第5章　統計的仮説検定の考え方と方法

5.1 仮説検定とは

いま，1つのサイコロを3回投げる実験を考えよう．もし3回とも1の目が出たとする．そのとき実験者は“これは偶然だ”とか“少しおかしいぞ”と考えるだろう．たとえ偶然だと考えた人も，もし10回投げ続けて10回とも1の目が出ると“これはおかしい”とサイコロが正常であることを疑うであろう．さて3回の実験ですべて1の目が出たとき，サイコロが正常かどうかを問題にしたとする．まずこのサイコロは正常だという仮説をたてる．1回目に出る目の数を X_1，2回目に出る目の数を X_2，3回目に出る目の数を X_3 と確率変数で表わすとき

$$(5.1)\quad Z = X_1 + X_2 + X_3$$

なる確率変数 Z を考えよう．このとき Z の確率分布表は表 5.1 のようになる．

表 5.1　Z の確率分布表

Z のとる値	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
その確率	$\frac{1}{216}$	$\frac{3}{216}$	$\frac{6}{216}$	$\frac{10}{216}$	$\frac{15}{216}$	$\frac{21}{216}$	$\frac{25}{216}$	$\frac{27}{216}$	$\frac{27}{216}$	$\frac{25}{216}$	$\frac{21}{216}$	$\frac{15}{216}$	$\frac{10}{216}$	$\frac{6}{216}$	$\frac{3}{216}$	$\frac{1}{216}$

この表からもわかるように，サイコロが正常であるという仮定のもとでは，すべての回に出る目の数が1である，すなわち $Z=3$ となる確率は 1/216 となる．したがって，同様な実験を216回くり返すとき平均1回起こる結果なのである．その結果がただ1回の実験で起こったのだから“あまりにも偶然だ”とか“おかしいぞ”と思うわけである．いいかえればサイコロが正常であるという仮説のもとで起こる確率が非常に小さい事象がただ1回の実験で起こったとき“あまりにも偶然だ”とか‘サイコロがおかしいのではないか”と考えサイコロが正常であるという仮説を疑うのである．

そこで一般的に1つの仮説（普通，H_0 で示す）をたて，この仮説のもと

で実験，調査をおこなう．そして得られた結果（サンプル）から作られる統計量 T の実現値を得る．この T は仮説 H_0 のもとでは1つの確率分布を持っており，その確率分布によって T の実現値の現われる確率がきまる．このとき T の実現値の現われる全領域のうち，確率 α で現われる領域をつくる．もしその領域に T の実現値が現われれば，仮説 H_0 のもとで確率 α で現われることが現実に起こったということになる．いま $\alpha=0.01$ とすると，1/100 の確率でしか起こらないことが1回の実験，調査で起こったということになり，この仮説 H_0 は認められないであろうと考えるわけである．すなわち仮説 H_0 を棄却するのである．普通はこの確率 α で T の実現値が現われる領域は図5.1のように両側（または片側）にとるのが普通である．図 5.1 においては T の頻度分布（確率分布）を連続分布曲線（密度関数）$f(x)$ で近似したものである．

T の実現値の現われる確率が α の領域を，**危険率 100 α % の棄却域**という．これは，もし T の実現値が棄却域に入って仮説 H_0 を棄てた場合に，その仮説 H_0 が正しいにもかかわらず棄却する確率が α となるわけであるから，これだけの危険をおかしているということからこのように呼ばれるわけである．このような手続きでおこなう検定のことを**統計的仮説検定**という．検定の対象となる仮説として H_0 のことを“**帰無仮説**”と呼んでいる．以後この書物でも，この言葉を用いることにする．

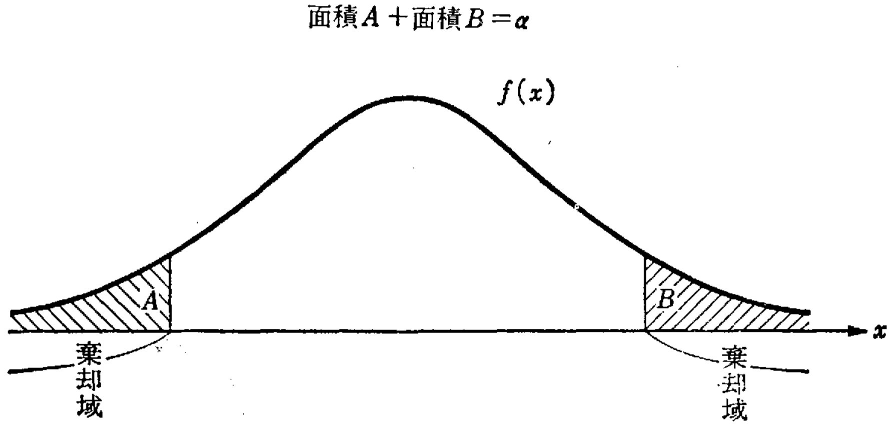

図 5.1　危険率 100 α % の棄却域

例　サイコロを3回投げる実験で，帰無仮説を“サイコロが正常であるとする”とき，式(5.1)にもとづき危険率5％以下の棄却域（片側）を作る．

解　いま式（*5.1*）で示した Z の確率分布は，Z が実現値 z をとる確

率を $\mathrm{P}(\boldsymbol{Z}=z)$ と表わすと表 5.1 から

$$\mathrm{P}\{\boldsymbol{Z}=z\}=\begin{cases} 1/216 & (z=\ 3) \\ 3/216 & (z=\ 4) \\ 6/216 & (z=\ 5) \\ 10/216 & (z=\ 6) \\ \vdots & \\ 10/216 & (z=15) \\ 6/216 & (z=16) \\ 3/216 & (z=17) \\ 1/216 & (z=18) \end{cases}$$

となり，危険率 4.6 % (5 % 以下)(*)の棄却域は $z=3,\ z=4,\ z=5$ となる．すなわち

$$\mathrm{P}\{\boldsymbol{Z}=3\}+\mathrm{P}\{\boldsymbol{Z}=4\}+\mathrm{P}\{\boldsymbol{Z}=\ 5\ \}=0.046<0.05$$

となる．だからサイコロを 3 回投げてもし，出た目の数の和が 3 か 4 か 5 なら危険率 4.6 % で仮説は棄却される．したがって，正常なサイコロとはみなされず，小さい目の出やすいサイコロではないかと疑う．

この場合，サイコロが正しくてもこの棄却域にはいる確率は 0.046 だけあるわけで，帰無仮説が正しくても棄却するという誤りを 0.05 以下の確率で犯していることになる．この 5 % のことをこれ以上危険率が大きくならない基準として**有意水準**と呼んでいる．(**) 普通，有意水準としては 5 % 1 % が用いられる．また，仮説が棄却されるとき検定の結果は**有意**という．

さて，一般的に帰無仮説 H_0 が正しいにもかかわらずこれを棄却する誤りを H_0 に対する**第 1 種の誤り**という．第 1 種の誤りを犯す確率が危険率である．また H_0 が正しくないにもかかわらずこれを採択する誤りを**第 2 種の誤り**と呼んでいる．

〈**質問**〉 図 5.1 のように棄却域を両側または片側にとるのはどのような理由によるものですか．

(*) この場合，危険率が丁度 5 % となる棄却域は存在しない．

(**) この例の場合，**有意水準** 5 % **の棄却域**は $z=3,\ z=4,\ z=5$ となり，これは危険率 4.6% の棄却域と一致する．

答　我々は第1種の誤りと第2種の誤りのどちらも犯さないように棄却域を作りたいわけですが，それは不可能ですね．そこで第1種の誤りを犯す確率を一定にしておいて第2種の誤りをできるだけ小さくする棄却域を作ることを考えるわけです．いま第1種の誤りを犯す確率を一定の α にしておくと棄却域は図5.2のように片側でも，両側でも，中央にでもどこにでも作ることができます．

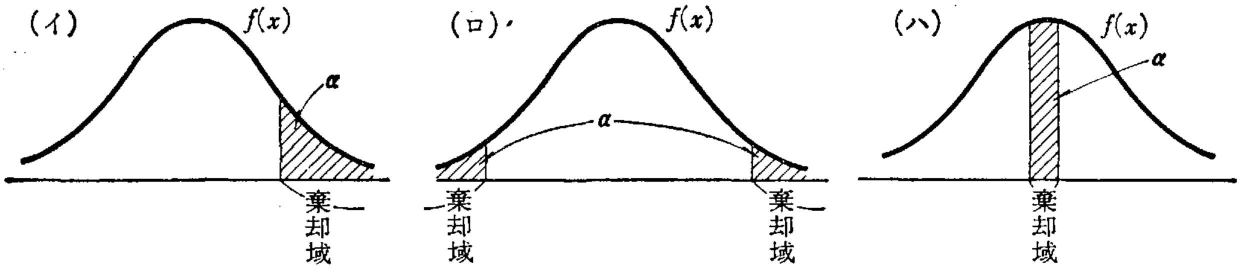

図 5.2　第1種の誤りを犯す確率 α の棄却域

次に帰無仮説 H_0 が正しくないとき，別の正しい仮説 H_1 を考えます(*)．たとえばサイコロの例ですと H_1 として

1 の目の出る確率	$\frac{1}{5}$
2 の　〃	$\frac{1}{5}$
3 の　〃	$\frac{1}{5}$
4 の　〃	$\frac{2}{15}$
5 の　〃	$\frac{2}{15}$
6 の　〃	$\frac{2}{15}$

なる仮説を考えることはできますね．一般に別の正しい仮説 H_1 のもとでの T の確率分布を連続曲線で近似したものを $g(x)$ で表わすことにします（図 5.3で点線で

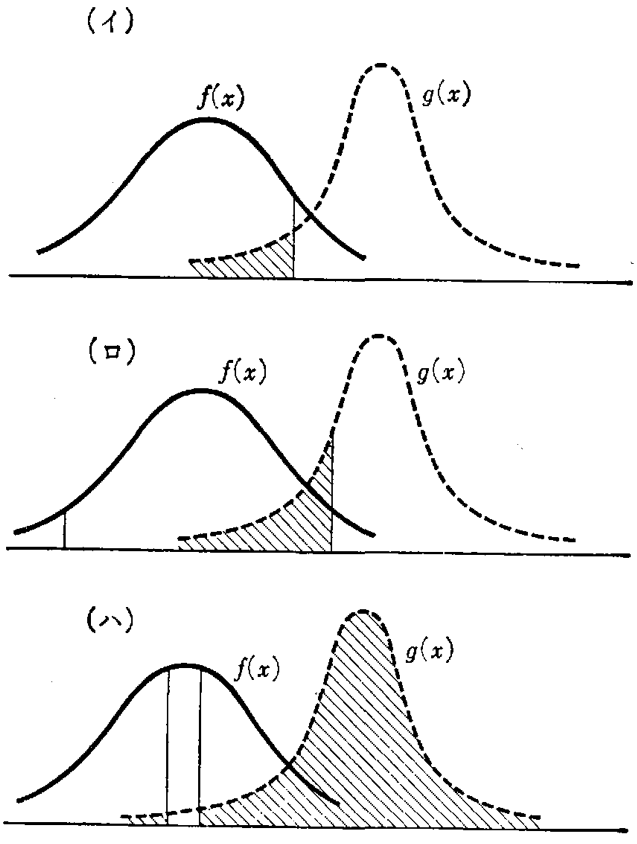

図 5.3　第2種の誤りを犯す確率
（斜線の部分）

(*)　H_1 を**対立仮説**と呼んでいる．

表わす）．図 5.3 で第 1 種の誤りを犯す確率 α を一定にしたいとき，第 2 種の誤りを犯す確率が（イ）（ロ）（ハ）のどれが 1 番小さくなるか考えてみましょう．図 5.2 の棄却域以外のところが仮説 H_1 を受け入れる領域ですから図 5.3 において斜線で示した部分が仮説 H_1 を受け入れる確率となるわけです．これが第 2 種の誤りの確率で（イ）（ロ）に比べ（ハ）が非常に大きいことがおわかりでしょう．したがって第 2 種の誤りを小さくするには棄却域は片側または両側にとることが好ましいわけです．

5.2 適合度検定法（その 1）

10 円の銅貨を投げて表が出れば 1，裏が出れば 0 と記入する実験を考えてみよう．いま A，B 2 つの銅貨を 100 回ずつ投げた結果が次表のようであった．

表 5.2 A，B 2 つの銅貨を投げた結果

A	B
1 0 0 1 0 1 1 1 0 1	0 1 0 0 0 1 0 1 1 0
0 1 0 1 1 0 0 1 1 0	1 0 0 0 0 1 0 1 1 0
0 1 1 1 0 1 0 1 0 1	0 1 0 1 0 0 0 1 1 0
0 0 1 1 0 0 0 1 0 1	1 1 0 1 0 1 0 0 0 0
1 0 1 1 0 0 1 1 0 1	0 1 0 0 1 1 0 1 0 0
0 0 0 1 0 1 1 1 0 1	1 1 0 0 0 1 0 0 0 1
1 0 1 1 1 0 0 0 1 1	1 0 1 1 0 0 0 0 0 0
0 1 0 0 1 1 0 1 0 0	1 1 0 0 0 0 0 0 1 0
1 0 0 1 1 0 1 0 1 1	0 1 0 0 1 0 0 1 0 0
0 1 1 0 1 0 1 1 1 0	0 0 1 0 0 0 0 1 0 1

この結果，この銅貨 A，B が正しい銅貨であるかどうか検定したい．そこでまず次のような帰無仮説をたてる．

帰無仮説：“表と裏の出現率が等しい”

この仮説のもとでは当然 100 回銅貨を投げたとき表と裏が 50 回ずつ出ることが期待される．この期待される値のことを**期待度数**と呼んでおり，実際実験で得られた結果のことを**観測度数**と呼んでいる．表 5.2 において観測度数は次のようになる．

銅貨 A $\begin{cases} 1 \text{ の観測度数} & 54 \\ 0 \text{ の}\quad〃 & 46 \end{cases}$

銅貨 B $\begin{cases} 1 \text{ の観測度数} & 36 \\ 0 \text{ の}\quad〃 & 64 \end{cases}$

この値をみただけで，銅貨 B の方は少し裏の出やすいゆがんだ銅貨ではないかと疑問を持つであろう．この疑問を確率的に裏ずけるための検定方式を考えたいわけである．すなわち，どれくらい小さい確率のものが実際に起こっているかを判断するわけである．いま，奇数が出れば 1，偶数が出れば 0 と記録する実験を 1 桁の 100 個の乱数を用いておこなうとちょうど正しい銅貨を 100 回投げて表が出れば 1，裏が出れば 0 と記録する実験に相当する．このような実験を 200 回くりかえして

$$(5.2)\qquad \frac{(0\text{ の観測度数}\ -50)^2}{50}+\frac{(1\text{ の観測度数}\ -50)^2}{50}$$

の値の頻度分布を作ってみると図 5.4 のようになる．この頻度分布は次式で示す自由度1のカイ 2 乗分布（χ^2-分布とも書く）と呼ばれる曲線で近似されることが知られている．その事実を図 5.4 に示す．

$$(5.3)\qquad f(x)=\frac{1}{\sqrt{2\pi x}}e^{-\frac{x}{2}}\quad (0<x<\infty)$$

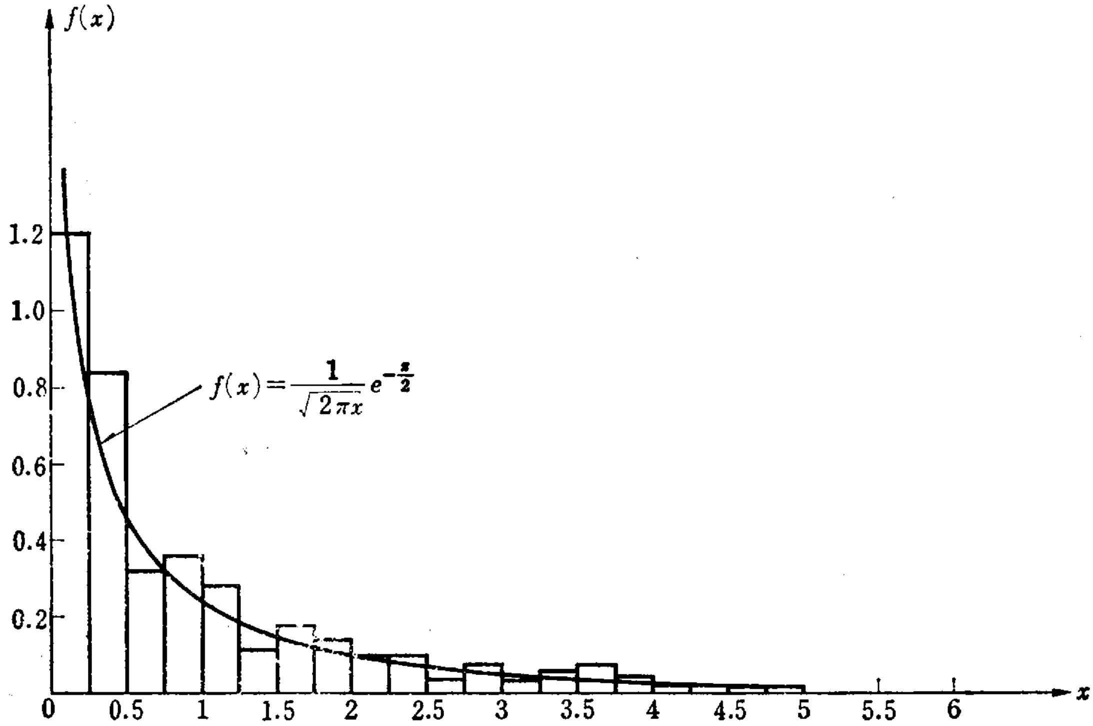

図 5.4　実験の結果の頻度分布とカイ 2 乗分布曲線の近似

式 (5.2) の値は観測度数がその期待値に完全に一致したときに0となり，期待値からの差が大きくなればなるほど大きくなる．したがって棄却域は右側の片側にとればよく，たとえば有意水準5％としたときの棄却域(*)は巻末付表4より x が 3.84 より大きい領域となる．表 5.2 の銅貨 A と B の式 (5.2) の実現値は

A:　0.64

B:　7.84

となり，銅貨 B の方は棄却域にはいり，帰無仮説は棄却される（否定される）．有意水準を1％としても棄却域は 6.63 より大きい領域となって，やはり銅貨 B の場合棄却域にはいることになる．このことは前節でも述べた通り，確率 $\frac{1}{100}$ 以下でしか起こらないことがただ1回の実験で起こったことになり，表と裏の出現率が等しいという帰無仮説は認められないと判断する．このことは，はじめに人間の直感で銅貨 B が少しゆがんでいるのではないかと疑ったことが確率的にも裏ずけられたという意味で有用であろう．

さて，一般に n 回の実験の起こりうる結果が $k+1$ 個の階級 C_1, C_2, ……, C_{k+1} に分類されていて，その階級の観測度数と期待度数が表5.3のようになっているとする．

表 5.3 階級の観測度数と期待度数

階　級	C_1, C_2………C_{k+1}	計
観測度数	n_1, n_2………n_{k+1}	n
期待度数	m_1, m_2………m_{k+1}	n

このとき

$$(5.4)\quad \sum_{i=1}^{k+1} \frac{(n_i - m_i)^2}{m_i}$$

の値と自由度 k のカイ2乗分布表の棄却域とを用いることによって次のような帰無仮説を検定することができる．

帰無仮説: "n 回の実験の結果が階級 C_1, C_2, ……, C_{k+1} に確率 p_1, p_2, ……, p_{k+1} でふり分けられている．" ただし $m_i = np_i$ $(i = 1, 2, \cdots\cdots, k+1)$

(*) 前節の例の z（離散変量）と異なって，x は連続変量なので危険率5％の棄却域と一致する．このような場合，この章では有意水準5％の棄却域ということにする．

これは前の銅貨の例と同様で，n 回の実験を何回もくりかえし行なうと式 (5.4) の実現値の頻度分布が**自由度 k のカイ 2 乗分布**（χ^2-**分布**）と呼ばれる次式のような曲線（密度関数）で近似されることを利用したものである．ただし m_i ($i=1, 2, \cdots\cdots, k+1$) はある程度大きいとする ($m_i \geqq 10$ ぐらいならばよい)．

$$(5.5)\quad f_k(x)=\frac{1}{2^{\frac{k}{2}}\Gamma\left(\frac{k}{2}\right)}x^{\frac{k}{2}-1}e^{-\frac{x}{2}} \qquad (0<x<\infty)$$

ただし $\Gamma\left(\frac{k}{2}\right)$ はガンマ関数と呼ばれるものである．(補注 [9])

参考のために，k の値と $f_k(x)$ のグラフの略図を描いておく．(図 5.5)

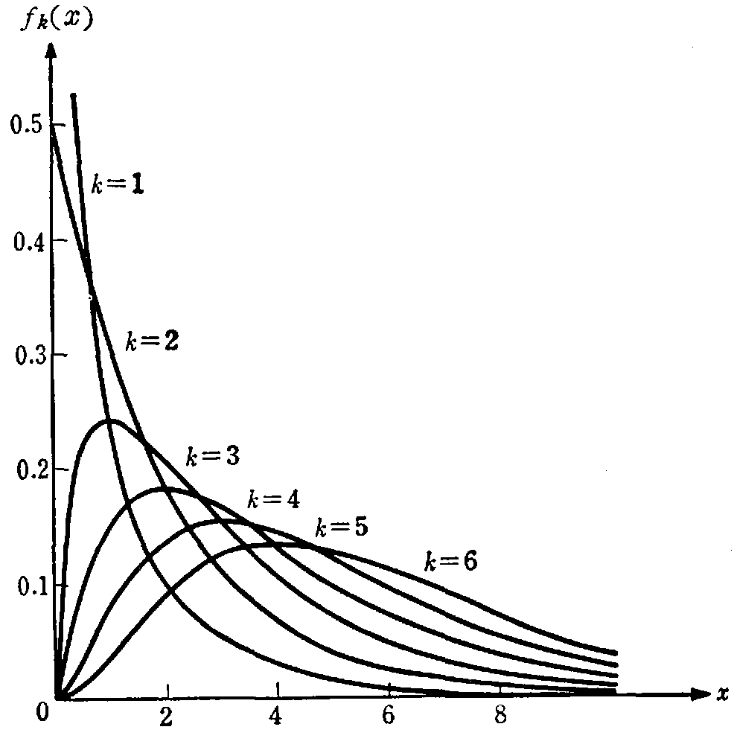

図 5.5　種々の自由度のカイ 2 乗分布曲線

例 1　サイコロを 60 回投げる実験において 1 から 6 までの出た目の観測度数と期待度数は表 5.4 のようであった．

この結果，このサイコロはどの目の出る確率も $\frac{1}{6}$ とみなしてよいか検定する．

解　表 5.4 を一目見ただけで誰しも 1 と 6 の目が出やすいと疑問をもつであろう．さて，これを確率的に仮説検定の立場から裏ずけてみよう．

表 5.4

サイコロの目 (i)	1	2	3	4	5	6	計
観測度数 (n_i)	16	6	5	10	7	16	60
期待度数 (m_i)	10	10	10	10	10	10	60

帰無仮説："どの目の出る確率も $\frac{1}{6}$" とする．

表 5.4 から式 (*5.4*) は

$$\sum_{i=1}^{6}\frac{(n_i-m_i)^2}{m_i}=\frac{(16-10)^2}{10}+\frac{(6-10)^2}{10}+\frac{(5-10)^2}{10}+\frac{(10-10)^2}{10}+\frac{(7-10)^2}{10}+\frac{(16-10)^2}{10}=12.2$$

となり，自由度は 5 であるから巻末付表 4 から有意水準 5％ の棄却域は 11.07 以上の領域となり，実現値は棄却域にはいっており，仮説は棄却される（否定される）.（図 5.6 参照）

いいかえれば確率 0.05 以下でしか起こらないことが 1 回の実験で起こったことになるわけで，仮説は認められないというわけである.

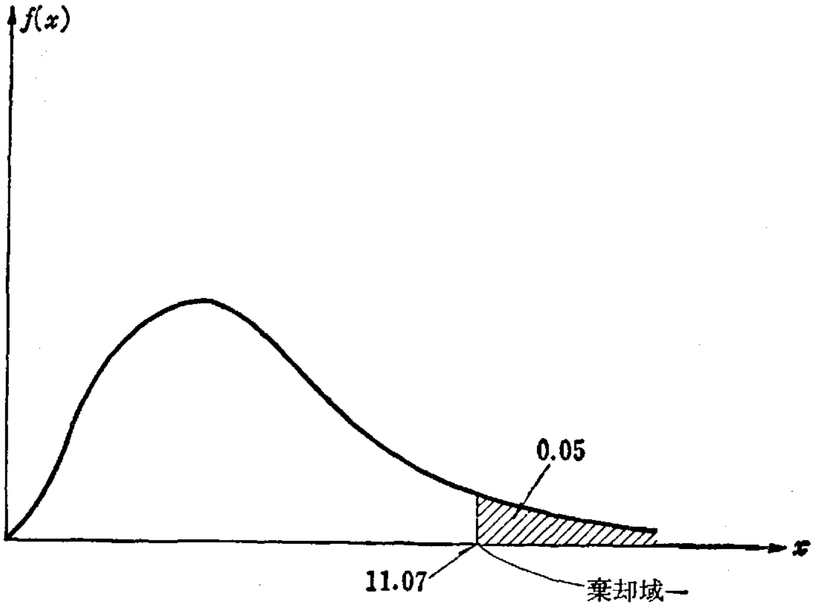

図 5.6　自由度 5 のカイ 2 乗分布曲線の有意水準 5％ の棄却域

例 2　第 2 章で述べた円周率 π の小数部分のはじめから 10000 個の数字に対して

帰無仮説: “0 から 9 までの数字の出現率は等しい”

を検定する.

解　いまはじめから 1000 個までの数を第 1 ブロック，1001 個から 2000 個までの数を第 2 ブロック，……，9001 個から 10000 個までの数を第 10 ブロックとする．そのときブロック j ($j=1, 2, \cdots\cdots, 10$) に含まれる数字 i ($i=0, 1, 2, \cdots\cdots, 9$) の個数と式 (*5.4*) の値は，表 5.5 のようであった.

表 5.5　ブロック j に含まれる数字 i の個数と式 (5.4) の値

j \ i	0	1	2	3	4	5	6	7	8	9	$\sum_{i=0}^{9}\frac{(n_i-m_i)^2}{m_i}$ の値
1	93	116	103	102	93	97	94	95	101	106	4.74
2	89	96	104	86	102	108	106	102	101	106	4.94
3	77	97	96	77	123	110	102	90	108	120	22.80
4	103	120	105	103	87	102	96	90	95	99	7.58
5	104	103	88	91	103	108	115	111	87	90	9.38
6	91	94	98	113	105	97	106	118	90	88	9.28
7	100	107	98	114	89	108	89	88	98	109	7.84
8	97	100	119	95	107	104	108	92	84	94	8.80
9	101	103	100	103	101	99	98	97	90	108	1.98
10	113	90	110	90	102	113	107	87	94	94	9.32

各ブロックとも自由度は9で，有意水準 5％ の棄却域は，巻末付表4から16.92 より大きい領域となる．そこで表 5.5 を見ると第3ブロックの値が22.80 と極端に大きい．このブロックにおいては帰無仮説は認められず，0と3が少なく4と9が多く現われるという結果を得る．

例 3　ある町の，1月から6月までの自動車事故の件数は，

1月	2月	3月	4月	5月	6月
22	25	20	35	30	18

であった．この結果どの月も同じ割合で事故が起こると言えるか．

解

帰無仮説："どの月も事故の発生率は同じであるとする"

このとき観測度数と期待度数は次表のようになる．

表 5.6

月 (i)	1	2	3	4	5	6	計
観測度数 (n_i)	22	25	20	35	30	18	150
期待度数 (m_i)	25	25	25	25	25	25	150

この表より式 (5.4) の値は

$$\sum_{i=1}^{6}\frac{(n_i-m_i)^2}{m_i}=\frac{(22-25)^2}{25}+\frac{(25-25)^2}{25}+\frac{(20-25)^2}{25}$$
$$+\frac{(35-25)^2}{25}+\frac{(30-25)^2}{25}$$
$$+\frac{(18-25)^2}{25}=8.32$$

となる．巻末付表4の自由度5のカイ2乗分布表より有意水準5％の棄却域は11.07より大きい領域となるので帰無仮説は棄却されない．したがって4月と5月が直感的には事故が多いように見えるが，これぐらいの数となる確率はあまり小さくないのでどの月も同じ割合で事故が起こっているとみなしてもさしつかえない．

〈質問〉 表5.2の1枚の銅貨投げにおいてたとえば

$$\underbrace{1\ \ 1\ \ 1\cdots\cdots\cdots1}_{48\text{ 個}}\qquad\qquad \underbrace{0\ \ 0\ \ 0\cdots\cdots\cdots0}_{52\text{ 個}}$$

というように1が続けて48個，0が続けて52個出た場合

帰無仮説："表と裏の出現率は等しい"

は式（*5.4*）の実現値が棄却域にはいらないために棄却されない（否定されない）ですね．しかしこの場合明らかに，はじめは表が出るようにして途中から裏が出るように意識的に銅貨を曲げたか何か人為的なことが加わっていると思います．そこでこの並び方までも考慮した検定法についてはどのようにすればよいのですか．

答 銅貨投げの結果から，銅貨が正しいものであるかどうかを検定する場合は，表と裏の等出現性だけ調べたのでは駄目で，もう1つ並んでいる数の無規則性を調べなければなりません．質問は，この無規則性の検定法をたずねられており，大変大切な問題であると思います．

無規則性の検定にはいろいろありますが，よく用いられる代表的な方法である連の検定について述べておきましょう．

たとえば

A A B A A A A B A

とか

$+\ -\ -\ +\ -\ +\ -\ -\ +\ +$

のように，相異なる2つの文字または記号の列があったとき，同じ文字または同じ記号のひと続きを**連**（run）といいます．またひと続きの個数を**連の長さ**といいます．はじめの文字の例では A の連が3つあって，その長さは 2，4，1，また B の連が2つあって，その長さはいずれも1です．あとの記号の例では長さ1の ＋ の連が3つ，長さ2の ＋ の連が1つ，

長さ2の － の連が2つ，長さ1の － の連が1つあることになります．

いまn個のAとBの列においてAの個数をn_A，Bの個数をn_Bとしましょう．そのとき帰無仮説として

"n_A個のAとn_B個のB ($n_A+n_B=n$) の並び方が無規則である"

としますと，この帰無仮説のもとで得られる連の数の確率分布ができます．たとえば，$n_A=3$，$n_B=3$のときすべての並び方と連の数は次表のようになります．

表 5.7　並び方と連の数

並	び	方				連の数
A	A	A	B	B	B	2
A	A	B	A	B	B	4
A	A	B	B	A	B	4
A	A	B	B	B	A	3
A	B	A	A	B	B	4
A	B	A	B	A	B	6
A	B	A	B	B	A	5
A	B	B	A	A	B	4
A	B	B	A	B	A	5
A	B	B	B	A	A	3
B	A	A	A	B	B	3
B	A	A	B	A	B	5
B	A	A	B	B	A	4
B	A	B	A	A	B	5
B	A	B	A	B	A	6
B	A	B	B	A	A	4
B	B	A	A	A	B	3
B	B	A	A	B	A	4
B	B	A	B	A	A	4
B	B	B	A	A	A	2

このとき，連の数の度数分布表は次のようになります．

表 5.8　連の数の度数分布表

連の数	2	3	4	5	6	計

$$平均値 = \frac{2\,n_{\mathrm{A}}\,n_{\mathrm{B}}}{n_{\mathrm{A}} + n_{\mathrm{B}}} + 1$$

$$分\quad 散 = \frac{2\,n_{\mathrm{A}}\,n_{\mathrm{B}}\,(2\,n_{\mathrm{A}}\,n_{\mathrm{B}} - n_{\mathrm{A}} - n_{\mathrm{B}})}{(n_{\mathrm{A}} + n_{\mathrm{B}})^2\,(n_{\mathrm{A}} + n_{\mathrm{B}} - 1)}$$

なる正規分布曲線で近似できることが一般に知られています．したがって帰無仮説のもとでは

$$(5.6)\qquad \frac{連の数 - \left(\dfrac{2\,n_{\mathrm{A}}\,n_{\mathrm{B}}}{n_{\mathrm{A}} + n_{\mathrm{B}}} + 1\right)}{\sqrt{\dfrac{2\,n_{\mathrm{A}}\,n_{\mathrm{B}}\,(2\,n_{\mathrm{A}}\,n_{\mathrm{B}} - n_{\mathrm{A}} - n_{\mathrm{B}})}{(n_{\mathrm{A}} + n_{\mathrm{B}})^2\,(n_{\mathrm{A}} + n_{\mathrm{B}} - 1)}}}$$

の頻度分布は平均値 0，分散 1 の規準型正規分布曲線で近似することができます．したがって有意水準 5％ の棄却域

$$(-\infty,\ -1.96)\qquad (1.96,\ \infty)$$

に式（*5.6*）の値がはいるかどうかを見ればよいことになります．たとえば，この節のはじめの銅貨投げの実験で銅貨 A について，この無規則性の検定を行なってみましょう．

帰無仮説: "54 個の 1 と 46 個の 0 の並び方が無規則である".

このとき，表 5.2 より連の数は 64 で

$n_{\mathrm{A}} = 54$　（1 の個数）

$n_{\mathrm{B}} = 46$　（0 の個数）

と考えればよいから式（*5.6*）の値は 2.70 となり，これは有意水準 5％ の棄却域にはいってしまいます．

したがって表 5.2 の銅貨 A の実験結果の起る確率は 0.05 よりも小さく，小さい確率のものが 1 度の実験で起こったのですから帰無仮説は有意水準 5％ で認められないことになります．したがって表と裏の等出現性は認められても無規則性が認められなくなり，正常な銅貨であることが保証されなくなります．

もちろん質問の中の

$$\overbrace{1\quad 1\quad 1\cdots\cdots\cdots 1}^{48\ 個}\qquad\qquad \overbrace{0\quad 0\quad 0\cdots\cdots\cdots 0}^{52\ 個}$$

の場合は連の数が 2 で式（*5.6*）の値が有意水準 5％ でも 1％ でも棄却

域にはいることは明らかでしょう.

無規則性の検定になるとなかなか人間の直感では判断することが困難となり，統計的仮説検定法の威力が認識できるわけです.

問 題 12

1. メンデルはえんどうの交配実験において，次の表の結果を得た．度数分布は理論的には 9:3:3:1 の割合になるべきものである．この実験の結果はこの理論によく適合しているといえるか.

表 5.9 えんどうの交配実験

種子の種類	円くて黄色	角ばって黄色	円くて緑色	角ばって緑色	計
観測度数	315	101	108	32	556

2. 第2章で述べた円周率πの小数部分のはじめから 100 個の数字において 0から9までの数の出現率が等しいかどうかを有意水準 5% で検定せよ．数字は第2章に示されているものを利用せよ.

3. 円周率πの小数部分のはじめから 100 個の数字について0から4までを A, 5 から9までを B とみなして連の検定を有意水準 5% で行なえ.

5.3 適合度検定法 (その 2)—分割表

例 1 人口 15000 人の地域（母集団）でインフルエンザの予防注射の効果があるかどうかを調べるために 200 人をランダムに選び，インフルエンザにかかったかどうかと同時に，予防注射を受けたかどうかについて調べて，その結果（観測度数）を表 5.10 にまとめた．この表のことを 2×2 **分割表**という.

表 5.10 2×2 分割表

	インフルエンザにかかった人	インフルエンザにかからなかった人	計
予防注射を受けた人	35	45	80
予防注射を受けなかった人	80	40	120
	115	85	200

この表から見ると直感的には予防注射のききめはありそうだが，これを確率的に検定法の立場から吟味してみよう.

一般の母集団において2つの属性 B, C がそれぞれ2つの項目 B_1, B_2,

C_1, C_2 に分類されているとする．この母集団からの大きさ n のランダムサンプルにもとづき表 5.11 のような 2×2 分割表にその観測度数が示してある．

表 5.11 2×2 分割表

B＼C	C_1	C_2	計
B_1	a	b	$a+b$
B_2	c	d	$c+d$
計	$a+c$	$b+d$	n

このとき次のような帰無仮説をたてる．

帰無仮説: “属性 B と C は無関係である”

この帰無仮説のもとで，母集団から大きさ n のランダムサンプルを多数回くり返し

$$(5.7) \quad \frac{(a-\alpha)^2}{\alpha}+\frac{(b-\beta)^2}{\beta}+\frac{(c-\gamma)^2}{\gamma}+\frac{(d-\delta)^2}{\delta}$$
$$=\frac{n(ad-bc)^2}{(a+b)(c+d)(a+c)(b+d)}$$

$$\text{ただし，}\ \alpha=\frac{(a+c)(a+b)}{n},\quad \beta=\frac{(b+d)(a+b)}{n}$$
$$\gamma=\frac{(a+c)(c+d)}{n},\quad \delta=\frac{(b+d)(c+d)}{n}$$

の値を計算して，その頻度分布をつくると $\alpha, \beta, \gamma, \delta$ がある程度大きければ（目安としてすべて 10 以上），この頻度分布は自由度 1 のカイ 2 乗分布で近似できることが知られている．そこで有意水準を定めると自由度 1 のカイ 2 乗分布表（巻末付表 4 ）より棄却域が定まり，式（5.7）の値が棄却域にはいるかどうかを調べれば帰無仮説が棄却される（否定される）かどうかがわかる．

さて，前の予防注射とインフルエンザの例にもどろう．

帰無仮説: “予防注射の有無とインフルエンザにかかったか，かからなかったかは無関係である”

表 5.10 より式（5.7）の値は

$$\frac{200\times(35\times 40-45\times 80)^2}{80\times 120\times 115\times 85}=10.32$$

となってこれは自由度 1 のカイ 2 乗分布の有意水準 5％ の棄却域 $(3.841, \infty)$ にはいり，帰無仮説が正しいとすると表 5.10 のような観測

度数となることは確率 0.05 以下でしか起こらないことになり，帰無仮説は認められないと判断する．このような意味で予防注射はききめがあると考えてよかろう．

表 5.11 において a, b, c, d のうちあまり大きくないもの（10 以下となるもの）があるときは，式（5.7）の頻度分布の自由度1のカイ 2 乗分布への近似度が悪くなり，式（5.7）に修正をほどこして次式の頻度分布が自由度1のカイ 2 乗分布で近似できることを利用しなければならない．

$$(5.8)\qquad \frac{n\left(ad-bc\pm\dfrac{n}{2}\right)^2}{(a+b)(c+d)(a+c)(b+d)}$$

ただし，複号 ± は分子の括弧内の絶対値が小さくなるようにする．これを**イェツ**（Yates）**の修正**と呼んでいる．

例 2　表 5.12 のような観測度数が得られたとする．

このとき

帰無仮説：“属性 B と C は無関係である”

を有意水準 5％ で検定する．

表 5.12

$B \backslash C$	C_1	C_2	計
B_1	80	5	85
B_2	100	15	115
計	180	20	200

解　式（5.8）より

$$\frac{n\left(ad-bc\pm\dfrac{n}{2}\right)^2}{(a+b)(c+d)(a+c)(b+d)}$$

$$=\frac{200\times\left(80\times15-100\times5-\dfrac{200}{2}\right)^2}{85\times115\times180\times20}=2.05$$

となり，有意水準 5％ の棄却域（3.841, ∞）にはいらない．したがって帰無仮説が正しいと考えたとき観測度数が表 5.12 のようになることは妥当であると判断する．

属性 B の項目が k 個（$k>2$），C の項目が l 個（$l>2$）のときも 2×2 分割表が $k\times l$ 分割表となるだけで適合度検定ができる（補注［10］）．

問　題　13

ある県の中学1年生の男子300人をランダムに選び，家庭で牛乳を飲んでいるかいないかを尋ねると同時に身長を測定したら，次のような観測度数が得られた．

	148 cm 未満	148 cm 以上	計
牛乳を飲んでいる	65	105	170
牛乳を飲んでない	72	58	130
計	137	163	300

このとき

帰無仮説：“牛乳を飲んでいるかいないかと身長の大小は無関係である”を有意水準 5％ で検定せよ．その結果何がわかるか．

5.4　母集団の平均値に関する検定

例 1　ある県の市街地域の中学1年生の男子（母集団 A）の身長の頻度分布は過去のデータから平均値 146.7 cm の正規分布曲線で近似できることがわかっている．いまこの県の農村地域の中学1年生男子（母集団 B，身長の頻度分布は正規分布曲線で近似できる）から 30 人をランダムに選び，身長を測定した結果，サンプル平均値 143.2 cm，サンプルからの母集団 B の分散の推定値 $(7.63\ \mathrm{cm})^2$ を得た．この結果，農村地域の中学1年生男子の身長は市街地域の中学1年生男子に比べて劣っていると考えられるか．

解　農村地区の中学1年生男子の身長のサンプル平均値 143.2 cm とサンプル分散 $(7.63\ \mathrm{cm})^2$ から母集団 B の平均値の信頼度 95％ の信頼区間は第3章の式（*3.22*）より図 5.7 のようになる．

$$\bar{x}_n \pm t_\alpha\sqrt{\frac{u_n{}^2}{n}} = 143.2 \pm 2.05\sqrt{\frac{(7.63)^2}{30}}$$

$$= 143.2 \pm 2.9$$

この図から明らかなように母集団 A の平均値 146.7 cm を信頼区間の中に含んでない．このことから農村地域の中学1年生男子の

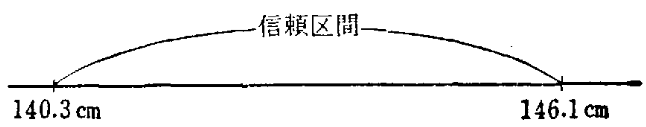

図 5.7　信頼度 95％ の信頼区間

身長は市街地域に比べて劣っているといってよく，母集団 A と母集団 B の平均値の間に差が認められと考える．

これを検定の立場からは次のように考えるが，その意味するところはほとんど同じである．

帰無仮説："母集団 A と母集団 B の平均値が等しい"．

のもとで，第3章の式 (*3.21*)

$$t=\frac{\bar{x}_n-146.7}{\sqrt{\dfrac{u_n{}^2}{n}}}$$

の頻度分布が自由度 $n-1$ の t 分布曲線で近似できることから，この例では

$$t=\frac{143.2-146.7}{\sqrt{\dfrac{(7.63)^2}{30}}}=-2.51$$

となり，この場合，自由度29だから有意水準5％の棄却域 $(-\infty,\ -2.05)$ $(2.05,\ \infty)$ にはいる．したがって帰無仮説が正しいとすると30人のランダムサンプルのサンプル平均値が143.2となる確率は0.05以下であることがわかり，仮説を否定するわけである．有意水準を1％とすると巻末付表5の t-分布表より棄却域は区間

$$(-\infty,\ -2.76)\qquad(2.76,\ \infty)$$

となり，この例では t の値は棄却域にはいらないので帰無仮説を否定することはできない．これは帰無仮説が正しいとすると，30人のランダムサンプルのサンプル平均値が143.2となる確率は0.01以下にはならないことを意味している．有意水準5％で棄却域にはいり，有意水準1％で棄却域にはいらないということは，帰無仮説が正しいときサンプル平均143.2となる確率が0.05より小さく0.01より大きいことを意味している．

注　$t^2=\dfrac{(\bar{x}_n-\mu)^2}{\dfrac{u_n{}^2}{n}}$ の頻度分布が自由度対 $[1,\ n-1]$ の F-分布曲線（補注[12]）で近似できることが知られており，巻末付表6の F-分布表を用いてもよい．

例 2　ある県の市部の12歳の男子（母集団 A）からランダムに選んだ30人の身長の平均値は149.5 cmであり，郡部の12歳の男子（母集団 B）

からランダムに選んだ 20 人の身長の平均値は 145.2 cm であった.

またサンプルからの母集団 A, B の分散の推定値はそれぞれ $(7.82\ \mathrm{cm})^2$, $(7.75\ \mathrm{cm})^2$ であった. もともと母集団 A, B の分散は等しいとみなしてよいことが過去のデータからわかっており, また 2 つの母集団の身長の頻度分布はいずれも正規分布曲線で近似できることもわかっている. このとき母集団 A と母集団 B の身長の平均値を等しいとみなしてよいか.

解 母集団 A, B の身長の頻度分布は正規分布曲線で近似できることから, 3 章の式 (*3.22*) と巻末 t-分布表より, 母集団 A, B の平均値の信頼度 95 % の信頼区間は図 5.8 のようになる.

$$\text{母集団 } A:\quad 149.5 \pm 2.05 \times \sqrt{\frac{(7.82)^2}{30}} = 149.5 \pm 2.93$$

$$\text{母集団 } B:\quad 145.2 \pm 2.09 \times \sqrt{\frac{(7.75)^2}{20}} = 145.2 \pm 3.62$$

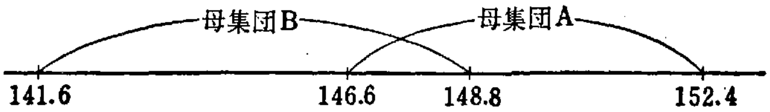

図 5.8 信頼度 95 % の信頼区間

この図をみると, 信頼区間に重複しているところがあり [区間 (146.6, 148.8)], この区間に母集団 A, B の真の平均値がはいっている可能性がある. サンプル平均だけ

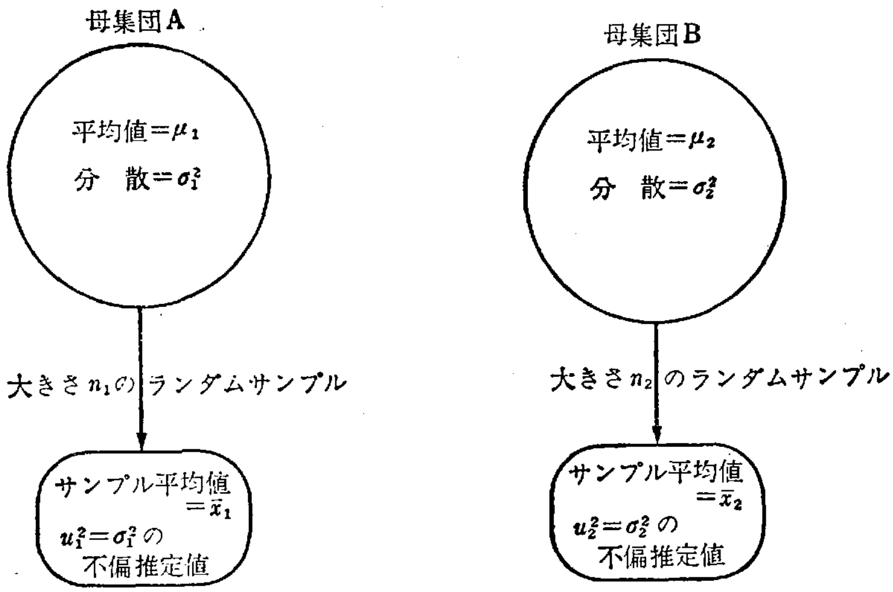

図 5.9 母集団とランダムサンプル

からは母集団 A の平均値の方が大きいように見えるが必ずしもそうはいえないことがわかろう．これを検定の立場から考えてみよう．まず，検定の手順を一般的に書いてみよう（図 5.9 参照）．

(母集団における仮定)

① 母集団 A, B の頻度分布が正規分布曲線で近似できる．

② $\sigma_1{}^2 = \sigma_2{}^2$ とみなせる．

(検定の手順)

(イ) 帰無仮説: "$\mu_1 = \mu_2$" をたてる．

(ロ) 帰無仮説のもとで

$$(5.9)\quad t = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\dfrac{(n_1-1)u_1{}^2 + (n_2-1)u_2{}^2}{n_1+n_2-2} \times \left(\dfrac{1}{n_1} + \dfrac{1}{n_2}\right)}}$$

の頻度分布が自由度 $n_1 + n_2 - 2$ の t-分布曲線（補注 [11]）で近似できる．

(ハ) したがって t の値を母集団 A, B からのランダムサンプルの値から計算し，巻末付表5（t-分布表）からつくった有意水準 $\alpha\%$ の棄却域に t の値がはいるかどうかを見て，もしはいれば有意水準 $\alpha\%$ で仮説を否定する．

さて，この例の場合

$$t = \frac{149.5 - 145.2}{\sqrt{\dfrac{29 \times 7.82^2 + 19 \times 7.75^2}{48} \times \left(\dfrac{1}{30} + \dfrac{1}{20}\right)}} = 1.91$$

となり，巻末付表5より有意水準 5% の棄却域にはいらず仮説: $\mu_1 = \mu_2$ は有意水準 5% では否定できないことになる．

注　この場合も例1と同様

$$t^2 = \frac{(\bar{x}_1 - \bar{x}_2)^2}{\dfrac{(n_1-1)u_1{}^2 + (n_2-1)u_2{}^2}{n_1+n_2-2} \times \left(\dfrac{1}{n_1} + \dfrac{1}{n_2}\right)}$$

の頻度分布が自由度対 $[1, n_1+n_2-2]$ の F-分布曲線（補注 [12]）で近似できることを利用して，巻末付表6の F-分布表を用いてもよい．

〈質問〉　例1，例2ともに母集団の頻度分布が正規分布曲線で近似できることを仮定しているのですが，正規分布曲線で近似できないとき，いい

かえれば母集団の頻度分布がどんな形をしているかわからないときの検定はどのようにすればよいのですか.

答 これは分布型によらない検定法（ノンパラメトリックな検定法）として,

メディアン検定

順位を用いる検定

ウィルコクソンの U 統計量を用いる検定

などが知られていますが, これらについては, たとえば巻末参考書 [9] を参照して下さい. ただ, p. 108 の例 2 においてサンプル数 n_1, n_2 が非常に大きければ母集団の頻度分布が正規分布曲線で近似できなくても, また $\sigma_1^2 = \sigma_2^2$ とみなすことができなくても, 中心極限定理によって $\bar{x}_1$, $\bar{x}_2$ の頻度分布は, それぞれ平均値 μ_1, μ_2, 分散 $\dfrac{\sigma_1^2}{n_1}$, $\dfrac{\sigma_2^2}{n_2}$ の正規分布曲線で近似できますので

帰無仮説: "$\mu_1 = \mu_2$"

のもとで

$$z = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\dfrac{\sigma_1^2}{n_1} + \dfrac{\sigma_2^2}{n_2}}}$$

の頻度分布が平均値 0, 分散 1 の正規分布曲線で近似できることを用いればよろしい. σ_1^2, σ_2^2 は未知なので u_1^2, u_2^2 を代入して, 実際には

$$z' = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\dfrac{u_1^2}{n_1} + \dfrac{u_2^2}{n_2}}}$$

の頻度分布が平均値 0, 分散 1 の正規分布曲線で近似できることを利用するわけです. 多少の代入誤差ははいりますが, サンプルが大きければ無視できますので有意水準 5% の棄却域は, 巻末付表 2 の正規分布表から定まり, z' の値がこの棄却域にはいれば帰無仮説を否定するわけです.

問　　題 14

1. ある工場でつくった電球のうちから 18 個のランダムサンプルをとり出し寿命時間を測定したところ次の結果を得た.

1680	1780	1760	1800	1720	1830	1720	1760
1740	1790	1680	1590	1820	1660	1720	1750
1730	1760						

この工場でつくった電球の平均寿命が 1800 時間であるといっているが，これは認められか．ただしこの工場でつくった電球の寿命時間の頻度分布は正規分布曲線で近似できることがわかっているとする．

2．ある大学1年生の男子の中から 12 人を，2年生の男子の中から 10 人をランダムに選び，50 m 走の記録をとったところ次のようになった．

1年生男子（単位：秒）　6.9, 7.7, 7.3, 7.2, 7.8, 7.6, 7.3, 6.8, 7.4, 7.3, 7.6, 7.7

2年生男子（単位：秒）　7.1, 7.0, 6.7, 7.0, 6.9, 7.5, 7.6, 7.4, 6.9, 6.9

この結果，1年生の男子と2年生の男子の 50 m 走の平均値に有意水準 5% で差があるといえるか．ただし1年生，2年生の 50 m 走の記録の頻度分布は過去のデータから正規分布曲線で近似できることがわかっており，また両者の分散は等しいと考えてよい．

5.5　2つの母集団の比率の差の検定

例　ある市においてテレビの視聴率調査を男女別に行なった．ある番組に対して男性は 200 人のランダムサンプルのうち，20 人が見ていると答えたのに対し，女性は 300 人のランダムサンプルのうち 25 人が見ていると答えた．このとき，男性と女性の視聴率に差があるかどうかを調べる．

解　前節の2つの母集団の平均値の差の検定と考え方は同じであるが，まず男性と女性（2つの母集団）の番組の視聴率 p_1，p_2 の信頼度 95% の信頼区間を第3章の式（*3.29*）によりつくってみよう．（図 5.10），ただし母集団の大きさは非常に大きいと考え，式（*3.29*）で $\dfrac{N-n}{N-1} \fallingdotseq 1$ とみなしている．

$$\text{男性:}\quad \hat{p}_1 \pm 1.96 \times \sqrt{\frac{\hat{p}_1(1-\hat{p}_1)}{n_1}} = 0.1 \pm 0.042$$

$$\text{女性:}\quad \hat{p}_2 \pm 1.96 \times \sqrt{\frac{\hat{p}_2(1-\hat{p}_2)}{n_2}} = 0.083 \pm 0.031$$

ただし　$\hat{p}_1$，$\hat{p}_2$ は男性，女性のサンプル比率，n_1，n_2 は男性，女性のランダムサンプルの数

この図をみると男性と女性の視聴率の信頼区間に重複している部分が多く，サンプル比率だけから男性の方が視聴率が高いということはできない．これを検定の立場から考えてみよう．一般に p_1, p_2 を2つの母集団の比率とし，$\hat{p}_1$, $\hat{p}_2$ をそれぞれのサンプル比率とするとき検定の手順は次のようになる．

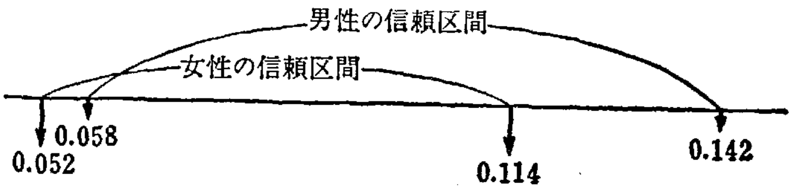

図 5.10　信頼度 95％ の視聴率の信頼区間

（イ）　帰無仮説: "$p_1 = p_2$" をたてる

（ロ）　帰無仮説のもとで，n_1，n_2 がある程度大きいとき
(目安として $n_1, n_2 \geqq 20$)

$$(5.10)\quad z = \frac{\hat{p}_1 - \hat{p}_2}{\sqrt{\hat{p}(1-\hat{p}) \times \left(\dfrac{1}{n_1} + \dfrac{1}{n_2}\right)}},$$

ただし　$\hat{p}_1 = \dfrac{k_1}{n_1}$，$\hat{p}_2 = \dfrac{k_2}{n_2}$，$\hat{p} = \dfrac{k_1 + k_2}{n_1 + n_2}$

の頻度分布が平均値0，分散1の正規分布曲線で近似できる．

（ハ）　したがって z の値を2つの母集団のランダムサンプルから計算し，巻末付表2（正規分布表）からつくった有意水準 α％ の棄却域に z の値がはいるかどうかを見て，もしはいれば有意水準 α％ で仮説を否定する．

さて，この例の場合に式（5.10）から

$$z = \frac{0.1 - 0.083}{\sqrt{\dfrac{45}{500} \times \left(1 - \dfrac{45}{500}\right) \times \left(\dfrac{1}{200} + \dfrac{1}{300}\right)}} = 0.65$$

となる．有意水準 5％ の棄却域は巻末付表2の正規分布表より区間 $[-\infty, -1.96]$，$[1.96, \infty]$ となるから z の値はこの棄却域にはいらない．したがって帰無仮説は有意水準 5％ で否定することはできないと判断をくだすわけである．もちろん図 5.10 を見ただけで直感的にもこの事実はわかるであろう．

〈質問〉 有意水準を 5% とか 1% にする場合が多いのですが，5%，1% を特に使うのは何か理由がありますか．

答　この章でもほとんど有意水準 5% として考えてきましたが，これは信頼区間をつくるときの信頼度と同様で特別な理由はなく，伝統的な基準として使われているにすぎません．目的に応じて有意水準 4%，6% などとしていっこうにかまいません．ただ，巻末付表に有意水準 5%，1% については棄却域をつくるときの数値が示してあるので 5%，1% を用いると便利でしょう．

問　題　15

ある市の 2 つの地区 A，B の住民に対して市立図書館の利用度を調べるために，それぞれ 200 人，150 人のランダムサンプルを選び利用しているかどうかについてたずねたところ，A 地区は 200 人のうち 10 人が，B 地区は 150 人のうち 18 人が利用していると答えた．この結果，A 地区と B 地区の市立図書館の利用度に差があるといえるかどうか，有意水準 5% で検定せよ．ただし，2 つの地区の住民の数はサンプルに比べて非常に大きいとする．