$2\times C$ 分割表の等質性検定における棄却域の頑健性について

須 川 和 明

目 次

1．はじめに

分割表における等質性検定の統計的方法としては，χ^2 検定法，Yates の修正 χ^2 検定法，Fisher の直接確率計算法等の方法が知られ広く利用されている。また田口によりF近似を用いたF検定法も提案されている。これらの方法の多くは χ^2 近似やF近似による連続近似により棄却域を導くものであり，このことは特に標本数の少ない場合に近似誤差の程度が各検定法により異なることなどから検定法の選択を困難なものにしている。近似的に適用された結果がどの程度信頼できるかを解析的に知ることは難しいが，各検定法の棄却域の大きさ，第1種の過誤の確率の真値からのずれ，第2種の過誤の大きさ等を計算機シミュレーションによって比較することで，検定法の頑健性や検出力をある程度数量的に知ることができる。

χ^2 検定法については，R.C.Lewontin et al. (1965)が 2×5 分割表と 2×10 分割表について両側の周辺値を固定した超幾何分布モデルを用いてモンテカルロシミュレーションを行い，各セルの期待値が少なくとも1

以上であれば第1種の過誤の確率に対する近似は十分良いことを報告している。G.Taguchi (1975)は2×2分割表の場合に，棄却域と第1種の過誤の確率を χ^2 検定法，修正 χ^2 検定法，直接確率計算法，F検定法の間で比較した結果，棄却域については χ^2 検定法とF検定法，修正 χ^2 検定法と直接確率計算法がそれぞれほぼ等しい大きさを持つこと，第1種の過誤の確率の近似については χ^2 検定法とF検定法が優れていること等を報告している。しかし，これらの報告は χ^2 検定法についてのモンテカルロシミュレーションであったり，2×2分割表に限定された比較であるため十分であるとは云えない。

本報告では，$2\times C(C=2,\ 3,\ 4)$ 分割表の場合について，χ^2 検定法，修正 χ^2 検定法，F検定法の棄却域の大きさ，第1種の過誤の確率 α 値，第1種の過誤の確率 α 値の近似誤差を周辺値のすべての組合せについて計算することにより，各検定法の棄却域の比較と第1種の過誤の確率 α 値の比較を行い，これらの検定法を用いる場合の指針を求めた[1]。また，現在はほとんど利用されることのないF検定法ではあるが他の検定法に較べて優れた点を持つことも示す。

2．$2\times C$ 分割表と検定モデル

c 個の集団からそれぞれランダムに抽出した要素数 n_1，n_2，……，n_c の標本の集合があるとき，i 番目の集合の中で特定の要因を有する標本の数を $x_i(i=1,\ 2,\ \cdots\cdots,\ C)$ とすれば，表1に示す $2\times C$ 分割表が得られる。

表1　$2\times C$ 分割表

クラス	1	2	3	・・・	C	計
要因＋	x_1	x_2	x_3	・・・	x_c	x
要因－	n_1-x_1	n_2-x_2	n_3-x_3	・・・	n_c-x_c	$n-x$
計	n_1	n_2	n_3	・・・	n_c	n

$n=\sum_{i=1}^{c} n_i$ は全標本数であり，　$x=\sum_{i=1}^{c} x_i$ は要因を有する標本全体の数である。

この分割表に対する検定モデルとしては，標本数 n を固定した多項分布モデル，周辺値 n_i を固定した c 個の独立な 2 項分布モデル，周辺値 n_i と x を固定した超幾何分布モデルがある。ある要因の有無に関してこれら c 個の集団の等質性を検定する等質性検定法は 2 番目の c 個の独立な 2 項分布モデルの場合であり，i 番目の集団からの標本が要因を有する確率の母集団値を p_i とすれば，帰無仮説 H_0 は $H_0: p_1=p_2=\cdots\cdots=p_c$ と表せる。c 個の集団が独立であると仮定し各 n_i を固定すれば，$(x_1,\ x_2,\ \cdots\cdots,\ x_c)$ の同時確率分布は c 個の独立な 2 項分布の積となり，その同時確率密度関数は次式で表される。

$$p(x_1, x_2, \cdots\cdots, x_c | n_1, n_2, \cdots\cdots, n_c ; p_1, p_2, \cdots\cdots, p_c)=\prod_{i=1}^{c}\binom{n_i}{x_i}p_i^{x_i}(1-p_i)^{n_i-x_i} \tag{2.1}$$

この 2 項分布モデルのもとで χ^2 検定法と F 検定法の棄却域を求める。検定に用いる統計量は次の 3 つである。

(1)　χ^2 検定

$$X^2=\frac{\sum_{i=1}^{c} n_i(\bar{p}_i-\bar{p})^2}{\bar{p}(1-\bar{p})} \qquad (c=2,\ 3,\ 4) \tag{2.2}$$

(2)　修正 χ^2 検定

$$X_c^2=\frac{n\{|x_1n_2-x_2n_1|-\frac{1}{2}n\}^2}{n_1n_2x(n-x)} \qquad (c=2) \tag{2.3}$$

(3)　F 検定

$$F=\frac{\sum_{i=1}^{c} n_i(\bar{p}_i-\bar{p})^2}{c-1} \Bigg/ \frac{\sum_{i=1}^{c} n_i\bar{p}_i(1-\bar{p}_i)}{n-c} \qquad (c=2,\ 3,\ 4) \tag{2.4}$$

ここで $n_i,\ x_i,\ n,\ x$ は表 1 で定義されたものであり，$\bar{p}_i=x_i/n_i$，$\bar{p}=x/n$ と定義する。

3. 棄却域と第1種の過誤の確率

棄却域は，標本数 n が20，30，40の場合について求めた。各周辺値の組 n_i については，要因+の確率を p_i とすると要因−の確率は $1-p_i$ であることから p_i は $0<p_i\leqq 0.5$ としてよい，また各 x_i の期待値 $n_i p_i$ が少なくとも1以上であることを仮定し，$n_i p_i \geqq 1$ と p_i の平均としての $p_i=0.25$ から $n_i \geqq 4$ とする。

以下に，棄却域を決定し，第1種の過誤の確率とその近似誤差を計算する手順を示す。

(1) 分割表の生成

各標本数 n ごとに，$n_i \geqq 4$ を満たす n_i の組を決定する。決定した各 n_i の組に対して，$0 \leqq x_i \leqq n_i$ を満たす x_i の組を決定し，ひとつの分割表を生成する。

(2) 棄却域の決定

決定された n_i の組ごとに，可能な分割表の全て，つまり $0 \leqq x_i \leqq n_i$ を満たす x_i の組を生成し，式(2.2)，(2.3)，(2.4)に示した X^2，X_c^2，F の値を計算し，有意水準5％と1％の棄却域を決定する。

(3) 第1種の過誤の確率

帰無仮説 $H_0: p_1=p_2=\cdots\cdots=p_c$ のもとでパラメータ p_i に特定の値を仮定して，(2)で決定した棄却域に対して第1種の過誤の確率 α 値を式(2.1)から計算する。ここでは p_i の値として0.1，0.2，0.3，0.4，0.5を用いる。

(4) 第1種の過誤の確率 α 値の近似誤差

パラメータ p_i の上記5つの値に対して計算した真の α 値を，それぞれ α_1，α_2，……，α_5 とし，棄却域を決定する際に設定した α 値を α_0 とするとき，第1種の過誤の確率 α 値の近似誤差を次の式で評価する。

$$\mathrm{EMQ}=\left\{\frac{1}{5}\sum_{k=1}^{5}(\alpha_k-\alpha_0)^2\right\}^{\frac{1}{2}} \tag{2.5}$$

$$\mathrm{EMQ}(k)=\{E[(\alpha_k-\alpha_0)^2]\}^{\frac{1}{2}} \quad (k=1,\ 2,\ \cdots\cdots,\ 5) \tag{2.6}$$

（ここで $E[\cdot]$ は，棄却域の決定に用いた周辺値 n_i の組全体での期待値である。）

4．棄却域の比較

棄却域の大きさの比較は，それぞれの有意水準で帰無仮説が棄却される分割表の個数を較べることで行うことができる。例えば，$n=40$ のときの 2×2 分割表では，周辺値を $(n_1, n_2)=(11, 29)$ としたとき可能な分割表の個数は全部で360あるが，そのうち 5 %の有意水準で帰無仮説が棄却される分割表の個数は，修正 χ^2 検定法，χ^2 検定法，F検定法でそれぞれ154，182，180であり，1 %の有意水準では108，138，140である。棄却される分割表の個数が多いほど棄却域が広く，少ないほど狭い。このようにして各検定法の棄却域の大きさを較べると，修正 χ^2 検定法の棄却域はすべての場合で他の 2 検定法よりも狭くなり，χ^2 検定法とF検定法では，χ^2 検定法の方が広い場合，両方等しい場合，F検定法の方が広い場合がある。この様子は 2×2 分割表と 2×C 分割表 ($C=3, 4$) で異なるので，それぞれを別に示す。

(1) 2×2 分割表

2 検定法について，棄却域の大きさを比較した結果を表 2 に示す。修正 χ^2 検定法の棄却域はすべて他の 2 法の棄却域より狭くなったので表には含めていない。

表 2　棄却域の比較（2 × 2 分割表）

有意水準	5%			1%		
標本数	20	30	40	20	30	40
$\chi^2 \supset$ F	0	2	7	0	0	0
$\chi^2=$ F	7	10	10	6	6	9
$\chi^2 \subset$ F	0	0	0	1	6	8

$\chi^2 \supset$ F は χ^2 検定の棄却域がF検定の棄却域を真に含むことを示す。他も同様。

5 %の有意水準では，n の値が20のとき χ^2 検定法とF検定法はすべて等しい棄却域を持つが，30以上では 2 法の棄却域は等しいか χ^2 検定法の方がF検定法よりも広い棄却域を持ち，F検定法が広い棄却域を持つこ

とはない。これに対して１％の有意水準では，２法の棄却域は等しいかまたはＦ検定法の方が広い棄却域を持ち，逆に χ^2 検定法が広い棄却域を持つことはない。

(2) $2\times C$ 分割表 ($C=3$, 4)

２検定法について，2×2 分割表と同様に棄却域を比較した結果を表３に示す。2×3 分割表，2×4分割表では，周辺値 n_i のすべての組合せについて χ^2 検定法がＦ検定法より広い棄却域を持つことはない。有意水準５％では χ^2 検定法とＦ検定法が同じ棄却域を持つことがあるが，その数は，2×3 分割表に較べて 2×4 分割表では大きく減少している。有意水準１％では周辺値のすべての組合せについてＦ検定法が χ^2 検定法より広い棄却域を持つ。

表３　棄却域の比較（２×２ 分割表，２×３ 分割表）

	2×3 分割表					
有意水準	5%			1%		
標本数	20	30	40	20	30	40
$\chi^2 \supset$F	0	0	0	0	0	0
$\chi^2 =$F	7	28	53	0	0	0
$\chi^2 \subset$F	3	9	27	10	37	80
	2×4 分割表					
$\chi^2 \supset$F	0	0	0	0	0	0
$\chi^2 =$F	1	0	3	0	0	0
$\chi^2 \subset$F	4	47	166	5	47	169

$\chi^2 \supset$F は χ^2 検定の棄却域がＦ検定の棄却域を真に含むことを示す。
他も同様。

５．第１種の過誤の確率の比較

棄却域の広さは，棄却域が広ければ帰無仮説が棄却され対立仮説が採択される割合が大きくなることから，検定法の検出力と大きく関わって

いるが，このことは第1種の過誤の確率 α 値が設定した値を越えないことが保証されて始めて意味を持つ。棄却域の広さと第1種の過誤の確率 α 値の関係を見るために，標本数 $n=30$ の場合について，各検定法で棄却域の広さにもっとも大きな差があった周辺値の組を選んで，その第1種の過誤の確率 α 値，第1種の過誤の確率 α 値の近似誤差の平均EMQ（2.5式），棄却域の広さ（棄却された分割表の個数）を表4に示す。2×2分割表，2×3分割表では，有意水準5％，1％とも同一の周辺値の組で棄却域の広さの差が最大になったが，2×4分割表では5％と1％で差が最大となる周辺値の組が異なるため両方の組を示してある。

この例は棄却域の広さの差が最大になるものであるが，有意水準5％では2×3分割表を除きほぼ設定した値に収まっている。有意水準1％では，F検定法の α 値が設定した値を大きく越えているのに対して χ^2 検定法では逆に設定した値を大きく下回っていることが判る。また，修正 χ^2 検定法の α 値はすべて設定した値の半分にも達していないことが判る。パラメータ p 値の変化による変動は $p=0.5$ で最も近似が良く，p の値が0または1に近づくにつれ悪くなることは良く知られた結果である。

表4　第1種の過誤の確率 α 値の比較

	2×2分割表 $(n_1, n_2)=(14, 16)$													
有意水準	5%							1%						
パラメータ p_i	0.1	0.2	0.3	0.4	0.5	EMQ	棄却域	0.1	0.2	0.3	0.4	0.5	EMQ	棄却域
修正 χ^2 検定	0.255	1.062	1.604	1.848	1.844	3.727	100	0.005	0.096	0.218	0.252	0.252	0.841	68
χ^2 検定	2.811	5.831	5.434	4.757	4.555	1.089	126	0.189	0.695	0.952	1.045	1.067	0.390	90
F検定	2.712	4.985	4.762	4.612	4.534	1.064	124	0.255	1.031	1.394	1.572	1.650	0.540	98
	2×3分割表 $(n_1, n_2, n_3)=(8, 10, 12)$													
χ^2 検定	2.461	4.782	5.134	5.021	5.035	1.141	760	0.244	0.555	0.733	0.836	0.873	0.421	522
F検定	2.823	5.062	5.285	5.255	5.268	0.996	766	0.756	1.091	1.109	1.228	1.277	0.204	570
	2×4分割表 $(n_1, n_2, n_3, n_4)=(5, 5, 5, 15)$													
χ^2 検定	3.321	3.877	3.834	3.951	4.126	1.209	1928	0.818	1.133	0.867	0.704	0.653	0.235	1294
F検定	5.067	4.637	4.711	4.846	4.735	0.250	2778	2.396	1.533	1.212	1.019	0.924	0.676	1412
	2×4分割表 $(n_1, n_2, n_3, n_4)=(6, 8, 8, 8)$													
χ^2 検定	3.164	4.405	4.828	4.832	4.780	0.875	3150	0.355	0.618	0.692	0.696	0.695	0.410	2028
F検定	3.164	4.407	4.862	4.987	5.031	0.865	3168	0.389	0.887	1.179	1.282	1.308	0.344	2348

表4はそれぞれ最も極端な例であるが，第1種の過誤の確率 α 値の近似の程度をすべての周辺値の組合せで比較するために，各標本数 n と帰無仮説のパラメータ p 値ごとに周辺値の全組合せに対して，α 値の平均値と α 値の近似の精度を表す尺度として α 値の平均2乗誤差 EMQ（k）(2.6式)を計算した。表5は2×2分割表の α 値の平均，表6は同じく2×2分割表の α 値の平均2乗誤差である。

表5からは，有意水準5％では χ^2 検定法が設定値を越える傾向があること，逆に有意水準1％ではＦ検定法が設定値を越える傾向があるが，ともに標本数 n の値の増加とともに設定値に近づく。修正 χ^2 検定法は5％，1％ともに大きく設定値を下回っている。表6からは，α 値の近似の精度が修正 χ^2 検定法で著しく悪いことが判るが，χ^2 検定法とＦ検定法の間では特に大きな違いは見られない。

表5　第1種の過誤の確率 α 値の平均

2×2分割表										
有意水準	5%					1%				
標本数(n=20)	0.1	0.2	0.3	0.4	0.5	0.1	0.2	0.3	0.4	0.5
修正 χ^2 検定	0.3012	0.6401	0.9240	1.0935	1.1451	0.0217	0.0805	0.1343	0.1663	0.1762
χ^2 検定	3.2045	4.1100	4.5397	4.9204	5.0882	0.5688	0.9003	1.0319	1.0905	1.1180
F検定	3.2045	4.1100	4.5397	4.9204	5.0882	0.6522	0.9904	1.0633	1.0959	1.1189
標本数(n=30)	0.1	0.2	0.3	0.4	0.5	0.1	0.2	0.3	0.4	0.5
修正 χ^2 検定	0.6504	1.1193	1.4296	1.5621	1.5705	0.1282	0.1920	0.2054	0.2049	0.2027
χ^2 検定	3.6816	4.4981	4.8781	5.0453	5.1677	0.9306	0.9081	0.9156	0.9504	0.9621
F検定	3.6733	4.4271	4.8096	4.9602	5.0416	0.9365	0.9608	1.0344	1.1092	1.1473
標本数(n=40)	0.1	0.2	0.3	0.4	0.5	0.1	0.2	0.3	0.4	0.5
修正 χ^2 検定	0.8687	1.4840	1.7583	1.8196	1.8195	0.1543	0.2156	0.2463	0.2535	0.2502
χ^2 検定	4.1039	4.5062	4.7998	5.0252	5.1080	1.0325	0.9559	0.9228	0.9108	0.8850
F検定	4.0993	4.4001	4.6465	4.8750	4.9237	1.0485	1.0166	1.0060	0.9758	0.9524

表6　第1種の過誤の確率 α 値の平均2乗誤差

2×2分割表										
有意水準	5%					1%				
標本数(n=20)	0.1	0.2	0.3	0.4	0.5	0.1	0.2	0.3	0.4	0.5
修正 χ^2 検定	4.7113	4.3638	4.0807	3.9149	3.8649	0.9786	0.9205	0.8663	0.8360	0.8281
χ^2 検定	2.8912	1.1815	0.8961	0.6590	0.6967	0.8117	0.5300	0.2173	0.2063	0.2679
F検定	2.8912	1.1815	0.8961	0.6590	0.6967	0.7440	0.4625	0.1681	0.2038	0.2679
標本数(n=30)	0.1	0.2	0.3	0.4	0.5	0.1	0.2	0.3	0.4	0.5
修正 χ^2 検定	4.3731	3.8833	3.5823	3.4713	3.4687	0.8845	0.8132	0.7954	0.7983	0.8025
χ^2 検定	1.8661	0.8694	0.6639	0.3271	0.4406	0.9245	0.2859	0.1397	0.1755	0.2346
F検定	1.8760	0.8360	0.6530	0.3141	0.4423	0.9198	0.2617	0.1713	0.2714	0.3445
標本数(n=40)	0.1	0.2	0.3	0.4	0.5	0.1	0.2	0.3	0.4	0.5
修正 χ^2 検定	4.1460	3.5224	3.2663	3.2150	3.2197	0.8551	0.7860	0.7549	0.7512	0.7576
χ^2 検定	1.2914	0.7369	0.6985	0.4925	0.4724	0.7388	0.2411	0.1317	0.1720	0.2272
F検定	1.2989	0.8392	0.7134	0.4747	0.4831	0.7329	0.2510	0.0984	0.1607	0.2232

同様に，2×3分割表，2×4分割表について計算した結果を表7と表8に示す。

表7　第1種の過誤の確率 α 値の平均

2×3分割表										
有意水準	5%					1%				
標本数(n=20)	0.1	0.2	0.3	0.4	0.5	0.1	0.2	0.3	0.4	0.5
χ^2 検定	2.4771	3.9372	4.5572	4.9576	5.1340	0.2666	0.6658	0.7687	0.7511	0.7302
F検定	2.4817	3.9693	4.6130	5.0185	5.1939	0.8224	1.0737	1.1509	1.1994	1.2113
標本数(n=30)	0.1	0.2	0.3	0.4	0.5	0.1	0.2	0.3	0.4	0.5
χ^2 検定	3.9790	4.6270	4.7891	4.9558	5.0630	0.6632	0.8097	0.8182	0.7790	0.7536
F検定	3.9897	4.6501	4.8280	4.9912	5.0898	0.9357	1.1649	1.1136	1.0738	1.0558
標本数(n=40)	0.1	0.2	0.3	0.4	0.5	0.1	0.2	0.3	0.4	0.5
χ^2 検定	4.3806	4.5624	4.7342	4.9097	4.9790	1.1322	0.9777	0.8770	0.8128	0.7814
F検定	4.3968	4.5867	4.7586	4.9276	4.9933	1.2693	1.1386	1.0610	1.0082	0.9838
2×4分割表										
標本数(n=20)	0.1	0.2	0.3	0.4	0.5	0.1	0.2	0.3	0.4	0.5
χ^2 検定	2.3895	3.3814	4.0298	4.5947	4.8187	0.1728	0.4711	0.6126	0.6002	0.5698
F検定	2.4632	3.7213	4.6068	5.1761	5.3338	0.5621	1.2677	1.4454	1.4816	1.4933
標本数(n=30)	0.1	0.2	0.3	0.4	0.5	0.1	0.2	0.3	0.4	0.5
χ^2 検定	3.3818	4.2408	4.5007	4.6691	4.7476	0.6945	0.8277	0.7506	0.6849	0.6569
F検定	3.7444	4.5800	4.8285	5.0147	5.0930	1.2777	1.2678	1.1893	1.1335	1.1045
標本数(n=40)	0.1	0.2	0.3	0.4	0.5	0.1	0.2	0.3	0.4	0.5
χ^2 検定	4.3147	4.4845	4.5862	4.6927	4.7475	1.0726	0.9174	0.8085	0.7303	0.6964
F検定	4.4578	4.7290	4.8288	4.9362	4.9877	1.2893	1.2032	1.1133	1.0399	1.0064

有意水準5％では，2×3分割表の α 値は x^2 検定法とF検定法で特に大きな違いはみられないが，2×4分割表では x^2 検定法の α 値が設定値より小さくなりF検定法との差がやや大きくなっている。平均2乗誤差もほぼ同様で，2×3分割表ではほとんど違いがみられないが2×4分割表ではF検定法の方が小さくなっている。有意水準1％では，2×3分割表，2×4分割表共に x^2 検定法は設定値を大きく下回り，F検定法では逆に上回っている。平均2乗誤差は，2×3分割表，2×4分割表共にF検定法の方が小さい値をとることが多いが，一部逆に大きくなることもあり5％の場合ほど安定した傾向はみられない。

表8　第1種の過誤の確率 α 値の平均2乗誤差

					2×3分割表					
有意水準	5%					1%				
標本数(n=20)	0.1	0.2	0.3	0.4	0.5	0.1	0.2	0.3	0.4	0.5
x^2 検定	2.6799	1.2396	0.6110	0.4093	0.4469	0.7401	0.3753	0.2568	0.2651	0.2959
F検定	2.6763	1.2270	0.5997	0.4107	0.4644	0.6698	0.3957	0.1839	0.2252	0.2671
標本数(n=30)	0.1	0.2	0.3	0.4	0.5	0.1	0.2	0.3	0.4	0.5
x^2 検定	1.6931	0.5746	0.3717	0.3169	0.3766	0.5073	0.2537	0.1912	0.2372	0.2788
F検定	1.6785	0.5581	0.3539	0.3171	0.3835	0.3763	0.2746	0.1335	0.1301	0.1759
標本数(n=40)	0.1	0.2	0.3	0.4	0.5	0.1	0.2	0.3	0.4	0.5
x^2 検定	1.1427	0.5341	0.3991	0.3158	0.3109	0.6242	0.1895	0.1311	0.2087	0.2584
F検定	1.1424	0.5161	0.3825	0.3048	0.3103	0.6914	0.2373	0.0766	0.1119	0.1640
					2×4分割表					
標本数(n=20)	0.1	0.2	0.3	0.4	0.5	0.1	0.2	0.3	0.4	0.5
x^2 検定	2.7683	1.7486	1.0164	0.4555	0.2713	0.8284	0.5298	0.3919	0.4069	0.4397
F検定	2.6971	1.4042	0.4608	0.3901	0.4377	0.4557	0.2908	0.4560	0.5025	0.5299
標本数(n=30)	0.1	0.2	0.3	0.4	0.5	0.1	0.2	0.3	0.4	0.5
x^2 検定	1.7853	0.7958	0.5383	0.4119	0.3471	0.4316	0.2410	0.2536	0.3203	0.3536
F検定	1.4156	0.4916	0.2495	0.2076	0.2385	0.5678	0.3179	0.1937	0.1554	0.1601
標本数(n=40)	0.1	0.2	0.3	0.4	0.5	0.1	0.2	0.3	0.4	0.5
x^2 検定	1.2128	0.5550	0.4364	0.3738	0.3483	0.4333	0.1584	0.1941	0.2783	0.3193
F検定	1.1181	0.3485	0.2305	0.2308	0.2448	0.5186	0.2629	0.1177	0.0996	0.1312

各検定法の第1種の過誤の確率 α 値の近似の精度を2×C 分割表(C=2，3，4)のCの値の変化による傾向や標本数 n の値の変化による傾向として見るために，第1種の過誤 α 値の近似誤差であるEMQの値を標

本数 n と各周辺値の組み合わせごとに平均した結果を表 9 に示す。

各検定法と有意水準で標本数 n の増加とともに近似誤差の平均値が減少しているが，このことは分布の漸近理論から明らかなことである。2×2分割表では修正 x^2 検定法の α 値の近似が著しく悪いことが判る。x^2 検定法とＦ検定法の間には大きな違いは見られない。2×3分割表，2×4分割表では，有意水準１％，５％ともにＦ検定法の方が x^2 検定法よりも近似の誤差が小さく，この違いは2×3分割表よりも2×4分割表でより大きくなっている。

表 9　近似誤差 EMQ の平均値

2×2 分割表						
有意水準	5%			1%		
標本数	20	30	40	20	30	40
修正 x^2 検定	4.1979	3.7682	3.4845	0.8874	0.8190	0.7814
x^2 検定	1.4649	0.9752	0.7674	0.4581	0.3799	0.3325
F検定	1.4649	0.9715	0.7811	0.4054	0.3982	0.3219
2×3 分割表						
x^2 検定	1.2923	0.7532	0.5795	0.4219	0.3039	0.2829
F検定	1.2896	0.7416	0.5694	0.3635	0.2191	0.2771
2×4 分割表						
x^2 検定	1.4610	0.8979	0.6280	0.5434	0.3226	0.2760
F検定	1.3194	0.6446	0.4927	0.4463	0.2869	0.2280

6．考　察

各検定法の比較結果から，Ｆ検定法は2×2分割表の有意水準５％検定の一部を除き，他の全ての場合で x^2 検定法や修正 x^2 検定法よりも広いかまたは同等の棄却域を持つことが判る。特に１％検定ではその差はより大きいものとなる。x^2 検定法がＦ検定法よりも広い棄却域を持つ2×2分割表の５％検定でもその差は僅かであることから，Ｆ検定法は他の２

検定法に較べて同程度かまたは有意な検定結果を出しやすい検定法であると云える。このことが直ちに検出力の高さを示すものでないことは既に述べたが，F検定法において第1種の過誤の確率 α 値が設定値を越えることが多いにもかかわらず，第1種の過誤の確率の近似誤差が平均的に χ^2 検定法よりも小さいことや，逆に χ^2 検定法の α 値がほとんど全ての場合で設定値よりも小さくなること等を合わせると，検定結果が有意に出すぎる危険よりも見過ごしてしまう危険の方がより致命的となる問題，例えば臨床医学や公衆衛生学のような分野に多くみられるが，このような分野の問題においてはF検定法の方が χ^2 検定法よりも有効であると考えられる。

2×2分割表において，修正 χ^2 検定法は他の2法に較べて常に狭い棄却域を持ち，第1種の過誤の確率 α 値の近似誤差も大きい値となっている。このことについては，R.L.Plackett (1964)が，2×2分割表の一方の周辺値だけを固定した二つの独立な2項分布モデルに対しては連続修正が適当ではないことを報告している。今回の結果からも，2項分布モデルを仮定した分割表の検定に対しては標本数が少ない場合でも連続修正を行うことは適当ではないと考えられる。

今回の報告では，$C=2, 3, 4$ の場合の2×C 分割表について棄却域の広さと第1種の過誤の確率 α 値を比較したが，既に見たように，$C=2$ の場合と $C \geqq 3$ の場合ではその差に相違があると思われる。この違いは $C=3$ の場合よりも $C=4$ の場合でさらに大きくなるように見えるが，今回の報告では計算機環境の制約もあり $C \geqq 5$ の場合の比較をすることができなかった。この点については今後の課題としたい。また，検出力については適当な対立仮説のもとで第2種の過誤の確率 β 値の計算を一部試みてはいるが不十分であるため，この点についても今後の課題としたい。

[謝辞]

本研究を始めるにあたり，久留米大学医学部福田勝洋教授及び Univ. Bordeaux I Prof.Lucas Pun の両氏から有益な示唆を与えていただきました。ここに記して深甚の謝意を表します。

［注］

(1) 本報告の数値計算は，北星学園大学情報処理センターの FACOM M340S を使用して行った。

［参考文献］

Brier, S. (1980). Analysis of contingency tables under cluster sampling. *Biometrika* **67**,591-596.

Donner,A. (1989). Statistical methods in ophthalmology : An adjusted chi-square approach.
Biometrics **45**,605-611.

Gart,J.J. (1971). The comparison of proportions : a reveiw of significance tests,confidence intervals,and adjustments for stratification. *Review International Statistical Institute* **39**, 148-149.

Lewontin,R.C. and Felsenstein,J. (1965). The robustness of homogeneity test in 2×N tables. *Biometrics* **21**,19-33.

Plackett,R.L. (1964). The continuity correction in 2x2 tables.
Biometrika **51**,327-337.

Sillito,G.P. (1949). Note on approximations to the power function of 2×2 comparative trial.
Biometrika **36**,347-352.

Taguchi,G. (1975). *A new statistical analysis method for clinical data,the accumulating analysis, in contrast with the chisquare test*. Shinjuku shobo.

田口玄一（1970）.『統計解析』丸善

竹内啓・藤野和建（1981）.『2 項分布とポアソン分布』東京大学出版会

山内二郎編（1972）.『統計数値表』日本規格協会

The Robustness of the Critical Regions for Homogeneity Test in $2 \times C$ Tables

Kazuaki SUGAWA

Most of the statistical methods for homegeneity test in $2 \times C$ tables determine the critical regions by continuous approximations to the chi square distribution. In the case of small samples, the errors induced by the approximations are considerably large and these errors vary according to which method is used. To evaluate these errors and differences, the critical regions and type I errors of three methods (the chi square test, the Yates' corrected chi square test and the F test) are determined by numerical calculation on a computer. The results are that the critical regions of the F test have a same or larger area than those of the other two methods and the approximation errors of the F test are slightly less than the other two.