Scholars Academic Journal of Pharmacy
Abbreviated Key Title: Sch Acad J Pharm
ISSN 2347-9531 (Print) | ISSN 2320-4206 (Online)
Journal homepage: http://saspublishers.com

Pharmetric Laboratory

Construction of 2x2 Contingency Tables from Normally Distributed Data. Analysis of the Contingency Tables Using a Modified Binomial Test. Accounting for Differentiating and Integrating Factors in the Calculation of Asymptotic Significance

Stanislav V. Yefimov[1*]

[1]Pharmetric Laboratory, 11880 28th St N #210, 33716, St. Petersburg, FL, USA

DOI: 10.36347/sajp.2022.v11i06.001 | **Received:** 06.05.2022 | **Accepted:** 14.06.2022 | **Published:** 17.06.2022

***Corresponding author:** Stanislav V. Yefimov
Pharmetric Laboratory, 11880 28th St N #210, 33716, St. Petersburg, FL, USA

Abstract **Original Research Article**

The work is devoted to determining the comparative efficiency of analyzers with a binary scale: 1/0, +/-, yes/no. Such a task arises, for example, when validating a new method of microbiological analysis for sterility when comparing the results of the analysis of two methods, a new one, usually more sensitive, and a traditional one. We will analyze the procedure for constructing contingency tables for normally distributed data, and then analyze the constructed tables using the modified Binomial Test. We will point out the factors that are critical for the correct calculation of the asymptotic significance (p-value).

Keywords: Paired dichotomous data, 2×2 contingency table, asymptotic significance (p-value), Binomial Test; Yefimov method for p-value, Validation procedure for a new Rapid Microbiological Method.

INTRODUCTION

Two sets of dichotomous data are collected during the validation procedure for a new rapid microbiological method (USP 1223, 2008; Technical Report, 2013), the first data set is the data obtained by the reference method, and the second data set is the data obtained by the new method. The data are organized in 2x2 contingency tables (Felsenstein J., 2010), and the asymptotic significance (p-value) is calculated using either the Binomial test, McNemar's test, or Fisher's exact probabilistic test (Felsenstein J., 2010; 2x2 contingency Table, 1999; Abdi H., 2007). Based on the analysis of p-values, it is concluded which method is more sensitive. However, if the currently accepted procedures for validating a rapid microbiological method are followed, it is very common for a p-value >0.05, which indicates statistical equivalence of the methods tested, although according to other, independent assessments the sensitivity of the methods differ significantly. In this work, we will identify the reasons that affect the effectiveness of the analysis, modify the statistical method of analyzing 2x2 contingency tables, and give examples of the successful application of the modified method.

MATERIALS AND METHODS

Experiment Design

Let us justify the method for determining the sensitivity of an instrument using the example of weighing sand particles on two balances. We will construct contingency tables based on the result of hypothetical weighing randomly selected grains of sand from three populations of sand (light, medium, and heavy) on control and more sensitive balance. We will determine the probability of a positive result by weighing each of the groups of sand on the control balance. Then we will consider the reverse process, namely, using the constructed 2x2 contingency tables, using the binomial test with our modification, we will determine which of the balance is more sensitive. In the end, we will discuss the generalization of the obtained results to any test devices and point out the factor influencing the determination of the mutual sensitivity of testers. Finally, we will show how the method works by comparing two competing luminescent sterility analyzers.

Chemicals and instruments

Incubator 32.5°C ± 2°C; Celsis® Advance II; Celsis® Ampiscreen Reagent Kits; Vacuum manifold;

Biological Safety Cabinet; Eppendorf BioPur pipette tips; PALL micro funnels – GN6 membrane 0.45 microns; Refrigerator; Freezer; Eppendorf Centrifuge 5415D; Eppendorf Centrifuge Tubes 1.5 mL; Bio balls (Biomerieux); Fluid Thioglycollate Medium (FTM); Tryptic Soy Broth (TSB); PC HP Windows10, Free software "LibreOffice" version 6.0.0.3.

RESULTS AND DISCUSSION

Let's say we have three grades of sand (A, B, C). Suppose the grains of sand in each of the varieties have a normal distribution by weight (N (0.5,1), N(2,1), N(4,1)) (Figure 1). Let's say we have two balances: more sensitive (2) and less sensitive, control (1). The balance shows a positive result (+ or 1) if the weight exceeds its Limit of Detection (LOD), otherwise, the result is (0). The balance has two values 0 and 1. Let's say we weigh randomly chosen grains of sand from three populations (A, B, and C) on control (1), and then on a tested (2) balance. The results of the hypothetical weighing of the grains of sand are recorded in the 2x2 Contingency Table 1.

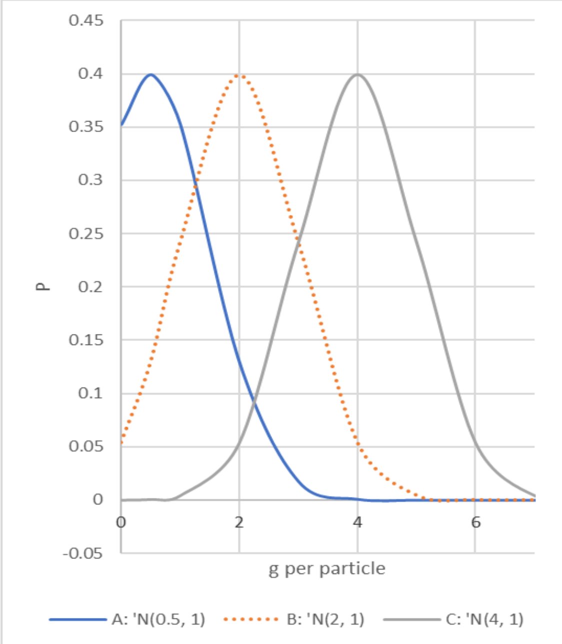

Fig-1: Probability density functions of the three populations of sand (A, B, C).

Table-1: 2x2 Contingency Table

	Balance 1 (+)	**Balance 1 (-)**	
Balance 2 (+)	a	b	a+b
Balance 2 (-)	c	d	c+d
	a+c	b+d	N=a+b+c+d

Let's prepare the data for filling in the 2x2 Contingency table. For the beginning, let us calculate the probabilities that the control balance -1, and balance -2 react to weight (+), and do not react to weight (-) when weighing a randomly selected grain of sand of a certain population N(μ, σ^2). Formulas are written in Microsoft

Excel notation

(1) p2+= 1-NORMDIST(LOD2, μ, σ, 1)
(2) p2-= NORMDIST(LOD2, μ, σ, 1) - NORMDIST(0, μ, σ, 1)
(3) p1+= 1-NORMDIST(LOD1, μ, σ, 1)
(4) p1-= NORMDIST(LOD1, μ, σ, 1) - NORMDIST(0, μ, σ, 1)

Note that this model is suitable for scales under normal laboratory conditions, that is, the weight of the particles is not negative. If the balance is in a vacuum, then we can weigh particles whose density is less than the density of air, and the weight scale becomes unlimited (-∞, +∞).

Let's fill in Table 1, but instead of numbers, let's start with their probabilities. Let's call this table a 2x2 Probability Table. So, for example, in the cell for value "a" we put the probability that the control balance -1 and the balance -2 show a positive result when weighing a grain of sand, that is, it reacts to weight. We will use the equations (*):

(*) a=p2+·p1+; b= p2+·p1-; c= p2-·p1+; d= p2-·p1-

Let's consider the fact that the probability (P) that the control balance responds to the weight of a grain of sand from different populations (A, B, and C) is not the same. Let's estimate this probability as:

(3a) P=p1+=1-NORMDIST (LOD1, μ, σ, 1). Having done the necessary calculations, we get Table 2.

Table-2: Example of 2x2 probability tables. LOD1=2, LOD2=0.5

A	A	B	B	C	C
0.0334	0.3123276	0.4666	0.44537	0.97702	0.02271
0.01279	0.119598	0.02203	0.02103	0.0002	4.6E-06
P=	0.0668072	P=	0.5	P=	0.97725
p2+ =	0.5	p2+ =	0.93319	p2+ =	0.99977
p2- =	0.1914625	p2- =	0.04406	p2- =	0.0002
p1+ =	0.0668072	p1+ =	0.5	p1+ =	0.97725
p1- =	0.6246553	p1- =	0.47725	p1- =	0.02272

Note that the sum of all cells of the 2x2 probability table is not equal to one. For population C, it is close to 1 (0.99994). This inequality is a consequence of the fact that the weight scale is bounded on the left. If the scale is not limited, the sum is 1, and obviously, the system has 3 degrees of freedom.

Now we are ready to calculate contingency tables (Table 3.) using the Probability Tables, for this, we multiply each cell of the Probability Table by a constant factor Q and the result is rounded up to an integer value. Thus, each contingency table (CT) we have built has 4 parameters: CT(N(μ, σ^2), LOD1, LOD2, Q).

Table-3: Example of 2x2 Contingency tables, derived from Table 2. Q=30, LOD1=2, LOD2=0.5. Probability (P) was calculated by the formula (3a).

A	A	B	B	C	C
1	9	14	13	29	1
0	4	1	1	0	0
P=	0.06681	P=	0.5	P=	0.97725

Evaluation of the sensitivity of a tester according to the Contingency table. Yefimov method for p-value

Now let's go in the opposite direction using a set of contingency tables for various LOD2. We will analyze the resulting tables using the Binomial test with our modification. The analysis will allow us to determine which of the balances is more sensitive. The modification, as we will show below, is very significant, it consists in using a probability (P) close to the real one, that is, to the one that we calculated using formula (3a). But now we can estimate this probability only based on the Contingency tables. We will calculate this probability according to the formula (5): (5) P=(a+c)/N, if a+c≠0, where N=a+b+c+d. If a + c = 0, then P, has a zero value, but this has no physical meaning, so we assign a small but non-zero value to P=0.1. Let's prove that the sensitivity of balances 1 and 2 is not the same using the Binomial test.

Binomial test (Felsenstein J., 2010; Abdi H., 2007). In statistics, the Binomial test is an exact test of the statistical significance of deviations from a theoretically expected distribution of observations into two categories. One common use of the binomial test is in the case where the null hypothesis is that two categories are equally likely to occur. A binomial test can be used, where b (Table 1) is compared to a binomial distribution (Figure 2.): pdf(b,n,P)= $C^n_b P^b (1-P)^{n-b}$ with size parameter n = b + c, integer variable b from 0 to n, and P = 0.5. Effectively, the exact binomial test evaluates the imbalance in the discordant *b* and *c*. In this case, the Null Hypothesis (**H_o**) is $P_b=P_c= 0.5$. The goal is to calculate the p-value (or asymptotic significance) using a 2x2 contingency table and Binomial distribution. A p-value ≤ 0.05 indicates a statistically significant difference, and strong evidence against the null hypothesis, so the null hypothesis should be rejected. The traditional advice has been to use the exact Binomial test when b + c < 25.

The two-sided two-tailed p-value is calculated by the formula 6 in Excel notation:

(6) p-value(b, n, P)= IF(n=0,1,IF(b=0,BINOMDIST(b,n,P,1),MIN(BINOMDIST(b,n,P,1),1-BINOMDIST(b-1,n,P,1))))

In the present work, we use the right-tailed p-value (formula 7), because we assume that the second tester is more sensitive, and b≥c in the 2x2 contingency table:

(7) p-value (b, n, P)= 1-BINOMDIST(b-1,n,P,1)

We set right-tailed p-value =1 if n=0, and right-tailed p-value = 1 if b=0 to escape error marks.

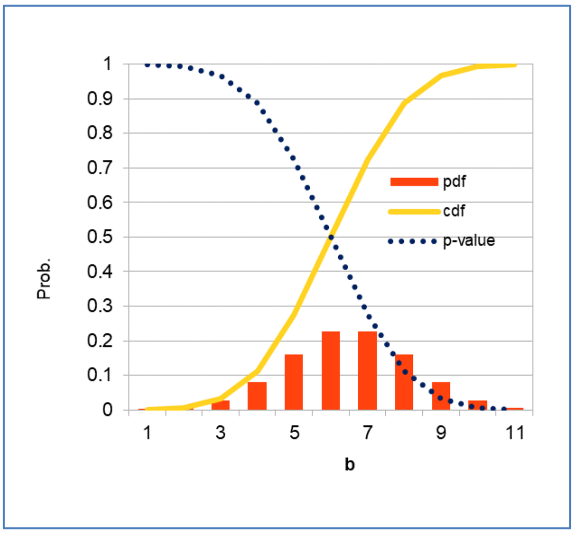

Fig-2: Binomial distribution, pdf, cdf, and right-tailed p-value. pdf=BINOMDIST(b,n,p,0) where n=11, p=0.5.

Modification of the Exact Binomial test

For an adequate assessment of testers, it is necessary to refuse to fix the probability (P=0.5). As we saw in the three types of sand example, this probability can vary. In this case, we consider the Null Hypothesis (**H_o**) is written as **Pb=Pc=P** against the one-sided alternative hypothesis **(H1) Pb>Pc**. The value of **P** in each case is different; it depends on the tested population and can be estimated from the data of the contingency table, by formula (5). Recall that the probability P is the probability that the control tester (1) will respond to testing a randomly selected object from a given population.

The dependence of the one-sided right-tail p-value from b at fixed b+c value, built for the light population of sand N (0.5,1), and for the population of sand N(2, 1) is presented in Figure 3. The corresponding 2x2 contingency tables are presented in Table 3.

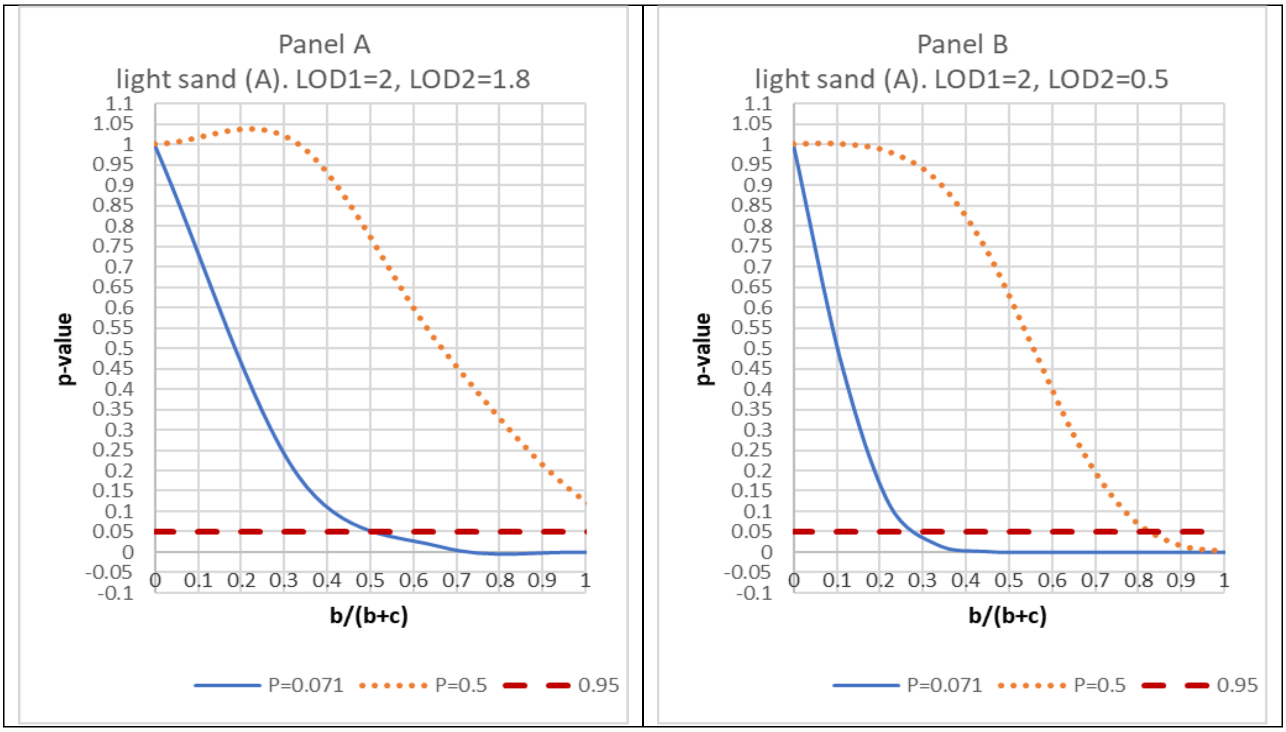

Fig-3 A, B: The right-tailed p-value. Blue line: b+c=9, P=(a+c)/N=0.071; Orange dotted line: b+c=14, P=0.5. The red dushed line is the significance level of 95%.

As can be seen from Figure 3A, if we use the probability (P=(a+c)/N) corresponding to a given population (A) rather than a fixed value (P=0.5), then the advantage of the more sensitive instrument is statistically confirmed even with a relatively small difference in sensitivities (2/1.8).

The result of calculating one-sided right-tailed asymptotic significance for five LODs of the 2nd tester (2, 1.8, 1.5, 1.0, and 0.5) for three populations of sand light (A), medium (B), heavy(C), and 5x3=15 contingency tables are presented in Table 4. p-values calculated by the Yefimov method are the bold numbers in rows 2, 5, 8, 11, and 14. The numbers marked by (*) are the p-values calculated by the Binomial test. The elements of contingency tablets b and n=b+c for A, B, and C are presented on the left side of Table 4.

For light sand (A), in all cases, when the given sensitivity of balance -2 exceeds the sensitivity of the balance -1 (LOD2<LOD1), the p-value calculated by the Yefimov method is less than the significance level, which indicates against the Null hypothesis and favor to Alternative hypothesis N1. In other words, the analysis confirms the different sensitivity of the scales. Moreover, it can be argued that balance -2 is more sensitive since in these cases b>c.

When testing balances using medium sand (B), the advantage of balance-2 is revealed only when the sensitivity ratio LOD1/LOD2=2/1 or more (Table 4).

The leveling effect occurs if we use the heavy sand (C) to test the balances. The test does not reveal the statistical difference between the scales.

Table-4: The bold numbers are the p-values by the Yefimov method.

LOD1=2 LOD2=2	**A**	**B**	**C**	**A**	**light**	**B**	**medium**	**C**	**heavy**	**1**
b=	1	7	1	**0.14**	*0.75**	**0.65**	*0.60**	**1.00**	*0.75**	**2**
n=	2	14	2	P=0.07	P=0.5	P=0.52	P=0.5	P=0.97	P=0.5	3
LOD1=2 LOD2=1.8				A	light	B	medium	C	heavy	4
b=	2	8	1	**0.01**	*0.50**	**0.45**	*0.40**	**0.97**	*0.50**	**5**
n=	3	14	1	P=0.07	P=0.5	P=0.52	P=0.5	P=0.97	P=0.5	6
LOD1=2 LOD2=1.5				A	light	B	medium	C	heavy	7
b=	3	10	1	**0.00**	*0.31**	**0.09**	*0.09**	**0.97**	*0.50**	**8**
n=	4	14	1	P=0.07	P=0.5	P=0.52	P=0.5	P=0.97	P=0.5	9
LOD1=2 LOD2=1				A	light	B	medium	C	heavy	10

b=	6	12	1	**0.00**	*0.06**	**0.01**	*0.01**	**0.97**	*0.50**	**11**
n=	7	14	1	P=0.13	P=0.5	P=0.52	P=0.5	P=0.97	P=0.5	12
LOD1=2 LOD2=0.5				A	light	B	medium	C	heavy	13
b=	9	13	1	**0.00**	*0.00**	**0.00**	*0.00**	**0.97**	*0.50**	**14**
n=	9	14	1	P=0.07	P=0.5	P=0.52	P=0.5	P=0.97	P=0.5	15

The 2x2 contingency tables we have built have 3 independent parameters: b, c, and P. The value of d can be calculated by the formula: $d= a\cdot(1 - p) / p – (b + c) + c / p$.

Generalization of the obtained results to any testers with a binary scale

We have no reason to believe that the Yefimov method for determining a more sensitive tester is suitable only for balances with a binary scale. On the contrary, our reasoning was quite general, and, consequently, this method is suitable for detecting more sensitive devices, regardless of the principle of detection. We will show this by comparing two luminometers when determining the biological contamination of solutions.

The equivalence of two test methods that detect microbiological contamination was evaluated by comparing the rate of positive and negative results obtained from identical samples. The methods were: the Rapid Adenylate Kinase-amplified ATP bioluminescence method (AK- method) which is used in the Celsis® instrument (The Celsis Advance II™ system, 2000), and ATP bioluminescent method without AK (ATP -method) which is used in Pallchek™ instrument (Pallchek™ Rapid Microbiology System, 2021).

For sample preparation, we used Bioballs (Biomerieux) microorganism standards (Bioball (Biomerieux), 2022). In the first case, Staphylococcus cells at concentrations of 10, 1, and 0.1 CFU in the Fluid Thioglycollate Medium were incubated for 4 days at 32.5°C. After incubation, the suspensions were tested by two competitive methods, the AK- method, and the ATP- method.

In the second case, aqueous cell suspensions of slowly growing bacteria - Propionibacterium acnes, concentrations of 1, 0.5, and 0.1 CFU were tested immediately after the preparation of the suspensions. The test results are presented in Table 5.

Table-5: Two 2x2 contingency tables illustrate the leveling and differentiating effect of testing objects.

4 days		0 days	
8	**0**	**0**	**10**
0	**7**	**0**	**6**
P= 0.5		P=0.1	
p-value=1.00		p-value=0.00	

The left 2x2 contingency table is an example of a leveling effect. Our previous work showed that the AK- method is about 1000 times more sensitive than the ATP- method, but the test result does not reveal it. Why? The reason is that the number of microorganisms in 4 days of incubation is increased by about 1,000,000 times (Yefimov S. 2022). The right 2x2 contingency table is an example of a differentiating effect. A small initial concentration of bacteria is still constant. Statistical analysis by the Yefimov method proved the advantage of the AK- method in comparison with the ATP- method.

CONCLUSION

To analyze 2x2 contingency tables to identify statistically significant superiority, namely, greater sensitivity, of the binary tester we are interested in over the reference one, we have developed a new method, the Yefimov Method. The Yefimov method includes 2 main components.

The first component refers to the properties of the set of tested objects such as concentration, optical density, weight, noise level, electrical potential, and others. The average property value (M) for a random sample from the general population of tested objects should be less than the sensitivity limit (LOD1) of the reference tester and approximately equal to the expected sensitivity limit of the tester of interest (LOD2), $LOD2 \cong M < LOD1$. If $M > LOD1$, the tested object is diluted, separated, or noise, light filters, and electrical resistances are applied to reduce M.

The second component is to estimate the probability (P) of triggering the reference tester. This probability is approximately estimated by the formula $P=(a+c)/N$ based on the data of the contingency table constructed for the M=M sample.

Statistically significant superiority of the tester we are interested in is confirmed if the p-value calculated by the Modified Binomial test is $p\text{-value}<0.05$.

REFERENCES

1. USP 1223 Validation of Alternative Microbiological. In United States Pharmacopeia, 31st Rev.; U.S. Pharmacopeial Convention. 2008; Rockville, MD. https://www.drugfuture.com/pharmacopoeia/usp32/pub/data/v32270/usp32nf27s0_c1223.html

2. Evaluation, Validation, and Implementation of Alternative and Rapid Microbiological Methods. Technical Report No 33. 2013. Parenteral Drag Association, INC. https://store.pda.org/TableOfContents/TR3313_TOC.pdf
3. Felsenstein J. Binomials, contingency tables, chi-square tests. 2010. https://evolution.gs.washington.edu/gs560/2010/lecture4.pdf
4. 2x2 contingency Table 1999. Calculator http://vassarstats.net/tab2x2.html
5. Abdi, H. (2007). The Binomial Distribution The Binomial and Sign Tests In: Neil Salkind (Ed.). Encyclopedia of Measurement and Statistics. Thousand Oaks (CA): Sage. https://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.77.3055&rep=rep1&type=pdf
6. The Celsis Advance II™ system. (2000). https://www.criver.com/products-services/qc-microbial-solutions/microbial-detection/microbial-detection-instruments/celsis-advance-ii-system?region=3601
7. Pallchek™ Rapid Microbiology System. 2021. https://www.labmakelaar.eu/wp-content/uploads/2021/03/5656.pdf
8. Bioball (Biomerieux). (2022). https://www.biomerieux-industry.com/en-us/products/bioball-standardized-strains
9. Yefimov, S. V. (2022). Sterility Test: Comparison of the sensitivity of rapid microbiological methods based on ATP bioluminescence with compendial plate count method. Establishment of a scale of sensitivity - Transition from qualitative microbiological method to quantitative method. International Journal of Chemical and Pharmaceutical Sciences, *13*(1). http://www.ijcps.com/files/vol13issue1/2.pdf