二項分布の正規分布による近似条件と比率の差の検定（大標本）について

石　川　忠　孝

1　は じ め に

二つの標本の大きさ n_1, n_2 が大きいとき，それぞれの特定の性質をもつ数を k_1, k_2 とすると，比率の差の検定には，次の二つの方法がある。

(1)　$\hat{p}_1=\dfrac{k_1}{n_1}$, $\hat{p}_2=\dfrac{k_2}{n_2}$ として

$$z_1=\frac{\hat{p}_1-\hat{p}_2}{\sqrt{\dfrac{\hat{p}_1(1-\hat{p}_1)}{n_1}+\dfrac{\hat{p}_2(1-\hat{p}_2)}{n_2}}}$$

(2)　$\hat{p}_1=\dfrac{k_1}{n_1}$, $\hat{p}_2=\dfrac{k_2}{n_2}$, $\hat{p}=\dfrac{k_1+k_2}{n_1+n_2}$ として

$$z_2=\frac{\hat{p}_1-\hat{p}_2}{\sqrt{\hat{p}(1-\hat{p})\left(\dfrac{1}{n_1}+\dfrac{1}{n_2}\right)}}$$

ところで，これらの公式を適用するためには，考えられる二つの二項分布 $B(n_1,\ p_1)$, $B(n_2,\ p_2)$ が，正規分布 $N(n_1p_1,\ n_1p_1q_1)$, $N(n_2p_2,\ n_2p_2q_2)$ に近似可能でなければならない。二項分布の正規分布による近似条件は経験法則としてたくさんある。普通一般に使用されている近似条件「$np \geqq 5$ かつ $nq \geqq 5$」を用いて，比率の差の検定を行い，その考察をする。

なお，比率の差の検定は 2×2 分割表による χ^2 検定（独立性の検定）が適用できる。この χ^2 の値は上記の z_2 の値の 2 乗に等しい（微差は丸めの誤差)。

	B_1	B_2	計
A_1	f_{11}	f_{12}	$f_{1.}$
A_2	f_{21}	f_{22}	$f_{2.}$
計	$f_{.1}$	$f_{.2}$	N

(3)　$\chi^2 = \dfrac{N(f_{11}f_{22}-f_{12}f_{21})^2}{f_{1.}f_{2.}f_{.1}f_{.2}}$，$N = f_{11}+f_{12}+f_{21}+f_{22}$

この検定法でも例題を解いて，二項分布の正規分布による近似条件と比率の差の検定について考察を深め，よりよい検定について探求する。

2　比率の差の検定について

例題 1　ある中学校における近視者の数は表一1の通りである。男女により近視のなりやすさに差異が認められるか。有意水準5％で検定せよ。

表一 1

	近　視	近視でない	計
男子	62	273	335
女子	73	245	318
計	135	518	653

〈解 1〉　まず，普通使用される正規分布による近似条件「$n_ip_i \geqq 5$，$n_iq_i \geqq 5 (i=1, 2)$」を満たしているかを調べる。

$$n_1p_1 = 335\times\frac{62}{335} = 62>5,\quad n_1q_1 = 335\times\frac{273}{335} = 273>5$$

$$n_2p_2 = 318\times\frac{73}{318} = 73>5,\quad n_2q_2 = 318\times\frac{245}{318} = 245>5$$

よって，正規分布による近似が可能であるといえる。

H_0: $p_1 = p_2$（男女の近視率に差がない）

H_1: $p_1 \neq p_2$（男女の近視率に差がある）

両側検定，有意水準5％

$$\hat{p}_1 = \frac{62}{335} = 0.185$$

$$\hat{p}_2 = \frac{73}{318} = 0.230$$

$$z_1 = \frac{0.185-0.230}{\sqrt{\dfrac{0.185\times0.815}{335}+\dfrac{0.230\times0.770}{318}}} = \frac{-0.045}{0.0317} = -1.420 > -1.960$$

したがって，H_0 は採択され，男女の近視率に差があるとはいえない。

〈解 2〉　上記のように正規分布による近似が可能であるといえる。

母比率の点推定値 $\hat{p}=\dfrac{k_1+k_2}{n_1+n_2}=\dfrac{62+73}{335+318}=\dfrac{135}{653}=0.2067$ より

$$z_2=\frac{0.185-0.230}{\sqrt{0.2067\times0.7933\times\left(\dfrac{1}{335}+\dfrac{1}{318}\right)}}=\frac{-0.045}{0.0317}=-1.420>-1.960$$

よって，H_0 は採択され，男女の近視率に差があるとはいえない。

〈注〉　$|z_1|=|z_2|$ で両検定とも採択である。手元の文献では解答にあたり正規分布による近似条件を満たしているかを調べているのは，1冊だけである。本稿では厳密さと，その後の考察のため各例題とも正規分布による近似条件を満たしているかを調べた。一般の推定や検定の計算を手計算で正確に行うことは困難な場合が多く，途中で末位を丸めるため誤差が生じる。したがって，棄却域すれすれの値が出たときは注意したほうがよい。なお，検定統計量の値は有効数字が 3 桁か 4 桁くらいがよいといわれる。

〈解 3〉　2 × 2 分割表による χ^2 検定で解く。

H_0: $p_1=p_2$,　H_1: $p_1\neq p_2$

有意水準 5 ％

自由度 $f=(2-1)\times(2-1)=1$

$$\chi^2=\frac{653\times(62\times245-273\times73)^2}{335\times318\times135\times518}=2.059<3.841=\chi_1^2(0.05)$$

よって，H_0 は採択され，男女の近視率に差があるとはいえない。

〈注〉　z_1 による検定，z_2 による検定及び χ^2 検定とも採択である。

$$z_2=\frac{\hat{p}_1-\hat{p}_2}{\sqrt{\hat{p}\times(1-\hat{p})\times\left(\dfrac{1}{n_1}+\dfrac{1}{n_2}\right)}}\ \text{ただし}\ \hat{p}_1=\frac{k_1}{n_1},\ \hat{p}_2=\frac{k_2}{n_2},\ \hat{p}=\frac{k_1+k_2}{n_1+n_2}$$

を用いれば，このとき $z_2^2=\chi^2$ となり両者は完全に等しくなる。すなわち，$k_1=f_{11}$, $k_2=f_{12}$, $n_1=f_{\cdot1}$, $n_2=f_{\cdot2}$ となっていることに注意して，この z_2 にこれらの値を代入し z_2^2 を計算すると

	B_1	B_2	計
A_1	f_{11}	f_{12}	$f_{1\cdot}$
A_2	f_{21}	f_{22}	$f_{2\cdot}$
計	$f_{\cdot1}$	$f_{\cdot2}$	N

$$z_2^2=\frac{N(f_{11}f_{22}-f_{12}f_{21})^2}{f_{1.}f_{2.}f_{.1}f_{.2}},\quad N=f_{11}+f_{12}+f_{21}+f_{22}$$

となる。なお，これを用いた方が計算は容易である。

例題 2 あるテレビの音楽番組の視聴率調査で，男子は任意抽出された320人のうち40人が，女子は任意抽出された180人のうち36人が見ていた。男女の視聴率に差があるといえるか。有意水準5％で検定せよ。

表一 2

	見ている	見ていない	計
男子	40	280	320
女子	36	144	180
計	76	424	500

〈解1〉 まず，正規分布による近似条件「$n_ip_i \geqq 5$, $n_iq_i \geqq 5 (i=1, 2)$ を満たしているかを調べる。

$$n_1p_1=320\times\frac{40}{320}=40>5,\quad n_1q_1=320\times\frac{280}{320}=280>5$$

$$n_2p_2=180\times\frac{36}{180}=36>5,\quad n_2q_2=180\times\frac{144}{180}=144>5$$

よって，正規分布による近似が可能であるといえる。

H_0: $p_1=p_2$（男女の視聴率に差がない）

H_1: $p_1 \neq p_2$（男女の視聴率に差がある）

両側検定，有意水準5％

$$\hat{p}_1=\frac{40}{320}=\frac{1}{8}=0.125$$

$$\hat{p}_2=\frac{36}{180}=\frac{1}{5}=0.2$$

$$z_1=\frac{0.125-0.2}{\sqrt{\frac{0.125\times0.875}{320}+\frac{0.2\times0.8}{180}}}=\frac{-0.075}{0.0351}=-2.137<-1.960$$

したがって，H_0 は棄却され，男女の視聴率に差があるといえる。

〈解2〉 上記のように正規分布による近似が可能であるといえる。

母比率の点推定値 $\hat{p}=\frac{k_1+k_2}{n_1+n_2}=\frac{40+37}{320+180}=\frac{76}{500}=0.152$ より

$$z_2=\frac{0.125-0.2}{\sqrt{0.152\times0.848\times\left(\frac{1}{320}+\frac{1}{180}\right)}}=\frac{-0.075}{0.0334}=-2.246<-1.960$$

よって，H_0 は棄却され，男女の視聴率に差があるといえる。

〈解３〉　２×２分割表による χ^2 検定で解く。

$H_0: p_1=p_2$,　$H_1: p_1\neq p_2$

有意水準５％

自由度 $f=(2-1)\times(2-1)=1$

$$\chi^2=\frac{500\times(40\times144-280\times36)^2}{320\times180\times76\times424}=5.027>3.841=\chi_1^2(0.05)$$

よって，H_0 は棄却され，男女の視聴率に差があるといえる。

〈注〉　$|z_1|<|z_2|$ であり，z_1 による検定，z_2 による検定及び χ^2 検定とも棄却である。

例題３　新薬が開発されたときには，その薬が本当に効くかどうかを次のようにテストする。すなわち，新薬と偽薬（プラシボという）のどちらかを，被験者に教えずに与えて，その効能があったかどうかを調べるのである。

いま，200人の被験者に対するテストの結果が表―３のようであったとする。果たしてこの薬は有効であったといえるだろうか。有意水準５％で検定せよ。

表―３

	効能あり	効能なし	計
新薬	79	31	110
偽薬	52	38	90
計	131	69	200

〈解１〉　まず，正規分布による近似条件「$n_ip_i\geqq5$,　$n_iq_i\geqq5(i=1,\ 2)$」を満たしているかを調べる。

$$n_1p_1=110\times\frac{79}{110}=79>5,\quad n_1q_1=110\times\frac{31}{110}=31>5$$

$$n_2p_2=90\times\frac{52}{90}=52>5,\quad n_2q_2=90\times\frac{38}{90}=38>5$$

よって，正規分布による近似が可能であるといえる。

$H_0: p_1 = p_2$（薬の種類と効能とは無関係である）

$H_1: p_1 \neq p_2$（薬の種類と効能とは関係がある）

両側検定，有意水準5％

$$\hat{p}_1 = \frac{79}{110} = 0.718$$

$$\hat{p}_2 = \frac{52}{90} = 0.578$$

$$z_1 = \frac{0.718 - 0.578}{\sqrt{\dfrac{0.718 \times 0.282}{110} + \dfrac{0.578 \times 0.422}{90}}} = \frac{0.140}{0.0675} = 2.074 > 1.960$$

よって，H_0は棄却され，薬の種類と効能には関係がある。すなわち，この新薬は確かに効果があったといえる。

〈解2〉　上記のように正規分布による近似が可能であるといえる。

母比率の点推定値 $\hat{p} = \dfrac{k_1 + k_2}{n_1 + n_2} = \dfrac{79+52}{110+90} = \dfrac{131}{300} = 0.655$ より

$$z_2 = \frac{0.718 - 0.578}{\sqrt{0.655 \times 0.345 \times \left(\dfrac{1}{110} + \dfrac{1}{90}\right)}} = \frac{0.140}{0.0676} = 2.071 > 1.960$$

よって，H_0は棄却され，薬の種類と効能には関係がある。すなわち，この新薬は確かに効果があったといえる。

〈解3〉　2×2分割表によるχ^2検定で解く。

$H_0: p_1 = p_2$，$H_1: p_1 \neq p_2$

有意水準5％

自由度 $f = (2-1) \times (2-1) = 1$

$$\chi^2 = \frac{200 \times (79 \times 38 - 31 \times 52)^2}{110 \times 90 \times 131 \times 69} = 4.318 > 3.841 = \chi_1^2(0.05)$$

よって，H_0は棄却され，薬の種類と効能には関係がある。すなわち，この新薬は確かに効果があったといえる。

〈注〉　$z_1 \fallingdotseq z_2$である。z_1による検定，z_2による検定及びχ^2検定とも棄却である。

例題 4　ある予備校で有名T大選抜クラスを設置した。表一4から選抜クラスの在籍者が有利だったといえるだろうか。有意水準5％で検定せよ。

表一 4

	合　格	不合格	計
在　籍	7	50	57
非在籍	6	135	141
計	13	185	198

〈解 1〉　まず，正規分布による近似条件「$n_ip_i \geqq 5$，$n_iq_i \geqq 5(i=1,\ 2)$」を満たしているかを調べる。

$$n_1p_1 = 57\times\frac{7}{57} = 7>5,\quad n_1q_1 = 57\times\frac{50}{57} = 50>5$$

$$n_2p_2 = 141\times\frac{6}{141} = 6>5,\quad n_2q_2 = 141\times\frac{135}{141} = 135>5$$

よって，正規分布による近似が可能であるといえる。

H_0: $p_1 = p_2$（合否は選抜クラス在籍とは無関係である）

H_1: $p_1 \neq p_2$（合否は選抜クラス在籍とは関係がある）

両側検定，有意水準5％

$$\hat{p}_1 = \frac{7}{57} = 0.123$$

$$\hat{p}_2 = \frac{6}{141} = 0.043$$

$$z_1 = \frac{0.123-0.043}{\sqrt{\dfrac{0.123\times0.877}{57}+\dfrac{0.043\times0.957}{141}}} = \frac{0.080}{0.0467} = 1.713<1.960$$

よって，H_0 は採択され，合否は選抜クラス在籍とは関係があるとはいえない。すなわち，選抜クラス在籍者の方が有利だったとはいえない。

〈解 2〉　上記のように正規分布による近似が可能である。

$$\text{母比率の点推定値}\ \hat{p} = \frac{k_1+k_2}{n_1+n_2} = \frac{7+6}{57+141} = \frac{13}{198} = 0.066$$

$$z_2 = \frac{0.123-0.043}{\sqrt{\dfrac{0.123\times0.877}{57}+\dfrac{0.043\times0.957}{141}}} = \frac{0.080}{0.0390} = 2.051>1.960$$

よって，H_0 は棄却され，合否は選抜クラス在籍とは関係がある。すなわち，選

抜クラス在籍者の方が有利だったということになる。

〈解3〉　2×2分割表による χ^2 検定で解く。

$H_0: p_1 = p_2$, $H_1: p_1 \neq p_2$

有意水準5％

自由度 $f = (2-1)\times(2-1) = 1$

$$\chi^2 = \frac{198\times(7\times135-50\times6)^2}{57\times141\times13\times185} = 4.262 > 3.841 = \chi_1^2(0.05)$$

よって，H_0 は棄却され，合否は選抜クラス在籍とは関係がある。すなわち，選抜クラス在籍者の方が有利だったということになる。

ところで，表中のすべての数値がなるべくなら10以上が望ましいが，10未満の数があるから，イェツの補正を行うと

$$\chi^2 = \frac{198\times(|7\times135-50\times6|-198\div2)^2}{57\times141\times13\times185} = 3.054 < 3.841 = \chi_1^2(0.05)$$

よって，H_0 は採択され，合否は選抜クラス在籍とは関係があるとはいえない。すなわち，選抜クラス在籍者の方が有利だったとはいえない。

〈注〉　$z_1 < z_2$ であり，z_1 による検定は採択，z_2 による検定は棄却，また χ^2 検定は棄却，イェツの補正をした χ^2 検定は採択である。一般に z_1 と z_2 のどちらを使用してもよく同じ結果となるのだが，この例題では両者の結果が異なる。

例題5　A，B社からの納入品の標本調査で，A社は任意抽出された355個のうち4個が，B社は任意抽出された250個のうち9個が不良品であった。会社によって不良品の数に違いがあるといえるか。有意水準5％で検定せよ。

表－5

	良品	不良品	計
A社	351	4	355
B社	241	9	250
計	592	13	605

〈解1〉　まず，正規分布による近似条件「$n_ip_i \geqq 5$, $n_iq_i \geqq 5(i=1, 2)$」を満たしているかを調べる。

$$n_1p_1 = 355\times\frac{4}{355} = 4 < 5, \quad n_1q_1 = 355\times\frac{351}{355} = 351 > 5$$

$$n_2p_2 = 250\times\frac{9}{250} = 9>5,\quad n_2q_2 = 250\times\frac{241}{250} = 241>5$$

よって，A会社は正規分布による近似が不可能，B会社は正規分布による近似が可能であるといえる。したがって，z_1 と z_2 による検定は行えない。そこで，2×2 分割表による χ^2 検定を行う。

H_0: $p_1 = p_2$（会社によって不良品の数に違いがない）

H_1: $p_1 \neq p_2$（会社によって不良品の数に違いがある）

有意水準 5 %

自由度 $f = (2-1)\times(2-1) = 1$

$$\chi^2 = \frac{605\times(351\times 9-4\times 241)^2}{355\times 250\times 592\times 13} = 4.268>3.841 = \chi_1^2(0.05)$$

よって，H_0 は棄却され，会社によって不良品の数に違いがあるといえる。

ところで，表中のすべての数値がなるべく 10 以上が望ましいが，10 未満の数とくに 5 未満の数があるから，イェツの補正を行うと

$$\chi^2 = \frac{605\times(|351\times 9-4\times 241|-605\div 2)^2}{355\times 250\times 592\times 13} = 3.172<3.841 = \chi_1^2(0.05)$$

よって，H_0 は採択され，会社によって不良品の数に違いがあるとはいえない。

〈注〉　A会社は正規分布による近似が不可能であるから，z_1 と z_2 による検定は行えないが，考察のためこれらの検定を試みると次のようになる。

○ z_1 による検定

H_0: $p_1 = p_2$，H_1: $p_1 \neq p_2$

両側検定，有意水準 5 %

$$\hat{p}_1 = \frac{4}{355} = 0.011,\quad \hat{p}_2 = \frac{9}{250} = 0.036$$

$$z_1 = \frac{0.011-0.036}{\sqrt{\frac{0.011\times 0.989}{355}+\frac{0.036\times 0.964}{250}}} = \frac{-0.025}{0.0130} = -1.923>-1.960$$

よって，H_0 は採択され，会社によって不良品の数に違いがあるとはいえない。

○ z_2 による検定

母比率の点推定値 $\hat{p}=\dfrac{k_1+k_2}{n_1+n_2}=\dfrac{4+9}{355+250}=\dfrac{13}{605}=0.021$

$$z_2=\frac{0.011-0.036}{\sqrt{0.021\times0.979\times\left(\dfrac{1}{355}+\dfrac{1}{250}\right)}}=\frac{-0.025}{0.0118}=-2.119<-1.960$$

よって，H_0 は棄却され，会社によって不良品の数に違いがあるといえる。

$|z_1|<|z_2|$ であり，z_1 による検定は採択，z_2 による検定は棄却，また χ^2 検定は棄却，イェツの補正をした χ^2 検定は採択である。一般に z_1 と z_2 のどちらを使用しても，同じ結果がでるのだが，この場合も両者の結果は異なる。

ところで，教科書では項目ごとに問題が与えられているから，その解法はおのずから分かる。すなわち，例題１，例題２，例題３は z_1 と z_2 による検定の項目中に，例題４は２×２分割表による χ^2 検定の項目中に，例題５は２×２分割表によるイェツの補正を加えた χ^2 検定の項目中に，それぞれ与えられている。しかし，問題が任意に与えられたときは，その問題の適切な解法を見付けねばならない。

3　二項分布の正規分布による近似条件について

例題の解答にあたり，厳密を期すためとその後の考察のために二項分布の正規分布による近似条件「$np\geqq5$ かつ $nq\geqq5$」を満たしているかを調べた。ところが，二項分布の正規分布による近似条件は，経験法則としていろいろあり，これらのいずれかを満たしていれば，近似的に二項分布が正規分布をするものと見なして取り扱っても，実用上差し支えないといわれている。手元の文献を調査すると二項分布の正規分布による近似条件の経験法則は次の14個がある。

① $p>0.1$，$n>50$，　$p>0.3$，$n>20$ で $np>5$

② $np>5$ かつ $nq>5$

③ $np\geqq5$ かつ $nq\geqq5$

④ $n>30$ で $np>5$ かつ $nq>5$

⑤ $np\geqq5$，$p\leqq0.5$

⑥　$np \geqq 5$

⑦　$p \leqq \dfrac{1}{2}$ ならば $np>5$, $p>\dfrac{1}{2}$ ならば $nq>5$

⑧　$npq \geqq 3$ または $np \geqq 5$

⑨　$npq \geqq 3$

⑩　$\min(np,\ nq)>10$

⑪　$0.1 \leqq p \leqq 0.9$ かつ $npq>5$

⑫　$npq \geqq 25$

⑬　$\min(np,\ nq) \geqq 5$

⑭　$\dfrac{9}{n+9} \leqq p \leqq \dfrac{n}{n+9}$

なお，⑭は次のように成り立つことが数学的に証明できる。

X が二項分布に従うとき，X を標準化した変数 $Z=\dfrac{X-np}{\sqrt{npq}}$ は n が十分大きいならば標準正規分布に従う。それゆえ，ほとんど確実に(約 0.9974 の確率で)

$$-3 \leqq \frac{X-np}{\sqrt{npq}} \leqq 3$$

であり，従って

$$np-3\sqrt{npq} \leqq X \leqq np+3\sqrt{npq}$$

が成り立つ。X がとる値は，0 から n までの整数だから

$$np-3\sqrt{npq} \geqq 0$$

$$np+3\sqrt{npq} \leqq n$$

でなければならない。これらをそれぞれ p について解けば

$$\therefore \quad \frac{9}{n+9} \leqq p \leqq \frac{n}{n+9}$$

ここで，いままでに取り扱った五つの例題にでてきた二項分布が，14 個の二項分布の正規分布による近似条件を満たしているかを調べると表—6 の通りで

表—6　二項分布の正規分布による近似条件の判定

近似条件＼問題	例題1	例題2	例題3	例題4	例題5
① $p>0.1,\ n>50$ $p>0.3,\ n>20$ で $np>5$	○　○	○　○	○　○	○　×	×　×
② $np>5$ かつ $nq>5$	○　○	○　○	○　○	○　○	×　○
③ $np \geqq 5$ かつ $nq \geqq 5$	○　○	○　○	○　○	○　○	×　○
④ $n>30$ で $np>5$ かつ $nq>5$	○　○	○　○	○　○	○　○	×　○
⑤ $np \geqq 5$　$p \leqq 0.5$	○　○	○　○	×　×	○　○	×　○
⑥ $np \geqq 5$	○　○	○　○	○　○	○　○	×　○
⑦ $p \leqq \frac{1}{2}$ ならば $np>5$ $p>\frac{1}{2}$ ならば $nq<5$	○　○	○　○	○　○	○　○	×　○
⑧ $npq \geqq 3$ または $np \geqq 5$	○　○	○　○	○　○	○　○	○　○
⑨ $npq \geqq 3$	○　○	○　○	○　○	○　○	○　○
⑩ $\min(np,\ nq)>10$	○　○	○　○	○　○	×　×	×　×
⑪ $0.1 \leqq p \leqq 0.9$ かつ $npq>5$	○　○	○　○	○　○	○　×	×　×
⑫ $npq>25$	○　○	○　○	×　×	×　×	×　×
⑬ $\min(np,\ nq) \geqq 5$	○　○	○　○	○　○	○　○	×　○
⑭ $\frac{9}{n+9} \leqq p \leqq \frac{n}{n+9}$	○　○	○　○	○　○	×　×	×　○

ある。表中○印は近似条件を満たしていることを表し，×印は近似条件を満たしていないことを表す。

例題1，例題2は，全部の近似条件を満たし，検定結果は先に述べたように三つの検定とも同じで，例題1は採択，例題2は棄却である。例題3は近似条

件⑤「$np \geqq 5$，$p<0.5$」と⑫「$npq>25$」を満たしていないが，検定結果は三つの検定とも棄却である。ところが，例題4と5は，表―6のように近似条件の判定は一定でなくいろいろであり，検定結果は検定方法により異なる。そこで，普通よく使用される近似条件⑤「$np \geqq 5$ かつ $nq \geqq 5$」と数学的に成り立つことが証明できる⑭「$\frac{9}{n+9} \leqq p \leqq \frac{n}{n+9}$」を用いて，例題4と5を考察する。例題1，2，3はこの両者を満たしている。例題4は近似条件「$np \geqq 5$ かつ $nq \geqq 5$」を満たしているが，「$\frac{9}{n+9} \leqq p \leqq \frac{n}{n+9}$」については $\frac{9}{57+9}>p_1=\frac{7}{57}<\frac{57}{57+9}$，$\frac{9}{141+9}>p_2=\frac{6}{141}<\frac{141}{141+9}$ となり，満たしていない。そして，z_1 による検定とイェツの補正を加えた χ^2 検定は採択，z_2 による検定と χ^2 検定は棄却である。すなわち，z_1 による検定と z_2 による検定の結果は異なる。したがって，例題4は χ^2 検定を行うのが適切である。例題5は近似条件「$np \geqq 5$ かつ $nq \geqq 5$」と「$\frac{9}{n+9} \leqq p \leqq \frac{n}{n+9}$」の両方とも，B社の二項分布が満たしてなくて，「$\frac{9}{n+9} \leqq p \leqq \frac{n}{n+9}$」については $\frac{9}{355+9}>p_1=\frac{4}{355}<\frac{355}{355+9}$，$\frac{9}{250+9}<p_2=\frac{9}{250}<\frac{250}{250+9}$ である。そして，例題4と同じように，z_1 による検定とイェツの補正を加えた χ^2 検定は採択，z_2 による検定と χ^2 検定は棄却である。すなわち，z_1 による検定と z_2 による検定の結果は異なる。したがって，例題5は 2×2 分割表の表中に5未満の数があるから，イェツの補正を加えた χ^2 検定を行うのが適切である。

近似条件「$np \geqq 5$ かつ $nq \geqq 5$」を満たしただけでは，例題4のように z_1 による検定と z_2 による検定の結果が異なる場合がある。よって，二項分布の正規分布への近似条件は「$np \geqq 5$ かつ $nq \geqq 5$」よりも「$\frac{9}{n+9} \leqq p \leqq \frac{n}{n+9}$」を使用する方が安全であるといえる。

4 ま と め

二項分布の正規分布による近似条件「$np \geqq 5$ かつ $nq \geqq 5$」が成り立てば，近似的に二項分布 $B(n,\ p)$ が正規分布 $N(np,\ npq)$ をするものとみなして取り扱っても，実用上差し支えないといわれている。ところが，例題4のように近似条件「$np \geqq 5$ かつ $nq \geqq 5$」を満たしていても，z_1 による検定と z_2 による検定の結果が異なる場合がある。したがって，安全で正確な比率の検定を行うためには，近似条件「$\frac{9}{n+9} \leqq p \leqq \frac{n}{n+9}$」を使用して，次のように行うのがよいと考える。

① 近似条件「$\frac{9}{n+9} \leqq p \leqq \frac{n}{n+9}$」が成り立つ場合は，$z_1$ と z_2 のどちらで検定してもよい。もちろん，χ^2 検定を行ってもよい。

② 近似条件「$\frac{9}{n+9} \leqq p \leqq \frac{n}{n+9}$」が成り立たない場合には，$\chi^2$ 検定を行う。ただし，2×2 分割表の表中に5未満の数があるときは，イェツの補正を加えた χ^2 検定を行う。また，2×2 分割表中に5以上10未満の数があるときは，χ^2 検定とイエツの補正を加えた χ^2 検定の両方を行って安全を期する。

参 考 文 献

1) 林知己夫編「統計学の基本」朝倉書店，1994年，p 126～p 128
2) 石井博昭・塩出省吾・新森修一共著「確率統計の数理」裳華房，1995年，p 118～p 119，p 125～p 128
3) 近藤良夫・安藤貞一編「統計的方法百問百答」日科技連，1976年，p 60～p 61，p 75～p 76
4) 久保応助・藤沢偉作共著「確率・統計演習」聖文社，平成元年，p 198～p 199
5) 村上正康著「高等学校における確率・統計の指導法について」統計学会チュトリアル・セミナー，1995年，p 85～p 86
6) 成実清松・坂井忠次共著「数理統計学要説」培風館，昭和33年，p 126～p 127，p 145～p 149
7) 西平重喜著「統計調査法」培風館，昭和55年，p 159～p 162，p 175～p 179
8) 小寺平治著「新統計入門」裳華修，1996年，p 110～p 111，p 116～p 121

9）小寺平治著「明解演習数理統計」共立出版，1997年，p 114，p 146
10）岡田泰栄著「統計学概論」共立出版，1995年，p 124～p 126，p 130～p 133
11）宇田川正友著「教師のための初等数学講座　確率と統計」岩崎書店，昭和34年，p 157～p 158
12）和達三樹・十河清共著「キーポイント確率・統計」岩波書店，1994年，p 120～p 123