Power and Sample Size Investigation for Correlated Binary Data

Dar Shong Hwang[1], James Lee[2], Chyi-Hung Hsu[3]
[1]B.R.S.I., Florham Park, NJ 07932
[2] Daiichi Sankyo Pharma Development, Edison, NJ 08837
[3] Janssen Research & Development, Raritan, NJ 08869

Abstract

In repeated measurements or RBD with correlated binary response data, power and sample size determination is of great practical interest. The main stumbling block to the solution: the test distribution involves hard to interpret multinomial parameters, not the desired marginal binomial parameters P1, P2, …, Pk. Lennox and Sherman (2009) document the four decade-long endeavours to the solution of the simplest case- the matched pairs design or repeated measures with k=2 time points. Hwang and Lee(2009) reported that the 2-dimentional multinomial parameters Pij involved in the Exact McNemar test may be expressed as a function of the 1-dimentional or binomial parameters P1, P2 (hence odds ratio) and correlation. This multinomial to Binomial transformation enables researchers to specify P1 and P2 of interest and compute the power and sample size for each specified value of correlation. When correlation is 0, the sample size obtained is reduced to that of Fisher's exact test. Section I presents the power and sample size tables based on this transformation. Section II extends the methodology to the k=3 treatments or time points case which is analyzed by Modified Cochran's Q test (Hwang, Lee and Hsu (2004)). Here the 3 dimensional multinomial parameters Pijk of the test distribution are transformed first to 2-dim multinomial parameters and correlations, and then to 1-dim Binomial parameters P1, P2, P3 and correlations.

Keywords: Multinomial to binomial transform, repeated measurements, RBD, correlated binary data, power and sample size, exact McNemar's test, modified Cochran's Q test.

Section I. Matched Pair Design or Repeated Measure Design with k=2 Times

Let X_1 and X_2 be two correlated binomials with $X_i \sim$ Bin (n, P_i), i= 1, 2.
Let P_{ij}, i =0, 1; j=0, 1 be the corresponding 2-dimensional multinomial parameters with

$$P_{00} + P_{01} + P_{10} + P_{11} = 1$$

1. Brief Outline of Methodology

Hwang and Lee (2009) proposed a 3- step process:

(i) Re-expressing H_0 in terms of Multinomial parameters $P_{01} / P_{10} = 1$

(ii) Show the existence and actual derivation of a distribution which depends on the new multinomial parameters. The distribution turns out to be a binomial with parameter $\Pi = P_{01}/ (P_{01} + P_{10})$ and sample size equal to the number of discordant pairs in the sample.

(iii) Transformation of the new multinomial parameters back to marginal binomial parameters.

Although the null distribution in (ii) does not involve any nuisance parameter and, in particular, is independent of correlation, the parameter under the non-null hypothesis, namely $P \equiv P_{01} / (P_{10} + P_{01})$ can be expressed as or transformed to a close-end function of marginal binomial parameters P_1, P_2 (and hence odds ratio ϕ) and correlation ρ.

$$P = (\phi - \rho \sqrt{\phi}) / [1 + \phi - 2 \rho \sqrt{\phi}] \quad \ldots\ldots\ldots\ldots\ldots\ldots (1)$$

Where ρ = correlation between X_1 and X_2,
ϕ = odds ratio of X_2 over X_1 = ($P_2 / (1-P_2)$) / ($P_1 / (1-P_1)$),
$0 < \phi$, $-1 < \rho < 1$, and $\rho\sqrt{\phi} < 1$.

The transformation enables one to explicitly specify P_1, P_2 of interest under non–null hypothesis, instead of specifying P_{01}, P_{10} as many authors did.

The relationship (1) also means different combinations of Odds ratio and correlation can lead to the same value of P (Hwang and Lee (2009). Table 1, p. 4319).

With the transformation (1), the conditional power **for any fixed value of correlation ρ** can be computed for any pair of P_1, P_2 (and hence odds ratio ϕ) under non-null hypothesis.

The unconditional power is obtained by weighing the conditional power by a probability distribution which turns out to be binomial with the parameter $P'' = P_{01} + P_{10}$. Hwang and Lee (2009) proved that the parameter P'' can also be transformed to a function of P_1, P_2 and correlation ρ. P'' was termed "nuisance parameter" (Lennox and Sherman (2009)).

2. Presentation of Sample Size Tables for Use with the Exact McNemar Test for Matched Pairs Design or Repeated Measures Design

The present paper computes and presents, **for each pre-assigned value of correlation** ρ, the sample size required to obtain a specified power when type one error $\alpha = 0.01$ or 0.05 (One sided). The sample size corresponding to 3 different specified power is provided: power = 0.90, 0.80, and 0.50. The values of P_1 and P_2 under the alternative hypothesis are: $P_1 = 0.1, 0.2, \ldots, 0.9, 0.95$; $P_2 = 0.05, 0.1, 0.2, 0.3, \ldots, 0.8, 0.9$. Note the sample size here is the number of pairs in a matched pair design, number of blocks in a randomized block design, and number of subjects or patients in a repeated measurement design.

The above sample size table is generated separately for each selected correlation value, $\rho = 0.1, 0.2, 0.3, \ldots, 0.9$. Also $\rho = -0.1, -0.2, -0.3, \ldots, -0.9$.

We purposely present the resultant tables in the same format as Haseman's (1978) sample size tables for Fisher's exact test. Of course his tables are for 2 independent binomial samples. In essence, for each of his table, we generate several tables, one for each pre-specified value of correlation ρ. This may serve as an aid for practitioners to compare and select the most economical design in his /her situation.

As in the estimation of sample size required for 2 independent samples to be analyzed by Fisher's exact test, some preliminary information about a proposed study must be

obtained .Besides the likely size of P_1 and P_2, one should also have some idea about the magnitude of correlation, either from pertinent literatures or pilot study.

(1) Tables of Power and Sample Size Computations

Table 1: Number of Patients (Or Blocks) Required to Achieve a Specified Power
Where $\alpha = 0.05$ (One Tailed) and Correlation $\rho = 0.2$

Upper Figure : Power = 0.9
Middle Figure : Power = 0.8
Lower Figure : Power = 0.5

	P_1									
P_2	0.95	0.9	0.8	0.7	0.6	0.5	0.4	0.3	0.2	0.1
0.9	407									
	303									
	153									
0.8	74	188								
	58	141								
	35	73								
0.7	36	61	275							
	29	48	203							
	20	28	99							
0.6	23	33	79	330						
	19	27	60	245						
	14	17	33	117						
0.5	16	22	39	90	358					
	14	18	31	68	265					
	11	12	19	35	126					
0.4		15	25	43	94	358				
		13	20	33	70	265				
		10	13	20	36	126				
0.3		12	17	26	43	90	330			
		10	14	21	33	68	245			
		8	10	13	20	35	117			
0.2			12	17	25	39	79	275		
			10	14	20	31	60	203		
			8	10	13	19	33	99		
0.1				12	15	22	33	61	188	
				10	13	18	27	48	141	
				8	10	12	17	28	73	
0.05						16	23	36	74	407
						14	19	29	58	303
						11	14	20	35	153

Note: The sample sizes for some of the lower triangular cases are blank because $\rho\sqrt{\phi} > 1$. For those cases, the sample sizes above the blank cases could be used to achieve greater than the specified power since ϕ is greater.

Table 2: Number of Patients (Or Blocks) Required to Achieve a Specified Power
Where $\alpha = 0.05$ (One Tailed) and Correlation $\rho = 0.4$

Upper Figure : Power = 0.9
Middle Figure : Power = 0.8
Lower Figure : Power = 0.5

	P_1									
P_2	0.95	0.9	0.8	0.7	0.6	0.5	0.4	0.3	0.2	0.1
0.9	310									
	234									
	128									
0.8	57	144								
	47	109								
	32	62								
0.7		48	208							
		38	155							
		25	78							
0.6		25	60	252						
		22	47	186						
		16	28	92						
0.5			31	69	273					
			25	52	202					
			16	30	98					
0.4			19	33	71	273				
			16	27	54	202				
			12	17	31	98				
0.3				19	33	69	252			
				16	27	52	186			
				12	17	30	92			
0.2					19	31	60	208		
					16	25	47	155		
					12	16	28	78		
0.1							25	48	144	
							22	38	109	
							16	25	62	
0.05									57	310
									47	234
									32	128

Table 3: Number of Patients (Or Blocks) Required to Achieve a Specified Power
Where $\alpha = 0.05$ (One Tailed) and Correlation $\rho = 0.6$

Upper Figure : Power = 0.9
Middle Figure : Power = 0.8
Lower Figure : Power = 0.5

	P_1									
P_2	0.95	0.9	0.8	0.7	0.6	0.5	0.4	0.3	0.2	0.1
0.9	210									
	167									
	102									
0.8		97								
		78								
		50								
0.7			140							
			106							
			61							
0.6			40	168						
			33	127						
			24	68						
0.5				46	183					
				37	138					
				24	71					
0.4					48	183				
					39	138				
					25	71				
0.3						46	168			
						37	127			
						24	68			
0.2							40	140		
							33	106		
							24	61		
0.1									97	
									78	
									50	
0.05										210
										167
										102

Table 4: Number of Patients (Or Blocks) Required to Achieve a Specified Power
Where $\alpha = 0.05$ (One Tailed) and Correlation $\rho = -0.2$

Upper Figure : Power = 0.9
Middle Figure : Power = 0.8
Lower Figure : Power = 0.5

	P_1									
P_2	0.95	0.9	0.8	0.7	0.6	0.5	0.4	0.3	0.2	0.1
0.9	597									
	442									
	211									
0.8	104	277								
	78	204								
	43	99								
0.7	49	87	402							
	38	66	297							
	23	35	141							
0.6	31	46	115	486						
	25	35	86	357						
	15	20	43	168						
0.5	22	29	56	133	528					
	18	23	43	98	387					
	11	14	23	48	181					
0.4	17	21	34	61	138	528				
	13	17	26	46	102	387				
	9	10	15	24	50	181				
0.3	13	16	23	36	61	133	486			
	11	13	18	27	46	98	357			
	7	8	11	16	24	48	168			
0.2	10	12	17	23	34	56	115	402		
	9	10	14	18	26	43	86	297		
	6	7	8	11	15	23	43	141		
0.1	8	9	12	16	21	29	46	87	277	
	7	8	10	13	17	23	35	66	204	
	6	6	7	8	10	14	20	35	99	
0.05	7	8	10	13	17	22	31	49	104	597
	6	7	9	11	13	18	25	38	78	442
	5	6	6	7	9	11	15	23	43	211

Table 5: Number of Patients (Or Blocks) Required to Achieve a Specified Power Where $\alpha = 0.05$ (One Tailed) and Correlation $\rho = -0.4$

Upper Figure : Power = 0.9
Middle Figure : Power = 0.8
Lower Figure : Power = 0.5

	P_1									
P_2	0.95	0.9	0.8	0.7	0.6	0.5	0.4	0.3	0.2	0.1
0.9	687									
	508									
	242									
0.8	118	318								
	89	236								
	46	113								
0.7	55	100	466							
	42	74	342							
	24	38	162							
0.6	34	52	134	564						
	27	40	99	413						
	16	22	48	192						
0.5	24	33	65	152	612					
	19	25	48	113	448					
	12	15	25	54	207					
0.4	18	23	39	71	159	612				
	15	18	30	53	118	448				
	9	11	16	27	56	207				
0.3	14	17	26	41	71	152	564			
	11	14	20	31	53	113	413			
	7	9	12	17	27	54	192			
0.2	11	14	18	26	39	65	134	466		
	9	11	15	20	30	48	99	342		
	6	7	9	12	16	25	48	162		
0.1	9	10	14	17	23	33	52	100	318	
	8	9	11	14	18	25	40	74	236	
	6	6	7	9	11	15	22	38	113	
0.05	8	9	11	14	18	24	34	55	118	687
	6	8	9	11	15	19	27	42	89	508
	5	6	6	7	9	12	16	24	46	242

Table 6: Number of Patients (Or Blocks) Required to Achieve a Specified Power
Where $\alpha = 0.05$ (One Tailed) and Correlation $\rho = -0.6$

Upper Figure : Power = 0.9
Middle Figure : Power = 0.8
Lower Figure : Power = 0.5

	P_1									
P_2	0.95	0.9	0.8	0.7	0.6	0.5	0.4	0.3	0.2	0.1
0.9	777									
	573									
	271									
0.8	132	360								
	99	266								
	50	127								
0.7	62	112	530							
	47	84	388							
	26	42	181							
0.6	37	58	151	641						
	29	44	112	469						
	17	23	54	216						
0.5	26	37	73	173	696					
	20	28	54	128	508					
	12	16	27	62	234					
0.4	19	25	43	80	179	696				
	16	20	33	60	133	508				
	10	12	17	30	63	234				
0.3	15	19	29	45	80	173	641			
	12	15	22	34	60	128	469			
	8	9	13	18	30	62	216			
0.2	12	14	20	29	43	73	151	530		
	10	12	15	22	33	54	112	388		
	6	8	9	13	17	27	54	181		
0.1	9	11	14	19	25	37	58	112	360	
	8	9	12	15	20	28	44	84	266	
	6	6	8	9	12	16	23	42	127	
0.05	8	9	12	15	19	26	37	62	132	777
	6	8	10	12	16	20	29	47	99	573
	5	6	6	8	10	12	17	26	50	271

Due to the limit of space available for this paper, the tables for other correlations and $\alpha = 0.01$ cannot be presented here. These additional tables can be provided upon request.

Section II. Power and Sample Size Investigation for a RBD Design or Repeated Measures Design with k=3 Time Points

When there are k = 3 treatments or time points, power and sample size determination depends on the study objectives. There are 2 major types of objectives in practice:

1. Pairwise treatment comparison or pairwise time point comparison is the only interest.

1.1 Example of Focusing on Pairwise Comparisons Only

In a new drug clinical trial, 3 treatments are involved: Active, Placebo, and the Reference drug from the competitor. The pairwise treatment comparisons are the only interest:

(1) Active vs Placebo: FDA requirement for marketing approval of a new drug.
(2) Reference vs Placebo: verify the study is valid.
(3) Active vs Reference: Proving Active is better than competitor's drug should contribute to market dominance.

There is no interest in whether the 3 drugs are the same.

1.2 Power and Sample Size Determination for the Pairwise Comparison Case

In this case, one can estimate power and sample size for each pair of treatments using the methodology of Section I. Suppose the sample sizes are N_a, N_b and N_c, respectively, then one can simply use the largest sample size of the three as the required sample size.

2. Overall treatment comparison simultaneously or time point comparison simultaneously is of interest.

For this objective, Cochran's Q test (1950) was often employed to simultaneously compare the marginal binomial parameters P_1, P_2, …, P_k among k correlated Bernoulli r.v.'s.

2.1 Cochran's Q test - A Randomization Test

Since Cochran's explanation of his nonparametric, randomization test under the null hypothesis only is rather brief, Plackett (1974), Patil (1975) tried to fill the gap of his theory and independently derived the same Exact Cochran's Q test distribution under the null hypothesis, again using the randomization test approach. Patil then attempted to compute power of the test but as we'll see later, apparently ran into the same problem of erroneously relating multinomial parameters to the desired marginal binomial parameters under investigation.

2.2 Modified Q-Test - A Non-Randomization Test

Hwang, Lee and Hsu (2004) abandoned the randomization test approach and proposed a Modified Cochran's Q test based on the population model (Lehmann, 1998).

The approach yields a distribution which is exactly the same as Plackett and Patil under the null hypothesis but has a general distribution under the non-null hypothesis which involves multinomial parameters. As in Section I, power and sample size will be investigated thru this non-null distribution by transforming its multinomial parameters into the desired marginal binomial parameters P_1, P_2, P_3 and correlations which can be used to specify treatment difference for power computation purpose.

(a) Study Design and Related Probability Distributions

In a RBD design with 3 treatments or repeated measures design with k= 3 time points:

	Time 1	Time 2	Time 3
Subject #1	X_1	X_2	X_3
	(P_1)	(P_2)	(P_3)

Where X_1, X_2, and X_3 are three correlated Bernoulli random variables for the same Subject with $X_i \sim \text{Bin}(n=1, P_i)$, i = 1, 2, 3.

Suppose there are m patients in the sample.
Let S_{ijk} = Number of patients whose outcome is (i, j, k). Then the probability distribution of $\{S_{ijk}, i=0,1; j=0,1; k=0,1\}$ is Multinomial with

$$(S_{000}, S_{001}, S_{010}, S_{100}, S_{011}, S_{101}, S_{110}, S_{111}) \sim M(n=m; P=(P_{000}, P_{001}, P_{010}, P_{100}, P_{011}, P_{101}, P_{110}, P_{111}))$$

Where $S_{000}+ S_{001}+ S_{010}+ S_{100}+ S_{011}+ S_{101}+ S_{110}+ S_{111} = m$,
And $P_{000}+ P_{001}+ P_{010}+ P_{100}+ P_{011}+ P_{101}+ P_{110}+ P_{111} = 1$.

(b) The 3-Step Process
(i) H_0: $P_1 = P_2 = P_3$ is true if

(1) $H_0^{(1)}$: $\theta_1 \equiv P_{100} / P_{001} = 1$, and $\theta_2 \equiv P_{010} / P_{001} = 1$ ……(2.1)
(2) $H_0^{(2)}$: $\lambda_1 \equiv P_{110} / P_{011} = 1$, and $\lambda_2 \equiv P_{101} / P_{011} = 1$ ……..(2.2)

Are both true.

(2.1) and (2.2) means:

(1) $H_0^{(1)}$: $P_{100} = P_{001} = P_{010}$ ……(2.1')
(2) $H_0^{(2)}$: $P_{110} = P_{011} = P_{101}$ ……..(2.2')

(ii) Existence and Actual Derivation of a (conditional) Distribution which Depends on the New Multinomial Parameters Only

The distribution is found to be a conditional distribution:

$P\{(S_{010},S_{100},S_{110},S_{101}) = (s_{010},s_{100},s_{110},s_{101}) \mid (S_{001}+S_{010}+S_{100})=k_1, (S_{011}+S_{110}+S_{101})= k_2, S_{000}= s_{000}\}$

$$= \binom{k_1}{S_{010}, S_{100}, k_1 - S_{010} - S_{100}} \left(\theta_2/(1+\theta_1+\theta_2)\right)^{S_{010}} \left(\theta_1/(1+\theta_1+\theta_2)\right)^{S_{010}}$$

$$\bullet \binom{k_2}{S_{110}, S_{101}, k_2 - S_{110} - S_{101}} \left(\lambda_1/(1+\lambda_1+\lambda_2)\right)^{S_{110}} \left(\lambda_2/(1+\lambda_1+\lambda_2)\right)^{S_{101}}$$

$$\bullet \left(1/(1+\theta_1+\theta_2)\right)^{k_1 - S_{010} - S_{100}} \left(1/(1+\lambda_1+\lambda_2)\right)^{k_2 - S_{110} - S_{101}} \qquad (1)$$

Which depends on the parameters useful for testing H_0 (i.e. (2.1) and (2.2)) only. All nuisance parameters are eliminated or conditioned out.

Under the null hypothesis, the conditional distribution (1) becomes

$P\{(S_{010},S_{100},S_{110},S_{101}) = (s_{010},s_{100},s_{110},s_{101}) \mid (S_{001}+S_{010}+S_{100})=k_1, (S_{011}+S_{110}+S_{101})= k_2, S_{000}= s_{000}\}$

$$= \binom{k_1}{S_{010}, S_{100}, k_1 - S_{010} - S_{100}} (1/3)^{k_1}$$

$$\bullet \binom{k_2}{S_{110}, S_{101}, k_2 - S_{110} - S_{101}} (1/3)^{k_2} \qquad (2)$$

= product of 2 independent trinomial distributions
= $\pi 1 \ \pi 2$
Which is the same as Plackett (1974, p. 102) and Patil (1975, p. 187, (3.1)).

The distribution (1) is multivariate in nature, we shall use Cochran's Q (Cochran(1950)) as the testing statistic for (1) :

$$Q = k(k-1)\sum_{j=1}^{k} \left(T_j - \overline{T}\right)^2 \Big/ \left(k\sum_i U_i - \sum_i U_i^2\right) \qquad (3)$$

where k = number of treatments or time points.
Tj = Total number of successes for the j-th treatment or time point across all subjects
Or blocks, j=1, 2, …, k

$\overline{T} = \Sigma \ Tj \ / k$
Ui = Number of success among the 3 time points X_1, X_2, X_3 for the i-th subject.

In our present case, k = 3 time points. Ui is to be obtained from each patient, i = 1, 2, …, m. Tj, j=1, 2, 3 is to be obtained from the multinomial observation **S** ≡ (S_{000}, S_{001},S_{010},S_{100},S_{110},… ,S_{111}) as outlined below:

Formula (3) transforms multivariate observations **S** to a univariate statistic Q. For any value of Q, its probability is the probability of the corresponding multinomial observation **S.**

In this way the distribution of Q is obtained, both under H_0 and under H_1. Power and sample size will be based on the null and non-null distribution of Q.

(iii)Transformation of New Multinomial Parameters Back to Binomial Parameters

Although the null distribution (2) is free of any nuisance parameter and in particular, is independent of correlations, the non-null distributions (1) does depend on the hard to interpret functions of 3-dimensional multinomial parameters **θ1, θ2,** λ1 and λ2. As in Section I, it is found by tedious algebra that each of them can be transformed to P_1, P_2, P_3 and correlations. We first transform a 3-dimensional multinomial parameter to a function of 2-dimensional multinomial parameters and correlations. After this, one is in the situation of Section I, and can further transform each of the 2-dimensional multinomial parameters to 1-dimensional or binomial parameters P_1, P_2, P_3 and correlations. As an example:

$\theta_1 \equiv P_{100} / P_{001}$ can be transformed to a function of P_1, P_2, P_3 and correlations $\rho(X_1,X_2)$, $\rho(X_1,X_3)$, $\rho(X_2,X_3)$ and conditional correlation $\rho = \rho(X_2,X_3|\ X_1=0)$, $\rho = \rho(X_1,X_3|\ X_2=0)$. Specifically:

The numerator of θ_1 is

$P_{100}/(Q_2) = P_{100}/(1 - P_2)$

$= (P_{10\bullet}/Q_2)\,(P_{00\bullet} / Q_2) - \rho\,[(P_{10\bullet} / Q_2)\,(P_{00\bullet} / Q_2)\,(P_{\bullet 01}/ Q_2)\,(P_{\bullet 00} / Q_2)]^{0.5}$

Where $\rho = \rho(X_1,X_3|\ X_2=0)$, $Q_2 = 1- P_2$

And where the 2-dim multinomial parameters $P_{\bullet 01}$, $P_{\bullet 00}$, $P_{10\bullet}$, $P_{00\bullet}$ can be expressed as functions of P_1, P_2, P_3 and correlations as follows:

Lemma 1.3' $P_{\bullet 01} = (1-P_2)\,P_3 - \rho(X_2,X_3)\,[\,P_3(\,1 - P_3)\,P_2\,(\,1-P_2)\,]^{0.5}$,

Lemma 1.4' $P_{\bullet 00} = (1-P_2)\,(1-P_3) - \rho(X_2,X_3)\,[\,P_3(\,1 - P_3)\,P_2\,(\,1-P_2)\,]^{0.5}$,

Lemma 1.2* $P_{10\bullet} = P_1\,(1-P_2) - \rho(X_1,X_2)\,[\,P_1(\,1 - P_1)\,P_2\,(\,1-P_2)\,]^{0.5}$,

Lemma 1.4* $P_{00\bullet} = (1-P_1)\,(1-P_2) - \rho(X_1,X_2)\,[\,P_1(\,1 - P_1)\,P_2\,(\,1-P_2)\,]^{0.5}$,

Similarly, the Denominator of θ_1 is

$P_{001}/(Q_1) = (P_{0\bullet 1}/Q_1)\,(P_{00\bullet} / Q_1) - \rho\,[(P_{01\bullet} / Q_1)\,(P_{00\bullet} / Q_1)\,(P_{0\bullet 1}/ Q_1)\,(P_{0\bullet 0} / Q_1)]^{0.5}$

Where $\rho = \rho(X_2,X_3|\ X_1=0)$, $Q_1 = 1- P_1$

And where the 2-dimensional multinomial parameters $P_{0\bullet 0}$, $P_{0\bullet 1}$, $P_{00\bullet}$, $P_{01\bullet}$ can be further expressed as functions of 1-dim or binomial parameters P_1, P_2, P_3 and correlations as follows:

Lemma 1.4+ $P_{0\bullet 0} = (1-P_1)\,(1-P_3) - \rho(X_1,X_3)\,[\,P_1(\,1 - P_1)\,P_3\,(\,1-P_3)\,]^{0.5}$,

Lemma 1.3+ $P_{0\bullet 1} = (1-P_1)\,P_3 - \rho(X_1,X_3)\,[\,P_1(1-P_1)\,P_3\,(1-P_3)\,]^{0.5}$,

Lemma 1.4* $P_{00\bullet} = (1-P_1)\,(1-P_2) - \rho(X_1,X_2)\,[\,P_1(1-P_1)\,P_2\,(1-P_2)\,]^{0.5}$,

Lemma 1.3* $P_{01\bullet} = (1-P_1)\,P_2 - \rho(X_1,X_2)\,[\,P_1(1-P_1)\,P_2\,(1-P_2)\,]^{0.5}$.

(iii.1) Comment on Patil's "Power Evaluation" under Alternative Hypothesis

Under the non-null hypothesis:

$\pi 1 \sim M(n=K_1, P = P_1', P_2', P_3')$,

Where $P_1' = \theta_1/(1+\theta_1+\theta_2)$, $P_2' = \theta_2/(1+\theta_1+\theta_2)$, and $P_3' = 1/(1+\theta_1+\theta_2)$

Note: $P_1' + P_2' + P_3' = 1$,

And $\pi 2 \sim M(n=K_2, P = P_1'', P_2'', P_3'')$,

Where $P_1'' = \lambda_1/(1+\lambda_1+\lambda_2)$, $P_2'' = \lambda_2/(1+\lambda_1+\lambda_2)$, and $P_3'' = 1/(1+\lambda_1+\lambda_2)$

Note: $P_1'' + P_2'' + P_3'' = 1$

Patil (1975, p. 188-189, table 2a) attempts to compute the probability distribution of Q under non-null hypothesis in his randomization test approach. It is apparent that his parameters of p1, p2, p3 with p1 + p2 + p3 = 1 (p. 188) corresponds to our P_1', P_2', P_3' here, and his parameters q1, q2, q3 with q1 + q2 + q3 = 1 corresponds to our P_1'', P_2'', P_3'' here. Hence the parameters p1, p2, p3 and q1, q2, q3 he employed in the computation of distribution of Q under the alternative hypothesis are a mixture of binomial parameters P_1, P_2, P_3 and nuisance parameters, namely the correlations. They may not be used for power and sample size evaluation for the comparison of binomial parameters P_1, P_2, P_3.

(c) Conditional Test and Conditional Critical Region

Let α = Type 1 error (with $\alpha = 0.01$ or 0.05).
Under the interchangeability hypotheses of

$H_0^{(1)}: \theta_1 \equiv P_{100}/P_{001} = 1$, and $\theta_2 \equiv P_{010}/P_{001} = 1$ ……(2.1)

$H_0^{(2)}: \lambda_1 \equiv P_{110}/P_{011} = 1$, and $\lambda_2 \equiv P_{101}/P_{011} = 1$ ……..(2.2)

Which guarantees the homogeneity hypothesis

$H_0: P_1 = P_2 = P_3$ to be true,

We have the null distribution (2) where we replace the multinomial observation S by the univarite r.v. for Cochran's Q of (3). From now on, we shall refer to this as the null distribution.

The alternative hypothesis is $P_1 = P_1^*, P_2 = P_2^*, P_3 = P_3^*$
Where P_1^*, P_2^*, P_3^* are arbitrary but fixed value between 0 and 1, and not all equal.
The critical region is located on the right tail of the null (conditional) distribution (2). Let C be the (largest) number assumed by Cochran's Q such that the cumulative distribution of Q under $H_0^{(1)}$, $H_0^{(2)}$ is $\leq \alpha$, i.e.

Σ_{ω} $P\{(S_{010},S_{100},S_{110},S_{101}) = (s_{010},s_{100},s_{110},s_{101}) \mid (S_{001}+S_{010}+S_{100})=k_1, (S_{011}+S_{110}+S_{101})= k_2,$
$S_{000}= s_{000}, H_0^{(1)}, H_0^{(2)} \}$
$\leq \alpha$

Where $\omega = \{Q: c < Q\}$ is the critical region.

(d) Conditional Power of the Test

Suppose the alternative hypothesis is $P_1= P_1^*, P_2= P_2^*, P_3 = P_3^*$
Where P_1^*, P_2^*, P_3^* are arbitrary but fixed value between 0 and 1, and not all equal. The conditional power for this specific alternative is the probability of the critical region ω under the non-null distribution (1), where the multinomial sample **S** has been replaced by Cochran's Q, and the multinomial parameters there can be replaced by P_1, P_2, P_3 and correlations.

If we have some information or pilot study which gives us some idea about the magnitude of the correlations, then the non-null distribution will be function of P_1, P_2, P_3 only. In this way the conditional power for the specific alternative hypothesis $P_1= P_1^*$, $P_2= P_2^*$, $P_3 = P_3^*$ may be obtained. And sample size for the detection of the specific non-null hypothesis $P_1= P_1^*$, $P_2= P_2^*$, $P_3 = P_3^*$ may be obtained.

(e) Unconditional Power

The weighted sum of the conditional power weighted by the distribution of the conditioning r.v.'s is then the unconditional power of testing:

$H_0: P_1= P_2 = P_3$
vs.
$H_1: P_1= P_1^*, P_2= P_2^*, P_3 = P_3^*$

From (1), the conditioning r.v. is $(S_{000}, S_{001}+S_{010}+S_{100}, S_{011}+S_{101}+S_{110})$ whose distribution is again multinomial with

$(S_{000}, S_1 \equiv S_{001}+S_{010}+S_{100}, S2 \equiv S_{011}+S_{101}+S_{110}, S_{111})$
$\sim M (n=m, P = (P_{000}, (P_{001} + P_{010}+P_{100}), (P_{011}+P_{101} +P_{110}), P_{111})$

Where the multinomial parameters may be re-expressed as a function of P_1, P_2, P_3 and correlations as shown before. If the correlations are known or estimated, then the unconditional power for testing $H_0: P_1= P_2 = P_3$ vs. $H_1: P_1= P_1^*, P_2= P_2^*, P_3 = P_3^*$ may be computed.

References

Cochran, W. G. 1950. "The comparison of percentages in matched samples." Biometrika 37, pp. 256-66.

Frisen, M. 1980. "Consequences of the use of conditional inference in the Analysis of a correlated contingency table." Biometrika, 67, 1, pp. 23-30.

Haseman, J. K. 1978. "Exact sample size tables for use with the Fisher-Irwin test for 2x2 tables". Biometrics 34, pp. 106-109.

Hwang, John D. S. and J. Lee. 2009. “Sample size and power investigation for exact McNemar test”. Proceedings of 2009 ASA Meetings (Washington, D.C.), Biopharmaceutical Subsection, pp. 4314-4323.

Hwang, John D. S., J. Lee, and C-H. Hsu. 2004. “Exact analysis of repeated measurements of correlated binary and nominal response data- A Population Model approach.” Proceedings of 2004 ASA Meetings (Toronto), Biopharmaceutical Subsection, pp. 681-688.

Lehmann, E. L. 1998. Nonparametrics: Statistical Methods Based on Ranks.

Lennox and Sherman. 2009. “Efficient experimental design for binary matched pairs data”. Statistics in Medicine, 28: 2952-2966.

Madansky, A. 1963. “Tests of Homogeneity for correlated samples”. Journal of the American Statistical Association. pp. 97-119.

Patil, K. D. 1975. “Cochran’s Q Test: Exact Distribution.” Journal of the American Statistical Association, vol. 70, No. 349, pp. 186- 189.

Plackett, R. L. 1974. “The Analysis of Categorical Data”. p. 102.