A C T A U N I V E R S I T A T I S L O D Z I E N S I S
FOLIA OECONOMICA 131, 1993

*Malte Bismarck**

THE STATISTICAL PLANNING OF EXPERIMENTS YIELDING TO CONTINGENCY TABLE ANALYSIS

1. INTRODUCTION

In medical applications, e.g. clinical statistics: relatively few patients are spread over a large number of cells in a contingency table. Then independence, homogeneity or other hypotheses (esp. in higher - dimensional tables) become hardly to reject with reasonable confidence because small accidental changes in the observed frequencies will dramatically change the pattern of frequencies at all.

E x a m p l e. In an investigation in the relation of the prognosis of Multiple Sclerosis (classified into four categories) to the results of a liquor analysis (classified into six categories) the distribution of n = 51 patients has arisen as shown in Table 1.

The often used in practice asymptotic X^2 test of independence gives a nonsignificant result ($P(\underline{X^2} \leq 13.933) = 0.469$ on the commonly used significance levels.

In the Neyman-Pearson theory of statistical testing the relation between significance level, sample size, and some measure of the "distance" between null and alternative hypotheses determines the power of the test, i.e. the probability of rejecting the null hypothesis if the alternative is true. As the power of a

* Doctor at the Institute of Biostatistics and Medical Informatics, Martin Luther University Halle - Wittenberg.

Table 1

Results of liquor analysis and prognosis of 51 patients with multiple sclerosis

Result	Prognosis				n_i
	1	2	3	4	
1	2	1	0	0	3
2	0	0	2	1	3
3	3	6	1	1	11
4	10	9	5	2	26
5	2	1	0	0	3
6	2	2	1	0	5
$n._j$	19	19	9	4	n = 51

Source: The author's calculations.

statistical test in general is increasing with increasing sample size one could ask how large should the sample size be to make the detection of an association of the two responses safe enough.

Most simply one could say: if the observed relative frequencies remained stable also with larger sample sizes then, at least for $n = 51 \times 24.996/13.933 \approx 92$ (24.996 being the upper 5% value of the chisquared distribution with 15 degrees of freedom) the observed relations in the table would yield a significant association between the responses. This is equivalent to the determination of the required sample size to achieve the prespecified length of a confidence interval (e.g. Bristol 1989). Here, the value of power is not taken into account. As it indicates the probability of right rejection of the null hypothesis is additionally used for determining the necessary sample size.

This paper considers the calculation of power for tests of independence and homogeneity in two-way contingency tables. After an overview of the known theoretical results a discussion of practical aspects connected with the power calculations on the introductory example follows. The special results obtained for 2 x 2 tables will not be discussed here (cf., e.g., Suissa and Shuster 1985). Also higher - dimensional tables will not be considered (cf. Oler 1985), and other designs like the McNemar test or the Mantel-Haenszel procedure are excluded from discussion.

2. TESTS IN TWO - WAY TABLES

The calculation of power for the test of independence or homogeneity in a two - dimensional contingency table depends on the sampling situation from which the table arises. If the observational units are independent from each other three different situations are traditionally distinguished. Probably, for the first time B a r n a r d (1947) described them for a 2 x 2 table, and . R o y and M i t r a (1956) extended the concept to a x b tables. The respective formulations will be given now for the general a x b table, a, b $\geq$ 2.

CASE 1. One sample of size n is drawn, and the objects are classified corresponding to two categorical responses A and B. That means for the cited example that at first 51 patients have been randomly chosen from all accessible patients with Multiple Sclerosis and then they have been classified according both to results and prognosis. The mathematical model is a multinomial distribution

$$P((\underline{n_{ij}}) = (n_{ij})) = (n_{11} \overset{n}{\cdots} n_{ab}) \prod_{i=1}^{a} \prod_{j=1}^{b} p_{ij}^{n_{ij}}, (n_{11} \overset{n}{\cdots} n_{ab}) =$$

$$= \frac{n!}{n_{11}!n_{12}!\ldots n_{16}!n_{21}!\ldots n_{ab}!}, \quad \sum_{i=1}^{a} \sum_{j=1}^{b} p_{ij} = 1$$

"multinominal coefficient" (2.1)

Under the hypothesis H_{o1} of independence of A and B we have

$$p_{ij}^{o} = p_{i.}\, p_{.j}, \qquad i = 1(1)a, \qquad j = 1(1)b \tag{2.2}$$

where the $(p_{i.})$ and $(p_{.j})$ have to be specified otherwise (determined from further information or estimated from the data). Therefore, they are called nuisance parameters.

The probabilities (2.1) with (2.2) may be cumulated according to their increasing values. Then the "unconditional" critical region consists of all points (n_{ij}) on the a x b - dimensional lattice which fulfill the side condition $\sum_{i=1}^{a} \sum_{j=1}^{b} n_{ij} = n$ with cumulative probability smaller then or equal to the chosen significance level alpha. Of course, this critical region depends on the values of the nuisance parameters $(p_{i.}, p_{.j})$. The exact un-

conditional power is then the sum of (1.1) with $p_{ij} = p_{ij}^{1}$, $\sum\sum p_{ij}^{1} = 1$, over the critical region.

Really, the hypothesis tested in this way is not exactly H_{o1} but a somewhat more restrictive H_{o1}:

$$p_{ij}^{o} = p_{i.}\, p_{.j},$$

and

$$p_{i.} = p_{i.}^{o}, \quad p_{.j} = p_{.j}^{o}, \quad i = 1(1)a, \quad j = 1(1)b, \tag{2.3}$$

$(p_{i.}^{o})$ and $(p_{.j}^{o})$ being the specified values of the nuisance parameters.

The practical determination of the critical region and the calculation of power is a difficult task in all but the simplest cases because of the large computational amount. So almost always approximations with the chisquared distribution are used.

CASE 2. For each category A_i of A one sample of size $n_{i.}$ is drawn ($n_{i.}$ fixed before sampling) independent from the others and classified according to the response B. For the cited medical situation in this case before conducting the experiment it has been decided to sample 3 patients with results "1", 3 patients with results "2" etc., and 5 patients with results "6". On sampling they all would be classified into the categories of B. Mathematically, the common distribution of all $(\underline{n_{ij}})$ is the product of "a" multinomial distributions

$$P((\underline{n_{ij}}) = (n_{ij})) = \prod_{i=1}^{a} \binom{n_{i.}}{n_{i1}\ldots n_{ib}} \prod_{j=1}^{b} p_{ij}^{n_{ij}}, \quad \sum_{j=1}^{b} p_{ij} = 1,$$

$$i = 1(1)a \tag{2.4}$$

One can test the homogeneity of the distributions over B for the categories of A. The hypothesis H_{o2} may be formulated as

$$p_{ij}^{o} = p_{.j}, \quad j = 1(1)b \tag{2.5}$$

Now the $(p_{.j})$ are nuisance parameters for which the respective statements are valid as in Case 1. With specified values $(p_{.j}^{o})$ of $(p_{.j})$ an analogous procedure is possible as in the former case: construct a critical region from (2.4) with (2.5) and calculate the exact power from (2.4) with

$$p_{ij} = p_{ij}^{1}, \qquad \sum_{j=1}^{b} p_{ij}^{1} = 1, \qquad i = 1(1)a.$$

Similarly the null hypothesis then contains not only the homogeneity of distributions but also the specification of the parameter values $(p_{.j}^{o})$. The amount of computations involved is not so extensive as in Case 1 but yet too large for most practical situations. So it is again necessary to use approximations.

CASE 3. Another possibility in Cases 1 and 2 for determining the values of $(p_{i.})$ and/or $(p_{.j})$ is the estimation from the sampled data. For instance, maximum likelihood estimation yields

$$\hat{p}_{i.} = \frac{n_{i.}}{n}, \qquad \hat{p}_{.j} = \frac{n_{.j}}{n}, \qquad i = 1(1)a, \qquad j = 1(1)b \tag{2.6}$$

Then one can consider the distribution of $(\underline{n_{ij}})$ conditional on the observed $(n_{i.})$ and/or $(n_{.j})$.

This is the same distribution as if the values $(n_{i.})$ and $(n_{.j})$ were fixed in advance and the objects are sampled so that exactly $n_{i.}$ of them fall in class A_i and at the same time $n_{.j}$ in class B_j for all i and j. Practical applications of this sampling situation are seldom found, the classical example being the "lady tasting tea" problem cited by F i s h e r (1966).

The distribution is called hypergeometric distribution because its generating function is connected with the hypergeometric function. Under the hypothesis H_{o3} of independence of A and B the probabilities are given by

$$P((\underline{n_{ij}}) = (n_{ij}) \,|\, (n_{i.}), (n_{.j})) = \frac{\prod_{i=1}^{a} n_{i.}! \prod_{j=1}^{b} n_{.j}!}{n! \prod_{i=1}^{a} \prod_{j=1}^{b} n_{ij}!} \tag{2.7}$$

Again a critical region can be constructed from these probabilities, and it is a conditional one. Even in this case the testing procedure is often practically not feasible because of the same reason as before. The fastest known algorithm for calculation of the actual size of the test (ba M e h t a and P a t e l 1986, and further developed by J o e 1988) uses some shortcuts to reduce the amount of computation.

The conditional power function is determined by the extended hypergeometric distribution which may be written as

$$P((\underline{n_{ij}}) = (n_{ij})) = P((\underline{n_{rs}}) = (n_{rs}) | (n_{i.}), (n_{.j}), (\lambda_{rs}{}^{1})) =$$

$$= \{\exp\left(\sum_{r=1}^{a-1}\sum_{s=1}^{b-1} n_{rs}\lambda_{rs}{}^{1}\right) / \prod_{i=1}^{a}\prod_{j=1}^{b} n_{ij}!\} /$$

$$/ \{\sum_{r=1}^{a-1}\sum_{s=1}^{b-1} \exp\left(\sum_{r=1}^{a-1}\sum_{s=1}^{b-1} n_{rs}\lambda_{rs}{}^{1}\right) / \prod_{i=1}^{a}\prod_{j=1}^{b} n_{ij}!\} \qquad (2.8)$$

The noncentrality parameters $(\lambda_{rs}{}^{1})$ are connected with the $(p_{ij}{}^{1})$ in the following way:

$$\lambda_{rs}{}^{1} = \ln\left(p_{rs}{}^{1}\, p_{ab}{}^{1} / p_{rb}{}^{1}\, p_{as}{}^{1}\right), \quad r = 1(1)a - 1, \quad s = 1(1)b - 1 \qquad (2.9)$$

and called "odds ratios". The null hypothesis H_{o3} may be expressed as

$$\lambda_{rs} = 0, \quad r = 1(1)a - 1, \quad s = 1(1)b - 1 \qquad (2.10)$$

The power calculation is even more extensive than the calculation of the size of the test, and no shortcuts for speeding up the computations as in the latter case are known till now. Therefore the situation is analogous to those in the other both cases.

One of the most commonly used form of asymptotic tests for the null hypothesis in the three distinguished cases is that based on the chisquared statistic

$$\underline{X^2} = \sum_{i=1}^{a}\sum_{j=1}^{b} \frac{(n_{ij} - e_{ij})^2}{e_{ij}} \qquad (2.11)$$

with

$$e_{ij} = n\, p_{ij}{}^{o} \qquad \text{for } H_{o1}{}^{(')} \qquad (2.12)$$

$$e_{ij} = n_{i.}\, p_{.j}{}^{o} \qquad \text{for } H_{o2}{}^{(')} \qquad (2.13)$$

and

$$e_{ij} = n_{i.}\, n_{.j}/n \quad \text{for} \quad H_{o3}, \quad i = 1(1)a, \; j = 1(1)b \qquad (2.14)$$

This statistic is asymptotically distributed as a central chisquared variate under the null hypothesis. The degrees of freedom are equal to

$a \times b - 1$ in the case of H'_{o1} (2.15)

$a(b - 1)$ in the case of H'_{o2} (2.16)

$(a - 1)(b - 1)$ in the case of H_{o3} (2.17)

The likelihood ratio statistic

$$\underline{Y^2} = \sum_{i=1}^{a} \sum_{j=1}^{b} 2\, n_{ij} \ln (n_{ij} / e_{ij}) \tag{2.18}$$

and the Freeman-Tukey statistic

$$\underline{T^2} = 4n \sum_{i=1}^{a} \sum_{j=1}^{b} (\sqrt{n_{ij}/n} - \sqrt{e_{ij}/n})^2 \tag{2.19}$$

are asymptotically equivalent to the statistic $\underline{X^2}$.

The following discussion will be concerned with the statistic $\underline{X^2}$, but analogous results are valid for the other two statistics. The limiting chisquared distribution of $\underline{X^2}$ under H_{o3} can be derived as conditional or as marginal distribution under the hypotheses of independence in Cases 1 and 2. B i s m a r c k (1988) gave a systematic overview over the existing literature on the relation between the exact and asymptotic methods cited above together with the results of own investigations concerned with this problem.

3. APPROXIMATE POWER CALCULATIONS

A general formulation of the asymptotic power of tests based on X^2 was given by M i t r a (1958). Mitra considered alternatives of Pitman type, i.e. that tend to the null hypothesis with sample size increasing to infinity.

M e n g and C h a p m a n (1966) published a general formulation covering the Cases 1 and 2. Although presented without proof, the formulas given in C o h e n (1977) refer to Case 3. An explicit derivation of the noncentral chisquared distribution with the respective noncentrality parameter as the limiting distribution for the extended hypergeometric distribution is not known to the author from the literature.

The results given in the above cited papers may be formulated as follows:

Under the alternative hypothesis of Cases 1, 2, and 3, resp., the statistic $\underline{X^2}$ is asymptotically distributed as a noncentral chisquared variate with degrees of freedom ν as in the central case and noncentrality parameter Λ. This parameter is given by the following expressions:

CASE 1.

$$\Lambda = n \left[\sum_{i=1}^{a} \sum_{j=1}^{a} \frac{(p_{ij}^{1} - p_{i.}^{o}\, p_{.j}^{o})^2}{p_{i.}^{o}\, p_{.j}^{o}} - \sum_{i=1}^{a} \frac{(p_{i.}^{1} - p_{i.}^{o})^2}{p_{i.}^{o}} - \right.$$

$$\left. + \sum_{j=1}^{b} \frac{(p_{.j}^{1} - p_{.j}^{o})^2}{p.j^{o}} \right] \qquad (3.1)$$

with

$$p_{i.}^{o/1} = \sum_{j=1}^{b} p_{ij}^{o/1}, \quad i = 1(1)a$$

$$p_{.j}^{o/1} = \sum_{i=1}^{a} p_{ij}^{o/1}, \quad j = 1(1)b.$$

CASE 2.

$$\Lambda = n \cdot \left\{ \sum_{j=1}^{a} \frac{1}{p_{.j}^{o}} \left[\sum_{i=1}^{a} Q_i \, (p_{ij}^{1} - p_{.j}^{o})^2 + \right. \right.$$

$$\left. \left. - \left[\sum_{i=1}^{a} Q_i \, (p_{ij}^{1} - p_{.j}^{o}) \right]^2 \right] \right\} \qquad (3.2)$$

$$Q_i = n_{i.}/n, \qquad i = 1(1)a, \qquad \sum_{j=1}^{b} p_{.j}^{o} = 1.$$

CASE 3.

$$\Lambda = n^2 \left[\sum_{i=1}^{a} \sum_{j=1}^{b} \frac{(p_{ij}^{1} - n_{i.}\, n_{.j}/n^2)^2}{n_{i.}\, n_{.j}} \right]. \qquad (3.3)$$

The probability function of a noncentrally chisquared distributed variate $\underline{\chi^{12}}$ may be computed from that of a centrally distributed one through

$$P(\underline{\chi^{12}} \leq \chi^{12}[\ \nu,\ \Lambda) = P(\chi^{12}]\ \nu,\ \Lambda) =$$

$$= \sum_{r=0}^{\infty} e^{-\Lambda/2} \cdot \frac{(\Lambda/2)^r}{r!} P(\chi^{12}|\ \nu + 2r) \qquad (3.4)$$

(A b r a m o w i t z and S t e g u n 1966).

Tables of the noncentral chisquared distribution for several values of ν and Λ have been published by H a y n a m, G o v i n d a r a j u l u and L e o n e (1970) and again by H a y n a m et al. (1982, 1983).

From the formula of the probability function of a noncentral chisquared variate or these tables it can be seen that the power of the considered tests is generally increasing with increasing value of Λ. In all of the above cited cases the noncentrality parameter depends on the sample size n and some measure of the "distance" between null and alternative hypotheses related to the measure given by the statistic X^2. That is power is really increasing with increasing n and/or increasing "distance".

L a c h i n (1977) examined the determination of sample sizes given (asymptotic) power for a x b contingency tables in cases 1 and 2. In detail he considered the situation that

$$\text{(S1)} \qquad p_{.j}^{\,o} = \sum_{i=1}^{a} p_{ij}^{\,1}/a \qquad (3.5)$$

$$\text{(S2)} \qquad p_{.j}^{\,o} = p_{kj}^{\,1} \quad \text{for some } k,\ 1 \leq k \leq a \qquad (3.6)$$

and

$$\text{(S3)} \qquad p_{.j}^{\,o} = \sum_{j=1}^{a} p_{ij}^{\,1}\, Q \qquad (3.7)$$

for Case 2. In the latter situation the $p_{.j}^{\,o}$ are also functions of the sample fractions so that the sample allocation must be optimized which yields a nonlinear programming problem. In clinical trials the situation may be additionally constrained by the demand for the comparison of one placebo group with a-1 equally sized treatment groups. Then the total number assigned to the a-1 treatment groups ought to be minimized. Lachin offered the solutions as analytical procedures and APL programs.

Unfortunately, even the known computer program for calculating sample sizes in the analysis of two - dimensional contingency

tables does not seem to distinguish between the several possible sampling situation (S a y n and M e r k l e 1989) what is necessary from the theoretical arguments cited above.

4. ABOUT PRACTICAL POWER CALCULATIONS

The value of the asymptotic power depends on the values of sample size and noncentrality parameter, i.e. the specified alternative hypothesis. Little is known about guidelines for the determination of Λ under practical aspects. It may be determined from further information about the practical situation in which the statistical test should be applied. But in the medical context this is generally a difficult task and can only be done in cooperation between the statistician and the doctor.

C o h e n (1977) gave some hints what small or large deviations from H_o mean. He introduced a so - called "effect size index" $w = \sqrt{\Lambda/n}$ which is a function of the distance between both hypotheses. Then he proposed the following scale:

distance	effect size index	contingency coefficient
small	$w = 0.10$	$C = 0.100$
medium	$w = 0.30$	$C = 0.287$
large	$w = 0.50$	$C = 0.447$

The "effect size index" and Pearson´s coefficient of contingency C are related by

$$C = \sqrt{\frac{X^2}{X^2 + n}} = \sqrt{\frac{w^2}{w^2 + 1}} \qquad (4.1)$$

Now the consequences of the determination of Λ will be explained on the introductory example. What kind of hypothetical values should be chosen as alternatives to independence between liquor results and prognosis?

If the observed frequencies were taken as alternative to independence then $\Lambda = X^2 = 13.933$, i.e. $w = 0.523$. For $\alpha = 0.05$ this gives a power value of 0.635 from formula (3.4) of the noncentral chisquared distribution (cf. Table 2).

T a b l e 2

Observed and expected (under independence) relative frequencies

Result	Prognosis								Σ
	1		2		3		4		
	obs.	exp.	obs.	exp.	obs.	exp.	obs.	exp.	
1	3.9	2.2	2.0	2.2	0	1.0	0	0.5	5.9
2	0	2.2	0	2.2	3.9	1.0	2.0	0.5	5.9
3	5.9	8.0	11.8	8.0	2.0	3.8	2.0	1.7	21.6
4	19.6	19.0	17.6	19.0	9.8	9.0	3.9	4.0	51.0
5	3.9	2.2	2.0	2.2	0	1.0	0	0.5	5.9
6	3.9	3.7	3.9	3.7	2.0	1.7	0	0.8	9.8
Σ	37.3		37.3		17.6		7.8		100.0

S o u r c e: The author's calculations.

The following minimal sample sizes are necessary to achieve higher power values:

Power	α = 5%	α = 1%
2/3	60	85
0.80	77	104
0.90	96	126

(taken from Cohen's table for the conservative values $\nu = 16$ and $w = 0.5$).

In this case the power is relatively fastly increasing with increasing sample size so that a not very much larger sample would yield better or even sufficient power. But a larger power value does not automatically induce a significant test result for the relations among the observed frequencies projected to the larger sample size as can be shown for

$\alpha = 5\%, \quad \Lambda = X_1^2 = 13.933$:

n	$1 - \beta \approx$	X^2	$P(X^2)$
60	2/3	$60/51 \times X_1^2 = 16.391$	$0.643 < 0.95$
77	0.80	$77/51 \times X_1^2 = 21.035$	$0.864 < 0.95$
96	0.90	$96/51 \times X_1^2 = 26.226$	$0.964 > 0.95$

The extreme case for the alternative, i.e. strong dependence would indicate that each class of liquor result is connected with one and only one class of prognosis. Assuming that in this case results "1" and "2" together yield prognosis "1" and results "5" and "6" yield prognosis "4" the probabilities are concentrated on the main diagonal of a 4 x 4 table.

If rows 1 and 2 and 5 and 6, resp., are merged the degrees of freedom are $\nu = 9$, the observed $X^2 = 4.988$ with $P(X^2) = 0.165$ under H_{o3}. For the observed relative frequencies as percentages under the alternative in Case 2 the following noncentrality parameters and "effect size indexes" result for several choices of the marginal distribution $(p_{.j}^{o})$ and n = 51:

	$(p_{.j}^{o})$	Λ	w	$1 - \beta$ $\alpha = 0.05$
I	(19/51, 19/51, 9/51, 4/51)	4.988	0.313	0.280
II	(6/51, 11/51, 26/51, 8/51)	6.060	0.345	0.342
III	(0.25, 0.25, 0.25, 0.25)	4.830	0.308	0.271

This is an example for medium "effect size indexes" in the sense of Cohen.

Under the alternative of strong dependence the following table of percentages would have been expected:

Result	Prognosis				Σ	$n_{i.}$
	1	2	3	4		
1+2	1.0	0.0	0.0	0.0	1.0	6
3	0.0	1.0	0.0	0.0	1.0	11
4	0.0	0.0	1.0	0.0	1.0	26
5+6	0.0	0.0	0.0	1.0	1.0	8

The observed margins contradict the conditions of Case 3 above so that only sampling situations of the other two cases are possible. In Case 2 we get for this alternative:

$(p_{.j}^{o})$	Λ		w
I	$n\left[\sum_j \frac{Q_j}{p_{.j}^{o}} - 1 - \sum_j \frac{(Q_j - p_{.j}^{o})^2}{p_{.j}^{o}}\right]$	= 195.591	1.958
II	$n(b-1)$	= 153	1.732
III	$n\left[(b-1) - \sum_j \frac{(Q_j - 1/b)^2}{1/b}\right]$	= 133.647	1.619

The last table indicates that under strong dependence as extreme form of alternative the given sample size is large enough for reaching sufficiently high power. Indeed, the second line of the table shows as "effect size index" $\sqrt{b-1}$ which is at least 1 for $b > 2$ and therefore "large" in the sense of Cohen as the other "effect size indexes", too.

Perhaps a more realistic notion of dependence assigns every 50% to the cells in the main diagonal and 50/3 = 16.7% to the cells outside the main diagonal:

Result	Prognosis 1	2	3	4	Σ
1+2	0.50	0.167	0.167	0.167	1.0
3	0.167	0.50	0.167	0.167	1.0
4	0.167	0.167	0.50	0.167	1.0
5+6	0.167	0.167	0.167	0.50	1.0

In the three considered cases of hypothetical marginal distribution of prognosis we find:

$(p_{.j}^{o})$	Λ		w	$1 - \beta$ $\alpha = 0.05$
I		21.732	0.653	0.929
II	$n \cdot \frac{(c-2)^2}{4(c-1)}$	= 17	0.577	0.839
III	$n \cdot (\frac{c-2}{c-1})^2 \cdot \frac{c}{4} \cdot (1 - \sum_j p_{.j}^{2\,o}) =$	14.850	0.540	0.774

Even now the "effect size index" is "large", but the power values are not so high as in the former case.

An analogous discussion with corresponding results might be carried out for Case 1.

This example indicates that further examination of the relations between dimensions of the table, "effect size index" and sample size seems to be necessary resulting in more detailed suggestions for practical applications.

REFERENCES

Abramowitz M., Stegun I. A. (1964), *Handbook of Mathematical Functions*, National Bureau of Standards, Washington, D. C.

Barnard G. A. (1947), *Significance Tests for 2 x 2 Tables*, "Biometric" No. 34, p. 123-128.

Bismarck M. (1988), *Vergleichende Untersuchungen zu exakten und asymptotischen Tests in der Kontingenztafelanalyse*, Ph. D. Thesis, Martin Luther University Halle - Wittenberg.

Bristol D. R. (1989), *Sample Sizes for Constructing Confidence Intervals and Testing Hypotheses*, "Statistics in Medicine", No. 8, p. 803-811.

Cohen J. (1977), *Statistical Power Analysis for the Behavioral Sciences*, 2nd ed., New York, San Francisco, London.

Fisher R. A. (1966), *The Design of Experiments*, 8th ed., Oliver and Boyd, Edinburgh-London.

Haynam G. E., Govindarajulu Z., Leone F. C. (1970), *Tables of the Cumulative Non-central Chi-square Distribution*, [in:] H. L. Harter, D. R. Owen, *Selected Tables in Mathematical Statistics*, Vol. 1, Markham, Chicago.

Haynam G. E., et al. (1982, 1983), *Tables of the Cumulative Non-central Chi-square Distribution*, "Mathematische Operationsforschung und Statistik", Ser. Statistics, Nr. 13, p. 413-443, 517-634, Nr. 14, p. 75-139.

Joe H. (1988), *Extreme Probabilities for Contingency Tables under Row and Column Independence with Application to Fisher's Exact Test*, "Communications in Statistics. Theory and Methods", No. 17, p. 3677-3685.

Kendall M. G., Stuart A. (1966), *The Advanced Theory of Statistics*, Vol. 2, 2nd ed., Griffin, London.

Lachin J. M. (1977), *Sample Size Determinations for r x c Comparative Trials*, "Biometrics", No. 33, p. 315-324.

M e h t a C. R., P a t e l N. R. (1986), *FEXACT. A Fortran Subroutine for Fisher's Exact Test on Unordered r x c Contingency Tables*, "Transactions ACM", Mathematical Software, No. 12, p. 154-161.

M e n g R. C., C h a p m a n D. G. (1966), *The Power of Chi-square Tests for Contingency Tables*, "Journal of the American Statistical Association", No. 61, p. 965-975.

M i t r a S. K. (1958), *On the Limiting Power Function of the Frequency Chi-square Test*, "Annals of Mathematical Statistics", No. 29, p. 1221-1233.

O l e r J. (1985), *Noncentrality Parameters in Chi-squared Goodness - of - Fit Analyses with an Application to Log-linear Procedures*, "Journal of the American Statistical Association", No. 80, p. 181-189.

R o y S. N., M i t r a S. K. (1956), *An Introduction to some Non-parametric Generalizations of Analysis of Variance and Multivariate Analysis*, "Biometrika", No. 43, p. 361-376.

S a y n H., M e r k l e W. (1989), *Statistical Software for Sample Size Estimation*, "Statistical Software Newsletter, No. 15, p. 56-59.

S u i s s a S., S h u s t e r J. J. (1985), *Exact Unconditional Sample Sizes for the 2 x 2 Binomial Trial*, "Journal of Royal Statistical Society", No. 148(A), p. 317-327.

Malte Bismarck

STATYSTYCZNE PLANOWANIE EKSPERYMENTÓW ZWIĄZANE Z ANALIZĄ TABLIC WIELODZIELNYCH

Artykuł zawiera przegląd znanych teoretycznych wyników związanych z obliczeniami mocy testów niezależności i jednorodności w tablicach wielodzielnych a x b. Przedstawiono również wyniki obliczeń obserwowanych i oczekiwanych względnych częstości, przy założeniu niezależności elementów próby losowej. Podano także wyniki obliczeń empirycznych mocy ww. testów przy poziomie istotności 1% i 5%.