研究ノート

カイ自乗統計量について

木　村　　　等

1. 2×2分割表

カイ自乗統計量をもちいて検定の問題をあつかうものとして，分割表の独立性の検定がある。最初に，最も簡単な2×2分割表について述べる。2×2分割表は，表1のようにデータが A, B 2つの軸について，それぞれ2つの組に分類されているものである。A_1B_1, A_1B_2, A_2B_1, A_2B_2 の4つの組にふくまれる度数を x_{11}, x_{12}, x_{21}, x_{22} とかく。このようなデータがえられる状況としては，つぎのような3つの場合が考えられる。すなわち，

表　1

	B_1	B_2	計
A_1	x_{11}	x_{12}	r_1
A_2	x_{21}	x_{22}	r_2
計	c_1	c_2	

(a) 一つの母集団から抽出された n 個の標本を，A, B について，2重分類した場合：この場合 x_{ij} は組 A_iB_j にふくまれる標本数であり，この組におちる母集団確率を p_{ij} とすれば，表1の分割表に対応するサンプルをうる確率は，

$$p(x_{11},\ x_{12},\ x_{21},\ x_{22})=\frac{n!}{x_{11}!\,x_{12}!\,x_{21}!\,x_{22}!}p_{11}{}^{x_{11}}p_{12}{}^{x_{12}}p_{21}{}^{x_{21}}p_{22}{}^{x_{22}} \qquad (1)$$

となる。ここで当然 $p_{11}+p_{12}+p_{21}+p_{22}=1$ がなりたつ。周辺確率を

$$p_{1\cdot}=p_{11}+p_{12},\qquad p_{2\cdot}=p_{21}+p_{22}$$

$$p_{\cdot 1}=p_{11}+p_{21},\qquad p_{\cdot 2}=p_{12}+p_{22}$$

とかくことにすれば，A, B が独立であるという仮説は

$$H_0\colon\ p_{ij}=p_{i\cdot}p_{\cdot j}\quad (i,\ j=1,\ 2) \qquad (2)$$

とかくことができる。このとき（1）は

$$p_0(x_{11},\ x_{12},\ x_{21},\ x_{22})=\frac{n!}{x_{11}!\,x_{12}!\,x_{21}!\,x_{22}!}(p_{1\cdot}p_{\cdot 1})^{x_{11}}(p_{1\cdot}p_{\cdot 2})^{x_{12}}(p_{2\cdot}p_{\cdot 1})^{x_{21}}(p_{2\cdot}p_{\cdot 2})^{x_{22}}$$

$$=\frac{n!}{x_{11}!\,x_{12}!\,x_{21}!\,x_{22}!}p_{1\cdot}{}^{x_{11}+x_{12}}p_{2\cdot}{}^{x_{21}+x_{22}}p_{\cdot 1}{}^{x_{11}+x_{21}}p_{\cdot 2}{}^{x_{12}+x_{22}}$$

$$=\frac{n!}{x_{11}!\,x_{12}!\,x_{21}!\,x_{22}!}p_{1\cdot}{}^{r_1}p_{2\cdot}{}^{r_2}p_{\cdot 1}{}^{c_1}p_{\cdot 2}{}^{c_2} \qquad (3)$$

となる。

ここで，

$p_{1\cdot}+p_{2\cdot}=1$，　$p_{\cdot 1}+p_{\cdot 2}=1$

であるから，$p_{1\cdot}$，$p_{2\cdot}$ のうちの１つ，$p_{\cdot 1}$，$p_{\cdot 2}$ のうちの１つ合計２つが定まれば他はこの条件からもとまる。すなわち，未知パラメーターが２つある。また，x_{ij} については，

$x_{11}+x_{12}+x_{21}+x_{22}=n$

という条件を満足しなければならない。

（b）　２つの母集団A_1，A_2 からそれぞれ独立な標本 r_1，r_2 個を抽き出したとき，B_1，B_2 に分類される個数がそれぞれ，x_{11}，x_{12}；x_{21}，x_{22} である場合：母集団 A_1 においてB_1，B_2 におちる確率をそれぞれ p_1，q_1，母集団 A_2 において，B_1，B_2 におちる確率をそれぞれ p_2，q_2 とすれば，表１の分割表をうる確率は，

$$p(x_{11},\ x_{12};\ x_{21},\ x_{22})=\frac{r_1!}{x_{11}!\,x_{12}!}\ \frac{r_2!}{x_{21}!\,x_{22}!}p_1{}^{x_{11}}q_1{}^{x_{12}}p_2{}^{x_{21}}q_2{}^{x_{22}}$$

となる。仮説 H_0 は，２つの母集団における B_1，B_2 におちる確率がそれぞれ相等しい，すなわち，

$$H_0:\quad p_1=p_2=p,\qquad q_1=q_2=q \qquad (4)$$

である。

このとき，

$$p_0(x_{11},\ x_{12};\ x_{21},\ x_{22})=\frac{r_1!}{x_{11}!\,x_{12}!}\ \frac{r_2!}{x_{21}!\,x_{22}!}p^{x_{11}}q^{x_{21}}p^{x_{21}}q^{x_{22}}$$

$$=\frac{r_1!}{x_{11}!\,x_{12}!}\ \frac{r_2!}{x_{21}!\,x_{22}!}p^{c_1}q^{c_2} \qquad (5)$$

をうる，ここで，$p+q=1$であるから，未知パラメーターは１つであり，x_{ij} は $x_{11}+x_{12}=r_1$，$x_{21}+x_{22}=r_2$ の２つの条件を満足しなければならない。

（a）の場合と（b）の場合の相異は，（a）においては，r_1，r_2 は確率変数であるに対して，（b）においては定数であることである。いま（a）において，n 個の標本の抽出のくり返えしを考え，行和がそれぞれr_1，r_2 となる標本のみに着目する。このような標本をうる確率

は

$$p_0(r_1,\ r_2)=\frac{n!}{r_1!\,r_2!}p_{1\cdot}{}^{r_1}p_{2\cdot}{}^{r_2} \tag{6}$$

である。したがって，行和がそれぞれ r_1，r_2 であるという条件のもとで表1の分割表をうる条件つき確率は

$$p_0(x_{11},\ x_{12},\ x_{21},\ x_{22}\mid r_1,\ r_2)=\frac{n!}{x_{11}!\,x_{12}!\,x_{21}!\,x_{22}!}p_{1\cdot}{}^{r_1}p_{2\cdot}{}^{r_2}p_{\cdot 1}{}^{c_1}p_{\cdot 2}{}^{c_2}\Big/\frac{n!}{r_1!\,r_2!}p_{1\cdot}{}^{r_1}p_{2\cdot}{}^{r_2}$$

$$=\frac{r_1!}{x_{11}!\,x_{12}!}\ \frac{r_2!}{x_{21}!\,x_{22}!}p_{\cdot 1}{}^{c_1}p_{\cdot 2}{}^{c_2} \tag{7}$$

となる。ここで，$p_{\cdot 1}(=p_{11}+p_{21})$，$p_{\cdot 2}(=p_{12}+p_{22})$ はそれぞれ，B_1，B_2 におちる母集団確率であるから，（7）は（5）に一致する。

（c）　第3の場合は，すべての周辺度数 r_1，r_2，c_1，c_2 が固定されている場合である。このような場合の例としては，フィッシャーの紅茶のテストがある。この場合，A は1杯の紅茶が紅茶を先に入れたものであるか，ミルクを先に入れたものであるかという事実の分類であり，B は例の女性が，紅茶を先に入れたものであるといったか，ミルクを先に入れたものであるといったかという証言についての分類である。彼女には，それぞれの場合の数は知らされているから，当然彼女は $c_1=r_1$，$c_2=r_2$ となるように証言するであろう。したがって，この実験のくりかえしにおいては，周辺度数は全て固定されていると考えてよいであろう。この場合，A，B の分割が独立であると仮定すれば，表1の分割表をうる確率は（3）すなわち

$$p_0(x_{11},\ x_{12},\ x_{21},\ x_{22})=\frac{n!}{x_{11}!\,x_{12}!\,x_{21}!\,x_{22}!}p_{1\cdot}{}^{r_1}p_{2\cdot}{}^{r_2}p_{\cdot 1}{}^{c_1}p_{\cdot 2}{}^{c_2}$$

となる。ここで標本抽出のくりかえしにおいて周辺度数が r_1，r_2；c_1，c_2 である標本をうる確率は，A，B が独立であることから，

$$p_0(r_1,\ r_2;\ c_1,\ c_2)=p_0(r_1,\ r_2).p_0(c_1,\ c_2)$$

$$=\frac{n!}{r_1!\,r_2!}\ \frac{n!}{c_1!c_2!}p_{1\cdot}{}^{r_1}p_{2\cdot}{}^{r_2}p_{\cdot 1}{}^{c_1}p_{\cdot 2}{}^{c_2} \tag{8}$$

となる。したがって，周辺度数があたえられたときの，x_{11}，x_{12}，x_{21}，x_{22} をうる条件つき確率を，$p(x_{11},\ x_{12},\ x_{21},\ x_{22}\mid r_1,\ r_2;\ c_1,\ c_2)$ とかくことにすれば，一般に

$$p(x_{11},\ x_{12},\ x_{21},\ x_{22})=p(x_{11},\ x_{12},\ x_{21},\ x_{22}\mid c_1,\ r_2;\ c_1,\ c_2)p(r_1,\ r_2;\ c_1,\ c_2)$$

がなりたつから，仮説 H_0 のもとでは

$$p_0(x_{11},\ x_{12},\ x_{21},\ x_{22}\mid r_1,\ r_2;\ c_1,\ c_2)=\frac{n!}{x_{11}!\,x_{12}!\,x_{21}!\,x_{22}!}p_{1\cdot}{}^{r_1}p_{2\cdot}{}^{r_2}p_{\cdot 1}{}^{c_1}p_{\cdot 2}{}^{c_2}$$

$$\Big/\frac{n!}{r_1!\,r_2!}\ \frac{n!}{c_1!\,c_2!}p_{1\cdot}{}^{r_1}p_{2\cdot}{}^{r_2}p_{\cdot 1}{}^{c_1}p_{\cdot 2}{}^{c_2}$$

$$=\frac{r_1!\,r_2!\,c_1!\,c_2!}{n!\,x_{11}!\,x_{12}!\,x_{21}!\,x_{22}!} \qquad (9)$$

をうる。ここで，確率 $p_{i\cdot}$，$p_{\cdot j}$ は分母，分子で約分されてしまうから，未知パラメーターはなくなり，x_{ij} は

$x_{11}+x_{12}=r_1$，$x_{21}+x_{22}=r_2$

$x_{11}+x_{21}=c_1$，$x_{12}+x_{22}=c_2$

（ただし，$r_1+r_2=c_1+c_2=n$）を満足する。容易にわかるようにこの中で一次独立な条件は3個である。すなわち，x_{ij} は一次独立な3つの一次の条件式を満足しなければならない。したがって，x_{ij} のうちの1つが定まれば，他はそれに応じて決定されることとなる。

（a）および（b）の場合には未知パラメーターが存在することは前に述べた通りである。未知パラメーターの問題については，（a），（b）いづれも同様であるから，（a）の場合について述べる。（3）の $p_0(x_{11},\ x_{12},\ x_{21},\ x_{22})$ にふくまれるパラメーターは母集団周辺確率 $p_{i\cdot}$，$p_{\cdot j}$ である。この値は当然未知であるから，$p_0(x_{11},\ x_{12},\ x_{21},\ x_{22})$ はもとめることができない。しかしながら，どのような状況のもとでもそうであるかといえば，必ずしもそうとは限らない。n が大きくなって行けば，標本周辺比率 r_i/n，c_j/n は母集団周辺確率 $p_{i\cdot}$，$p_{\cdot j}$ に確率収束するという事実がある。このことから，n が大なるときには，$p_{i\cdot}$，$p_{\cdot j}$ は r_i/n，c_j/n によって充分よい精度で推定でき，したがって，$p_0(x_{11},\ x_{12},\ x_{21},\ x_{22})$ もまた推定できるのである。当然 n が大きい程推定の精度は良くなり，n が無限大になったとき，一意に決定するわけである。充分大きな標本数 n_0 の場合の $r_i^0/n_0, c_j^0/n_0$ を $p_{i\cdot}$，$p_{\cdot j}$ に近いものとし，比 r_i/n，c_j/n をこの値に固定して，n を無限大にとった極限値は，上の値に充分近いものになるであろう。ところが一方，比 r_i/n，c_j/n を固定するのであれば，これを未知パラメーターとするのでなく，周辺度数の比をこの値に固定した標本のみに着目して，n についての極限をとること，すなわち，（c）の場合の条件つき確率 $p(x_{11},\ x_{12},\ x_{21},\ x_{22}\mid r_1,\ r_2;\ c_1,\ c_2)$ の極限をとることと同じことになる。このように，大標本の場合には，（a）の場合も結局（c）の場合として取り扱つかえ

ば充分であることがわかる。

2. カイ自乗分布

(c)の場合の極限確率分布を計算することを考える。x_{ij}はそのうちの1つが定まれば，他はすべて定まってしまう。すなわち，独立変数は1個しかない。いまこれを x_{11} であると考えることとし，$p_0(x_{11}, x_{12}, x_{21}, x_{22} \mid r_1, r_2; c_1, c_2)$ を $p_0(x_{11} \mid r_1, r_2, c_1, c_2)$ とかくこととする。$n\to\infty$，$x_{11}\to\infty$の場合は，スターリングの公式が利用できる，すなわち，

$$x! \sim \sqrt{2\pi}\, x^{x+\frac{1}{2}} e^{-x}$$

をもちいて，(9)を変形し両辺の対数をとると

$$\begin{aligned}
&\log p_0(x_{11} \mid r_1, r_2, c_1, c_2)\\
&\quad=\log\sqrt{2\pi}+\left(r_1+\frac{1}{2}\right)\log r_1-r_1+\log\sqrt{2\pi}+\left(r_2+\frac{1}{2}\right)\log r_2-r_2\\
&\quad+\log\sqrt{2\pi}+\left(c_1+\frac{1}{2}\right)\log c_1-c_1+\log\sqrt{2\pi}+\left(c_2+\frac{1}{2}\right)\log c_2-c_2\\
&\quad-\log\sqrt{2\pi}-\left(n+\frac{1}{2}\right)\log n+n\\
&\quad-\log\sqrt{2\pi}-\left(x_{11}+\frac{1}{2}\right)\log x_{11}+x_{11}-\log\sqrt{2\pi}-\left(x_{12}+\frac{1}{2}\right)\log x_{12}+x_{12}\\
&\quad-\log\sqrt{2\pi}-\left(x_{21}+\frac{1}{2}\right)\log x_{21}+x_{21}-\log\sqrt{2\pi}-\left(x_{22}+\frac{1}{2}\right))\log x_{22}+x_{22}\\
&=-\log\sqrt{2\pi}+\left(x_{11}+x_{12}+\frac{1}{2}\right)\log r_1+\left(x_{21}+x_{22}+\frac{1}{2}\right)\log r_2\\
&\qquad+\left(x_{11}+x_{21}+\frac{1}{2}\right)\log c_1+\left(x_{12}+x_{22}+\frac{1}{2}\right)\log c_2\\
&\qquad-\left(x_{11}+x_{12}+x_{21}+x_{22}+\frac{1}{2}\right)\log n\\
&\qquad-\left(x_{11}+\frac{1}{2}\right)\log x_{11}-\left(x_{12}+\frac{1}{2}\right)\log x_{12}\\
&\qquad-\left(x_{21}+\frac{1}{2}\right)\log x_{21}-\left(x_{22}+\frac{1}{2}\right)\log x_{22}
\end{aligned}$$

$$(ここで,\ \log x_{11}=\log\left(x_{11}-\frac{r_1c_1}{n}+\frac{r_1c_1}{n}\right)=\log\frac{r_1c_1}{n}\left(1+\frac{x_{11}-\frac{r_1c_1}{n}}{\frac{r_1c_1}{n}}\right)$$

$$=\log r_1+\log c_1-\log n+\log\left(1+\frac{x_{11}-\frac{r_1c_1}{n}}{\frac{r_1c_1}{n}}\right)$$

$$=\log r_1+\log c_1-\log n+\frac{\left(x_{11}-\frac{r_1c_1}{n}\right)}{\frac{r_1c_1}{n}}-\frac{1}{2}\frac{\left(x_{11}-\frac{r_1c_1}{n}\right)^2}{\left(\frac{r_1c_1}{n}\right)^2}+\cdots$$

となるから,

$\xi=x_{11}-\frac{r_1c_1}{n}$ とおく,

$$\left(x_{12}-\frac{r_1c_2}{n}\right)+\left(x_{11}-\frac{r_1c_1}{n}\right)=x_{11}+x_{12}-\frac{r_1}{n}(c_1+c_2)$$

$$=r_1-r_1\frac{n}{n}=r_1-r_1=0$$

したがって,

$$\left(x_{12}-\frac{r_1c_2}{n}\right)=-\xi$$

同様にして,

$$x_{21}-\frac{r_2c_1}{n}=-\xi,\qquad x_{22}-\frac{r_2c_2}{n}=\xi$$

をうる。ここで, ξ の平均および分散をもとめる。

$$E(\xi)=E\left(x_{11}-\frac{r_1c_1}{n}\right)$$

$$=\sum\left(x_{11}-\frac{r_1c_1}{n}\right)p(x_{11}\mid r_1,\ r_2,\ c_1,\ c_2)$$

$$=\sum x_{11}\cdot p(x_{11}\mid r_1,\ r_2,\ c_1,\ c_2)-\frac{r_1c_1}{n},$$

$$\sum x_{11}p(x_{11}\mid r_1,\ r_2,\ c_1,\ c_2)$$

$$=\sum x_{11}\frac{r_1!\,r_2!\,c_1!\,c_2!}{n!\,x_{11}!\,x_{12}!\,x_{21}!\,x_{22}!}$$

$$=\left(\Sigma x_{11}\frac{r_1!}{x_{11}!\,x_{12}!}\ \frac{r_2!}{x_{21}!\,x_{22}!}\right)\frac{c_1!\,c_2!}{n!}$$

$$=\left(r_1\Sigma\frac{(r_1-1)!}{(x_{11}-1)!x_{12}!}\ \frac{r_2!}{x_{21}!x_{22}!}\right)^{(注)}\frac{c_1!\,c_2!}{n!}$$

$$=r_1\frac{(n-1)!}{(c_1-1)!\,c_2!}\cdot\frac{c_1!\,c_2!}{n!}$$

$$=\frac{r_1c_1}{n}$$

したがって,

$$E(\xi)=0$$

つぎに,

$$E(\xi^2)=E(x_{11}^2)-\left(\frac{r_1c_1}{n}\right)^2=E[x_{11}(x_{11}-1)]+E(x_{11})-\left(\frac{r_1c_1}{n}\right)^2$$

がなりたつから，まず $E[x(x-1)]$ をもとめる。

$$E[x_{11}(x_{11}-1)]=\left\{\Sigma x_{11}(x_{11}-1)\frac{r_1!}{x_{11}!\,x_{12}!}\ \frac{r_2!}{r_{21}!\,x_{22}!}\right\}\frac{c_1!\,c_2!}{n!}$$

$$=\left\{r_1(r_1-1)\Sigma\frac{(r_1-2)!}{(x_{11}-2)!\,x_{12}!}\ \frac{r_2!}{r_{21}!\,x_{22}!}\right\}\frac{c_1!\,c_2!}{n!}$$

$$=r_1(r_1-1)\frac{(n-2)!}{(c_1-2)!\,c_2!}\cdot\frac{c_1!\,c_2!}{n!}$$

$$=\frac{r_1(r_1-1)c_1(c_1-1)}{n(n-1)}$$

これを，上式に代入して

$$\sigma^2=E(\xi^2)=\frac{r_1(r_1-1)c_1(c_1-1)}{n(n-1)}+\frac{r_1c_1}{n}-\frac{r_1^2c_1^2}{n^2}$$

$$=\frac{r_1c_1}{n^2(n-1)}\{n(r_1-1)(c_1-1)+n(n-1)-r_1c_1(n-1)\}$$

注) $n=r_1+r_2$ がなりたつことから,

$$(p+q)^n=(p+q)^{r_1}(p+q)^{r_2}$$

をうる。両辺の $p^{c_1}q^{c_2}$ の係数を比較して,

$$\frac{n!}{c_1!\,c_2!}=\sum_{\substack{x_{11}+x_{21}=c_1\\x_{12}+x_{22}=c_2}}\frac{r_1!}{x_{11}!\,x_{12}!}\ \frac{r_2!}{x_{21}!\,x_{22}!}$$

をうる。

$$=\frac{r_1c_1(n-r_1)(n-c_1)}{n^2(n-1)}=\frac{r_1r_2c_1c_2}{n^2(n-1)}$$

をうる。n が充分大なる場合

$$\sigma^2=\frac{r_1r_2c_1c_2}{n^3}$$

としてよいから，今後この値を用いる。)

$$\begin{aligned}
&=-\log\sqrt{2\pi}+\left(x_{11}+x_{12}+\frac{1}{2}\right)\log r_1+\left(x_{21}+x_{22}+\frac{1}{2}\right)\log r_2\\
&\quad+\left(x_{11}+x_{21}+\frac{1}{2}\right)\log c_1+\left(x_{12}+x_{22}+\frac{1}{2}\right)\log c_2\\
&\quad-\left(x_{11}+x_{12}+x_{21}+x_{22}+\frac{1}{2}\right)\log n\\
&\quad-\left(x_{11}+\frac{1}{2}\right)\left(\log r_1+\log c_1-\log n+\frac{n}{r_1c_1}\xi-\frac{1}{2}\left(\frac{n}{r_1c_1}\right)^2\xi^2+\cdots\right)\\
&\quad-\left(x_{12}+\frac{1}{2}\right)\left(\log r_1+\log c_2-\log n-\frac{n}{r_1c_2}\xi-\frac{1}{2}\left(\frac{n}{r_1c_2}\right)^2\xi^2+\cdots\right)\\
&\quad-\left(x_{21}+\frac{1}{2}\right)\left(\log r_2+\log c_1-\log n-\frac{n}{r_2c_1}\xi-\frac{1}{2}\left(\frac{n}{r_2c_1}\right)^2\xi^2+\cdots\right)\\
&\quad-\left(x_{22}+\frac{1}{2}\right)\left(\log r_2+\log c_2-\log n+\frac{n}{r_2c_2}\xi-\frac{1}{2}\left(\frac{n}{r_2c_2}\right)^2\xi^2+\cdots\right)
\end{aligned}$$

$$\begin{aligned}
&=-\log\sqrt{2\pi}-\frac{1}{2}(\log r_1+\log r_2+\log c_1+\log c_2-3\log n)\\
&\quad-\left(x_{11}+\frac{1}{2}\right)\left\{\frac{n}{r_1c_1}\xi-\frac{1}{2}\left(\frac{n}{r_1c_1}\right)^2\xi^2+\cdots\right\}-\left(x_{12}+\frac{1}{2}\right)\left\{-\frac{n}{r_1c_2}\xi-\frac{1}{2}\left(\frac{n}{r_1c_2}\right)^2\xi^2+\cdots\right\}\\
&\quad-\left(x_{21}+\frac{1}{2}\right)\left\{-\frac{n}{r_2c_1}\xi-\frac{1}{2}\left(\frac{n}{r_2c_1}\right)^2\xi^2+\cdots\right\}-\left(x_{22}+\frac{1}{2}\right)\left\{\frac{n}{r_2c_2}\xi-\frac{1}{2}\left(\frac{n}{r_2c_2}\right)^2\xi^2+\cdots\right\}
\end{aligned}$$

(ここで，$x_{11}=\xi+\frac{r_1c_1}{n}$，$x_{12}=-\xi+\frac{r_1c_2}{n}$，$x_{21}=-\xi+\frac{r_2c_1}{n}$，$x_{22}=\xi+\frac{r_2c_2}{n}$ を代入する)

$$\begin{aligned}
&=-\log\sqrt{2\pi}-\frac{1}{2}\log\frac{r_1r_2c_1c_2}{n^3}\\
&\quad-\left(\xi+\frac{r_1c_1}{n}+\frac{1}{2}\right)\left\{\frac{n}{r_1c_1}\xi-\frac{1}{2}\left(\frac{n}{r_1c_1}\right)^2\xi^2+\cdots\right\}
\end{aligned}$$

$$-\left(-\xi+\frac{r_1c_2}{n}+\frac{1}{2}\right)\left\{-\frac{n}{r_1c_2}\xi-\frac{1}{2}\left(\frac{n}{r_1c_2}\right)^2\xi^2+\cdots\right\}$$

$$-\left(-\xi+\frac{r_2c_1}{n}+\frac{1}{2}\right)\left\{-\frac{n}{r_2c_1}\xi-\frac{1}{2}\left(\frac{n}{r_2c_1}\right)^2\xi^2+\cdots\right\}$$

$$-\left(\xi+\frac{r_2c_2}{n}+\frac{1}{2}\right)\left\{\frac{n}{r_2c_2}\xi-\frac{1}{2}\left(\frac{n}{r_2c_2}\right)^2\xi^2+\cdots\right\}$$

$$=-\log\sqrt{2\pi}-\frac{1}{2}\log\frac{r_1r_2c_1c_2}{n^3}$$

$$-\frac{1}{2}\xi\left(\frac{n}{r_1c_1}-\frac{n}{r_1c_2}-\frac{n}{r_2c_1}+\frac{n}{r_2c_2}\right)-\xi(1-1-1+1)$$

$$-\xi^2\left(\frac{n}{r_1c_1}+\frac{n}{r_1c_2}+\frac{n}{r_2c_1}+\frac{n}{r_2c_2}\right)+\frac{1}{4}\xi^2\left\{\left(\frac{n}{r_1c_1}\right)^2+\left(\frac{n}{r_1c_2}\right)^2+\left(\frac{n}{r_2c_1}\right)^2+\left(\frac{n}{r_2c_2}\right)^2\right\}$$

$$+\frac{1}{2}\xi^2\left(\frac{n}{r_1c_1}+\frac{n}{r_1c_2}+\frac{n}{r_2c_1}+\frac{n}{r_2c_2}\right)+\cdots$$

$$=-\log\sqrt{2\pi}-\frac{1}{2}\log\frac{r_1r_2c_1c_2}{n^3}-\frac{1}{2}\xi\frac{n^3}{r_1r_2c_1c_2}\cdot\frac{(r_1-r_2)(c_1-c_2)}{n^2}$$

$$-\frac{1}{2}\xi^2\frac{n^3}{r_1r_2c_1c_2}+\frac{1}{4}\xi^2\frac{n^6}{(r_1r_2c_1c_2)^2}\ \frac{(r_1^2+r_2^2)(c_1^2+c_2^2)}{n^4}+\cdots$$

（ここで，$\frac{r_1}{n}$，$\frac{r_2}{n}$，$\frac{c_1}{n}$，$\frac{c_2}{n}$ を固定して n を大きくするとき，ξ/σ は有限の値をとり，$\sigma=\sqrt{\frac{r_1}{n}\ \frac{r_2}{n}\ \frac{c_1}{n}\cdot c_2}$ はいくらでも大きくなることから，$\xi\cdot\frac{n^3}{r_1r_2c_1c_2}\cdot\frac{r_1-r_2}{n}\ \frac{c_1-c_2}{n}=\frac{\xi}{\sigma^2}\left(\frac{r_1}{n}-\frac{r_2}{n}\right)\left(\frac{c_1}{n}-\frac{c_2}{n}\right)=\frac{1}{\sigma}\ \frac{\xi}{\sigma}\left(\frac{r_1}{n}-\frac{r_2}{n}\right)\left(\frac{c_1}{n}-\frac{c_2}{n}\right)$ は0に近づく，同様にして，第5項は

$$\xi^2\cdot\frac{n^6}{(r_1r_2c_1c_2)^2}\left(\frac{r_1^2+r_2^2}{n^2}\right)\left(\frac{c_1^2+c_2^2}{n^2}\right)=\frac{\xi^2}{\sigma^2}\ \frac{1}{\sigma^2}\cdot\left(\frac{r_1^2}{n^2}+\frac{r_2^2}{n^2}\right)\left(\frac{c_1^2}{n^2}+\frac{c_2^2}{n^2}\right)$$

となるから0となる。以下の項もすべて同様であるから）

$$=-\log\sqrt{2\pi}-\log\sigma-\frac{1}{2}\frac{\xi^2}{\sigma^2} \tag{10}$$

したがって，

$$p_0(x_{11}\mid r_1,\ r_2,\ c_1,\ c_2)=\frac{1}{\sqrt{2\pi}\,\sigma}e^{-\frac{\xi^2}{2\sigma^2}} \tag{11}$$

をうる。同一の ξ^2 をあたえる x_{11} は $\frac{r_1c_1}{n}$ に対して対称な2つがあるから，$\frac{\xi^2}{\sigma^2}$ の分布を考えるためには

$$\frac{2}{\sqrt{2\pi}\,\sigma}e^{-\frac{\xi^2}{2\sigma^2}}dx_{11}=\frac{2}{\sqrt{2\pi}\,\sigma}e^{-\frac{\xi^2}{2\sigma^2}}d\xi=\frac{1}{\sqrt{2\pi}}e^{-\frac{\xi^2}{2\sigma^2}}d\frac{\xi}{\sigma}$$

$$=\frac{1}{\sqrt{2\pi}}e^{-\frac{\chi^2}{2}}2\,d\chi=\frac{1}{\sqrt{2\pi}}e^{-\frac{\chi^2}{2}}d\chi^2 \qquad (12)$$

これは自由度1の χ^2−分布である。このことから，n が充分大きいとき

$$\chi^2=\frac{\xi^2}{\sigma^2}=\xi^2\frac{n^3}{r_1r_2c_1c_2}=n\xi^2\frac{(r_1+r_2)(c_1+c_2)}{r_1r_2c_1c_2}=n\xi^2\frac{r_1c_1+r_1c_2+r_2c_1+r_2c_2}{r_1r_2c_1c_2}$$

$$=n\xi^2\left(\frac{1}{r_1c_1}+\frac{1}{r_2c_2}+\frac{1}{r_1c_2}+\frac{1}{r_2c_2}\right)$$

$$=\frac{\xi^2}{\frac{r_1c_1}{n}}+\frac{\xi^2}{\frac{r_1c_2}{n}}+\frac{\xi^2}{\frac{r_2c_1}{n}}+\frac{\xi^2}{\frac{r_2c_2}{n}}$$

$$=\frac{\left(x_{11}-\frac{r_1c_1}{n}\right)^2}{\frac{r_1c_1}{n}}+\frac{\left(x_{12}-\frac{r_1c_2}{n}\right)^2}{\frac{r_1c_2}{n}}+\frac{\left(x_{21}-\frac{r_2c_1}{n}\right)^2}{\frac{r_2c_1}{n}}+\frac{\left(x_{22}-\frac{r_2c_2}{n}\right)^2}{\frac{r_2c_2}{n}}$$

が自由度1の χ^2− 分布にしたがうことがわかる。

さきに，$n\to\infty$ とすれば，(a)の場合と(c)の場合が一致することを述べた。これをみるために，$p(x_{11},\ x_{12},\ x_{21},\ x_{22})$ の極限分布をもとめる。

$$p(x_{11},\ x_{12},\ x_{21},\ x_{22})=\frac{n!}{x_{11}!\,x_{12}!\,x_{21}!\,x_{22}!}p_{11}{}^{x_{11}}p_{12}{}^{x_{12}}p_{21}{}^{x_{21}}p_{22}{}^{x_{22}}$$

$$\log p(x_{11},\ x_{12},\ x_{21},\ x_{22})=\log n!-\log x_{11}!-\log x_{12}!-\log x_{21}!-\log x_{22}!$$
$$+x_{11}\log p_{11}+x_{12}\log p_{12}+x_{21}\log p_{21}+x_{22}\log p_{22}$$
$$=\log\sqrt{2\pi}+\left(n+\frac{1}{2}\right)\log n-n$$
$$-\log\sqrt{2\pi}-\left(x_{11}+\frac{1}{2}\right)\log x_{11}+x_{11}-\log\sqrt{2\pi}-\left(x_{12}+\frac{1}{2}\right)\log x_{12}+x_{12}$$
$$-\log\sqrt{2\pi}-\left(x_{21}+\frac{1}{2}\right)\log x_{21}+x_{21}-\log\sqrt{2\pi}-\left(x_{22}+\frac{1}{2}\right)\log x_{22}+x_{22}$$
$$+x_{11}\log p_{11}+x_{12}\log p_{12}+x_{21}\log p_{21}+x_{22}\log p_{22}$$

$$=-x_{11}(\log x_{11}-\log n-\log p_{11})-x_{12}(\log x_{12}-\log n-\log p_{12})$$
$$-x_{21}(\log x_{21}-\log n-\log p_{21})-x_{22}(\log x_{22}-\log n-\log p_{22})$$
$$+C'$$

（ここで，$x_{ij}-np_{ij}=u_{ij}$，$np_{ij}=n_{ij}$ とすれば，

$$\frac{x_{ij}}{np_{ij}}=\frac{n_{ij}+(x_{ij}-n_{ij})}{n_{ij}}=1+\frac{u_{ij}}{n_{ij}}$$

となるから，これを代入する）

$$=-\sum_{i,j}(n_{ij}+u_{ij})\log\left(1+\frac{u_{ij}}{n_{ij}}\right)+C'$$

$$=-\sum_{i,j}(n_{ij}+u_{ij})\left(\frac{u_{ij}}{n_{ij}}-\frac{1}{2}\frac{u_{ij}^2}{n_{ij}^2}+\cdots\right)+C'$$

$$=-\sum_{i,j}u_{ij}-\frac{1}{2}\sum_{i,j}\frac{u_{ij}^2}{n_{ij}}+\cdots+C'$$

（ここで，$\sum u_{ij}=\sum(x_{ij}-np_{ij})=\sum x_{ij}-n\sum p_{ij}=n-n=0$ であるから）

$$=-\frac{1}{2}\sum\frac{u_{ij}^2}{n_{ij}}+C' \qquad (13)$$

したがって，

$$p(x_{11},\ x_{12},\ x_{21},\ x_{22})=Ce^{-\frac{1}{2}\sum\frac{u_{ij}^2}{n_{ij}}} \qquad (14)$$

をうる。ここで，仮説 H_0 を考えれば，

$$\sum\frac{u_{ij}^2}{n_{ij}}=\sum\frac{(x_{ij}-np_{ij})^2}{np_{ij}}=\sum\frac{(x_{ij}-np_{i\cdot}p_{\cdot j})^2}{np_{i\cdot}p_{\cdot j}}$$

（n が充分大なるとき，$p_{i\cdot}\fallingdotseq\frac{r_i}{n}$，$p_{\cdot j}\fallingdotseq\frac{c_j}{n}$ となることから，これを代入すれば，）

$$=\sum\frac{\left(x_{ij}-\frac{r_ic_j}{n}\right)^2}{\frac{r_ic_j}{n}}=n\sum\frac{\left(x_{ij}-\frac{r_ic_j}{n}\right)^2}{r_ic_j}$$

（ここで，$x_{11}-\frac{r_1c_1}{n}=\xi$ とすれば，前に示した通り，ξ の分散は $\sigma^2=\frac{n^3}{r_1r_2c_1c_2}$ となり，）

$$=\xi^2\left(n\sum\frac{1}{r_ic_j}\right)=n\xi^2\left(\frac{1}{r_1c_1}+\frac{1}{r_1c_2}+\frac{1}{r_2c_1}+\frac{1}{r_2c_2}\right)$$

$$=n\xi^2\frac{(r_1+r_2)(c_1+c_2)}{r_1c_1c_1c_2}=\xi^2\frac{n^3}{r_1r_2c_1c_2}=\frac{\xi^2}{\sigma^2}$$

をうる。したがって

$$p_0(x_{11},\ x_{12},\ x_{21},\ x_{22})=C_0e^{-\frac{\xi^2}{2\sigma^2}} \tag{15}$$

がなりたつ，C_0 は，$\int p_0 d\xi^2=1$ から決定され，前に示した通り，自由度１の χ^2－分布となる。

3. $r\times s$ 分割表

まづ表２に示されるような２×３分割表について考える。表２の２×３分割表をうる確率は

$$p(x_{ij}\mid n)=\frac{n!}{x_{11}!\cdots x_{23}!}p_{11}{}^{x_{11}}\cdots p_{23}{}^{x_{23}} \tag{16}$$

であたえられる。独立性の仮説

表　２

	B_1	B_2	B_3	計
A_1	x_{11}	x_{12}	x_{13}	$t_{1\cdot}$
A_2	x_{21}	x_{22}	x_{23}	$t_{2\cdot}$
計	$t_{\cdot 1}$	$t_{\cdot 2}$	$t_{\cdot 3}$	n

$$H_0:\ p_{ij}=p_{i\cdot}\cdot p_{\cdot j}\qquad i=1,\ 2;\ j=1,\ 2,\ 3 \tag{17}$$

のもとでは，この確率は

$$p_0(x_{ij}\mid n)=\frac{n!}{\prod_{ij} x_{ij}!}p_{1\cdot}{}^{t_{1\cdot}}p_{2\cdot}{}^{t_{2\cdot}}p_{3\cdot}{}^{t_{3\cdot}}p_{\cdot 1}{}^{t_{\cdot 1}}p_{\cdot 2}{}^{t_{\cdot 2}} \tag{18}$$

となる。周辺度数 $t_{1\cdot}$，$t_{2\cdot}$，$t_{\cdot 1}$，$t_{\cdot 2}$，$t_{\cdot 3}$ を与えたときの条件付き確率は，前と同じ様にして，

$$p_0(x_{ij}\mid t_{i\cdot},\ t_{\cdot j})=\frac{n!}{\prod_{ij} x_{ij}!}\cdot\frac{t_{1\cdot}!\,t_{2\cdot}!}{n!}\ \frac{t_{\cdot 1}!\,t_{\cdot 2}!\,t_{\cdot 3}!}{n!}$$

$$=\frac{t_{1\cdot}!\,t_{2\cdot}!\,t_{\cdot 1}!\,t_{\cdot 2}!\,t_{\cdot 3}!}{n!\,x_{11}!\,x_{12}!\,x_{13}!\,x_{21}!\,x_{22}!\,x_{23}!} \tag{19}$$

ここで，

$$x_{11}+x_{12}=R_{12},\quad x_{21}+x_{22}=R_{22},\quad t_{\cdot 1}+t_{\cdot 2}=T_{22}=R_{12}+R_{22}$$

をもちいて補助表表３をつくる。またこの量

表　３

x_{11}	x_{12}	R_{12}
x_{21}	x_{22}	R_{22}
$t_{\cdot 1}$	$t_{\cdot 2}$	T_{22}

R_{12}	x_{13}	$t_{1\cdot}$
R_{22}	x_{23}	$t_{2\cdot}$
T_{22}	$t_{\cdot 3}$	n

をもちいて，(19)式を

$$p_0(x_{ij}\mid t_{i\cdot},\ t_{\cdot j})=\frac{R_{12}!\,R_{22}!\,t_{\cdot 1}!\,t_{\cdot 2}!}{x_{11}!\,x_{12}!\,x_{21}!\,x_{22}!\,T_{22}!}\cdot\frac{t_{1\cdot}!\,t_{2\cdot}!\,T_{22}!\,t_{\cdot 3}!}{R_{12}!\,x_{13}!\,R_{22}!\,x_{23}!\,n!}$$

とかく，この右辺の2つの確率は，表3に示した補助表のおのおのにおける周辺度数が与えられたときの条件つきの確率である。補助表のおのおのは 2×2 分割表であるから，$n\to\infty$，$x_{ij}\to\infty$ のとき自由度1の χ^2— 分布にしたがう統計量がえられることは，第2節に示した通りである。また(20) において，$p_0(x_{ij}\mid t_{i\cdot},\ t_{\cdot j})$ が2つの確率の積に分解されていることは，この統計量が極限において独立であることを示すものである。したがってそれらの和は，χ^2— 分布の性質から，自由度2の χ^2— 分布にしたがう。

おのおのの補助表における χ^2— 分布にしたがう統計量は，

$$\chi_1^2=\frac{\left(x_{11}-\frac{t_{\cdot 1}R_{12}}{T_{22}}\right)^2}{t_{\cdot 1}\frac{R_{12}}{T_{22}}}+\frac{\left(x_{12}-\frac{t_{\cdot 2}R_{12}}{T_{22}}\right)^2}{\frac{t_{\cdot 2}R_{12}}{T_{22}}}+\frac{\left(x_{21}-\frac{t_{\cdot 1}R_{22}}{T_{22}}\right)^2}{\frac{t_{\cdot 1}R_{22}}{T_{22}}}+\frac{\left(x_{22}-\frac{t_{\cdot 2}R_{22}}{T_{22}}\right)^2}{\frac{t_{\cdot 2}R_{22}}{T_{22}}}\tag{21}$$

$$\chi_2^2=\frac{\left(R_{12}-\frac{T_{22}t_{1\cdot}}{n}\right)^2}{\frac{T_{22}t_{1\cdot}}{n}}+\frac{\left(x_{13}-\frac{t_{\cdot 3}t_{1\cdot}}{n}\right)^2}{\frac{t_{\cdot 3}t_{1\cdot}}{n}}+\frac{\left(R_{22}-\frac{T_{22}t_{2\cdot}}{n}\right)^2}{\frac{T_{22}t_{2\cdot}}{n}}+\frac{\left(x_{23}-\frac{t_{\cdot 3}t_{2\cdot}}{n}\right)^2}{\frac{t_{\cdot 3}t_{2\cdot}}{n}}\tag{22}$$

である。この和が自由度2の χ^2— 分布にしたがうわけである。いま $\chi_1^2+\chi_2^2$ をもとめるために，(21)を変形する。ここで，

$$n_{11}'=\frac{t_{\cdot 1}R_{12}}{T_{22}},\quad n_{12}'=\frac{t_{\cdot 2}R_{12}}{T_{22}},\quad n_{21}'=\frac{t_{\cdot 1}R_{22}}{T_{22}},\quad n_{22}'=\frac{t_{\cdot 2}R_{22}}{T_{22}}$$

とかくこととする。

$$\chi_1^2=\sum_{i,j}\frac{(x_{ij}-n_{ij}')^2}{n'_{ij}}=\sum\frac{x_{ij}^2-2x_{ij}n'_{ij}+n'^2_{ij}}{n'_{ij}}$$

$$=\frac{x_{11}^2}{n_{11}'}-2x_{11}+n_{11}'+\frac{x_{12}^2}{n_{12}'}-2x_{12}+n_{12}'+\frac{x_{21}^2}{n_{21}'}-2x_{21}+n_{21}'+\frac{x_{22}^2}{n_{22}'}-2x_{22}+n_{22}'$$

$$=\frac{x_{11}^2}{n_{11}'}+\frac{x_{12}^2}{n_{12}'}-R_{12}+\frac{x_{21}^2}{n_{21}'}+\frac{x_{22}^2}{n_{22}'}-R_{22}\tag{23}$$

同様にして

$$\chi_2^2=\frac{R_{12}^2}{\frac{T_{22}t_{1\cdot}}{n}}+\frac{x_{13}^2}{\frac{t_{\cdot 3}t_{1\cdot}}{n}}-t_{1\cdot}+\frac{R_{22}^2}{\frac{T_{22}t_{2\cdot}}{n}}+\frac{x_{23}^2}{\frac{t_{\cdot 3}t_{2\cdot}}{n}}-t_{2\cdot}\tag{24}$$

をうる。周辺度数が与えられた場合，$n\to\infty$ となるとき，$t_{i\cdot}/n$，$t_{\cdot j}/n$ は一定値であり，

$x_{ij} \to \infty$, $t_{i.} \to \infty$, $t_{.j} \to \infty$ となる。そして，例えば，x_{13} は $t_{1.}t_{.3}/n$ を中心に変動し，

$$\left(x_{13}-\frac{t_{1.}t_{.3}}{n}\right)^2 \Big/ \frac{t_{1.}t_{.3}}{n}$$

は χ^2— 分布にしたがう統計量の一つの項として有界である。したがって，

$$\left(x_{13}-\frac{t_{1.}t_{.3}}{n}\right) \Big/ \frac{t_{1.}t_{.3}}{n} \to 0$$

すなわち，

$$x_{13} \Big/ \frac{t_{1.}t_{.3}}{n} \to 1, \qquad x_{13} \doteqdot \frac{t_{1.}t_{.3}}{n}$$

同様にして，

$$R_{12} \doteqdot \frac{T_{22}t_{1.}}{n}, \qquad R_{22} \doteqdot \frac{T_{22}t_{2.}}{n}$$

がなりたつから，

$$n_{11}'=\frac{t_{.1}R_{12}}{T_{22}} \doteqdot \frac{t_{.1}t_{1.}}{n}, \qquad n_{12}'=\frac{t_{.2}R_{12}}{T_{22}} = \frac{t_{.2}t_{1.}}{n}$$

$$n_{21}'=\frac{t_{.1}R_{22}}{T_{22}} \doteqdot \frac{t_{.1}t_{2.}}{n}, \qquad n_{22}'=\frac{t_{.2}R_{22}}{T_{22}} \doteqdot \frac{t_{.2}t_{2.}}{n}$$

をうるから，

$$\chi_1^2+\chi_2^2=\frac{x_{11}^2}{n_{11}'}+\frac{x_{12}^2}{n_{12}'}-R_{12}+\frac{x_{21}^2}{n_{21}'}+\frac{x_{22}^2}{n_{22}'}-R_{22}$$

$$+\frac{R_{12}^2}{\frac{T_{22}t_{1.}}{n}}+\frac{x_{13}^2}{\frac{t_{1.}t_{.3}}{n}}-t_{1.}+\frac{R_{22}^2}{\frac{T_{22}t_{2.}}{n}}+\frac{x_{23}^2}{\frac{t_{.3}t_{1.}}{n}}-t_{2.}$$

$$\doteqdot\frac{x_{11}^2}{\frac{t_{1.}t_{.1}}{n}}+\frac{x_{12}^2}{\frac{t_{1.}t_{.2}}{n}}+\frac{x_{13}^2}{\frac{t_{1.}t_{.3}}{n}}-t_{1.}+\frac{x_{21}^2}{\frac{t_{2.}t_{.1}}{n}}+\frac{x_{22}^2}{\frac{t_{2.}t_{.2}}{n}}+\frac{x_{23}^2}{\frac{t_{2.}t_{.3}}{n}}-t_{2.}$$

$$=\sum_{i=1}^{2}\sum_{j=1}^{3}\frac{\left(x_{ij}-\frac{t_{i.}t_{.j}}{n}\right)^2}{\frac{t_{i.}t_{.j}}{n}} \tag{30}$$

が自由度２の χ^2— 分布に したがう 統計量である。

つぎに，一般の $r\times s$ 分割表について考える。前と全く同様にして，独立性の仮説

H_0: $p_{ij}=p_{i.}\cdot p_{.j}$　$i=1, 2, \cdots, r$; $j=1, 2, \cdots, s$

表　4

	B_1	B_2	……	B_s	計
A_1	x_{11}	x_{12}	……	x_{1s}	$t_{1\cdot}$
A_2	x_{21}	x_{22}	……	x_{2s}	$t_{2\cdot}$
…	……	……	……	……	…
A_r	x_{r1}	x_{r2}	……	x_{rs}	$t_{r\cdot}$
計	$t_{\cdot 1}$	$t_{\cdot 2}$	……	$t_{\cdot s}$	n

のもとで，周辺度数 $t_{\cdot 1}$, $t_{\cdot 2}$, …, $t_{\cdot s}$; $t_{1\cdot}$, $t_{2\cdot}$, …, $t_{r\cdot}$ を与えたときの条件付確率は

$$p_0(x_{ij} \mid t_{i\cdot},\ t_{\cdot j})=\frac{\prod_i t_{i\cdot}!\ \prod_j t_{\cdot j}!}{n!\ \prod_{ij} x_{ij}!} \tag{31}$$

つぎに，つぎの記号を導入する。

$$R_{il}=\sum_{j=1}^{l} x_{ij}\ ,\quad i=1,\ 2,\ \cdots,\ r;\ l=2,\ 3,\ \cdots,\ s \tag{32}$$

$$C_{mj}=\sum_{i=1}^{m} x_{ij}\ ,\quad j=1,\ 2,\ \cdots,\ s;\ m=2,\ 3,\ \cdots,\ r \tag{33}$$

$$T_{ml}=\sum_{i=1}^{m}\sum_{j=1}^{l} x_{ij}=\sum_{i=1}^{m} R_{il}=\sum_{j=1}^{l} C_{mj},\quad m=2,\ 3,\ \cdots,\ r;\ l=2,\ 3,\ \cdots,\ s \tag{34}$$

とくに，$R_{is}=t_{i\cdot}$, $C_{rj}=t_{\cdot j}$, $T_{rs}=n$ である。これをもちいてつぎの補助表をつくる。

表　5

x_{11}	x_{12}	R_{12}	R_{12}	x_{13}	R_{13}		$R_{1,s-1}$	x_{1s}	R_{1s}	$(=t_{1\cdot})$
x_{21}	x_{22}	R_{22}	R_{22}	x_{23}	R_{23}	……	$R_{2,s-1}$	x_{2s}	R_{2s}	$(=t_{2\cdot})$
C_{21}	C_{22}	T_{22}	T_{22}	C_{23}	T_{23}		$T_{2,s-1}$	C_{2s}	T_{2s}	
C_{21}	C_{22}	T_{22}	T_{22}	x_{23}	R_{23}		$T_{2,s-1}$	C_{2s}	T_{2s}	
x_{31}	x_{32}	R_{32}	R_{32}	x_{33}	R_{33}	……	$R_{3,s-1}$	x_{3s}	R_{3s}	$(=t_{3\cdot})$
C_{31}	C_{32}	T_{32}	T_{32}	C_{33}	T_{33}		$T_{3,s-1}$	C_{3s}	T_{3s}	
……			……				……			
$C_{r-1,1}$	$C_{r-1,2}$	$T_{r-1,2}$	$T_{r-1,2}$	$C_{r-1,3}$	$T_{r-1,3}$		$T_{r-1,s-1}$	$C_{r-1,s}$	$T_{r-1,s}$	
x_{r1}	x_{r2}	R_{r2}	R_{r2}	x_{r3}	R_{r3}	……	$R_{r,s-1}$	x_{rs}	R_{rs}	$(=t_{r\cdot})$
C_{r1}	C_{r2}	T_{r2}	T_{r2}	C_{r3}	T_{r3}		$T_{r,s-1}$	C_{rs}	T_{rs}	$(=n)$
$(=t_{\cdot 1})$	$(=t_{\cdot 2})$			$(=t_{\cdot 3})$				$(=t_{\cdot s})$		

表5の $(r-1)\times(s-1)$ 個のおのおのの補助表に対応する，2×2 分割表の周辺度数を与えたときの条件つき確率は

$$\frac{R_{12}!\,R_{22}!\,C_{21}!\,C_{22}!}{x_{11}!\,x_{12}!\,x_{21}!\,x_{22}!\,T_{22}!},\quad \frac{R_{13}!\,R_{23}!\,T_{22}!\,C_{23}!}{R_{12}!\,x_{13}!\,R_{22}!\,x_{23}!\,T_{23}!},\quad \cdots\cdots$$

$$\cdots,\ \frac{T_{r-1,s}!\,R_{rs}!\,T_{r,s-1}!\,C_{rs}!}{T_{r-1,s-1}!\,C_{r-1,s}!\,R_{r,s-1}!\,x_{rs}!\,T_{rs}!} \tag{35}$$

である。$p_0(x_{ij}\mid t_{i\cdot},\ t_{\cdot j})$ はこれらの積に分解できることは容易にわかる。すなわち,

$$p_0(x_{ij}\mid t_{i\cdot},\ t_{\cdot j})=\frac{\prod_i t_{i\cdot}!\prod_j t_{\cdot j}!}{\prod_{i,j} x_{ij}}$$

$$=\frac{R_{12}!\,R_{22}!\,C_{21}!\,C_{22}!}{x_{11}!\,x_{12}!\,x_{21}!\,x_{22}!\,T_{22}!}\cdot\frac{R_{13}!\,R_{23}!\,T_{22}!\,C_{23}!}{R_{12}!\,x_{13}!\,R_{22}!\,x_{23}!\,T_{23}!}\cdot\cdots\cdots$$

$$\cdots\cdots\cdot\frac{T_{r-1,s}!\,R_{rs}!\,T_{r,s-1}!\,C_{rs}!}{T_{r-1,s-1}!\,C_{r-1,s}!\,R_{r,s-1}!\,x_{rs}!\,T_{rs}!} \tag{36}$$

がなりたつ,ここで,$n\to\infty$, $x_{ij}\to 0$ とするとき,おのおのの補助表から,自由度1の χ^2― 分布にしたがう統計量 χ_i^2 がえられるから,それらの和 $\sum\chi_i^2$ は自由度 $(r-1)\cdot(s-1)$ の χ^2― 分布にしたがうことがわかる。$\sum\chi_i^2$ をもとめることを考える。

$$\sum\chi_i^2=\frac{\left(x_{11}-\frac{C_{21}R_{12}}{T_{22}}\right)^2}{\frac{C_{21}R_{12}}{T_{22}}}+\cdots+\frac{\left(x_{rs}-\frac{C_{rs}R_{rs}}{T_{rs}}\right)^2}{\frac{C_{rs}R_{rs}}{T_{rs}}}$$

$$=\frac{x_{11}^2}{\frac{C_{21}R_{12}}{T_{22}}}-2x_{11}+\frac{C_{21}R_{12}}{T_{22}}+\frac{x_{12}^2}{\frac{C_{22}R_{12}}{T_{22}}}-2x_{12}+\frac{C_{22}R_{12}}{T_{22}}+\frac{R_{12}^2}{\frac{T_{22}R_{13}}{T_{23}}}$$

$$+\cdots+\frac{x_{rs}^2}{\frac{C_{rs}R_{rs}}{T_{rs}}}-2x_{rs}+\frac{C_{rs}R_{rs}}{T_{rs}}$$

(ここで,補助表を表6としたとき,前に2×3分割表において述べた通り,

$$b_{ij}\to\frac{t_{i\cdot}t_{\cdot j}}{t}$$

となるから,

$$\frac{R_{12}^2}{\frac{T_{22}R_{13}}{T_{23}}}\doteqdot\frac{R_{12}^2}{R_{12}}=R_{12},\cdots\cdots$$

としてよい)

表　6

b_{11}	b_{12}	$t_{1\cdot}$
b_{21}	b_{22}	$t_{2\cdot}$
$t_{\cdot 1}$	$t_{\cdot 2}$	t

$$=\frac{x_{11}^2}{\frac{C_{21}R_{12}}{T_{22}}}+\frac{x_{12}^2}{\frac{C_{22}R_{12}}{T_{22}}}+\cdots\qquad+\frac{x_{1s}^2}{\frac{C_{2s}R_{1s}}{T_{2s}}}-R_{1s}$$

$$+\frac{x_{21}^2}{\frac{C_{21}R_{22}}{T_{22}}}+\frac{x_{22}^2}{\frac{C_{22}R_{22}}{T_{22}}}+\cdots+\frac{x_{2s}^2}{\frac{C_{2s}R_{2s}}{T_{2s}}}-R_{2s}$$

$$+\cdots$$

$$+\frac{x_{r1}^2}{\frac{C_{r1}R_{r2}}{T_{r2}}}+\frac{x_{r2}^2}{\frac{C_{r2}R_{r2}}{T_{r2}}}+\cdots+\frac{x_{rs}^2}{\frac{C_{rs}R_{rs}}{T_{rs}}}-R_{rs}$$

$$=\frac{x_{11}^2}{\frac{C_{21}R_{12}}{T_{22}}}-x_{11}+\frac{x_{12}^2}{\frac{C_{22}R_{12}}{T_{22}}}-x_{12}+\cdots+\frac{x_{1s}^2}{\frac{C_{2s}R_{1s}}{T_{2s}}}-x_{1s}$$

$$+\frac{x_{21}^2}{\frac{C_{21}R_{22}}{T_{22}}}-x_{21}+\frac{x_{22}^2}{\frac{C_{22}R_{22}}{T_{22}}}-x_{22}+\cdots+\frac{x_{2s}^2}{\frac{C_{2s}R_{2s}}{T_{2s}}}-x_{2s}$$

$$+\frac{C_{21}^2}{\frac{C_{31}T_{22}}{T_{32}}}-C_{21}+\frac{C_{22}^2}{\frac{C_{32}T_{22}}{T_{32}}}-C_{22}+\cdots+\frac{C_{2s}^2}{\frac{C_{3s}T_{2s}}{T_{3s}}}-C_{2s}$$

$$+\cdots$$

$$+\frac{x_{r1}^2}{\frac{C_{r1}R_{r2}}{T_{r2}}}-x_{r1}+\frac{x_{r2}^2}{\frac{C_{r2}R_{r2}}{T_{r2}}}-x_{r2}+\cdots+\frac{x_{rs}^2}{\frac{C_{rs}R_{rs}}{T_{rs}}}-x_{rs}$$

（前と同様にして

$$\frac{C_{21}^2}{\frac{C_{31}T_{22}}{T_{32}}}\doteqdot\frac{C_{21}^2}{C_{21}}=C_{21}=x_{11}+x_{21},\ \cdots$$

として）

$$=\frac{x_{11}^2}{\frac{C_{21}R_{12}}{T_{22}}}+\frac{x_{12}^2}{\frac{C_{22}R_{12}}{T_{22}}}+\cdots+\frac{x_{1s}^2}{\frac{C_{2s}R_{1s}}{T_{2s}}}$$

$$+\cdots$$

$$-C_{r-1,1}-C_{r-1,2}\cdots-C_{r-1,s}$$

$$+\frac{x_{r1}^2}{\frac{C_{r1}R_{r2}}{T_{r2}}}-x_{r1}+\frac{x_{r2}^2}{\frac{C_{r2}R_{r2}}{T_{r2}}}-x_{r2}+\cdots+\frac{x_{rs}^2}{\frac{C_{rs}R_{rs}}{T_{rs}}}-x_{rs}$$

$$=\frac{x_{11}^2}{\frac{C_{21}R_{12}}{T_{22}}}+\frac{x_{12}^2}{\frac{C_{22}R_{12}}{T_{22}}}+\cdots+\frac{x_{1s}^2}{\frac{C_{2s}R_{1s}}{T_{2s}}}$$

$$+\cdots$$

$$+\frac{x_{r1}^2}{\frac{C_{r1}R_{r2}}{T_{r2}}}+\frac{x_{r2}^2}{\frac{C_{r2}R_{r2}}{T_{r2}}}+\cdots\qquad+\frac{x_{rs}^2}{\frac{C_{rs}R_{rs}}{T_{rs}}}-T_{rs}$$

（さらに，

$$\frac{C_{21}R_{12}}{T_{22}}\doteqdot\frac{C_{21}R_{13}}{T_{23}}\doteqdot\cdots\cdots\doteqdot\frac{C_{21}R_{1s}}{T_{2s}}\doteqdot\frac{C_{31}R_{1s}}{T_{3s}}\doteqdot\cdots\doteqdot\frac{C_{r1}R_{1s}}{T_{rs}},$$

$\cdots\cdots$

がなりたつことから）

$$=\frac{x_{11}^2}{\frac{C_{r1}R_{1s}}{T_{rs}}}+\frac{x_{12}^2}{\frac{C_{r2}R_{1s}}{T_{rs}}}+\cdots+\frac{x_{1s}^2}{\frac{C_{rs}R_{1s}}{T_{rs}}}$$

$$+\cdots\cdots$$

$$+\frac{x_{r1}^2}{\frac{C_{r1}R_{rs}}{T_{rs}}}+\frac{x_{r2}^2}{\frac{C_{r2}R_{rs}}{T_{rs}}}+\cdots+\frac{x_{rs}^2}{\frac{C_{rs}R_{rs}}{T_{rs}}}-T_{rs}$$

$$=\sum_{i,j}\frac{\left(x_{ij}-\frac{R_{is}C_{rj}}{T_{rs}}\right)^2}{\frac{R_{is}C_{rj}}{T_{rs}}}=\sum_{i,j}\frac{\left(x_{ij}-\frac{t_{i.}t_{.j}}{n}\right)^2}{\frac{t_{i.}t_{.j}}{n}} \tag{37}$$

をうる。

4. X^2 の不連続分布

これまでに論じたのは，カイ自乗統計量が n を大きくしていったとき，極限において χ^2- 分布に従うということである。したがって，充分大きな標本についての独立性の検定においては，連続分布である χ^2- 分布をもちいることで充分であろう。しかしながら，小さい標本についてはどうであろうか，これがここでの問題である。標本数 n が小さい場合にはカイ自乗統計量の値は有限個しか存在せず，分布は不連続となる。例えば，周辺度数が，

$t_{1.}=7$，$t_{2.}=5$，$t_{.1}=5$，$t_{.2}=7$

である場合，可能な２×２分割表の数は６であり，それぞれの表および対応するカイ自乗統計量および確率は表７のとうりである。

表 7

0	7	7
5	0	5
5	7	12

$\chi^2=12.0000000$
$p=0.0012626$

1	6	7
4	1	5
5	7	12

$\chi^2=5.1820408$
$p=0.0441919$

2	5	7
3	2	5
5	7	12

$\chi^2=1.1853061$
$p=0.2651515$

3	4	7
2	3	5
5	7	12

$\chi^2=0.0097959$
$p=0.4419192$

4	3	7
1	4	5
5	7	12

$\chi^2=1.6555102$
$p=0.2209596$

5	2	7
0	5	5
5	7	12

$\chi^2=6.1224490$
$p=0.0265152$

このような小さいnに対するカイ自乗統計量 X^2 の正確な分布と， 連続な χ^2ー分布の比較を考える。まづ，X^2 が特定の値 x をもつ確率 $p(X^2=x)$ に対して χ^2ー 分布をどう対応させるかについて考える。連続分布においては一点の確率は0であり，区間（正確には可測集合）に対してはじめて確率が定義される。すなわち，$p(a\leq\chi^2\leq b)$ はカイ自乗値が区間〔a, b〕におちる確率であり，$f(\chi^2)$ を χ^2ー 分布の密度関数とすれば，

$$p(a\leq\chi^2\leq b)=\int_a^b f(\chi^2)d\chi^2$$

とかける。いま，起りうるすべての X^2ー 値を，大きさの順序にならべ，$x_1<x_2<x_3<\cdots<x_k$ とする。$m_0=0$，$m_1=(x_1+x_2)/2$，$m_2=(x_2+x_3)/2$，……，$m_{k-1}=(x_{k-1}+x_k)/2$として，

$$p(X^2=x_1)\ :\ p(m_0\leq\chi^2<m_1)=\int_0^{m_1}f(\chi^2)d\chi^2$$

$$p(X^2=x_2)\ :\ p(m_1\leq\chi^2<m_2)=\int_{m_1}^{m_2}(\chi^2)d\chi^2$$

$$p(X^2=x_k)\ :\ p(m_{k-1}\leq\chi^2)\quad=\int_{m_{k-1}}^{\infty}f(\chi^2)d\chi^2$$

を対応させる（図参照）したがって，

$p(x_i\leq X^2)$ に対応する連続分布での確率は $p(m_{i-1}\leq\chi^2)=\int_{m_{i-1}}^{\infty}f(\chi^2)d\chi^2$である。

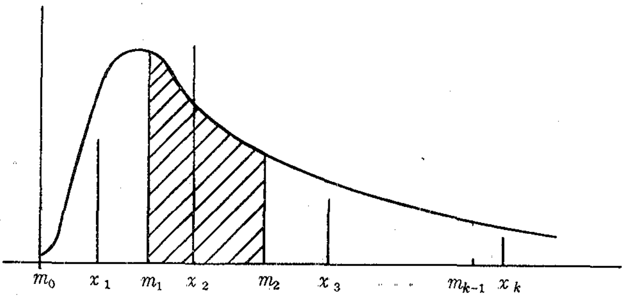

表7であげた例についてこの比較を示せば，表8のようになる。

表　8

x_i	$p(X^2 \geqq x_i)$	$p(\chi^2 \geqq m_i)$	m_i
0.00980	1.00000 (0.44192)*	1.00000 (0.56049)**	0.
1.18531	0.55808 (0.26515)	0.43951 (0.20618)	0.59755
1.65551	0.29293 (0.22096)	0.23334 (0.16888)	1.42041
5.18204	0.07197 (0.04419)	0.06446 (0.04703)	3.41878
6.12245	0.02778 (0.02652)	0.01743 (0.01482)	5.65224
12.00000	0.00126 (0.00126)	0.00261 (0.00261)	9.06122

*　括弧の中は $p(X^2=x_i)$

**　〃　$p(m_{i-1} \leqq \chi^2 < m_i)$

なお以下にいくつかの例について比較表をあげておく，テールではちがいが大きいのがみうけられるが，よく検定に用いられる5%のあたり近似はかなりよいと考えてよいであろう。

5　　　　となっているのは$n=10$，周辺度数が $t_{1.}=t_{2.}=t_{.1}=t_{.2}=5$ ということである。X2は表8のx_i，PEXACTは$P(X^2 \geqq x_i)$，PCONTは$P(\chi^2 \geqq m_i)$ に対応するものである。
5
10　5　5

5			
5			
10	5	5	
X2		PEXACT	PCONT
0.40000		1.00000	1.00000
3.60000		0.20635	0.15730
10.00000		0.00794	0.00912

10			
10			
20	10	10	
X2		PEXACT	PCONT
0.0		1.00000	1.00000
0.80000		0.65628	0.52709
3.20000		0.17890	0.15730
7.20000		0.02301	0.02259
12.80000		0.00109	0.00157
20.00000		0.00001	0.00005

15			
15			
30	15	15	
X2		PEXACT	PCONT
0.13333		1.00000	1.00000
1.20000		0.46609	0.41422
3.33333		0.14311	0.13218
6.53333		0.02684	0.02634
10.80000		0.00281	0.00324
16.13333		0.00015	0.00024
22.53333		0.00000	0.00001

20			
20			
40	20	20	
X2		PEXACT	PCONT
0.0		1.00000	1.00000
0.40000		0.75237	0.65472
1.60000		0.34307	0.31731
3.60000		0.11283	0.10686
6.40000		0.02564	0.02535
10.00000		0.00385	0.00419
14.40000		0.00036	0.00048
19.60000		0.00002	0.00004

25			
25			
50	25	25	
X2		PEXACT	PCONT
0.08000		1.00000	1.00000
0.72000		0.57214	0.52709
2.00000		0.25779	0.24354
3.92000		0.08874	0.08535
6.48000		0.02271	0.02259
9.68000		0.00420	0.00448
13.52000		0.00054	0.00066
18.00000		0.00005	0.00007
23.12000		0.00000	0.00001

30			
30			
60	30	30	
X2		PEXACT	PCONT
0.0		1.00000	1.00000
0.26667		0.79655	0.71500
1.06667		0.43891	0.41422
2.40000		0.19636	0.18799
4.26667		0.06985	0.06789
6.66667		0.01938	0.01938
9.60000		0.00412	0.00435
13.06667		0.00066	0.00076
17.06667		0.00008	0.00010
21.60000		0.00001	0.00001

35			
35			
70	35	35	
X2		PEXACT	PCONT
0.05714		1.00000	1.00000
0.51429		0.63294	0.59298
1.42857		0.33909	0.32432
2.80000		0.15102	0.14593
4.62857		0.05507	0.05395
6.91429		0.01622	0.01629
9.65714		0.00380	0.00400
12.85714		0.00070	0.00079
16.51429		0.00010	0.00013
20.62857		0.00001	0.00002

40			
40			
80	40	40	
X2		PEXACT	PCONT
0.0		1.00000	1.00000
0.20000		0.82325	0.75183
0.80000		0.50262	0.47950
1.80000		0.26347	0.25421
3.20000		0.11700	0.11385
5.00000		0.04351	0.04288
7.20000		0.01341	0.01352
9.80000		0.00339	0.00355
12.80000		0.00069	0.00078
16.20000		0.00011	0.00014
20.00000		0.00001	0.00002

45			
45			
90	45	45	
X2		PEXACT	PCONT
0.04444		1.00000	1.00000
0.40000		0.67355	0.63735
1.11111		0.39924	0.38472
2.17778		0.20568	0.19972
3.60000		0.09116	0.08919
5.37778		0.03444	0.03412
7.51111		0.01100	0.01113
10.00000		0.00295	0.00309
12.84444		0.00066	0.00073
16.04444		0.00012	0.00014
19.60000		0.00002	0.00002

50			
50			
100	50	50	
X2		PEXACT	PCONT
0.0		1.00000	1.00000
0.16000		0.81462	0.77730
0.64000		0.54874	0.52709
1.44000		0.31734	0.30782
2.56000		0.16120	0.15730
4.00000		0.07134	0.07013
5.76000		0.02731	0.02717
7.84000		0.00898	0.00912
10.24000		0.00252	0.00264
12.96000		0.00060	0.00066
16.00000		0.00012	0.00014
19.36000		0.00002	0.00003
23.04000		0.00000	0.00000

55			
55			
110	55	55	
X2		PEXACT	PCONT
0.03636		1.00000	1.00000
0.32727		0.70312	0.66982
0.90909		0.44577	0.43172
1.78182		0.25247	0.24607
2.94545		0.12676	0.12419
4.40000		0.05604	0.05531
6.14545		0.02169	0.02166
8.18182		0.00731	0.00744
10.50909		0.00214	0.00224
13.12727		0.00054	0.00059
16.03636		0.00012	0.00013
19.23636		0.00002	0.00003
22.72727		0.00000	0.00000

60			
60			
120	60	60	

X2	PEXACT	PCONT
0.0	1.00000	1.00000
0.13333	0.85523	0.79625
0.53333	0.58408	0.56370
1.20000	0.36139	0.35188
2.13333	0.20106	0.19671
3.33333	0.09996	0.09827
4.80000	0.04417	0.04374
6.53333	0.01726	0.01729
8.53333	0.00594	0.00606
10.80000	0.00179	0.00188
13.33333	0.00047	0.00051
16.13333	0.00011	0.00012
19.20000	0,00002	0.00003
22.53333	0.00000	0.00000

1630			
1033			
2663	2647	16	

X2	PEXACT	PCONT
0.01130	1.00000	1.00000
0.16673	0.79800	0.76544
0.38551	0.61518	0.59926
0.85180	0.44155	0.43155
1.28936	0.31145	0.30081
2.06651	0.19763	0.19520
2.72285	0.12459	0.12175
3.81087	0.06957	0.07069
4.68598	0.03734	0.03929
6.08487	0.01885	0.02031
7.17875	0.00778	0.01002
8.88850	0.00392	0.00459
10.20117	0.00102	0.00201
12.22178	0.00064	0.00081
16.08470	0.00008	0.00017
20.47726	0.00001	0.00002

8			
84			
92	20	72	

X2	PEXACT	PCONT
0.05476	1.00000	1.00000
0.43962	0.68108	0.61906
1.27930	0.36456	0.35389
2.43386	0.19320	0.17302
4.11323	0.06461	0.07041
8.55655	0.01106	0.01184
14.60926	0.00113	0.00067
22.27136	0.00006	0.00002

5			
37			
42	20	22	

X2	PEXACT	PCONT
0.13209	1.00000	1.00000
0.34880	0.65604	0.62389
1.73572	0.34647	0.30730
2.38585	0.17448	0.15113
5.15971	0.04918	0.05209
6.24324	0.01823	0.01695

20		
22		
50		
92	84	8

X2	PEXACT	PCONT
0.07766	1.00000	1.00000
0.45552	0.89141	0.87521
0.62727	0.78624	0.76285
1.05093	0.69110	0.65734
1.17688	0.61489	0.57295
1.46314	0.55328	0.51685
1.50894	0.49783	0.47568
2.59671	0.36066	0.35829
2.68831	0.32122	0.26680
2.82571	0.28617	0.25195
3.06617	0.25373	0.22924
3.20357	0.23063	0.20858
3.29517	0.20468	0.19697
3.92494	0.17387	0.16447
6.15394	0.15558	0.13269
4.21119	0.10953	0.12553
4.42874	0.08807	0.11533
4.51195	0.06997	0.10433
5.34476	0.05524	0.08298
6.31803	0.04325	0.05417
6.71879	0.03877	0.03842
7.18825	0.03476	0.03091
7.36000	0.02783	0.02633
7.81801	0.02206	0.02249
8.32182	0.01652	0.01769
8.56227	0.01383	0.01468
9.24929	0.00934	0.01164
9.52409	0.00608	0.00916
10.47446	0.00415	0.00674
11.21872	0.00377	0.00411
12.13474	0.00345	0.00291
12.73015	0.00319	0.00200
12.85610	0.00239	0.00167
14.36753	0.00141	0.00111
14.73394	0.00126	0.00069
14.91714	0.00080	0.00060
16.19957	0.00029	0.00042
19.54303	0.00019	0.00013
19.92089	0.00010	0.00005
22.34833	0.00006	0.00003
22.66894	0.00002	0.00001
27.87879	0.00000	0.00000

8				
8				
16	4	4	4	4

X2	PEXACT	PCONT
0.0	1.00000	1.00000
2.00000	0.89930	0.80125
4.00000	0.36224	0.39163
6.00000	0.24289	0.17180
8.00000	0.06387	0.07190
10.00000	0.03030	0.02929
16.00000	0.00047	0.00454