha$. □

Proof of Lemma 1 part b). Let $y \in \{1,\ldots,n-1\}$. By assumption there exists an $u \in \{0,\ldots,m-1\}$ such that $F(u,u+y,m,n) = 1-\alpha$, hence $\psi(u|u+y,m,n,\alpha) = 0$ and $\psi(u+1|u+y,m,n,\alpha) = 1$. The monotonicity property (A.3) yields $\psi(u|u+y,m+1,n,\alpha) = 0$ and $\psi(u+2|u+y+1,m+1,n,\alpha) = 1$ and by combining (A.3) and (A.4) we get $\psi(u|u+y+1,m+1,n,\alpha) = 0$. Together with (A.2) this yields $\psi(u+1|u+y+1,m+1,n,\alpha) = (\alpha - 1 + F(x|x+y,m+1,n))/f(x|x+y,m+1,n)$. Hence there exist an integer $v \in \{u+1,\ldots,m\}$ such that

$$\psi(x|x+y,m+1,n,\alpha) = \begin{cases} 0, & \text{for } 0 \leq x \leq u, \\ \dfrac{\alpha - 1 + F(x|x+y,m+1,n)}{f(x|x+y,m+1,n)}, & \text{for } u+1 \leq x \leq v+1, \\ 1, & \text{otherwise.} \end{cases}$$

Similarly as in the proof of part a) by making use of (A.3) and (A.1) we conclude that there exists a real number $r \in (0,1]$, such that

$$\psi(x|x+y,m,n,\alpha) = \begin{cases} 0, & \text{for } 0 \leq x \leq u, \\ \dfrac{\alpha - 1 + F(x|x+y,m,n)}{f(x|x+y,m,n)}, & \text{for } u+1 \leq x \leq v, \\ r, & \text{for } x = v+1, \\ 1, & \text{otherwise.} \end{cases}$$

Using (A.6) we get

$$\text{(A.11)} \qquad \beta_1(p|y,m,n,\alpha) = H(v+2|m,p) + rg(v+1|m,p) + \sum_{x=u+1}^{v} W(x|y,m,n,p,\alpha)$$

and

$$\text{(A.12)} \qquad \beta_1(p|y,m+1,n,\alpha) = H(v+2|m+1,p) + \sum_{x=u+1}^{v+1} W(x|y,m+1,n,p,\alpha),$$

where W is defined by

$$W(x|y,m,n,p,\alpha) = \frac{g(x+y|m+n,p)}{g(y|n,p)}[\alpha - 1 + F(x|x+y,m,n)].$$

With (A.8) and (A.7) we obtain

$$\begin{aligned} W(x|y,m+1,n,p,\alpha) &= (1 - \frac{x+y}{n+m+1})\frac{g(x+y|m+n+1,p)}{g(y|n,p)}[\alpha - 1 + F(x|x+y,m,n)] \\ &\quad + \frac{x+y}{n+m+1}\frac{g(x+y|m+n+1,p)}{g(y|n,p)}[\alpha - 1 + F(x-1|x-1+y,m,n)] \end{aligned}$$

$$\begin{aligned}
&= (1-p)\frac{g(x+y|m+n,p)}{g(y|n,p)}[\alpha-1+F(x|x+y,m,n)] \\
&\quad +p\frac{g(x-1+y|m+n,p)}{g(y|n,p)}[\alpha-1+F(x-1|x-1+y,m,n)] \\
&= (1-p)W(x|y,m,n,p,\alpha)+pW(x-1|y,m,n,p,\alpha),
\end{aligned}$$

i.e., (A.12) can be written as

$$\text{(A.13)} \qquad \begin{aligned} \beta_1(p|y,m+1,n,\alpha) &= H(v+2|m+1,p)+\sum_{x=u+1}^{v} W(x|y,m,n,p,\alpha) \\ &\quad +(1-p)W(v+1|y,m,n,p,\alpha)+pW(u|y,m,n,p,\alpha) \end{aligned}$$

Noting that $W(u|y,m,n,p,\alpha)=0$ since $F(u|u+y,m,n)=1-\alpha$, combination of (A.11) and (A.13) and application of (A.5) results in

$$\text{(A.14)} \qquad \begin{aligned} \beta_1(p|y,m+1,n,\alpha)-\beta_1(p|y,m,n,\alpha) &= (p-r)g(v+1|m,p) \\ &\quad +(1-p)W(v+1|y,m,n,p,\alpha). \end{aligned}$$

If $v=m$ the r.h.s. of (A.14) is equal to $\alpha\binom{m+n}{m+y+1}p^{m+1}/\binom{n}{y}$, which is positive for all $p\in(0,1)$. So we can reduce attention to the case $v\le m-1$. Now the r.h.s. of (A.14) equals

$$g(v+1|m,p)\left[(p-r)+(1-p)\frac{\alpha-1+F(v+1|v+1+y,m,n)}{f(v+1|v+1+y,m,n)}\right].$$

This term equals zero for all $p\in(0,1)$ if and only if

$$p+(1-p)\frac{\alpha-1+F(v+1|v+1+y,m,n)}{f(v+1|v+1+y,m,n)}=r \qquad \forall p\in(0,1).$$

But this can happen if and only if $[\alpha-1+F(v+1|v+1+y,m,n)]/f(v+1|v+1+y,m,n)=1$ and $r=1$, or, equivalently $F(v|v+y+1,m,n)=1-\alpha$. This completes the proof. □

Acknowledgement. We would like to thank Werner Schachinger from the University of Vienna for some helpful hints concerning diophantine equations.

References

Finner, H. & Strassburger, K. (2000). Structural properties of UMPU-tests for 2×2 tables and some implications. *Submitted for publication.*

Fisher, R. A. (1934) Statistical methods for research workers. (originally published 1925, 14th ed. revised and enlarged 1973. *Oliver and Boyd, Edinburgh.*

Fisher, R. A. (1935). The logic of inductive inference. *J. R. Statist. Soc., n. Ser.* **98**, 39-82.

Fraenkel, A. S. (1971). Diophantine equations involving generalized triangular and tetrahedral numbers. *Computers in Number Theory, Proc. Atlas Sympos. No. 2, Oxford 1969,* 99-114.

Harborth, H. (1988). Fermat-like binomial equations. *In: A.N. Philippou et al. (eds.), Applications of Fibonacci Numbers, Kluwer Academic Publ., 1* 1-5.

Irwin, J. O. (1935). Tests of significance for difference between percentages based on small numbers. *Metron* **12**, 83-94.

Khatri, M. N. (1955). Triangular numbers and Phythagorean triangles. *Scripta Math* **21**, 94.

Sierpiǹski, W. (1962). Sur une propriètè des nombres tètraèdraux. *Elem. Math.* **17**, 29-30.

Sloane, N. J. A. (2000). The On-Line Encyclopedia of Integer Sequences. Published electronically at http://www.research.att.com/∼njas/sequences/.

Tocher, K. D. (1950) Extension of the Neyman-Pearson theory of tests to discontinuous variates. *Biometrika* **37**, 130-144.

Wunderlich, M. (1962). Certain properties of pyramidal and figurate numbers. *Math. Comp.* **16**, 482-486.

Yates, F. (1934). Contingency tables involving small numbers and the χ^2-test. *J. R. Statist. Soc. Supp.* **1**, 217-235.