2005. 7. 22

第4単元

連続量に関する推定と検定

目 次

第４単元

連続量に関する推定と検定

単元のねらい　第３単元では，離散量について，統計的検定と区間推定などの考え方と，解析手順を学んだ．そこでは２項分布が主要な分布であった．

この単元では，連続量について学ぶ．ここでは正規分布が基本の分布で，それから導かれるｔ分布，カイ２乗分布，Ｆ分布などが出てくる．これらの分布の一部については，第３単元ですでに出ていたが，詳しい説明は省略された．第４単元を学んだ後，もう一度第３単元を復習すると，理解が深まるであろう．

第３単元と第４単元で説明が重複する箇所があちこちに出てくる．これは，仮説検定，区間推定の基本的な考え方を十分に理解していただくためにあえて，繰り返したものである．

1　母平均に関する推定と検定

1.1　シミュレーション実験

(1)　事例

ある30歳代の男性サラリーマンは，自分の健康管理のために，毎日定時に血圧測定を続けている．その結果の一部を表示1.1に示す．

表示1.1: 血圧測定データ

		時間	最高	最低	備考
1/17		6:43	124	83	
1/18		6:40	117	85	結婚式打合せ
1/19	土	7:05	124	87	
1/20	日	7:28	117	83	＊＊氏結婚式
1/21		6:44	122	86	
1/22		7:15	130	94	朝寝坊した
1/23		6:44	128	86	
1/24		6:42	129	94	名古屋日帰り出張
1/25		6:48	125	89	
1/26	土	7:37	131	91	
1/27	日	7:37	120	88	
1/28		6:44	124	88	
1/29		6:48	128	93	飲み会
1/30		6:42	126	87	
1/31		6:44	119	91	

この人は，表示1.1に書かれているように,「前の晩酒を飲んだ」など，健康状態に影響するかもしれないことや，出勤日と休日のように血圧に影響するかもしれないことを記録している．

血圧は極めて不安定なもので，深呼吸をするだけで変化する．そこで，この人は極力条件を同じにするように注意深く測定をした．

データの取られた期間特に病気をしたという自覚はないが，それでも，血圧

の値は日々変化している．

血圧には収縮期（高い方）と拡張期（低い方）があるが，ここでは，収縮期血圧だけを見ることにする．

このようなデータについて，第2単元で学んだ方法を用いて，平均値と標準偏差を計算すると，

平均　$\overline{x} = 130$

標準偏差 $s = 10$

が得られたとする．

ある日，血圧が少し上がり，145 となった．これから，血圧が上がったと判断して良いであろうか？

あるいは，最近体の調子が少しおかしく感じられ，この1週間（7日間）の血圧の平均値を計算すると 138 となった．これから，血圧が上がったと考え，病院に行くべきだろうか？

この単元では，このような問題に対してどのように答えるかを学ぶ．

■　演習 1　　第1単元 §1.4 で集めた自分の健康管理データについて，平均値よりかなり高い数値があったら，これから学ぶ統計的な判断によって異常か検証してみよ．てもとにデータのない受講生は，第1単元演習12のデータを使用して検証してみなさい．　■

(2)　乱数の生成

第3単元ではコインを使って実験し，離散量の性質を調べた．

第4単元では連続量について同様の実験をする．

ある人の正常時血圧は平均が130で標準偏差が10であるとする．また，血圧の分布は正規分布に従う ものと仮定する．

毎日の血圧測定値の代わりに，このような性質をもつ乱数を生成して，そのデータを使って，上の問題にどう答えるかを考えることにする．

このような性質を持つ乱数が必要な場合に，昔は「正規乱数表」が用いられた．それは，標準正規分布（平均が0で標準偏差が1の正規分布）である乱数を大型コンピュータを用いて求め，それが標準正規分布に従い，また順序に規則性がないことを確かめた結果得られたものであった．

現在では，Excelの分析ツールを用いて比較的簡単に求めることができる．

ここでは，平均が130で標準偏差が10の正規分布に従う乱数を7個求めて1組とし，それを100組生成する．これによって700日分の血圧測定値が得られたことにする．

新しいExcelシートで,「トップメニュー」から「ツール」,「分析ツール」,「乱数発生」を選択すると，表示1.2の「乱数発生」のメニューが表示される．

表示1.2: 分析ツールによる正規乱数の生成

「変数の数」に7を,「乱数の数」に100を入力する．
「分布」の右の▼をクリックして,「正規」を選択する．
「パラメータ」の「平均」に130を，標準偏差に10を入力する．
「出力オプション」で，シートの左上のB6を指定する．
表示1.2は，これらの入力を終了した後の画面である．

「OK」をクリックすると，B6:H105 に乱数が生成される．

データの一部を表示1.3 に示す．

表示1.3: 生成された正規乱数

	1	2	3	4	5	6	7	平均	標準偏差
1	130	117	126	107	130	140	133	126.2	11.1
2	116	132	133	131	130	148	139	132.8	9.8
3	124	138	126	132	119	118	143	128.6	9.4
4	130	126	134	130	132	129	134	130.6	2.9
5	128	127	146	137	115	136	128	131.0	9.8
	以下省略								
平均							129.8	129.8	9.7
標準偏差							10.0	3.8	2.7
最小値							99.3	119.4	2.9
最大値							166.1	141.5	17.3

生成された乱数は小数点以下を含む値であるが，表示1.3 には四捨五入して整数とした値が示されている．

700 個の平均値，標準偏差，最小値，最大値が，表示1.3 の下に求められており，それぞれ，129.8，10.0，99.3，166.1 である．

正規乱数の右に，7 個の乱数の平均値と標準偏差が計算され，さらに下にそれらの平均値などが求められているが，これらの値は後に用いられる．乱数であるから，生成するごとに異なる値が得られる．したがって，受講生が自分で生成した乱数は表示1.3 とは異なるはずである．

(3)　$x \geq 145$ の確率

まず，最初の問題「ある日の血圧が 145 であった．これは平均値 130 よりもかなり高い．異常といえるか」を考えよう．

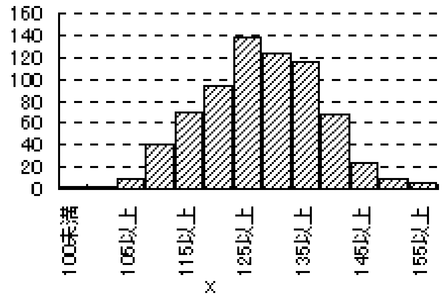
表示1.3 がある日の血圧と考えると，大部分は 145 未満であるが，145 以上の値もあちこちに見られる．これから，健康状態が正常であっても 145 以上になることは稀ではないことが分かる．

これを定量的に表わすために，表示1.3 の 700 個の値の度数表とヒストグラム

を作成すると，表示1.4が得られる（ヒストグラムの作成方法は第2単元の§3.2参照）．

表示1.4: 血圧 x の度数表とヒストグラム

		n
100	100未満	1
105	100以上	2
110	105以上	10
115	110以上	41
120	115以上	70
125	120以上	93
130	125以上	138
135	130以上	123
140	135以上	116
145	140以上	68
150	145以上	24
155	150以上	9
	155以上	5
		700



145以上（145を含む）の個数は38で，割合は$38/700 = 0.054$である．

これから，健康状態が正常でも血圧が145以上になることは5% 以上あることが分かる．第3単元で学んだ統計的な判断によると，これから，異常であるとはいえないことになる．

血圧 x が145であるということは，期待値 $\mu = 130$ から標準偏差 $\sigma = 10$ の1.5倍離れていることになる．すなわち，

$$z = \frac{x-\mu}{\sigma} = \frac{145-130}{10} = 1.5$$

である．1.5は，第2単元の用語を使えば,「ある日の血圧の偏差値が1.5であった」といえる．

標準正規分布において1.5以上である確率は，

```
=NORMSDIST(-1.5)
```

で求められる．NORMSDIST(z) は z 以下である確率を求める関数であるから，括弧内の数値に－をつけてある．この関数で計算した結果は0.0668である．

第3単元で学んだp値という用語を使うと,「上側のp値が 0.0668 である」ということになる．p値が 0.05 よりも大きいから，帰無仮説（H_0：$\mu = 130$）を棄却することはできない.

もし，血圧が 130 から 15 以上 下がった，すなわち，115 以下になった（$z = -1.5$）場合にも，健康状況が変化したと判断する場合は，上に求めたp値 0.0668 を2倍した，両側のp値（0.1336）で判断しなければならない.

■　演習 2　　シミュレーションで得られた値 0.054 と，正規分布から計算した値 0.0668 との差は，シミュレーションで生成した乱数が有限（700 個）であるために生じたものである．受講生の得た値とはどのくらい差があるかを調べてみよう．この差は乱数の個数を増やせば小さくなるであろう.

第3単元で学んだ結果を利用して，700 個の乱数の内145 以上の個数の期待値と標準偏差を計算してみよ．また，個数の約95% が含まれる区間（期待値±1.96*標準偏差）を求め，シミュレーション結果がこの区間に含まれているか確認せよ.　■

(4)　$\overline{x} \geq 138$ の確率

次の問題は「最近の7日間の血圧の平均が 138 である．これから，血圧が正常時よりも高くなったといえるか」であった.

表示1.3 の右の欄に，7 個の値の平均値が，下にその平均値などが計算されており，100 個の平均値は，119.4〜141.5 の範囲に分布し，標準偏差は 3.8 である.

前と同様，平均値の度数表とヒストグラムを作成すると，表示1.5 が得られる.

平均値の最大値は 141.5 であって，138 以上は100 個中わずか 2 個に過ぎない.

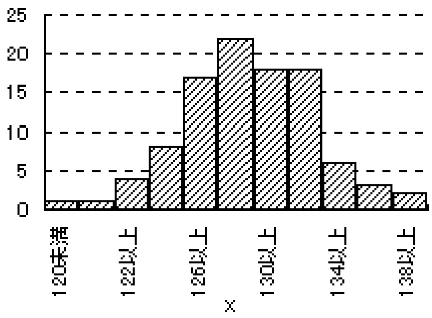
これから 7 日間の平均値が138 以上になったら，健康状態に異常があると判断しても良さそうである.

個々の血圧値では 15 の上昇には「有意差なし」とされたが，7 個の平均値ではその約半分の 8 の上昇で「有意差あり」という結論となった.

このような違いの原因は，表示1.4 と表示1.5 の2つのヒストグラムの目盛り

表示 1.5: 7 日間の平均血圧 $\overline{x}$ の度数表とヒストグラム

		n
120	120未満	1
122	120以上	1
124	122以上	4
126	124以上	8
128	126以上	17
130	128以上	22
132	130以上	18
134	132以上	18
136	134以上	6
138	136以上	3
	138以上	2
		100



の値に注意して比較すれば一目瞭然であるように，平均値の方がバラツキが小さくなっているためである．

シミュレーションでは100回中2回であったが，前項の最後に見たように，これはシミュレーションの回数が100回の結果であるから，精度が不十分である．以下 理論的に 平均値が 138以上になる確率を求めよう．

第1単元で学んだ誤差法則によると，n 個の平均 $\overline{x}$ の標準偏差 $\sigma_{\overline{x}}$ は個々の値の標準偏差 σ_x の $1/\sqrt{n}$ になる．この場合に当てはめると，

$$\sigma_{\overline{x}} = 10/\sqrt{7} = 3.780$$

となる．

この値を使って，7日間の平均値138と正常時の値 130 の基準化した距離は

$$z = \frac{138 - 130}{3.780} = 2.117$$

となる．標準正規分布で 2.117 以上の確率は 0.012 である．これは， 0.05 よりも小さいから，血圧が正常時よりも上昇したと判断できる．

■ 演習3 7日間の平均値が $\overline{x} = 135$ のとき，上の説明がどのように変化するかを計算して確かめよ． ■

以上，身近な例を取り上げて，x または $\overline{x}$ が異常かどうかを判断する基本的な考え方を説明した．

次節以降では，第3単元で学んだ，仮説検定，帰無仮説，棄却域，有意水準 α，p値などの統計用語を使って，詳しく説明する．

本日のまとめ

乱数によってデータの分布さらには平均の分布を実際に作成し，その分布がどのようになるかを体験できたと思う．このようにばらついているものから正しい判断をするための情報を得ることが統計的方法を利用することの主要な目的の一つである．今日は，この基本的な考え方を身につけてほしい．

1.2　検定と区間推定（σ既知の場合）

(1)　例

シャープペンシル用の芯を作っているある工場では，製造した芯の太さの平均値 μ (母平均) が $0.90\mathrm{mm}$ になるようにしないと，シャープペンシルに合わない芯が出てお客から苦情が出るので，母平均が $0.90\mathrm{mm}$ でなくなったときは機械を止めて調整することにしている．

これまでに得られたデータから，同じ日に作った芯の太さの標準偏差は $\sigma = 0.03\mathrm{mm}$ である ことが分かっている．これを,「σ既知の場合」という．

母平均が $0.90\mathrm{mm}$ になっているかどうかは，毎日できあがったたくさんの芯の中から，無作為に25本を抽出して，その太さを測り，その平均値 $\overline{x}$ を計算して調べている．

このとき，たとえ，母平均 μ が $0.90\mathrm{mm}$ で，標準偏差に変化がない状態であっても，それから得られた標本平均 $\overline{x}$ がちょうど 0.90 になるとはかぎらない．したがって，$\overline{x} = 0.91$ となったからといって，すぐ機械を止めて調整し直すとしたら，余計な手数をかけてむだなことをすることになるかもしれない．このよ

うな誤りを犯す危険のほとんどないような判断をするには，統計的検定が必要となる．

血圧が高めで，高血圧症にならないように注意している人が，最近の数日間の平均血圧が平常血圧から上昇したかどうかを確かめたい場合も，統計的な検定の考え方が適用できる．この場合は，最近の血圧が平常血圧から大きい方に外れた場合が問題である．

(2) 考え方

検定の基本的な考え方は，第3単元で学んだ離散量の場合と同じである．

まず，帰無仮説を決める．ここでは，$\mathrm{H}_0:\ \mu = 0.90$ である．

帰無仮説が正しいとき，$\overline{x}$ がどのような分布になるかを考え，起こりにくい領域（棄却域）を決める．

n 個の観測値の平均値 $\overline{x}$ は，x の分布が正規分布ではなくても，中心極限定理により正規分布に近づく．また，その標準偏差は，誤差法則により，$\sigma/\sqrt{n}$ である．

$\overline{x}$ のように推定に用いられる量を 推定量 (Estimator) といい，推定量の標準偏差を その推定量の 標準誤差 (Standard Error) と呼び，個々の値の標準偏差 (Standard Deviation) と区別することがある．以下 この値を se(推定量) で表わすことにする．

$$se(\overline{x}) = \frac{\sigma}{\sqrt{n}} \tag{1.1}$$

期待値から $\overline{x}$ までの距離をその標準誤差で割って基準化した

$$z = \frac{\overline{x}-\mu}{se(\overline{x})} = \frac{\overline{x}-\mu}{\sigma/\sqrt{n}} \tag{1.2}$$

z は標準正規分布になる．

標準正規分布が ± 1.96 の範囲内に含まれる確率が 0.95 であることは既に学んだ．すなわち，

$$-1.96 < z < 1.96 \tag{1.3}$$

$$\mu - 1.96se(\overline{x}) < \overline{x} < \mu + 1.96se(\overline{x}) \tag{1.4}$$

の範囲外に出る確率が0.05 である．これから，式(1.4) の範囲外が，$\alpha = 0.05$ とする仮説検定の棄却域となる．

それでは，$\sigma = 0.03$ が既知で，$\mu = 0.90$, $n = 25$, $\overline{x} = 0.91$ の場合について計算してみよう．

$$se(\overline{x}) = \frac{0.03}{\sqrt{25}} = 0.006$$
$$z = \frac{0.91 - 0.90}{0.006} = 1.667$$

で，

$$-1.96 < z < 1.96$$
$$0.90 - 1.96 \times 0.006 < \overline{x} < 0.90 + 1.96 \times 0.006$$
$$0.888 < \overline{x} < 0.912$$

の範囲外が棄却域となる．

25 個の平均値 $\overline{x} = 0.91$ は上の範囲内である（棄却域には入らない）ので，このデータから「芯の太さの母平均 μ が 0.90 ではなくなった」とはいえない．

また，$|z| = 1.667$ が 1.96 以下であることからも，同様の判断が下せる．

血圧の場合は，平均血圧 $\overline{x}$ が

$$\overline{x} < \mu(\text{平常血圧}) + 1.645 se(\overline{x})$$

の限界を超えたときに，血圧が有意に上昇したと結論できる．これが片側検定であることは，第3単元で既に学んだ．

母平均が0.90mm であれば，z は $\mathrm{N}(0, 1)$ に従うから，$|z| \geq 1.667$ となる確率は 0.096 である．これが，第3単元で説明した**p**値である．

p値は，標準正規分布が $|z|$ 以上の確率を2倍したもので，

```
=2*NORMSDIST(-ABS(Z))
```

という式で計算できる．`ABS` は絶対値を取る関数である．2 倍しているのは両側検定だからである．

棄却域に含まれるかどうか（有意かどうか）だけで結果を表わすよりも，**p**値で定量的に表わした方が，技術的判断に役立つ．

(3) Excelによる計算表

前項の計算では，個々の値は示さず，n, $\overline{x}$, σ だけが与えられて計算したが，現実の問題では，個々の値から平均値を計算することになる．

そこで，個々の値から前項の解析を自動的に実行するExcelの計算表を表示1.6に示す(入力項目を太字で表わす)．

表示1.6: 母平均の仮説検定

	A	B	C	D
3	σ既知	（z検定）		
4		x		
5	1	**158**		
6	2	**143**		
7	3	**139**		
8	4	**143**		
9	5	**135**		
10	6	**128**		
11	7	**118**		
12	8			
13	9			
14	10			
15				
16	個数	7	**25**	=COUNT(B5:B14)
17	平均	137.7	**0.91**	=AVERAGE(B5:B14)
18	H0: μ	**130**	**0.90**	
19	σ	**10.0**	**0.03**	
20	母平均の検定			
21	標準誤差	3.780	0.006	=B19/SQRT(B16)
22	z	2.041	1.667	=(B17-B18)/B21
23	α（両側）	**0.10**	**0.05**	
24	z(α)	1.645	1.960	=NORMSINV(1-B23/2)
25	棄却（下限）	123.783	0.888	=B18-B24*B21
26	（上限）	136.217	0.912	=B18+B24*B21
27	p値（下側）	0.979	0.952	=NORMSDIST(B22)
28	（上側）	0.021	0.048	=1-NORMSDIST(B22)
29	（両側）	0.041	0.096	=2*MIN(B27:B28)
30	母平均の区間推定			
31	下限	131.497	0.898	=B17-B24*B21
32	上限	143.931	0.922	=B17+B24*B21

30行目以下は次項で説明する．

(i) `B5:B14` にデータを入力する．ここでは，§1 の血圧データのシミュレー

ションから，任意の1行の値を入力した.

(ii)　`B16:B17` に観測値の個数と平均値が計算される.

(iii)　`B18:B19` に，帰無仮説に設定する母平均の値と，母標準偏差を入力する.

(iv)　`B21` に平均値の標準誤差が，`B22` に z が計算される.

(v)　`B23` に両側のαを入力する．血圧のデータは片側検定であるから，片側のαが 0.05 になるように，0.10 を入力する.

(vi)　`B25:B26` に下側と上側の棄却限界値が計算される.
血圧の例の場合は，片側検定であるから，$\overline{x}$ が上側の棄却限界値 136.2 を超えたとき，血圧が有意に上昇したといえる．$\overline{x} = 137.7$ はこの限界を超えている.

(vii)　`B27:B28` に下側と上側のp値が計算される．片側検定の場合は，このいずれかを用いて判断する．p値（両側）は，下側と上側のp値の内，小さいほうを2倍したものである．両側検定の場合は，これを使って判断する.

右には，B列のセルに入力された式が示されている.

n が10以内のときは，この計算表のデータを入れ替えるだけで，結果が自動的に再計算される.
n が11以上のときは，13行と14行の間に行を増やすと，個数と平均を求める範囲が自動的に広げられる.

シャープペンシルの芯の例の場合のように，n, $\overline{x}$ が別途に計算されている場合は，C列に示すように，`C16:C17` に n, $\overline{x}$ を入力する．この例は両側検定であるから，α（両側）には 0.05 を入力する．p値（両側）は 0.096 である.

(4)　母平均の区間推定

第3単元で2項分布の割合についての区間推定を学んだ.
連続量の場合も，母平均の信頼区間を求めることは極めて重要である.

$\overline{x}$ を挟む不等式 式(1.4) を，μ が2つの不等号で挟まれるように移項すると

$$\overline{x} - 1.96se(\overline{x}) < \mu < \overline{x} + 1.96se(\overline{x}) \tag{1.5}$$

が導かれる．

この不等式を使って，母平均の信頼区間を求めることができる．これは，信頼率が 95%である．信頼率は，係数の 1.96 を変えることにより，任意に変更できる．

シャープペンシルの芯の太さの区間推定は

$$0.91 - 1.96 \times 0.006 < \mu < 0.91 + 1.96 \times 0.006$$
$$0.898 < \mu < 0.922$$

となる．

片側検定に対応して，片側の信頼区間を求めると

$$\overline{x} - 1.645se(\overline{x}) < \mu \qquad \text{または}$$
$$\mu < \overline{x} + 1.645se(\overline{x}) \tag{1.6}$$

となる．

血圧データの場合は，片側検定で，平常値 130 よりも大きい方に外れているかどうかに関心があったので，式(1.6) の上の式が用いられる．

これらの計算は，表示1.6 の 30行以下で実行されている．

31, 31行に 下側と上側の信頼限界が求められる．片側のときは,「α（両側)」を 0.10 とし，どちらか一方のみを利用する．

$$131.5 < \mu \qquad \text{血圧の片側区間推定} \tag{1.7}$$
$$0.898 < \mu < 0.922 \qquad \text{シャープペンシルの芯の両側区間推定} \tag{1.8}$$

式(1.7) から，平常血圧 μ が 131.5 よりも大きいことを示している．この中には帰無仮説の値 130 は含まれていない．これは，片側の仮説検定の結果に対応する．

また，仮説検定では変化したかどうかだけを示すだけであるが，区間推定では，少なくとも $131.5-130=1.5$ 血圧が上昇したことを示してくれる．

■　演習4　　成分含有量の目標値を 1.54mg として，新しい錠剤を試験的に作り，そこから無作為に10個の錠剤を抽出し，分析した結果，次のデータが得られた．

`1.41 1.42 1.64 1.39 1.51 1.42 1.44 1.49 1.59 1.54`

$\sigma=0.08$ が既知として，帰無仮説 $\mathrm{H}_0:\ \mu=1.54\mathrm{mg}$ を検定せよ．
また，μ の95%信頼区間を求めよ．　■

芯の太さの母平均の信頼区間には基準値 $\mu=0.90$ が含まれており，これからは，製造工程に異常があるとはいえない．

仮説検定と区間推定の間には，次の関係がある．

仮説検定	区間推定
仮説が棄却される	信頼区間に帰無仮説が含まれない
仮説が棄却されない	信頼区間に帰無仮説が含まれる

仮説検定は，yes-no の結果しか与えないが，区間推定は定量的な値が得られるので，できるだけ，信頼区間を用いて判断し，報告書にも記載するのが良い．

これは，第3単元でも説明したが，大変重要であるので，繰り返して記した．

芯の太さの場合，このデータから太さが目標から外れているとはいえない（有意差がない）という結論となったけれども，区間推定を見ると，もしかすると太さの真の平均は 0.922 かもしれない．これを見逃すのは品質管理の面で問題であるとすると，無視することはできない．
この問題については§1.4 で取り上げる．

本日のまとめ

今日は，シャープペンシル用の芯を例に取り，芯の太さの母平均についての仮説検定と区間推定を学んだ．芯の太さの母標準偏差が既知である，すなわち「σ既知」であることに注意が必要である．多くの場合，「σ既知」であることは

少ないと思うが，基本を学ぶということから，今日はそのように設定してある．明日以降では，より一般的な「σ未知」の場合を学習する．

1.3　検定と区間推定（σ未知の場合）

(1)　t分布

§1.2 では，$\overline{x}$ の標準誤差 $se(\overline{x})$ を$\sigma/\sqrt{n}$ として，

$$z = \frac{\overline{x} - \mu}{\sigma/\sqrt{n}} \tag{1.9}$$

を用いて検定した．

現実には，真の標準偏差 σ は未知である．そこで，データから計算したサンプルの標準偏差 s で代用せざるを得ない．

式(1.9) の分母の σ を s に置き換えた値を t で表わすことにする．

$$t = \frac{\overline{x} - \mu}{s/\sqrt{n}} \tag{1.10}$$

z は標準正規分布に従い，その標準偏差は 1 である．t は分母もばらつくので，標準偏差は 1 よりも大きくなる．

表示1.3の乱数の各行には標準偏差 s を求められているので，それを用いて式(1.10)で t を計算する．同じシミュレーションデータについて z を計算した結果と比較すると下の結果が得られた．s の行には，表示1.3の下に求められている標準偏差 s の平均値などが転記されている．

	平均値	標準偏差	最小値	最大値
z	−0.04	0.99	−2.80	3.04
t	−0.04	1.25	−5.89	4.55
s	9.7	2.7	2.9	17.3

式(1.10)の分母の s の平均値は $\sigma = 10.0$ に近いが，標準偏差が 2.7 (変動係数が約 30%）で，2.9〜17.3 の広範囲にばらついていることが分かる．そのた

めに，t は z よりもバラツキが大きくなり，t の標準偏差 1.25 は，z の標準偏差 0.99 （理論値は 1.00）よりも大きくなる．

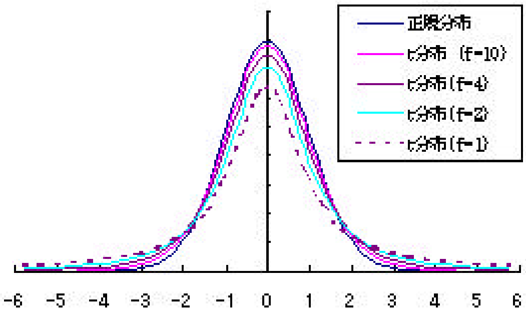
どのくらい大きくなるかは分母の s の精度に依存する．

s の精度については§3 で詳しく説明するが，s を求めるもととなった平方和 S の自由度 $f = n - 1$ によって決まる．

t の標準偏差は $\sqrt{f/(f-2)}$ となることが理論的に導かれている．自由度が2以下のとき，t は長く裾を引き，標準偏差は無限大になる．$n = 7, f = 6$ のとき，$\sqrt{6/(6-2)} = 1.22$ となる．

t の分布を **t 分布** と呼ぶ．t 分布の形は自由度によって変わり，自由度が大きくなると標準正規分布に近づく．その様子を表示1.7に示す．

表示1.7: t 分布の自由度による変化



(2)　検定

表示1.6 (p.13) の血圧データを，σが未知として解析する手順を表示1.8 に示す．

表示1.6 と比較し，σ既知の場合とどこが違うかを確認しながら読み進めてほしい．

B5:B18 はσ既知の場合と全く同じである．σが未知であるから，データから推定する．第2単元で学んだ手順に従って，平方和，自由度，平均平方を経て，

表示1.8: 平均値の検定と区間推定(σ 未知)

	A	B	C
3	σ未知	(t検定)	
4		x	
5	1	**158**	
6	2	**143**	
7	3	**139**	
8	4	**143**	
9	5	**135**	
10	6	**128**	
11	7	**118**	
12	8		
13	9		
14	10		
15			
16	個数	7	=COUNT(B5:B14)
17	平均	137.7	=AVERAGE(B5:B14)
18	H0: μ	**130.0**	
19	標準偏差の推定		
20	平方和	959.4	=DEVSQ(B5:B14)
21	自由度(f)	6	=B16-1
22	平均平方	159.9	=B20/B21
23	標準偏差	12.6	=SQRT(B22)
24	母平均の検定		
25	標準誤差	4.779	=B23/SQRT(B16)
26	t	1.614	=(B17-B18)/B25
27	α（両側）	**0.10**	
28	t(f, α)	1.943	=TINV(B27,B21)
29	棄却（下限）	120.713	=B18-B28*B25
30	（上限）	139.287	=B18+B28*B25
31	p値（下側）	0.921	=IF(B26>0,1-TDIST(B26,B21,1),TDIST(-B26,B21,1))
32	（上側）	0.079	=1-B31
33	（両側）	0.158	=TDIST(ABS(B26),B21,2)
34	母平均の区間推定		
35	下限	128.427	=B17-B28*B25
36	上限	147.002	=B17+B28*B26

標準偏差の推定値 $s = 12.6$ を求める．その過程が B19:B23 である．

前と同じ方法で，平均値の標準誤差を計算する．

式(1.10) によって，t値を計算する．

$$t = \frac{\overline{x} - \mu}{s/\sqrt{n}} = \frac{137.7 - 130.0}{12.6/\sqrt{7}} = \frac{7.7}{4.779} = 1.614$$

計算された t の値が有意かどうかは，t分布の両側確率が α となる t の%点

で判定する．

両側確率が5％および片側確率が5％（両側確率が10％）に対する t の％点表の一部を示す．

自由度	1	2	3	6	9	20	60	∞
t(0.05)	12.71	4.303	3.182	2.447	2.262	2.086	2.000	1.960
t(0.10)	6.314	2.920	2.353	1.943	1.833	1.725	1.671	1.645

表の値を見ると，自由度が∞のときは標準正規分布の値 1.960, 1.645 と一致しており，自由度が小さくなると値が大きくなる．特に，自由度が10以下になると急激に大きくなることが分かる．これから，自由度が小さい（n が小さい）とき，$|\overline{x}-\mu|$ がかなり大きくないと，有意差を見出しにくくなることが分かるであろう．

ｔ分布の両側の確率が α となる値は

=TINV(両側確率，自由度)

で求められる．

上に得られた $|t|=1.614$ は自由度 $f=7-1=6$ に対するｔ分布の片側5％点 1.943 よりも小さいから，帰無仮説は棄却できない（血圧が平常値から有意に外れているとはいえない）．

次に，ｐ値を使う仮説検定について説明する．

ｔ分布が |t| よりも大きい両側確率は

=TDist(|ｔの値|，自由度，2)

で求められる．最初のパラメータは，正の値でなければならない．最後のパラメータの 2 は両側確率を求めることを指定する．

片側確率を計算するには，ｔの値の正負によって，計算式を使い分ける必要があり，工夫が必要である．下側確率を求める計算式

=IF(ｔ>0,1-TDIST(ｔ,自由度,1),TDIST(-ｔ,自由度,1))

とする．上側確率は１から下側確率を引いて求められる．

以上の計算が，表示1.8 の B31:33 に示されている．

上側のｐ値 0.079 は 0.05 よりも大きいから，血圧が有意に上昇したとはいえない．

(3)　区間推定

σ 既知の場合と同様の考えで，母平均の区間推定ができる．

式(1.5) (p.15) の 1.96 を $t(f,\ 0.05)$ に，式(1.6) の 1.645 を $t(f,\ 0.10)$ に置き換えれば良い．

$$\overline{x} - t(f,\ 0.05)se(\overline{x}) < \mu < \overline{x} + t(f,\ 0.05)se(\overline{x})$$

$$\overline{x} - t(f,\ 0.10)se(\overline{x}) < \mu$$

$$\mu < \overline{x} + t(f,\ 0.10)se(\overline{x})$$

血圧データの場合の片側信頼区間は表示1.8 のB35 に求められており，

$$128.4 < \mu$$

となる．σ既知の場合とことなり，信頼区間に帰無仮説の値 130 が含まれており，上の検定の結果と一致する．

■　演習５　　演習４で，σ 未知として同様の解析をせよ．　■

本日のまとめ

昨日と同じ母平均についての仮説検定と区間推定であるが，今日は「σ 未知」の場合である．ここで，ｔ分布が登場し，それに基づく仮説検定（ｔ検定と呼ばれる）と区間推定を学んだ．基本的には，昨日の正規分布のところをｔ分布に置換えて考えれば良い．本日の内容を十分に理解できないときには昨日の内容を復習することをお勧めする．

1.4　検出力と n の決め方

(1)　仮説検定，区間推定の使い方

芯の太さの例で，$n = 25$ のサンプルの平均が 0.91 であるとき，帰無仮説 H_0： $\mu = 0.90$ を棄却することができなかった．これは，帰無仮説を肯定する（認める）ものではない．帰無仮説を否定する十分な証拠がないというだけである．

区間推定をすると，$0.898 < \mu < 0.922$ であって，$\mu = 0.90$ である可能性を否定することはできない．

このような場合には，次のように考えるのが良いであろう．ただし，両側検定，両側区間推定とする．

信頼区間の両側の値 $(0.898, 0.922)$ が本当の値であったとしたとき，どのようなアクション（対応）を取るべきかを考える．

下限の 0.898 は 目標の 0.90 に極めて近いから，何もする必要がない．しかし，上限の 0.922 は，目標から 0.022（標準偏差 0.03 の半分以上）離れているから，装置の再調整を必要とするかもしれない．もし，そうなら，有意差なしと済ませてしまうことはできないであろう．

このように，信頼区間の幅が広くて，両側で対応が異なる場合は，信頼区間の幅を狭くすることを検討する必要がある．

(2)　信頼区間の幅を考慮した n の決め方

以下，x の標準偏差 σ が既知である場合を考える．

95%信頼区間の半幅（信頼区間の 上限−下限 の半分）は

$$\text{信頼区間の半幅} = 1.96 se(\overline{x}) = 1.96 \frac{\sigma}{\sqrt{n}} \tag{1.11}$$

である．これを変形すると

$$n = \left(\frac{1.96\sigma}{\text{信頼区間の半幅}} \right)^2 \tag{1.12}$$

が得られる．この式に σ と信頼区間の希望する半幅を代入すると，必要とする n の近似値が得られる．

例えば，シャープペンシルの例で，$\sigma = 0.03$ で信頼区間の半幅を 0.01 にしたい場合は，

$$n = \left(\frac{1.96 \times 0.03}{0.01}\right)^2 = 5.88^2 = 34.57$$

となる．端数を切り上げた $n = 35$ が解となる．

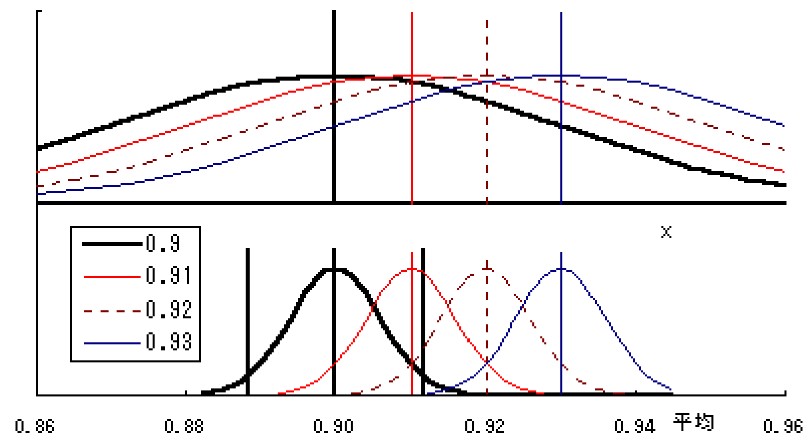
以上の計算は σ が既知であるとしている．

σ が未知である場合，式(1.11) は，σ がその推定値 s に変るため，一定数ではなくなる．また，1.96 は t 分布の%点になる．そのために，式(1.12)は正しくない．必要な n の概数を知るためにのみ使う．

(3) 検出力

製造工程に異常が発生し，芯の太さの母平均 μ が 0.90 から0.91，0.92 または 0.93 に変化した場合を考える．このような場合に，芯の太さ x と 25 本の芯の太さの標本平均 $\overline{x}$ の分布を示したのが，表示1.9 である．

表示1.9: x と $\overline{x}$ の分布(σ 既知)



上は x の分布（$\sigma = 0.03$）である．左の太線は $\mu = 0.90$ の場合の分布である．μ が $0.01, 0.02, 0.03$ だけ大きくなると，分布は右に移動する．このグラフから，母平均が標準偏差の $1/3, 2/3, 1$ だけ変化すると，分布がどのくらい変化するのかを知ることができるであろう．

下は $\overline{x}$ の分布（$\sigma/\sqrt{n} = 0.006$）である[1]．

左の太線は $\mu = 0.90$ の場合の分布である．

両側の太い縦線は（$\overline{x} = \mu \pm 1.96\sigma/\sqrt{n} = (0.888, 0.912)$ に引かれている．これは，帰無仮説 $\mu = 0.90$ のとき，$\overline{x}$ が含まれる確率が $1 - \alpha = 0.95$ となる範囲であり，この範囲外が棄却域である．

$\mu = 0.91$ の場合，平均値の分布のほぼ半分（正確には 38.5%）が棄却域の中である．したがって，$\mu = 0.91$ のときに，帰無仮説H_0：$\mu = 0.90$ が棄却されるのはほぼ半分である．すなわち，平均値が目標から外れたことが発見できる確率は約50%となる．

$\mu = 0.92$ の場合の分布はかなりの部分（正確には 91.5%）が棄却域に属し，帰無仮説が棄却される（平均値の変化が見つかる）確率はかなり高く，帰無仮説が棄却されない（平均値の変化を見逃す）確率はそれほど大きくない．

$\mu = 0.93$ の場合は分布のほとんど全部（正確には 99.9%）が棄却域に含まれ，母平均の変化を見逃すことはないであろう．

上に「平均値が目標値から外れたことが発見できる確率」とか,「平均値の変化が見つかる確率」という表現を使ったが，統計用語では,「帰無仮説が成立しないことが 検出された」 という．

第3単元で説明したように，検出される確率のことが 検出力，検出されない（差を見逃した）確率が 第2種の誤りの確率 β である．

検出力については §2.3 でさらに詳しく説明する．

[1] 平均値の分布は，標準偏差が $\sqrt{n}$ 分の1になるので，分布の高さは $\sqrt{n}$ 倍になる．表示1.9 は，図の大きさを配慮して，分布の高さを x の分布の高さに揃えている．

本日のまとめ

今日は，区間推定および仮説検定それぞれの観点から，標本数 n の決め方を学んだ．第 3 単元で学習した「検出力」が再度登場するが，今後も出てくるので忘れてしまった人は，第 3 単元の該当個所を復習してほしい．

2　母平均の差に関する推定と検定

2.1　母平均の差の推定と検定 (σ 既知の場合)

§1.2では，無作為に抽出した芯の太さを測り，その平均値 $\overline{x}$ から母平均 μ が 0.90mm といえるかどうかを検定したり，その母平均の値を推定する方法を学んだ．ここでは，それぞれ無作為に抽出した2つの標本の平均値 $\overline{x}_A$ と $\overline{x}_B$ から，それらに対応する母平均 μ_A と μ_B に差があるかどうかを検定したり，その差を推定する方法について学ぶことにしよう．

(1)　例

A, B 2社の同種の袋詰め製品を店頭から無作為に抽出して，正味の重量(g)を測定した．このような調査の場合は両社から同数個を抽出するのが普通であるが，同数でない場合にも解析できることを示すため，A社から10袋，B社から7袋を抽出し下表の結果が得られた．

従来の調査から，同じ会社の製品内での標準偏差が 0.8 であることが分かっている．

A, B 2社の製品の重量の母平均に差があるかどうかを検定したり，その差がいくらかを推定してみよう．

	データ										平均
A社	49.6	48.6	48.3	47.9	48.9	48.4	48.1	47.5	48.5	47.6	48.34
B社	47.3	48.2	47.6	48.4	46.6	47.7	47.0				47.54

(2)　考え方

母平均の差 $\mu_A - \mu_B$ の点推定値としては，$\overline{x}_A - \overline{x}_B = 0.80$ であり，これにもとづいて推定と検定をすることになる．そのためにはまず，$\overline{x}_A - \overline{x}_B$ の分布

を知ることが必要であるが，$\overline{x}_A - \overline{x}_B$ の期待値が $\mu_A - \mu_B$ であることは直感的に明らかであろう．

問題は，$\overline{x}_A - \overline{x}_B$ の分散を求めることであるが，$\overline{x}_A$, $\overline{x}_B$ はそれぞれ n_A個，n_B個のデータの平均値であるから，それらの分散は σ_A^2/n_A, σ_B^2/n_B である．

次に $\overline{x}_A$ と $\overline{x}_B$ は，互いに独立な標本から求められたものであるから，その差 $\overline{x}_A - \overline{x}_B$ の分散は，両者の分散の和であることが分かる，よって，

$$\mathrm{V}[\overline{x}_A - \overline{x}_B] = \frac{\sigma_A^2}{n_A} + \frac{\sigma_B^2}{n_B} \tag{2.1}$$

を得る．

平均値の差 d の標準誤差は式(2.1)の両辺の平方根を取って

$$se(d) = \sqrt{\frac{\sigma_A^2}{n_A} + \frac{\sigma_B^2}{n_B}} \tag{2.2}$$

となる．

さらに，2つの母集団の分布が正規分布ならば，$\overline{x}_A - \overline{x}_B$ の分布もまた正規分布で，$\mathrm{N}\left(\mu_A - \mu_B,\ \dfrac{\sigma_A^2}{n_A} + \dfrac{\sigma_B^2}{n_B}\right)$ になる．x_A, x_B の分布が正規分布でなくとも，中心極限定理によって n_A, n_B が大きければ，$\overline{x}_A - \overline{x}_B$ はほぼ正規分布に従うことはすでに学んだ通りである．そこで，

$$z = \frac{(\overline{x}_A - \overline{x}_B) - (\mu_A - \mu_B)}{se(d)} \tag{2.3}$$

と規準化すれば，z の分布は $\mathrm{N}(0,\ 1)$ になるから，§1.2 で $z = \dfrac{\overline{x} - \mu}{\sigma/\sqrt{n}}$ から $\overline{x}$ についての推定や検定を導いたときと同じ考え方で，差 $\overline{x}_A - \overline{x}_B$ の検定や推定の方法を導くことができる．

(3)　母平均の差の検定

例の場合の帰無仮説は，「2つの母平均に差がない」とする．すなわち，$\mu_A = \mu_B$ $(\mu_A - \mu_B = 0)$ が成り立つならば，

$$z = \frac{\overline{x}_A - \overline{x}_B}{se(d)} \tag{2.4}$$

が， $\mathrm{N}(0,1)$ **に従うから，** $|z| \geq 1.96$ **となる確率は**0.05**である．したがって標本から求めた** $\overline{x}_A,\ \overline{x}_B$ **の値をこの式に代入して** z **を計算し，** $|z| \geq 1.96$ **となったとき，有意水準** $\alpha = 0.05$ **で** $\mu_A = \mu_B$ **の仮説を否定して** $\mu_A \neq \mu_B$ **と判定することができる．または，p値を求めて結論を求めることができる．**

例についてExcelで計算した結果を表示2.1に示す．

表示2.1: **平均値の差の検定と区間推定**(*σ* **既知**)

	A	B	C	D	E
4		A	B		
5	1	**49.6**	**47.3**		
6	2	**48.6**	**48.2**		
7	3	**48.3**	**47.6**		
8	4	**47.9**	**48.4**		
9	5	**48.9**	**46.6**		
10	6	**48.4**	**47.7**		
11	7	**48.1**	**47.0**		
12	8	**47.5**			
13	9	**48.5**			
14	10	**47.6**			
15					
16	個数(n)	10	7	差	
17	平均	48.340	47.543	0.797	
18	H0: μ1-μ2	**0.0**			
19	σ	**0.8**	**0.8**		
20	母平均の差の検定				
21	差の標準誤差	0.394	=SQRT(B19^2/B16+C19^2/C16)		
22	z	2.022	=D17/B21		
23	α（両側）	**0.05**			
24	z(α)	1.960	=NORMSINV(1-B23/2)		
25	棄却（下限）	-0.773	=B18-B24*B21		
26	（上限）	0.773	=B18+B24*B21		
27	p値（下側）	0.978	=NORMSDIST(B22)		
28	（上側）	0.022	=1-B27		
29	（両側）	0.043	=2*MIN(B27:B28)		
30	母平均の差の区間推定				
31	下限	0.024	=D17-B24*B21		
32	上限	1.570	=D17+B24*B21		

(i)　**データを**`B5:C14`**の枠内に入力する．**

(ii)　**それぞれの個数と平均値および平均値の差 を計算する．**

(iii)　**帰無仮説** $\mu_A - \mu_B$ **の値（通常は** 0**）と既知の** $\sigma_A,\ \sigma_B$ **を入力する．**

(iv)　平均値の差の標準誤差 $se(d)$ を式(2.2) で計算する.

(v)　$z = d/se(d)$ を計算する.

(vi)　両側αを入力し，正規分布の両側α%点を計算する.

(vii)　棄却限界値を計算する.

(viii)　上側，下側，両側のｐ値を計算する.

ｐ値（両側）が 0.043 で，0.05 より小さいので，有意水準 5% で $\mu_A \neq \mu_B$ と判定する.

すなわち，A 社製品と B 社製品の重さの母平均に差がある.

(A 社製品の方が B 社製品より重い)

(4)　母平均の差の区間推定

μ_A, μ_B に何の仮説もおかないときには，式(2.4) ではなく，式(2.3) が $\mathrm{N}(0,1)$ に従うことを利用する.

$$z = \frac{d - (\mu_A - \mu_B)}{se(d)}$$

これを変形して，信頼率95%の信頼区間は次式で求める.

$$d - 1.96se(d) \leq \mu_A - \mu_B \leq d + 1.96se(d) \tag{2.5}$$

なお，信頼率99%の信頼区間は，式(2.5) の 1.96 の代わりに 2.576 を用いれば良い.

表示2.1の最後の3行は，両側信頼率が95% の区間推定をした結果を示している.

$$0.024 \leq \mu_A - \mu_B \leq 1.570$$

すなわち，A 社製品と B 社製品の重量の母平均の差は，0.024〜1.570 の区間内にある.

■　演習 6　血圧の測定値を，平日と休日に分け，次の表が得られた.

平日 ($n = 14$)	121	118	123	144	125	113	131
	135	119	135	140	131	139	125
休日 ($n = 7$)	128	133	116	127	126	113	125

標準偏差が 10 であるとして，平均値の差を検定せよ．

ヒント　表示2.1 を解答のシートにコピーして，演習のデータを入力する．ただし，$n = 14$ は表示2.1の表には入りきれない．そこで，データの部分の行数を増やさなければならない．

表示2.1 では，黒枠の中のデータについて計算するように計算式が入力されている．例えば，平均が計算されるセル B17 には =AVERAGE(B5:B14) が入力されている．黒枠の範囲内に 4 行追加すると，B17 の式は，自動的に，=AVERAGE(B5:B18) に修正される．

15 行以下に追加したのではこのような修正はされないので，正しく計算されない．

左のサンプル番号は付け直す必要がある．■

本日のまとめ

前章では，一つの母平均に関する仮説検定と区間推定を学習したが，本章では，2 つの母平均に関する仮説検定と区間推定を学ぶ．その中で，今日は「σ 既知」の場合における母平均の差の仮説検定と区間推定を学んだ．母平均の差を対象にしているのだから，その仮説検定と区間推定の基礎となる情報は標本平均の差である．したがって，標本平均の差の分布が必要になるが，これは，第 1 単元 §3 統計解析の基礎知識 から導かれるはずである．それが理解できれば，後は本単元の §1.2 と同様である．

2.2　母平均の差の推定と検定 (σ 未知の場合)

§2.1 では，2 つの母集団をそれぞれ A, B としたとき，それぞれの母標準偏差 σ_A および σ_B が既知の場合について，母平均の差の推定や検定の考え方と手順を学んだ．

ここでは，σ未知の場合の方法について学ぶことにしよう．基本的な考えは，§1.3 と同じである．

(1)　t 検定

前節の例で，標準偏差が未知であるときの仮説検定と区間推定を取り上げる．2 つの母標準偏差 σ_A と σ_B が既知の場合には，

$$z = \frac{(\overline{x}_A - \overline{x}_B) - (\mu_A - \mu_B)}{\sqrt{\dfrac{\sigma_A^2}{n_A} + \dfrac{\sigma_B^2}{n_B}}} \tag{2.6}$$

の分布が，$\mathrm{N}(0,\ 1)$ になることは既に学んだ．ここで，2 つの母標準偏差 σ_A と σ_B は未知であるが，$\sigma_A^2 = \sigma_B^2 = \sigma^2$ と仮定できるならば，

$$z = \frac{(\overline{x}_A - \overline{x}_B) - (\mu_A - \mu_B)}{\sqrt{\left(\dfrac{1}{n_A} + \dfrac{1}{n_B}\right)\sigma^2}} \tag{2.7}$$

となる．ここで，σ^2 は未知であるから，その推定値 s^2 で代用する．s_A^2 も s_B^2 も σ^2 の推定値であるが，できるだけ多くのデータを用いた方が推定精度が良くなるので，両者をあわせた量を考える．そこで，2 つの平方和 S_A, S_B を加えた $S = S_A + S_B$ をその自由度 $f = f_A + f_B = (n_A - 1) + (n_B - 1) = n_A + n_B - 2$ で割った量

$$s^2 = \frac{S}{f} = \frac{S_A + S_B}{n_A + n_B - 2} \tag{2.8}$$

を σ^2 の推定値として用いる．式 (2.7) において，σ^2 を s^2 で置き換えた

$$t = \frac{(\overline{x}_A - \overline{x}_B) - (\mu_A - \mu_B)}{\sqrt{\left(\dfrac{1}{n_A} + \dfrac{1}{n_B}\right)s^2}} \tag{2.9}$$

は，§1.3 と同様な考え方で，自由度 $f = n_A + n_B - 2$ の t 分布に従う．この性質を用いて，母平均の差の検定と推定の方法を導くことができる．

(検定)　帰無仮説は，「2 つの母平均に差がない」すなわち「$\mu_A = \mu_B$」である．このとき，式 (2.9) より，

$$t = \frac{(\overline{x}_A - \overline{x}_B) - (\mu_A - \mu_B)}{\sqrt{\left(\dfrac{1}{n_A} + \dfrac{1}{n_B}\right) s^2}} = \frac{(\overline{x}_A - \overline{x}_B) - (\mu_A - \mu_B)}{\sqrt{\left(\dfrac{1}{n_A} + \dfrac{1}{n_B}\right)\left(\dfrac{S}{f}\right)}} \tag{2.10}$$

は，自由度 $f = n_A + n_B - 2$ のt分布に従うから，$|t| \geq t(f,\ 0.05)$ となる確率は0.05である．したがって，標本から求めた $\overline{x}_A$, $\overline{x}_B$, S_A, S_B の値をこの式に代入して t を計算し，$|t| \geq t(f,\ 0.05)$ となったとき，有意水準 $\alpha = 0.05$ で $\mu_A = \mu_B$ の仮説を否定して$\mu_A \neq \mu_B$ と判定することができる．または，p値を求めて結論を求めることができる．

例について，Excelで計算した結果を表示2.2に示す．

表示2.1 のσ既知の場合との違いのみを説明する．

(i)　既知の標準偏差を入力する代わりに，データから標準偏差を求める．A, Bそれぞれについて平方和と自由度を求め，右に合計を計算する．それから平均平方，標準偏差を求める．

(ii)　B25 に平均値の差の標準誤差を計算する．$\sigma_A = \sigma_B$ とする．

(iii)　B26 に 差/差の標準誤差 を計算する．既知の場合は z であったが，未知の場合は t となる．

(iv)　棄却限界値の求め方は既知の場合と同じ．

(v)　p値は，t分布の確率を計算する関数 TDIST を用いる．使い方は，§1.3 と同じ．

両側のp値が 0.021 で 0.05 よりも小さいことから有意と判断される．すなわち，有意水準 5% でA社とB社の製品正味重量の母平均に差があると結論できる．

(推定)　上の検定結果より，A社とB社の製品正味重量の母平均に差があることが分かった．それではどの程度の差があるのであろうか．一つの答えは，標本平均の差

$$\overline{x}_A - \overline{x}_B = 48.34 - 47.54 = 0.80$$

表示2.2: 2つの平均値の差の検定（σ 未知）

	A	B	C	D	E	F
3	σ未知	(t検定)			(Welch)	
4		A	B			
5	1	**49.6**	**47.3**			
6	2	**48.6**	**48.2**			
7	3	**48.3**	**47.6**			
8	4	**47.9**	**48.4**			
9	5	**48.9**	**46.6**			
10	6	**48.4**	**47.7**			
11	7	**48.1**	**47.0**			
12	8	**47.5**				
13	9	**48.5**				
14	10	**47.6**				
15						
16	個数(n)	10	7	差		
17	平均	48.340	47.543	0.797		
18	H0: μ1-μ2	**0.0**				
19	標準偏差の推定					
20	平方和	3.504	2.437	5.941		
21	自由度(f)	9	6	15	12.885	
22	平均平方	0.389	0.406	0.396		
23	標準偏差			0.629		
24	母平均の差の検定					
25	差の標準誤差	0.310	(下に示す)		0.311	
26	t	2.570	=D17/B25		2.560	
27	α（両側）	**0.05**			**0.05**	
28	t(α)	2.131	=TINV(B27,D21)		2.179	
29	棄却（下限）	-0.661	=B18-B28*B25		-0.678	
30	（上限）	0.661	=B18+B28*B25		0.678	
31	p値（下側）	0.989	(下に示す)		0.987	
32	（上側）	0.011	=1-B31		0.013	
33	（両側）	0.021	=2*MIN(B31:B32)		0.025	
34	母平均の差の区間推定					
35	下限	0.136	=D17-B28*B25		0.119	
36	上限	1.458	=D17+B28*B25		1.476	
37	B25: =SQRT(1/B16+1/C16)*D23					
38	B31: =IF(B26>0,1-TDIST(B26,D21,1),TDIST(-B26,D21,1))					
39	E21: =(B22/B16+C22/C16)^2/((B22/B16)^2/B21+(C22/C16)^2/C21)					

である．これは今まで学習してきたように点推定である．区間推定は，式(2.9)(p.31) の

$$t = \frac{(\overline{x}_A - \overline{x}_B) - (\mu_A - \mu_B)}{\sqrt{\left(\dfrac{1}{n_A} + \dfrac{1}{n_B}\right) s^2}}$$

が自由度 $f = n_A + n_B - 2$ **の** t **分布に従うことから，この** t **が**

$$-t(f,\ 0.05) < t < t(f,\ 0.05)$$

となる確率は0.95**である．これより，以下のように，**$\mu_A - \mu_B$ **を挟む不等式を導くことができる．**

$$\begin{aligned}
-t(f,\ 0.05) &< t < t(f,\ 0.05)\\
-t(f,\ 0.05) &< \frac{(\overline{x}_A - \overline{x}_B) - (\mu_A - \mu_B)}{\sqrt{\left(\dfrac{1}{n_A} + \dfrac{1}{n_B}\right)s^2}} < t(f,\ 0.05)
\end{aligned}$$

$$(\overline{x}_A - \overline{x}_B) - t(f,\ 0.05)\sqrt{\left(\frac{1}{n_A} + \frac{1}{n_B}\right)s^2} < \mu_A - \mu_B <$$

$$(\overline{x}_A - \overline{x}_B) + t(f,\ 0.05)\sqrt{\left(\frac{1}{n_A} + \frac{1}{n_B}\right)s^2} \qquad (2.11)$$

この不等式を使って，$\mu_A - \mu_B$ **の信頼区間を求めることができる．式**(2.11)**は，信頼率が**95%**であるが，式中の** 0.05 **を変えることにより任意に指定できる．例では，**

$$\begin{aligned}
0.80 - 2.131 \times 0.310 < \mu_A - \mu_B &< 0.80 + 2.131 \times 0.310\\
0.136 < \mu_A - \mu_B &< 1.458
\end{aligned}$$

となる．

■　演習 7　　演習6 **の血圧データを，標準偏差未知として解析せよ．　■**

(2)　不等分散の場合

先の検定や推定では，「2**つの母標準偏差** σ_A **と** σ_B **は未知であるが，**$\sigma_A^2 = \sigma_B^2 = \sigma^2$ **と仮定できる」とした．しかし，**$\sigma_A^2 = \sigma_B^2$ **と仮定できない，あるいは仮定できるか分からないこともある．仮定できるかわからない場合には，**§3 **で学習する母分散の違いの検定をまず行い，**$\sigma_A^2 = \sigma_B^2$ **と仮定できるかできないかを判断することが多い．**

$\sigma_A^2 = \sigma_B^2$ と仮定できない場合（不等分散の場合）のために，ウェルチ の検定と呼ばれる方法が用意されている．$\sigma_A^2 = \sigma_B^2$ と仮定できないので，式(2.6)(p.31) の σ_A^2, σ_B^2 のそれぞれをそれらの推定値である s_A^2, s_B^2 で代用することになる．

$$t' = \frac{(\overline{x}_A - \overline{x}_B) - (\mu_A - \mu_B)}{\sqrt{\left(\dfrac{s_A^2}{n_A} + \dfrac{s_B^2}{n_B}\right)}} \tag{2.12}$$

式(2.12)の分母は，$n_A = n_B = n$ のとき，

$$\sqrt{\frac{s_A^2}{n_A} + \frac{s_B^2}{n_B}} = \sqrt{\frac{S_A}{n(n-1)} + \frac{S_B}{n(n-1)}} = \sqrt{\frac{2}{n}\frac{S}{f}}$$

となり，式(2.10)(p.32) で $n_A = n_B = n$ としたときの分母と等しくなる．すなわち，$n_A = n_B$ のときは $t' = t$ の関係が成立する．

この t' は，自由度

$$f' = \frac{(s_A^2/n_A + s_B^2/n_B)^2}{\dfrac{(s_A^2/n_A)^2}{f_A} + \dfrac{(s_B^2/n_B)^2}{f_B}} \tag{2.13}$$

の t 分布に近似的に従う[1]．

この性質を用いて，不等分散の場合における母平均の差の検定と推定の方法を導くことができる．

帰無仮説は，「2 つの母平均に差がない」すなわち「$\mu_A = \mu_B$ 」である．このとき，式(2.12) で計算される t' は，自由度 f' の t 分布に近似的に従うから，$|t'| > t(f', 0.05)$ となる確率はほぼ 0.05 である．したがって，標本から求めた $\overline{x}_A$, $\overline{x}_B$, s_A^2, s_B^2 の値をこの式に代入して t' を計算し，$|t'| > t(f', 0.05)$ となったとき，有意水準 α＝0.05 で $\mu_A = \mu_B$ の仮説を否定して $\mu_A \neq \mu_B$ と判定することができる．または，p 値を求めて結論を求めることができる．

表示2.2 (p.33) の E 列に ウェルチの検定の結果を示す．

---

[1] 自由度は サタースウェイトの方法と呼ばれる方法により求められている．この部分の理論は，本講座の想定する水準を越えた数理的知識が必要になるので，説明は省略する．

式(2.13) の自由度は，E21 で計算されており，12.885 となる．これは，元の自由度 15 に比べて小さくなり，小数点以下の端数を含む．

Excel 関数の TDIST, TINV 関数は小数自由度に対応しておらず，小数点以下の数値は切り捨てられてしまう．E28 の 2.179 は 自由度が 12 の t 分布の両側 5%点である[2]．この値は B28 の 2.131 よりも大きい．したがって，不等分散の場合は，棄却限界や信頼区間の範囲が広くなる．また，p 値も大きくなる．

両側の p 値は 0.025 で，既知の場合の 0.021 よりも大きい．

推定も式(2.12) にもとづき，今までと同様に信頼区間を構成することができる．

σ既知の場合の信頼区間 0.136 ～ 1.458 に比べて，0.119 ～ 1.476 と幾分広くなる．

■　演習 8

$f' \leq f$ の関係があり，$n_A = n_B$, $V_A = V_B$ のときに限り $f' = f$ となる．$V_A < V_B$, $n_A > n_B$ のとき，$f - f'$ の差が大きくなる．

表示2.2 のデータに次のような修正を加えて，この関係を数値で確かめよ．

(i)　B のデータに値を 3 個追加して，n を等しく 10 にする．さらに，両群の平均平方がほぼ等しくなるように，追加する値を試行錯誤で決める．

(ii)　(i) で得られたデータから，B のデータを 3 個除く．除く 3 個は，平均値に近いもの，あるいは，多いものを選ぶ．■

(3)　対応のある場合

(1)，(2)で，母平均の差の検定と推定の方法を学んだ．そこでは，解析の対象としたデータで，A 社の i 番目と B 社の i 番目の間には何の関係もなかった．したがって，各組の標本数 n_1, n_2 が等しいという制約はなかった．

---

[2] 自由度12.885 の t 分布の両側 5%点を，自由度が 12 と 13 の%点から補間によって求める方法もある．

$$\begin{aligned} t(12.885) &= (13 - 12.885) * t(12) + (12.885 - 12) * t(13) \\ &= 0.115 * 2.179 + 0.885 * 2.160 = 2.162 \end{aligned}$$

ここでは，$n_1 = n_2 = n$ で，i 番目のデータには対応のある場合を取り上げる．以下にいくつかの例をあげる．

2種類の測定方法を比較するために，n 個の品物を異なる方法で測定する．

2種類のクリームを比較するために，n 人の右手と左手に塗布して比較する．

降血圧剤の効果を比較するために，n 人に投与し，投与前と投与後の血圧を測定する．

以下次の例を取り上げて，解析の手順を説明する．

A, B の2つのハカリで10個の品物の重さを測ったところ，下の表の結果が得られた．これより，2つのハカリの指示に相違があるかどうか検定し，かつ，その相違の大きさを推定してみる．

品物	1	2	3	4	5	6	7	8	9	10
A	555	533	541	545	556	524	567	515	535	533
B	549	521	540	536	556	530	562	519	532	525
d	6	12	1	9	10	−6	5	−4	3	8

この例における縦に並んだ2つの測定値は独立ではない．同じ品物を測っているのだから，例えば，品物1のAのデータ 555 は品物1のBのデータ 549 とは関係があるが，品物2のBのデータ 521 とは関係がない．したがって，品物1のBのデータ 549 と品物2のBのデータ 521 を入れ替えることはできない．このようなデータを 対応のあるデータ と呼び，今までのデータとは区別する．もし，上のデータが一つの品物を2つのハカリA，Bでそれぞれ10回ずつ測定したというのならば，データの並び順には意味がなく入れ替えは自由なので，対応のあるデータではない．

対応のあるデータでは，それらの差を取って解析することが必要である．そうしないと，品物が異なることによるバラツキが誤差に含まれてしまうので，検定や推定の精度が悪くなり，母平均の差が見出しにくくなったり，信頼区間の幅が広くなったりする．

対応のあるデータの差を $d_i = x_{Ai} - x_{Bi}$ とする．帰無仮説 $\mu_A = \mu_B$ が成立するとき，d_i の母平均 μ_d は 0 になる．

これから，§1.3 の手順に従って，d を x として解析すれば良いことが分かる．このように考えると，Excel による解析手順は容易に想像できるであろう．

表示1.8 (p.19) の計算表で，データの右の2列にこのデータを入力する．データのセルに差 d を計算する．

帰無仮説値を 0，α（両側) を 0.05 とすると，表示2.3が得られる．

表示2.3: 2つの平均値の差の検定（対応のある場合)

σ未知　（t検定）

	A-B	A	B
1	6	**555**	**549**
2	12	**533**	**521**
3	1	**541**	**540**
4	9	**545**	**536**
5	10	**566**	**556**
6	-6	**524**	**530**
7	5	**567**	**562**
8	-4	**515**	**519**
9	3	**535**	**532**
10	8	**533**	**525**

個数	10
平均	4.4
H0: μ	**0.0**
標準偏差の推定	
平方和	318.4
自由度(f)	9
平均平方	35.4
標準偏差	5.9
母平均の検定	
標準誤差	1.881
t	2.339
α（両側）	**0.05**
t(f, α)	2.262
棄却（下限）	-4.255
（上限）	4.255
p値（下側）	0.978
（上側）	0.022
（両側）	0.044
母平均の区間推定	
下限	0.145
上限	8.655

結果の見方は§1.3 の場合と同じである．t が2.339 で t 分布の両側 0.05%点

2.262 より大きい，また，p 値（両側） 0.044 と 0.05 より小さいので，差の平均値は 0 ではない，すなわち，2 つの平均値間には有意差のあることが分かる．

母平均の信頼区間から，A，B 2 つのハカリの相違は，0.145g〜8.655g の範囲内の値であることが分かる．この範囲に 0 が含まれていないことから，2 つのハカリに差があるといえる．

■　演習 9

(i)　例のデータを対応があるデータではないとしたら解析結果はどうなるか．

(ii)　1 番目の品物を軽い物に替えた．その結果，ハカリ A では 455g，ハカリ B では 449g という結果が得られた．解析結果がどのように変化するか．

本日のまとめ

今日は「σ 未知」の場合における母平均の差の仮説検定と区間推定の種々の方法を学んだ．まず，昨日との違い，すなわち σ が「既知」なのか「未知」なのかの違いを明確にしてほしい．さらに，「σ 未知」の場合において，2 つの母標準偏差が等しいと仮定できる場合（等分散）と，できない場合（不等分散）とで方法が異なるので，その使い分けを明確にしてほしい．

また，同じようにまとめられたデータでも，データの収集法によって解析方法が変わってくる．今日の最後で「対応のある場合」について学習したが，この「対応」のあるなしの区別も明確に理解しなければならない．また，対応のある場合の解析は基本的に§1.3 と全く同じであることを再確認してほしい．

2.3　検出力と n の決め方

これまでは，得られたデータから，2 つの平均値の間には有意差があるかどうかを調べたり，母平均の差の信頼区間を求める考え方と手順を学んだ．

データを取る以前には，サンプルの個数 n をいくらにしたら良いかを検討す

る必要がある．n が小さいと，$\mu_A \neq \mu_B$ であっても，違いを見つけることができない．これが第2種の誤りであることは，すでに，第3単元で学んだ．

この考えを2つの平均値の差の検定問題に当てはめる．

まず，σ が既知の場合を取り上げ，ついで，σ が未知の場合の近似方法を説明する．

(1)　σ 既知の場合

2つの母集団があって，母標準偏差は共に $\sigma = 1$ で，母平均の差が $\delta = \mu_A - \mu_B$ であるとする．

それぞれの母集団から n 個のサンプルを取り出したとき，平均値の差 d をその標準誤差 $se(d)$ で割った比 z の分布を考える．z は

$$z = \frac{d}{se(d)} = \frac{\overline{x}_A - \overline{x}_B}{se(d)} = \frac{(\overline{x}_A - \overline{x}_B) - (\mu_A - \mu_B)}{se(d)} + \frac{\mu_A - \mu_B}{se(d)}$$

と分解でき，第1項は標準正規分布に従う．

したがって，z は 期待値が $0 + \dfrac{\mu_A - \mu_B}{se(d)} = \dfrac{\delta}{se(d)}$ で，標準偏差が 1 の正規分布に従う．

$se(d)$ は $n_A = n_B = n$ のとき，式(2.2) (p.27) より

$$se(d) = \sqrt{2/n}\,\sigma$$

である．

$n = 10,\ \delta = 1.5$ のときの，$x_A,\ x_B,\ d$ の分布の関係を表示2.4に示す．

表示2.4の上の2つの曲線は $x_A,\ x_B$ の分布を表わす．

平均値の差 $d = \overline{x}_A - \overline{x}_B$ の分布が表示2.4の下に示されている．

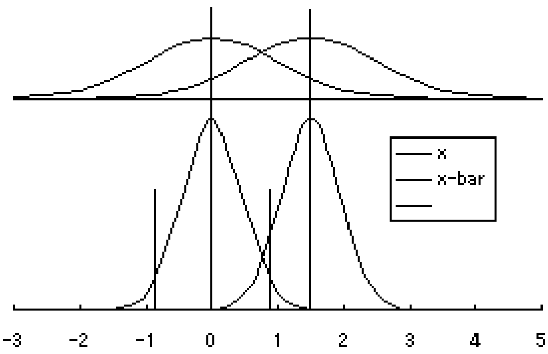
平均値の差 d の標準誤差は

$$se(d) = \sqrt{2/10} = 0.447$$

である．

左は $\delta = 0$ の場合，右は $\delta = 1.5$ の場合である．

表示2.4: 第2種の誤りと検出力



帰無仮説が正しいとき，平均値の差は左の分布となり，両裾の確率が 5% になる位置は，$1.96se(d) = 0.877$ となる．そこに縦の線が引かれている．平均値の差がこの範囲外に出たとき，有意差があると判断される．

それでは，$\delta = 1.5$ の場合どうなるであろうか？ 右の分布を見ると 0.877 より大きくなる確率はかなり大きいが100%ではない．0.877 より小さいほうにも裾を引いている．この面積に相当する確率で，差が見つからない．これが **第2種の誤り** の確率 β に相当する．

β は次のようにして計算される．

平均値の差 d の期待値は $\delta = 1.5$ で，棄却限界値 0.877 は期待値から $1.5 - 0.877 = 0.623$ 離れている．この距離は d の標準誤差 0.447 の $0.623/0.447 = 1.394$ 倍である．標準正規分布で 1.394 以上の確率を計算すると 0.082 となる．これが差を見逃す誤りの確率 β である．逆に，$1 - 0.082 = 0.918$ の確率で，差が発見できることを示している．これが **検出力** である．

表示2.4 のグラフの上のバーをクリックすると，δ が変化し，それに同期してグラフや，検出力などの数値がどのように変化するかを目で確かめることができる．

Excel を使って β と検出力を計算する手順を表示2.5 に示す．

σ, n, δ, α を入力する．検定の α は自由に設定できる．δ は正の値とする．

表示2.5: β と検出力の計算

	K	L	M	N	O
3	σ	1.0	6	6	
4	n	10	30	48	
5	δ	1.5	4	4	
6	α	0.05	0.05	0.05	
7	se(d)	0.4472	1.5492	1.2247	=SQRT(2/L4)*L3
8	棄却限界値	0.8765	3.0364	2.4005	=NORMSINV(1-L6/2)*L7
9	β	0.0816	0.2670	0.0958	=NORMSDIST((L8-L5)/L7)
10	検出力	0.9184	0.7330	0.9042	=1-L9

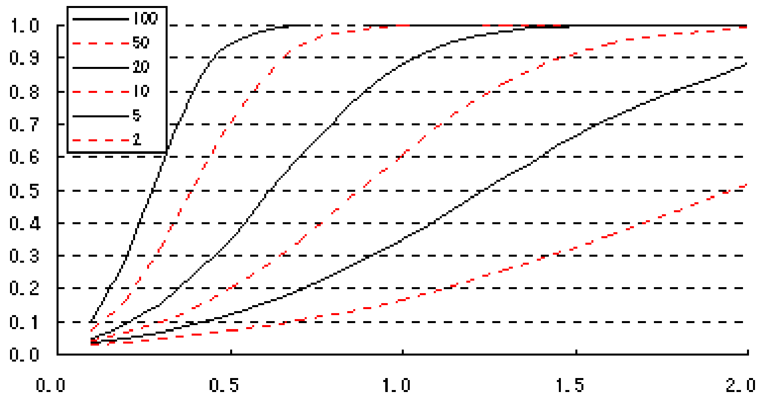
平均値の差の標準偏差 $se(d)$ と棄却限界値が計算される．これから，β と検出力が計算される．

検出したい差 δ を設定して，検出率が90% となるために必要なサンプル数 n を決めたいときにも表示2.5 の計算表を利用することができる．

表示2.5のM列で，σ, δ, α，例えば，$\sigma = 6$, $\delta = 4$, $\alpha = 0.05$ を入力する．n に適当な数値（例えば 20）を入力して，検出力の値を見る．検出力が期待する 0.90 よりも小さければ n を大きくする．このような試行錯誤を繰り返して，検出力が希望する値となる n を求めたのが表示2.5のN列である．この場合は $n = 48$ となる．(Excel のゴールシークを使うと試行錯誤をしないでも，この値が求められる）．

必要な n のおよその値を知るために，n を変えて，横軸に δ/σ を，縦軸に検出力を取ったグラフを描いたのが表示2.6 である．ただし，$\alpha = 0.05$ とする．

表示2.6: n と検出力の関係



(2) σ 未知の場合

σ が未知の場合の検出力は既知の場合に比べて，幾分低下する．検出力を正確に計算するためには，特別の分布（非心 t 分布）が必要となり，その理論根拠や計算法はこの講座のレベルを大幅に超えるので省略する．

しかし，n の決め方には，実用的に十分に使える近似方法がある．

σ が既知であるとして，予想される σ を用いて，必要な n を求める．得られた n に 1 を加えた値を必要な n とする．

表示2.5 の例では，$n = 48 + 1 = 49$ となる．

本日のまとめ

今日は，母平均の差の仮説検定の観点から，標本数 n の決め方を学んだ．実際，標本数 n をいくつにしたら良いかということは解析者がしばしば直面する問題である．他の場面への応用という意味でも，今日の内容はその基本となっているので，十分理解されることが望まれる．

2.4　ウィルコクソンの順位検定

(1)　ノンパラメトリック検定

平均値の差のｔ検定は,「2つの母集団それぞれの分布が正規分布に従う」という前提で導かれている．この前提から大きく外れているときには，分布について仮定を置かず，データの大きさの順序だけを使って検定することが考えられる．このような考えから導かれた一連の検定を，ノンパラメトリック検定 (Non-parametric Test) または，分布によらない検定 (Distribution Free Test) と呼ばれる．

この方法は次のような場合にも活用することができる．

- o　外れ値を含む場合や，等分散性が成立しない場合．
- o　官能評価などで，絶対値評価が困難で，順序しかつかない場合．
- o　2種類の蛍光灯の寿命を比較するために，各社の蛍光灯をそれぞれ10本を取り付け，寿命が来て取り替えた記録を取った．どの順序で寿命が尽きたかは分かるが，通算何時間点灯したかは不明である．

(2)　データの順位とグラフ化

まず，同じ値を含まない簡単な例（$n_A = 5,\ n_B = 4$）を取り上げて説明する．

表示2.7: データのグラフ化

```
  A    o    o o  o  o
----------------------------------
  B          x    x   x        x
      |   .  |  .   |  .  |   .  .

  A    o o   o o   o
----------------------------------
  B        x     x   x x


       o o x o o x o x x
順位   1 2 3 4 5 6 7 8 9
```

表示2.7の上のグラフは A B の観測値をそのままグラフ化したものである．A の位置を o，B の位置を x で表わしている．

表示2.7 の中のグラフは，観測値の値を無視し，順序情報だけをグラフ化したものである．A, B の行には，別々に，その下の行には，両者を一緒に示している．下に順位が記されている．以下，もとの観測値 x と区別するために順位（Rank）を r で表わすことにする．

平均値の差の大きさを変えた3組のデータについて，同様の表現をしたものが表示2.8 の上の2行である．

表示2.8: 3組のデータ

a)	b)	c)
`  oxooxooxx`	`  oooxoxoxx`	`  ooooxoxxx`
`r 123456789`	`r 123456789`	`r 123456789`
$T_A = 1+3+4+6+7 = 21$	$T_A = 1+2+3+5+7 = 18$	$T_A = 1+2+3+4+6 = 16$
$T_B = 2+5+8+9 = 24$	$T_B = 4+6+8+9 = 27$	$T_B = 5+7+8+9 = 29$

3つのデータを見ると，a) は有意差がないように，c) は有意差があるように見える．b) はどちらであろうか？

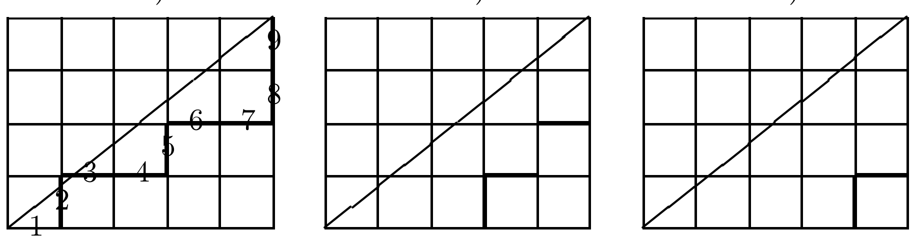
考えられるのは，順位合計を求め（表示2.8 の T_A, T_B），さらに順位の平均値を比較する方法である．その計算手順を説明するに先立って，その幾何学的意味を考えておく．

(3)　検定の考え方

表示2.9 のグラフに示すように，横に n_A，縦に n_B の方眼を描く．

表示2.9: 3組のデータのグラフ化

a)　　b)　　c)



$M_A = 6,\ M_B = 14$　　$M_A = 3,\ M_B = 17$　　$M_A = 1,\ M_B = 19$

左下を出発点とし，値の小さい方から順に，o であれば横に，x であれば上

に線を引く．最終的に，右上に到達する．

帰無仮説（2つの分布が等しい，2群の母平均に差がない）が正しければ，折れ線は左下と右上を結ぶ直線から大きく外れないと考えられる．折れ線の右下の面積と左上の面積を，それぞれ，M_A, M_B で表わすことにする．$M_A+M_B=n_An_B$ であるから，帰無仮説が正しいとき，$D=M_A-n_An_B/2$ または $2D=M_A-M_B$ は0を中心として分布する．

左下から右上に至る道順は ${}_{n_A+n_B}C_{n_A}={}_9C_5=126$ 通りある．

$M_A=0,\ 1,\ 2,\ \ldots,\ 11$ となる道順を数えた結果を，表示2.10の「組合わせ」の列に示す（T_A, T_B の欄については後に説明する）．

表示2.10: D の分布

M_A	D	組合わせ	上側組合わせ	両側組合わせ	両側確率	T_A	T_B
0	10	1	1	2	1.6%	15	30
1	9	1	2	4	3.1%	16	29
2	8	2	4	8	6.3%	17	28
3	7	3	7	14	10.9%	18	27
4	6	5	12	24	18.8%	19	26
5	5	6	18	36	28.1%	20	25
6	4	8	26	52	40.5%	21	24
7	3	9	35	72	54.7%	22	23
8	2	11	46	92	70.3%	23	22
9	1	11	57	114	85.9%	24	21
10	0	12		126	100.0%	25	20
11	−1	11					
			以下省略				

例えば，$M_A=3$ となる道順は，表示2.8 の b) `oooxoxoxx` の他に，`ooxoooxxx`，`ooooxxxox` があり，全部で3通りである．

$M_A\geq 11$ は 12 に対して対称であるから省略する．

「上側組合わせ」の欄には「組合わせ」を上から累積した値が示されている．「両側組合わせ」は「上側組合わせ」の2倍である．$M_A=10$ の両側組合わせは，$M_A=9$ の両側組合わせに $M_A=10$ の組合わせを加えたもので，全部の組合わせ数126に一致する．

「両側確率」の欄には「両側組合わせ」を 126 で割った % が求められている．

これから，「2 群の分布が等しいという帰無仮説が正しいとき，M_A が 1 以下（D が 9 以上）となる確率は 3.1%，すなわち，5% 以下である」ことが分かる．

すなわち，表示 2.8 の右の c) は $M_A = 1$ であるから，両側確率が 3.1% で，5% 以下であるから，予想通り有意差がある．中央の b) は $M_A = 3$，両側確率は 10.9% であるから，有意差は認められない．

これまでは，M, D を数えて有意かどうかを判断した．もっと簡単に順位の合計 T_A, T_B について考える．

a から b までの連続した整数の平均は $(a+b)/2$ であるから，順位（1 から $n_A + n_B$ まで）の総平均 $\overline{r}.$ は $\overline{r}. = \dfrac{1+n_A+n_B}{2} = \dfrac{1+5+4}{2} = 5.0$ である．帰無仮説が正しいとき，T_A, T_B の期待値は，それぞれ，$n_A\overline{r}. = 5 \times 5 = 25$, $n_B\overline{r}. = 4 \times 5 = 20$ である．

T_A, T_B と D との関係を表示 2.10 の右に示す．順位合計とその期待値の差が D となることが分かる．

T_B がこの期待値 20 から 9 以上離れたとき，両側の $\alpha = 0.05$ で有意差があると判断できる．

このような考えで導かれたのがウィルコクソンの順位検定である．

表示 2.10 のような計算をしないでも，有意差の有無を判断できるようにした統計数値表が準備されている[3]．

しかし，現実のデータの解析には以下に説明する近似的方法で十分であろう．

(4)　平均順位の差の標準誤差

順位を通常の数値と同様に考え，通常の平均値の差の検定の手順に従って，平均値 $\overline{r}_A$, $\overline{r}_B$ とその差

$$\begin{aligned} d &= \frac{T_A}{n_A} - \frac{T_B}{n_B} = \overline{r}_A - \overline{r}_B \\ &= \left(\frac{21}{5} - \frac{24}{4}\right) = -1.80 \qquad a) \end{aligned}$$

[3] 「簡約統計数値表」，日本規格協会 (1977)

$$= \left(\frac{18}{5} - \frac{27}{4}\right) = -3.15 \qquad b)$$
$$= \left(\frac{16}{5} - \frac{29}{4}\right) = -4.05 \qquad c)$$

を計算する．次に，平均値の差の標準誤差を求めて，平均値の差とその標準誤差の比を求めるのであるが，標準誤差の求め方が違う．

通常の t 検定では，平均値の差の分散

$$\mathrm{V}[\overline{x}_A - \overline{x}_B] = \left(\frac{1}{n_A} + \frac{1}{n_B}\right) V$$

を求める式の V として，$(S_A + S_B)/(n_A + n_B - 2)$ を用いた．これは，群内のバラツキの大きさを表わすものである．この値は，一方の群の全測定値に一定数を加えても変化しない．すなわち,「母平均の差に影響されない誤差」に比べて平均値の差が大きいかどうかを調べた．

しかし，順位の場合は一方の群の全測定値に一定数を加えてから順位変換すると，V は変化するが，総平方和は一定である．

$$N = n_A + n_B$$

とすると，総平方和 S_T は，

$$S_T = \sum_{r=1}^{N} \left(r - \frac{N+1}{2}\right)^2 = \frac{(N-1)N(N+1)}{12} \qquad (2.14)$$

として求められる[4]．ここで，$(N+1)/2$ は順位の総平均 $\overline{r}_{..}$ である．表示2.8の3つの場合は，いずれも，

$$S_T = \frac{8 \times 9 \times 10}{12} = 60$$

である．

そこで，S_T をその自由度 $f_T = n_A + n_B - 1$ で割った全体の平均平方 V_T

[4] $(r - \frac{N+1}{2})^2 = r^2 - 2r\frac{N+1}{2} + (\frac{N+1}{2})^2$ と展開し，それぞれを $\sum_{r=1}^{N}$ で加える．$\sum_{r=1}^{N} r^2 = \frac{N(N+1)(2N+1)}{6}$ の関係を使って，整理すると上の結果が導かれる．

$$V_T = \frac{S_T}{n_A + n_B - 1} = \frac{(n_A + n_B)(n_A + n_B + 1)}{12} = 7.50$$

を用いて，平均順位の差 $\overline{r}_A - \overline{r}_B$ の標準誤差を求める．

$$se(\overline{r}_A - \overline{r}_B) = \sqrt{\left(\frac{1}{n_A} + \frac{1}{n_B}\right) V_T} = \sqrt{\left(\frac{1}{5} + \frac{1}{4}\right) 7.50} = 1.84$$

平均順位の差とその標準誤差の比を計算する．

$$\begin{aligned} z &= \frac{\overline{r}_A - \overline{r}_B}{se(\overline{r}_A - \overline{r}_B)} \\ &= \left(\frac{-1.80}{1.84},\ \frac{-3.15}{1.84},\ \frac{-4.05}{1.84}\right) = (-0.980,\ -1.715,\ -2.205) \end{aligned}$$

帰無仮説が成立するとき z は近似的に標準正規分布に従うので，絶対値が得られた値よりも大きい確率を計算し，両側検定のp値とする．

上の z についてp値を計算すると，0.327, 0.086, 0.027 となる．これらの値は，表示2.10に求められた正確な確率 0.406, 0.109, 0.031 に比べて小さめである．

(5) 連続修正

2項分布で累積確率を正規分布で近似して求める方法が第3単元の §2.3 で説明された．そこでは，飛び飛びの値を取る x を連続な正規分布で近似するために，隣の値との中間に境界をずらして確率を求めた．

順位も飛び飛びの値を取るため，平均順位の差も不連続な値を取る．そのために，前項で求めたp値は小さめに出る．その補正方法もあるが，ここでは省略する．

(6) 順位変換

最初から順位としてデータが取られた場合はそのまま解析することができるが，通常の連続分布としてのデータが得られた場合は，順位に変換する必要がある．

順位を求める Excel 関数は

```
=RANK(順位をつけたい値，データの集まり,1)
```

である．RANK関数の最後のパラメータ 1 は，小さい方から 昇順に 順位をつけることを指示する．

表示2.11の1行目のデータについて，Excelで順位を求めた結果を表示2.11の2行目に示す．

表示2.11: 順位変換

データ	13	15	12	14	15	18
Excelの順位	2	**4**	1	3	**4**	6
平均順位	2	**4.5**	1	3	**4.5**	6

15 が2個あると，いずれも4位となる．すなわち，Excelの順位は，同順位に対して，小さい方の値を与える．次に大きい 18 は 6位となる．

順位を使って検定をするためには，$x = 15$ に対して，4位と5位の平均である 4.5 位に直す必要がある．

同順位の個数を求めるために，Excel 関数

=COUNTIF(データの集まり，同値を求めたい値)

が用いられる[5].

RANK 関数で求めた順位に COUNTIF 関数で求めた「同順位の個数－1」 の半分を加える．すなわち，

RANK(順位をつけたい値，データの集まり,1)

+(COUNTIF(データの集まり，同値を求めたい値)-1)/2

とする．

(7)　解析手順

表示2.2のデータ（計量値）を順位に変換し，平均順位の差を検定するための計算表を表示2.12に示す．

解析の手順は次の通りである．

[5] RANK と COUNTIF の2つの関数でパラメータの配列順序が逆になっていることに注意．

表示 2.12: 順位変換と検定のための計算

	A	B	C	D	E	F	G	H
2		元データ			順位変換値			
3		A	B		A	B		
4	1	49.6	47.3	1	17.0	3.0		
5	2	48.6	48.2	2	15.0	10.0		
6	3	48.3	47.6	3	11.0	5.5		
7	4	47.9	48.4	4	8.0	12.5		
8	5	48.9	46.6	5	16.0	1.0		
9	6	48.4	47.7	6	12.5	7.0		
10	7	48.1	47.0	7	9.0	2.0		
11	8	47.5		8	4.0			
12	9	48.5		9	14.0			
13	10	47.6		10	5.5			
14								
15				個数(n)	10	7	差	
16				平均	11.200	5.857	5.343	
17				平方和			407	=DEVSQ(E4:F13)
18				自由度(f)			16	=COUNT(E4:F13)-1
19				平均平方(V)			25.438	=G17/G18
20								
21				差の標準誤差	2.485	=SQRT((1/E15+1/F15)*G19)		
22				z	2.150	=G16/E21		
23				p値(両側)	0.0316	=NORMSDIST(-ABS(E22))*2		
24				E4: =RANK(B4,$B$4:$C$13,1)+(COUNTIF($B$4:$C$13,B4)-1)/2				

(i)　E,F 列に順位を求める．E4 に

```
E4: =RANK(B4,$B$4:$C$13,1)+(COUNTIF($B$4:$C$13,B4)-1)/2
```

を入力し，右と下にコピーする．

(ii)　15,16 行目に個数，平均と平均値の差 d を求める．

(iii)　G17 に総平方和を求め[6]，G18,G19 に自由度と平均平方を計算する．

(iv)　E21 に平均値の差の標準誤差 $se(d)$ を求める．

(v)　E22 に $z = d/se(d)$ を求める．

(vi)　E23 に z に対する両側 p 値を求める．

[6] 式 (2.14) に示す総平方和 S_T を求める式は，同順位のない場合である．Excel で計算する場合は，ここに示したように，DEVSQ 関数で求めるのが一般的で，簡単である．

本日のまとめ

今日は，正規分布という仮定を置かずに値の大きさの順序だけを使って検定を行う「ノンパラメトリック検定」の中で，母平均の差の仮説検定を行う「ウィルコクソンの順位検定」を学んだ．今までは正規分布を前提とした方法を中心に学習してきたのでとまどう点があるかもしれないが,「ノンパラメトリック検定」がしばしば使われる分野もあるので，基本的な考え方は理解してほしい．

3　バラツキに関する推定と検定

ここでは，次のような問題を扱う統計的方法を学ぶことにしよう．

例1　塩化ビニール系の床タイルは，寸法がそろっていないと，床に張ったとき，つぎ目にすき間ができてしまうので，寸法のバラツキを小さくする必要がある．いままでの製品は，母平均 μ が 305.0mm で，母標準偏差 σ が 0.2mm でバラツキが大きかった．そこで，1 枚ずつに切り取るときの方法を変えてみた．新しい方法で作った10 枚の製品の寸法を測り，バラツキを計算したところ，平方和 $S = 0.125$，標準偏差 $s = 0.118mm$ であった．この結果からバラツキが小さくなったと判断して良いだろうか．

例2　1kg と表示されている砂糖の量が，実際はいくらになっているかを調べるため，小売店で売られている製品を無作為に30 袋を抽出して重量 (単位g) を測った．その結果は 平均値 $\overline{x} = 1010$g，平方和 $S = 1240$，標準偏差 $s = 6.539$g であった．この結果から，製品全体の重量の母標準偏差 σ はいくらといえるか．

3.1　標準偏差の推定誤差

(1)　シミュレーション実験

§1.1 で実施した，平均値の分布のシミュレーション実験と同じように，バラツキについてのシミュレーション実験を実施してみよう．

平均 μ が5.0 で[1]，母標準偏差 σ が 0.2 の正規分布に従う乱数を10 個生成し，それから，平方和 S，平均平方 V，標準偏差 s を計算する．

このような乱数と計算を100 組について実施する．データの一部を表示3.1 に示す．

10 個の乱数の平方和，平均平方，標準偏差が100 組得られたので，それぞれ

[1] 計算結果の表示幅を狭くしたいので，$\mu - 300$ としてシミュレーションをした．

表示3.1: バラツキに関するシミュレーション実験（一部)

	1	2	3	4	5	6	7	8	9	10	S	V	s
1	4.84	4.96	4.94	4.94	4.88	5.00	5.40	4.80	5.35	5.00	0.366	0.041	0.202
2	5.02	4.95	4.97	4.82	5.01	4.74	5.00	5.20	5.12	4.85	0.169	0.019	0.137
3	4.85	4.85	4.71	5.16	5.31	4.75	5.20	4.80	5.00	4.87	0.380	0.042	0.206
4	5.21	5.00	5.28	4.53	5.23	5.24	4.91	4.74	4.99	5.15	0.545	0.061	0.246
5	4.98	5.12	5.04	5.00	5.01	5.21	4.66	5.01	4.98	4.99	0.174	0.019	0.139
6	5.07	5.17	5.39	5.38	5.02	4.96	5.14	4.85	5.07	5.06	0.263	0.029	0.171
7	5.05	5.05	4.79	4.88	5.23	5.02	5.00	4.90	5.18	5.23	0.205	0.023	0.151
8	4.62	5.05	4.93	5.01	5.07	4.94	5.13	5.04	5.13	4.89	0.206	0.023	0.151
9	5.15	5.25	5.13	4.86	5.31	5.08	4.93	4.94	4.86	4.85	0.259	0.029	0.170
10	4.74	4.65	5.23	4.65	4.74	4.72	5.13	4.88	5.07	5.08	0.429	0.048	0.218

の平均値，標準偏差，最小値，最大値を求めると，表示3.2 のようになる．

表示3.2: S, V, s の性質

	平均	標準偏差	最小値	最大値
平方和	**0.355**	**0.164**	0.104	0.749
平均平方	0.0395	0.0182	0.0115	0.0832
標準偏差	**0.193**	0.045	**0.107**	**0.288**

これから，母標準偏差が 0.2 であっても，$n = 10$ のデータから求めた標準偏差の最大値は最小値の2倍以上あり，かなりばらつくことが分かる．

したがって，計算された標準偏差 s が 0.2 よりも小さいというだけで，新しい方法の効果があったと判断するのは危険であることが分かる．s をそのまま σ の値であると信用することはできない．σ を含む信頼限界を計算する必要がある．

以下，その手順が導かれる過程を順次説明する．

(2)　平方和の分布

平方和 S の平均値 0.355 は $(n-1)\sigma^2 = (10-1)0.2^2 = 0.36$ に近い．ここでは100組のデータから平均値を求めたが，組の数を増やすと，0.36 にもっと近い値が得られる．すなわち，平方和の期待値 $\mathrm{E}[S]$ は

$$\mathrm{E}[S] = (n-1)\sigma^2 \tag{3.1}$$

となる.

この性質があるので，平方和を $n-1$ で割った平均平方 V の期待値は

$$\mathrm{E}[V] = \sigma^2 \tag{3.2}$$

となる．これが，バラツキの総量（平方和）から1個当たりのバラツキ（平均平方）を求めるときに，平方和を n でなく，自由度 $n-1$ で割る理論的根拠である.

平方和 S を母分散 σ^2 で割った値を χ^2 とおいたとき[2]，つまり，$\chi^2 = S/\sigma^2$ がどんな分布をするかを調べてみよう.

式(3.1) から，χ^2 の期待値は

$$\mathrm{E}[\chi^2] = n-1 = f \tag{3.3}$$

となる.

χ^2 の値は $f = n-1 = 9$ に近い値になるが，9 にはならないで，ばらついた値を取ることは，表示3.1 を見れば分かるであろう.

表示3.1 の平方和を $\sigma^2 = 0.04$ で割って，χ^2 を計算し，そのヒストグラムを作成すると表示3.3 が得られる.

このようにして得られた分布を カイ2乗分布 と呼ぶ.

カイ2乗分布は σ^2 には無関係で $f = n-1$ のみによって決まる．ここで，$f = n-1$ は自由度であって，カイ2乗分布は 自由度 のみによって決まる と言い換えることができる.

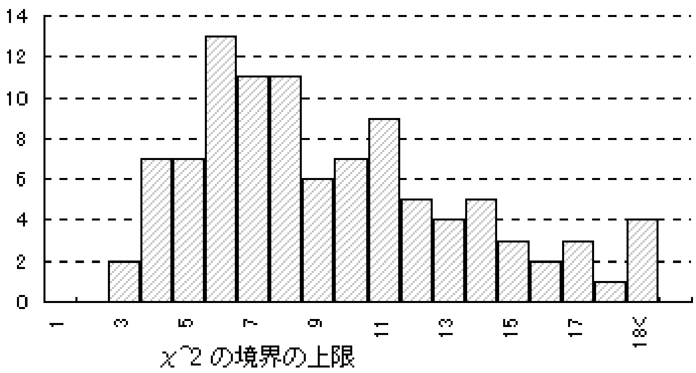
式(3.3) に示すように，カイ2乗分布の期待値は 自由度 に等しい．それでは，カイ2乗分布の標準偏差はどうなるであろうか.

$$\mathrm{D}[\chi^2] = \sqrt{2(n-1)} = \sqrt{2f}$$

---

[2] χ はギリシャ文字で カイ と読む．英語では chi．x ではないことに注意.

表示3.3: χ^2 のヒストグラム



であることが，数理統計学の理論から導かれる．

シミュレーションの結果 S の標準偏差 0.164 を $\sigma^2 = 0.04$ で割ると，カイ 2 乗分布の標準偏差の実験値 4.10 が得られる．シミュレーションの回数を増やすと，この値は $\sqrt{2f} = \sqrt{2 \times 9} = 4.22$ となる．

表示3.3 は自由度が9 のカイ 2 乗分布のおよその形を示している．自由度 9 のまわりに分布し，左右対称ではないことが分かる．

カイ 2 乗分布が自由度によってどのように変化するかを示しているのが表示 3.4 である．

表示3.4 を見ると，分布の形が自由度によって大きく変わることが分かる．

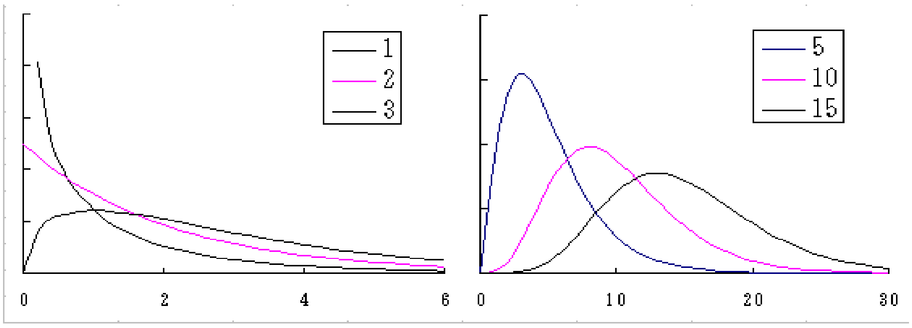
自由度が1のとき，$\chi^2 = 0$ で曲線は無限大になる．

自由度が2のとき，$\chi^2 = 0$ で有限の値を取り，下降線を取る．

自由度が3以上のとき，分布は山を持つようになる．

自由度が小さいと大きいほうに裾を引く，非対称の分布であるが，自由度が大きくなると，左右対称分布（正規分布）に近づく．

表示3.4: 自由度による χ^2 分布の変化



本日のまとめ

今日は，乱数によって平方和，平均平方，標準偏差の分布をシミュレーションで求め，バラツキに関する仮説検定と区間推定を行うための基礎となるカイ2乗分布を導入した．

バラツキに関する検定や区間推定では，カイ2乗分布が重要な役割を果たす．カイ2乗分布がどのように使われるかは明日以降で説明する．

数理統計学の理論から導かれる個所については，その結果だけを示しているので，その導出に興味のある受講生は専門書を参照されたい．

3.2 母分散の検定と推定

カイ2乗分布を用いて，母分散（母標準偏差）の推定や検定をすることができる．以下その方法について学ぶことにしよう．

(1)　母分散の検定

最初にあげた例1は，従来の母分散 $\sigma_0^2 = 0.2^2$ より小さくなったかどうかを検定する問題であった．もし，新しい方法で作っても $\sigma^2 = 0.2^2$ であったとする $(\mathrm{H}_0 : \sigma^2 = \sigma_0^2)$ と，$\chi^2 = S/\sigma_0^2$ はカイ2乗分布に従う．

自由度が f のカイ2乗分布で上側の確率が p である χ^2 の値を $\chi^2(f,\ p)$ で表わすことにする．

帰無仮説が正しいとき，χ^2 が

$$\chi^2(f,\ 0.975) \leq \chi^2 = \frac{S}{\sigma_0^2} \leq \chi^2(f,\ 0.025) \tag{3.4}$$

の範囲内に入る確率は 0.95 となる．

これから，上の範囲外が両側検定の棄却域となる．

この例の場合は，バラツキを小さくするように改善したのだから．σ^2 が大きくなることはないと考えて良い．すなわち，χ^2 が小さな値であるほど $\sigma^2 < 0.2^2$ が確かとなるから，片側検定となり，

$$\chi^2 < \chi^2(f,\ 0.95)$$

が棄却域となる（仮説を否定できることになる）．

§1.3 の表示1.8 (p.19) では，$n = 7$ の観測値から標準偏差を求め，12.6 が得られている．§1.2 の表示1.6 (p.13) では $\sigma = 10.0$ としている．

データから求めた標準偏差 12.6 は $\sigma = 10.0$ から有意に外れているか，また，σ の95%信頼区間を求めたい．

表示1.8 の計算表の下に，バラツキに関する検定と区間推定のための表を追加したのが表示3.5 である．

データの本体と，平均値に関する検定と推定の部分は省略してある．

(i)　データから $n,\ S$ を求め，それから，自由度，平均平方，標準偏差を計算する (第2単元で学習済み)．

(ii)　帰無仮説 $\mathrm{H}_0 :\ \sigma_0^2 = 10^2$ の値 σ_0 を入力する．

(iii)　$\chi^2 = 959.4/10^2 = 9.594$

表示3.5: バラツキに関する検定と区間推定の計算表

	A	B	C	D
3	σ未知	(t検定)		例1
4		x		
15				
16	個数	7	=COUNT(B5:B14)	10
17	平均	137.7	=AVERAGE(B5:B14)	
18	H0: μ	**130.0**		
19	標準偏差の推定			
20	平方和	959.4	=DEVSQ(B5:B14)	0.125
21	自由度(f)	6	=B16-1	9
22	平均平方	159.9	=B20/B21	0.014
23	標準偏差	12.6	=SQRT(B22)	0.118
37				
38	(χ2乗検定)			
39	母標準偏差の検定			
40	H0:σ	**10**		**0.2**
41	χ^2	9.594	=B20/B40^2	3.13
42	α (両側)	**0.05**		**0.1**
43	χ^2(1-α/2)	1.237	=CHIINV(1-B42/2,B21)	3.325
44	χ^2(α/2)	14.449	=CHIINV(B42/2,B21)	16.919
45	p値(下側)	0.857	=1-B46	0.0409
46	p値(上側)	0.143	=CHIDIST(B41,B21)	0.9591
47	p値(両側)	0.286	=2*MIN(B45:B46)	0.0817
48	母標準偏差の区間推定			
49	下限	8.149	=SQRT(B20/B44)	0.086
50	上限	27.846	=SQRT(B20/B43)	0.194

(iv)　α（両側）に 0.05 を入力する．片側検定でαを 5% にしたい場合は 0.10 を入力する．

(v)　$\chi^2(f, 1-\alpha/2)$ と $\chi^2(f,\alpha/2)$ を，Excel 関数 CHIINV で求める．1.237, 14.449 が得られる．$\chi^2 = 9.594$ はこの範囲に含まれるから，帰無仮説 $\sigma = 10$ は棄却できない．

(vi)　Excel 関数 =CHIDIST(χ^2, f) を使って上側のp値を計算する．下側p値は，1-上側p値で，両側p値は 小さい方の2倍とする．両側p値が 0.286 であるから，仮説は棄却できない．

例1のデータは，従来の標準偏差が$\sigma_0 = 0.20$ である製造工程で，改善した結果，標準偏差が小さくなったかどうかを確認したいものであった．したがって

片側検定である．

表示3.5のD列に示すように，$n,\ S,\ \sigma, \alpha$（両側）を入力する．

下側p値 0.0409 は 0.05 より小さいので，バラツキが有意に小さくなったことが分かる．

(2)　母分散・母標準偏差の区間推定

例1で，製造方法を改善することによって標準偏差が小さくなったことは確認されたが，σ はいくらになったかを知りたい．

σ の点推定値としては，$s = 0.118$ であったが，ここでは区間推定の方法を学ぶことにしよう．

仮説検定に用いた式

$$\chi^2(f, 0.975) \leq \chi^2 = \frac{S}{\sigma^2} \leq \chi^2(f, 0.025)$$

から，$\frac{S}{\sigma^2}$ がこの範囲内である確率が0.95であるから，この式を変形して，次のように σ^2 についての信頼率95%の信頼区間の式が求められる．

$$\frac{S}{\chi^2(f, 0.025)} \leq \sigma^2 \leq \frac{S}{\chi^2(f, 0.975)} \tag{3.5}$$

片側の信頼区間がほしいときには，

$$\sigma^2 \leq \frac{S}{\chi^2(f, 0.95)} \quad \text{または}$$
$$\sigma^2 \geq \frac{S}{\chi^2(f, 0.05)}$$

となる．

上の例についての区間推定が表示3.5の下に求められている．

分散の信頼区間を Excel関数 =CHIINV(上側確率, 自由度) を使って計算する．ただし上側確率は $\alpha/2,\ 1-\alpha/2$ とする．標準偏差の信頼区間は，分散の信頼区間の平方根を取って求められる．

血圧の場合　$8.149 < \sigma < 27.846$

例1の場合　$\sigma < 0.194$

という信頼区間が求められる．

■ **演習10**　例2のデータ $n = 30,\ S = 1240,\ s = 6.539$ から，重量の母標準偏差の信頼率が 90% の両側信頼区間を求めよ．■

本日のまとめ

今日は昨日導入したカイ2乗分布を用いて，母分散（母標準偏差）の仮説検定と区間推定を行う方法を学んだ．

演習 で 95%信頼区間を求めると，上限と下限の比は 2.5 倍以上であることが分かるであろう．このように広い区間では実用には不十分である．明日は，十分な精度で標準偏差を推定するためには，どのくらいの n が必要かを学ぶ．

3.3　検出力と標本数 n の決め方

(1)　検出力

$n = 10$ のデータで，帰無仮説 H_0 : $\sigma = \sigma_0$ を検定する場合に，$\sigma = 1.5\sigma_0$ または $= 2.0\sigma_0$ となったとき，どのくらいの割合でバラツキの変化が検出できるかを確かめてみよう．

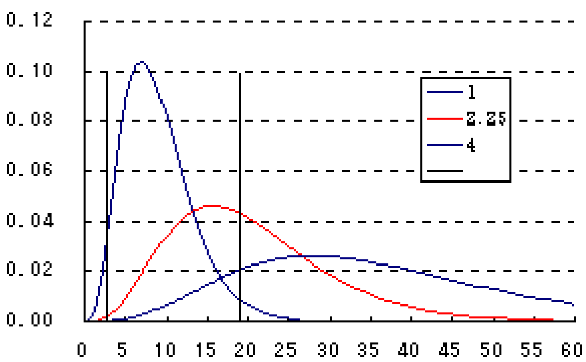
帰無仮説が正しいとき，$S/(\sigma_0^2)$ の分布を表示3.6の左の曲線で表わされる．これは，自由度 $f = 9$ のカイ2乗分布で，期待値は 9.0 である．

この分布の下側確率または上側確率が2.5%となる位置 $(2.70,\ 19.02)$ に縦線が引かれている．この線の両側が棄却域である．

$\sigma = 1.5\sigma_0$ となると，中央の曲線のように，横軸が $1.5^2 = 2.25$ 倍に広がり，縦軸は 1.5^2分の1に縮まる（曲線の下の面積は等しい）．この分布の期待値 $2.25 \times 9 = 20.25$ は棄却限界値に近く，棄却域に含まれる確率はほぼ半分である．

$\sigma = 2.0\sigma_0$ とすると，さらに右にずれた曲線が得られる．こんどはかなりの部分が棄却域に含まれる．

表示3.6: 分散の変化の検出力



以上の3つの分布が棄却域に含まれる確率を正確に計算すると，それぞれ，0.050, 0.489, 0.855 となり，グラフから見た傾向が正しいことが分かる．

検出力という言葉を用いると，$n = 10$ のとき，標準偏差が1.5倍の検出力は約5割であり，2.0倍の検出力は9割近くとなる．

任意の n, α, σ/σ_0 に対して，棄却域の下側に出る確率，上側に出る確率，その合計を求めることのできるExcel シートを表示3.7 に示す．

表示3.7: 分散の変化の検出力の計算シート

	R	S	T	U	V	W	X
2							
3						検出力	
4	n	α(両側)	σ/σ0	f	下側	上側	両側
5	10	0.05	1.5	9	0.001	0.489	0.490
6	24	0.05	1.5	23	0.000	0.813	0.813
7							
8		V5:	=1-CHIDIST(CHIINV(1-$S5/2,$U5)/$T5^2,$U5)				
9		W5:	=CHIDIST(CHIINV($S5/2,$U5)/$T5^2,$U5)				
10		X5:	=V5+W5				

5行目は表示3.6の中央の曲線に対応している．

$\sigma/\sigma_0 = 1.5$ を 80% の確率で検出するために必要な n は，5行目を6行目に

コピーし，`R6` **の** n **を試行錯誤的に変化させ，上側の検出力が** 0.80 **を超える最小の** n **を探すことにより得られる．結果は** $n = 24$ **となる．**

(2) 信頼区間の幅

表示3.5 **の例**1**で，**α **を** 0.05 **に修正すると，母分散と母標準偏差の信頼区間が次のように求められる．**

$$0.0058 < \sigma^2 < 0.0409$$

$$0.0762 < \ \sigma\ < 0.2022$$

上限と下限の比を取ってみると，それぞれ，7.04, 2.65 **倍である．**

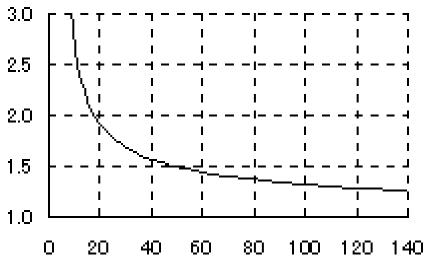
これから，$n = 10$ **程度の観測値から計算した標準偏差はかなりラフな推定値であることが分かる．**

ここに計算した分散の信頼区間の上限/下限の比は，式(3.5) (p.60) **から** $\dfrac{\chi^2(f, 0.025)}{\chi^2(f, 0.975)}$ **として求められる．すなわち，この比率は自由度** f**，すなわち，観測値の個数** n **だけで決まる．**

n **と標準偏差の信頼区間の上限/下限の比の関係を求めたのが表示**3.8 **である．**

表示3.8: **標準偏差の信頼区間の上限/下限の比**

n	下限	上限	比
5	0.35	1.67	4.80
10	0.55	1.45	2.65
15	0.63	1.37	2.15
20	0.68	1.31	1.92
30	0.74	1.26	1.69
40	0.78	1.22	1.57
60	0.82	1.18	1.44
100	0.86	1.14	1.32
140	0.88	1.12	1.27



このグラフから，標準偏差の信頼区間の上限/下限の比を 1.5 **以下にしたい場合には** n **は**50**以上が必要であることが分かる．**

標準偏差の信頼区間を求めたいときには，必要とされる精度とこのグラフを比べてサンプルの大きさを決める必要がある．

前項の演習 10 の答えは，

$$5.208 < \sigma < 8.790$$

で，上限と下限の比は $8.790/5.208 = 1.69$ となる．$n = 30$ のデータから求めた標準偏差の推定精度はこの程度である．

本日のまとめ

本日は，母標準偏差の仮説検定と区間推定の観点から，標本数 n の決め方を学んだ．母標準偏差の推定を精度良く行うためには，相当数の標本数 n が必要になることが確認できたはずである．

3.4　母分散の違いの検定と推定

ここでは，次のような問題を扱う統計的方法を学ぶことにしよう．

高価であるが精度の高い機械Aと，精度はやや落ちるが価格の安い機械Bがある．精度の違いと価格の違いを比較したうえで，どちらを買うかを決めたい．それぞれの機械で 16 個ずつの品物を作って，平方和を計算したところ，$S_A = 0.1190$, $S_B = 0.2924$ という値が得られた．この結果から精度に違いがあるといえるであろうか？　また，2 つの機械の母標準偏差の比 σ_A/σ_B について区間推定をしたい．

このような場合にも標本分散 s_A^2, s_B^2 を用いて推論することになるが，s_A^2, s_B^2 ともその母数 σ_A^2, σ_B^2 からばらついた値を取るから，統計的な推定・検定をすることが必要になってくる．

このような推定，検定のための統計手法として，伝統的に用いられているのはF分布を用いる手法である．他に，新しい手法としてレビンの検定がある．

以下この2つの方法を説明する．

(1) F分布

2つの平均平方の比を$F=\dfrac{V_1}{V_2}$とおくと，これは1に近い値を取りそうだと思われる．それでは，1からどのくらい離れるであろうか．それを数値で確かめるために，表示3.1のシミュレーション結果から，最初の10組と次の10組について，平均平方の比を計算した結果を表示3.9に示す.

表示3.9: 平均平方の比

	V1	V2	V3	V4	V5	V6	V7	V8	V9	V10
V11	1.02	2.21	0.98	0.68	2.14	1.42	1.82	1.81	1.44	0.87
V12	1.77	3.85	1.71	1.19	3.72	2.47	3.16	3.14	2.51	1.51
V13	0.52	1.13	0.50	0.35	1.09	0.73	0.93	0.92	0.74	0.44
V14	1.35	2.93	1.30	0.91	2.84	1.88	2.41	2.40	1.91	1.15
V15	0.73	1.59	0.70	0.49	1.54	1.02	1.30	1.30	1.03	0.62
V16	0.39	0.85	0.38	0.26	0.82	0.54	0.70	0.69	0.55	0.33
V17	0.55	1.20	0.53	0.37	1.16	0.77	0.98	0.98	0.78	0.47
V18	0.57	1.23	0.55	0.38	1.19	0.79	1.01	1.00	0.80	0.48
V19	0.99	2.16	0.96	0.67	2.09	1.39	1.77	1.76	1.41	0.85
V20	0.64	1.39	0.62	0.43	1.35	0.89	1.14	1.14	0.91	0.55

100個の比について，平均などを計算すると次のようになる．

平均	標準偏差	最小	最大	1以下	1より大
1.21	0.76	0.26	3.85	50	50

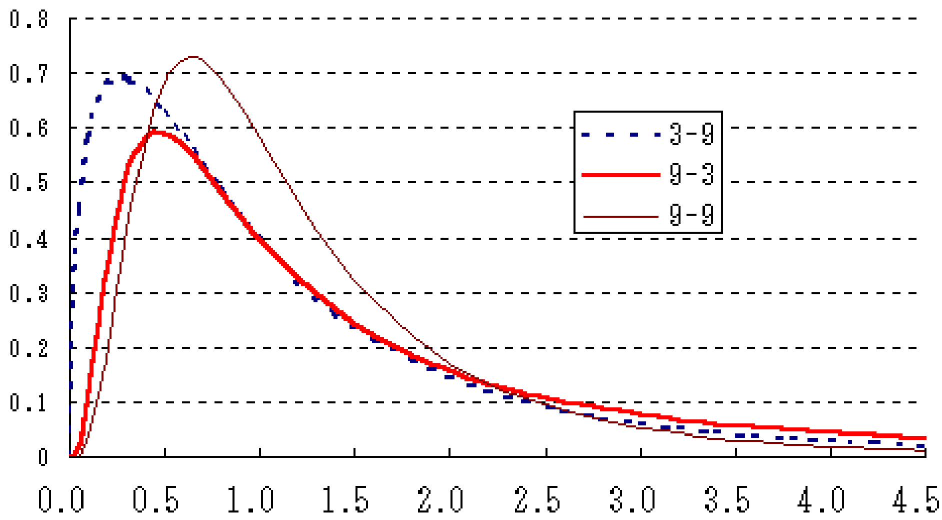
平均平方のバラツキが結構大きいことから予想されることであるが，標準偏差の比も1からかなり離れることが分かる．

分母が大きいか分子が大きいかは半々であるから，比が1以下となる確率と1以上となる確率は等しく 1/2 である．しかし，大きい方に裾を引くために平均値は 1 よりも大きくなっている．

ここに求めた平均平方の比の分布は **F 分布** と呼ばれる．

F 分布は，分子と分母の平均平方の自由度によって分布の形が変化する．表示3.10 に，自由度が (3,9), (9,3), (9,9) の場合のF 分布を示す．分母と分子の自由度が (3,9) と (9,3) は分布が異なることが分かる．

表示3.10: 自由度によるF 分布の変化



Excel ファイルでは，表示3.10 のグラフの左に，グラフを作成するために用いた表があり，表の上にある 2 つの自由度を修正することにより分布の形がどのように変化するかを見ることができる．

分子と分母の自由度が f_1, f_2 のF 分布で，上側の確率が p 以上となる F の値を $F(f_1,\ f_2;\ p)$ で表わすことにする．

$F(f_1,\ f_2;\ p)$ はExcel 関数=FINV(p, 分子の自由度, 分母の自由度) で計算される．

また，上側確率は=FDIST(F の値, 分子の自由度, 分母の自由度) で計算される．

(2)　母分散の違いのＦ検定

先に述べた例の最初の課題は，2つの機械で製品のバラツキ(σ^2)が違うかどうかを検定する問題であった．

もし，帰無仮説 $\mathrm{H}_0:\ \sigma_A^2=\sigma_B^2$（母分散に違いがない）が正しいとき，

$$F=\frac{V_A/\sigma_A^2}{V_B/\sigma_B^2}=\frac{V_A}{V_B} \tag{3.6}$$

となるから，$F=V_A/V_B$ はF分布となる．

これから，F が不等式

$$F(f_A,f_B;\ 0.975)<F<F(f_A,f_B;\ 0.025) \tag{3.7}$$

$$F(f_A,f_B;\ 0.95)<F \tag{3.8}$$

$$F<F(f_A,f_B;\ 0.05) \tag{3.9}$$

の範囲外となる確率はいずれも 5% となる．

上の不等式は両側検定，下の2つの不等式は片側検定の場合に用いられる．

標本から求めた F の値がこのような値を取ったとき，両側検定では $\sigma_A^2\neq\sigma_B^2$，片側検定では $\sigma_A^2<\sigma_B^2$ または $\sigma_A^2>\sigma_B^2$ と判断しても，判断を誤る危険率は5%である．

§2.2 の表示2.2(p. 33) では，A, B のそれぞれについて平均平方が計算され，0.389, 0.406 という値が得られている．この2つの値は極めて近く，検定するまでもなく，有意差があるとは思えないが，計算の手順の理解を深めるために，再び利用する．

α=0.05 の両側検定として，解析をする．

表示2.2 の計算表を元に表示3.11 の計算表を作成する．

23 行目に，各組の標準偏差を求める．

37 行目以降にＦ検定のための計算表が作られている．

39 行目にＦ比が計算されている．40 行目に，α（両側）を入力する．この場合は，$\alpha=0.05$ の両側検定であるから，0.05 を入力する．

41,42行目に，自由度が f_A，f_B のＦ分布の下側と上側の$\alpha/2$ %点が

`=FINV(1-B40/2,B21,C21)` =0.231

表示 3.11: 分散の違いの検定と区間推定

	A	B	C	D	E	F
3	σ未知	(t検定)				
4		A	B			
5	1	**49.6**	**47.3**			
14	10	**47.6**				
15						
16	個数(n)	10	7	差	**16**	**16**
17	平均	48.340	47.543	0.797		
18	H0: μ1-μ2	**0.0**				
19	標準偏差の推定					
20	平方和	3.504	2.437	5.941	**0.119**	**0.2924**
21	自由度(f)	9	6	15	15	15
22	平均平方	0.389	0.406	0.396	0.008	0.019
23	標準偏差	0.624	0.637	0.629	0.089	0.140
37		(F検定)				
38	母標準偏差の差の検定					
39	F	0.958			0.407	
40	α (両側)	**0.05**			**0.10**	
41	F(1-α/2)	0.231			0.416	
42	F(α/2)	5.523			2.403	
43	p値(下側)	0.458			0.046	
44	p値(上側)	0.542			0.954	
45	p値(両側)	0.916			0.092	
46	母標準偏差の比の区間推定					
47	下限	0.417			0.411	
48	上限	2.035			0.989	

=FINV(B40/2,B21,C21) =5.523

として求められている．

上に求めた $F = 0.958$ はこの範囲内であるから，バラツキに有意差があるとは認められない．

また，p値を用いて判断するために，p値（下側，上側，両側）が求められている．用いる関数は FDIST である．

p値（両側）は 0.916 で，当然のことながら，バラツキに有意差は認められない．

最初に説明した例は，A の機械は B の機械に比べて高価であるが，バラツキが小さい（$\sigma_A < \sigma_B$）ことを確かめるのが目的である．すなわち，片側検定で

ある．

この例のデータについて，Excelでこの検定を実行した結果を表示3.11のE,F列に示す．

n_A, n_B, S_A, S_B を入力する．α（両側）としては，α 5%の片側検定であるから 0.10 を入力する．

$F = 0.407$ は下側の限界値 0.416 よりも小さいから，または，p値(下側) 0.046 は0.05 より小さいから，帰無仮説は否定される．機械Aは機械B よりもバラツキが小さいことが確かめられた．

備考　古い統計のテキストでは，F比を計算するときは，「F比が1よりも大きくなるように分母と分子を決めなければならない」と書かれている．この例では，$F = V_B/V_A = 2.458$ として，$F(f_B, f_A; 0.05)$ と比較することになる．これは，昔はF分布の数値表が $p > 0.5$ に限られていたためであり，Excelで容易に計算できる現在ではこのような配慮は全く不要となった．

■ **演習11**　次の2組のデータがあるとき，各組の分散を推定し，さらに分散に有意差があるといえるかどうかを検定しなさい．

A: 18.3, 18.8, 18.5, 18.1, 18.2

B: 18.5, 19.6, 18.9, 19.4, 19.1, 18.7 ■

(3)　母分散の比の推定

2つの機械の選択の問題で，機械Aのバラツキは機械Bよりも小さいことは明らかになった．しかし，その違いがわずかであれば，高価な機械Aを選択することはできない．

機械Bの標準偏差に比べて機械Aの標準偏差は20%以上小さくないと，機械Aを導入する効果がないとする．標準偏差の比は，表示3.11 の 標準偏差 の行から，$0.089/0.140 = 0.638$，すなわち，20%以上小さくなった．これからどのように判断したら良いかを考える．

式(3.7) (p.67) の F に式(3.6)を代入し，

$$F(f_A, f_B;\ 0.975) < F = \frac{V_A/\sigma_A^2}{V_B/\sigma_B^2} < F(f_A, f_B;\ 0.025)$$

とし，これを変形すると，

$$\frac{V_A/V_B}{F(f_A, f_B;\ 0.025)} \leq \frac{\sigma_A^2}{\sigma_B^2} \leq \frac{V_A/V_B}{F(f_A, f_B;\ 0.975)} \tag{3.10}$$

が得られる．

式(3.10) は両側の信頼区間を求める式である．片側の信頼区間を求めるには，不等式の一方を除き，F の上側確率を 0.05 または 0.95 に修正する．

例の場合は，

$$\frac{\sigma_A^2}{\sigma_B^2} < \frac{V_A/V_B}{F(f_A, f_B;\ 0.95)} = \frac{0.4069}{0.4161} = 0.978$$

$$\frac{\sigma_A}{\sigma_B} < \sqrt{0.978} = 0.989$$

となる．

機械Aの標準偏差は機械Bの標準偏差の 0.989 倍以下であるということができる．片側信頼区間の上限が1以下であるということは，片側検定で有意になったことと同じ意味である．

しかし，希望である 0.8以下であることを確認するには，サンプル個数 n が足りない．必要なら，サンプル個数を増やして計測し，再度解析をすることになる．

以上の解析手順は，表示3.11の46行以下に示されている．ここでは，信頼率が 0.95 の片側信頼区間を求めたいので，α として0.1 が入力されている．両側信頼区間を求めたいのであれば 0.05 とする．

(4)　レビンの検定

母分散の違いのF検定は，平均平方を用いている．平均平方は平方和から求められるから，平均平方は外れ値が含まれるとその影響を大きく受けてしまう．そのようなとき，レビンの検定が有効である．

レビンの検定ではバラツキを偏差の絶対値として捉える．F検定では偏差の2乗を用いているが，偏差の絶対値を用いるため，外れ値の影響が小さくなると予想される．母分散に違いがなければ各標本において偏差の絶対値 $|x_{ij} - \bar{x}_{i\cdot}|$ の平均値はほぼ等しい値になると考えられるので，偏差の絶対値をデータとし

て母平均の差の t 検定を行う．その結果，偏差の絶対値の平均値に差がなければ母分散に違いはないと判断する．Excel でこの検定を実行した結果を表示 3.12 に示す．

表示 3.12: 分散の違いのレビンの検定

	A	B
1	1.26	0.24
2	0.26	0.66
3	0.04	0.06
4	0.44	0.86
5	0.56	0.94
6	0.06	0.16
7	0.24	0.54
8	0.84	
9	0.16	
10	0.74	

個数(n)	10	7	差
平均	0.460	0.494	-0.034
H0: μ1-μ2	0.0		
標準偏差の推定			
平方和	1.388	0.730	2.118
自由度(f)	9	6	15
平均平方	0.154	0.122	0.141
標準偏差	0.393	0.349	0.376
母平均の差の検定			
差の標準誤差	0.185		
t	-0.183		
α（両側）	0.05		
t(α)	2.131		
棄却（下限）	-0.395		
（上限）	0.395		
p値（下側）	0.429		
（上側）	0.571		
（両側）	0.857		

計算表は，平均の差の検定で用いた表示 2.2(p.33) がそのまま用いられる．

データの左上のセルに

```
=ABS(B5-B$17)
```

を入力して，右と下にコピーする．ただし，元のデータが空白のセルはコピーしない．

このようにして，偏差の絶対値 $|x_{ij} - \bar{x}_{i\cdot}|$ がデータ表に入力されると，解析

結果が自動的に求められる．p 値 (両側) は 0.857 となり，帰無仮説は否定されない．

本日のまとめ

今日は，2 つのバラツキの違いに関する仮説検定と区間推定が取り上げられた．そのためには，2 つの平均平方の比の分布が必要になり，F 分布が導入された．従来の統計のテキストでは，F 分布を使う検定や推定だけが説明されているが，本テキストでは，レビンの検定も取り上げた．この方法は 2 つの平均値の差の検定と同じであるから，考え方が理解し易いであろう．

4 補遺

第1単元から始まって，これまで多くのことを説明した．この章では，それらのまとめとしての説明や，話の流れから外れた話題を拾い集めて解説する．

4.1 解析手法のまとめ

第4単元では，かなりの数の解析手法が説明された．

現実の問題に遭遇したとき，どの方法を使ったら良いか迷うことが起こるであろう．そのような場合に役立つことを期待して，手法の一覧表を作成した．

表示4.1: 解析手法のまとめ

仮定分布/解析目的	解析手法	参照
連続量:正規分布		
母平均 μ の推定・検定(σ 既知)	z 検定	§1.2
母平均 μ の推定・検定(σ 未知)	t 検定	§1.3
母平均の差の検定・推定(σ 既知)	z 検定	§2.1
母平均の差の検定・推定(σ 未知，等分散)	t 検定	§2.2(1)
母平均の差の検定・推定(σ 未知，不等分散)	ウェルチの検定	§2.2(2)
母分散の検定・推定	χ^2 検定	§3.2
母分散の差の検定・推定	F 検定	§3.4
	レビンの検定	§3.4(4)
分布を仮定しない（ノンパラメトリック）		
母平均の差の検定・推定	順位検定	§2.4

この表は，復習や，手法の検索に役に立つであろう．

また，本文で説明した計算表を一つにまとめたのが,「第4単元.XLS」のシート「§4.1」である.

各計算表の関係が明らかになり，自分の解析に必要な部分を自分のファイルにコピーして用いることができるであろう.

本日のまとめ

個々の解析手法が理解できたとしても，実際の場面で適切な解析手法を選び,適用できなければ役に立たない．一つひとつの手法の適用場面を想定することができたであろうか？　ただし，これはあくまでも一つの指針であって，この場面ではこの手法を使えば良いといった機械的な使い方は控えるべきである．これまで数箇所で説明した,「対数変換」などと組合わせるなどの工夫が必要である.

今回紹介した手法は広い範囲で適用できる基本的なものであるが，ある特別な場合に特化した手法が存在する．例えば，第2単元 §3.5(4) で説明した 打ち切りデータ の解析は,「生存時間解析」として発展している．本講座での学習が最終のものではなく，今後の勉強の入口になることが期待される.

4.2　統計数値表と分布の関係

(1)　統計数値表

昔の統計のテキストでは，最後に統計数値表がついており，それを使って解析する方法が説明されていた.

このテキストでは Excel の統計関数を使って解析する手順を説明したので,数値表は必要がなくなった．しかし，テキストのあちこちに分散して示された Excel関数の説明は勉強するには良いが，後に参照するには不便である．そこで,いろいろの分布についての計算を一つにまとめた計算表を作成した.

表示4.2の計算表の黒枠のセルに数値を入力すると，残りのセルに結果が得られる．この計算表は「第4単元.XLS」のシート「§4.2」に記録されているので，数値を入れながら学習してほしい．

表示4.2: 統計数値表

確率の入力は 0.50 以下

正規分布

x	μ	σ	下側確率	両側確率	上側確率
90	80	8	0.8944	0.2113	**0.1056**
69.7476	80	8	0.1		
95.6797	80	8		0.05	
93.1588	80	8			0.05

標準正規分布

z			下側確率	両側確率	上側確率
1.250			0.8944	**0.2113**	**0.1056**
-1.2816			0.1		
1.9600				0.05	
1.6449					0.05

カイ2乗分布

χ2	f		下側確率	両側確率	上側確率
1.5625	1		0.7887	0.4226	**0.2113**
0.0010	1		0.025		
5.0239	1				0.025

t分布

t	f		下側確率	両側確率	上側確率
2	10		0.9633	**0.0734**	0.0367
-1.3722	10		0.1		
2.2281	10			0.050	
2.6338	10				0.050

F分布

F	f1	f2	下側確率	両側確率	上側確率
4	1	10	0.9266	0.1468	**0.0734**
0.0010	1	1000	0.025		
5.0391	1	1000			0.025

(2)　分布の関係

第4単元では正規分布，ｔ分布，カイ2乗分布，F分布の4つの分布が説明された．

これらの分布の関係を統計数値表（表示4.2） を使って復習する.

正規分布は統計の最も基本的な分布である．x が期待値 μ，分散 σ^2 の正規分布に従うものとする（$\mathrm{N}(\mu,\ \sigma^2)$）.

$\mu = 80,\ \sigma = 8$ の正規分布で，$x \geq 90$ の確率は，表示4.2 の「正規分布」の計算表の 1 行目で，黒枠の中にこの 3 つの値を入力すると，上側確率の欄に 0.1056 が得られる.

x から期待値 μを引き，標準偏差 σ で割って基準化した z は 標準正規分布 $\mathrm{N}(0,\ 1^2)$ に従う.

z の期待値は 0 であるから，分散の定義により，z の分散は z^2 の期待値に等しく，1 である.

$\mu = 80,\ \sigma = 8,\ x = 90$ に対する z は

$$z = \frac{x-\mu}{\sigma} = \frac{90-80}{8} = 1.25$$

である.「標準正規分布」の計算表で，z のセルに 1.25 を入力すると，前に「正規分布」の計算表で得た値と全く同じ結果（0.1056）が得られる.

z の 2 乗の分布は，自由度 1 のカイ 2 乗分布に従う.「χ 2 乗」の計算表で $z^2 = 1.25^2 = 1.5625, f = 1$ を入力すると，上側確率として 0.2113 が得られる．これは，$z = 1.25$ に対する両側確率（片側確率の 2 倍）に等しい.

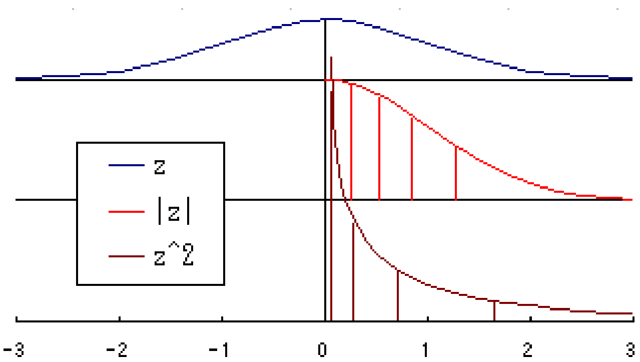
上に述べたように，z^2 の期待値であるから，自由度が 1 のカイ 2 乗分布の期待値は 1 である.

標準正規分布と自由度 1 のカイ 2 乗分布の関係を表示4.3 に示す.

上の曲線は標準正規分布に従う z の分布である．中央は $|z|$ の分布で，z の分布の負の部分を切り取り，高さを 2 倍する．4 本の縦線は確率が 20%となる 5 つの区間に分割したものである.

下は z^2 の分布で，自由度 1 のカイ 2 乗分布である．同様に，確率を 5 等分する区間を示す縦線が入れられている．これから，自由度 1 のカイ 2 乗分布は，0 で無限大になり，大きい方に長く裾を引くことが分かるであろう.

表示4.3: 正規分布と自由度1のカイ2乗分布の関係



■ 演習 12 上に述べた関係を，表示4.2 の統計数値表を使って確かめて見よ. ■

n 個の z_1, z_2, ..., z_n が 独立に標準正規分布に従うとき，

$$z_1^2 + z_2^2 + \ \dots \ + z_n^2$$

は自由度が n のカイ2乗分布に従う．自由度が大きくなると，分布は正規分布に近づく（表示3.4 (p.57) 参照）.

■ 演習 13 「自由度が大きくなると，カイ2乗分布は正規分布に近づく」という性質を，表示4.2 の統計数値表を使って数値的に確かめよ.

ヒント

$f = 50$ のカイ2乗分布で 60 以上となる確率をカイ2乗分布の表で求める.

自由度 f のカイ2乗分布の期待値は f，標準偏差は $\sqrt{2f}$ であることが §3.1(2) に書かれている．この性質を使って，統計数値表の正規分布の表で近似値を求める方法を工夫せよ. ■

「t 分布」で，自由度を大きくすると，標準正規分布に近づくことが確かめられるであろう.

t 分布と分子の自由度が1のF分布の間には

$$t^2(f) = F(1, f)$$

の関係がある．これは,「t 分布」で $t = 2$, $f = 10$ を入力し,「F 分布」で，F の値に t の値の 2 乗（=4）を，分子の自由度に 1 を，分母の自由度に t 分布の自由度 10 を入力すると，F 分布の上側確率 (0.0734) が t 分布の両側確率に等しくなることで確認できる．

また，F 分布とカイ 2 乗分布の間には

$$\chi^2(f)/f = F(f, \infty)$$

の関係がある．これは,「F 分布」で分子の自由度をカイ 2 乗分布の自由度に等しく取り，分母の自由度を大きくする（例えば 1000 とする）と，F の値は χ^2 の値の 1000 倍となることで確かめられる．

本日のまとめ

Excel の統計関数を用いれば，従来の統計数値表を参照する必要がなくなったことは，今までの学習から理解されたであろう．今日の内容は，それらをまとめたものであるから，それぞれの適用場面が浮かんできてはじめて使えるものになる．

また，これまでに学んだ正規分布，カイ 2 乗分布，F 分布，t 分布の関係を，数値表を使いながら理解しておくことも，将来役に立つであろう．

4.3 適合度検定

(1) ポアッソン分布かどうかの検定

n 個の測定値 $x_1,\ x_2,\ \ldots,\ x_n$ があるとき，それがポアッソン分布に従うかどうかを検定する方法を学ぶことにしよう．

もし，個々の現象（例えば，交通事故）が互いに独立に起こるときは，母平均 μ のポアッソン分布に従い，x_i の平均 $\overline{x}$ と分散 s^2 はほぼ等しくなるはずである(ポアッソン分布の性質 $\mu = \sigma^2$ から．⇒第3単元§4.3)．したがって s^2 が $\overline{x}$ より大き過ぎるとき，x_i の分布がポアッソン分布ではないと考えることができる．s^2 が $\overline{x}$ より大きいかどうかは，$s^2/\overline{x}$ の代わりに，その $n-1$ 倍である $\chi_0^2 = S/\overline{x}$ で判定する(正確には S/μ が カイ 2乗分布に従う)．

$$\chi_0^2 = \frac{S}{\overline{x}} \sim \chi^2(n-1) \tag{4.1}$$

第2単元 §3.1 の交通事故による死亡者数の分布が，ポアッソン分布になるかどうかを調べてみよう．必要な数値を，第2単元 §3.5 の表示3.20 から転記すると，

$$n = 305, \quad \overline{x} = 2.41, \quad S = 869.7,\ s^2 = 2.86$$

となる．これから，$s^2 = 2.86$ は $\overline{x} = 2.41$ よりいくぶん大きいように見える．その差が，統計的に有意かどうかを検定してみよう．

式(4.1)から，$\chi_0^2 = 869.7/2.41 = 361$ となる．自由度が 304 のカイ 2乗分布が 361 以上の確率を `CHIDIST` 関数で計算すると 0.014 となり，死亡者数は，ポアッソン分布よりバラツキの大きい分布をすることが分かる．

死亡者数がポアッソン分布にならないのは，1回の事故で数人が死ぬことによるものであろう．

もし，死亡者数の分布でなく，死亡事故数の分布を調べても，ポアッソン分布から期待されるバラツキよりも大きいバラツキが見られたときは，死亡事故が雨の日や月末に多かったり，休日にすくなかったりすることが原因であろう．これは，天候や曜日でデータを層別して解析することによって確かめることができる．

(2)　正規分布かどうかの検定

前項では，交通事故死亡者の分布にポアッソソ分布があてはまるかどうかを検定する方法を学んだが，ここでは，もっと広範囲の分布の適合度を検定する方法について学ぶことにしよう．

第2単元 §3.2 で中学生の体重の度数分布を求めた．この分布に，正規分布が当てはまるかどうかを検討してみよう．

第2単元 §3.5 の計算によれば，体重の平均値と標準偏差はそれぞれ50.9kg，5.06kgである．これから，もし正規分布が当てはまるならば，例えば46.95kg～48.95kg の間の人数は，標準正規分布で，

$$z = \frac{x - \overline{x}}{s} = \frac{46.95 - 50.9}{5.06} = -0.78$$

から，

$$z = \frac{48.95 - 50.9}{5.06} = -0.39$$

までの確率，すなわち，

$$\Pr[x < -0.39] - \Pr[x < -0.78] = 0.132$$

に，総人数110 人を乗じた人数(14.57 人) に近くなるはずである．このようにして計算した人数は，正規分布が当てはまると仮定したときに期待される人数であって，期待度数 と呼ばれる．

中学生の体重データについて以上の計算結果を表示4.4 の左に示す．

分布の裾の方で，期待度数が少ない部分はいくつかの級をプールして，期待度数が5 以上になるように併合する．その結果が表示4.4 の右に示す．

実測された度数と仮説のもとでの期待度数との差が大きいとき，仮説(この例では「正規分布が当てはまる」) を捨てることにする．両者の差の大きさは第3単元§4.1 で学んだ分割表の場合と同様，

$$\chi^2 = \sum \frac{(\text{実測度数} - \text{期待度数})^2}{\text{期待度数}} = \sum z^2$$

で測ることにする．

表示4.4: 中学生の体重の分布と正規分布の当てはめ

級上限	度数	z	期待度数	度数	期待度数	z
38.95	1	-2.37	0.99			
40.95	0	-1.97	1.70			
42.95	4	-1.57	3.66	5	6.35	-0.53
44.95	6	-1.18	6.77	6	6.77	-0.30
46.95	9	-0.78	10.73	9	10.73	-0.53
48.95	18	-0.39	14.57	18	14.57	0.90
50.95	21	0.01	16.94	21	16.94	0.99
52.95	16	0.40	16.89	16	16.89	-0.22
54.95	11	0.80	14.43	11	14.43	-0.90
56.95	10	1.19	10.56	10	10.56	-0.17
58.95	6	1.59	6.62	6	6.62	-0.24
60.95	4	1.99	3.56	8	6.14	0.75
62.95	3	2.38	1.64			
64.95	1	2.78	0.95			
総度数	110				カイ2乗	3.94
平均	50.9				自由度	7
標準偏差	5.06				p値	0.786

z_i と その2乗和 χ^2 が表示4.4の z の列に求められている．

χ^2 の自由度は級の数 k から3を引いたものである[1]．χ^2 に対するp値は 0.786 である．これから，帰無仮説は棄却できず，中学生の体重は正規分布に従うと考えて良い．

(3) χ^2 検定についての補足

[2]カイ2乗分布は前節で説明されたように，正規分布 $\mathrm{N}(\mu,\ \sigma^2)$ に従う母集団から得られた大きさ観測値 x があるとき，

$$z^2 = \left(\frac{x-\mu}{\sigma}\right)^2$$

またはその和の分布として導かれた．第4単元の§3.2 で説明された検定はこの基本的な性質から理解できるであろう．

[1] 自由度は級の数 k から (推定した分布の母数の数＋1) を引いたものになる．上の例ではデータから正規分布の平均と標準偏差を推定したので $k-(2+1)$ となる．

[2] この項は，解析手法の理論的な面に興味のない人は飛ばしても構わない

それに対して，第 3 単元の§4.2 の分割表や，前項で説明した分布についての検定でも χ^2 検定が用いられているが，なぜカイ 2 乗分布なのか理解に苦しまれたであろう．

分割表や分布についての検定では，各セルの期待度数が計算され，実測度数が期待度数からどれだけ離れているかを定量化するために χ^2 が計算された．

あるセルに出現する期待度数が決まると，実際の出現度数はポアッソン分布に従うと想像される．ここで，「ポアッソン分布では，期待値と分散が等しい」という性質を思い出して欲しい．標準偏差は期待値の平方根である．

これから，帰無仮説が正しいとき，

$$z = \frac{\text{実測度数} - \text{期待度数}}{\sqrt{\text{期待度数}}}$$

は標準正規分布で近似できることが分かる．

したがって，z を 2 乗して加えた値はカイ 2 乗分布に従うであろうことは理解できる．しかし，これは近似であって，期待度数がある程度大きくないとこの近似は良くない．そこで，分割表の χ^2 検定や，分布の検定では，「期待度数が 5 以上になることが望ましい」という条件がつけられている．

本日のまとめ

カイ 2 乗分布は，§3 のバラツキに関する検定や推定のために導入された．しかし，カイ 2 乗分布の適用範囲は広く，すでに第 3 単元の §4 における分割表の検定に用いられている．今日は，他の適用場所の一つを紹介した．今後統計を深く勉強すると，別の適用例に遭遇するであろう．

4.4 相関係数の検定と区間推定

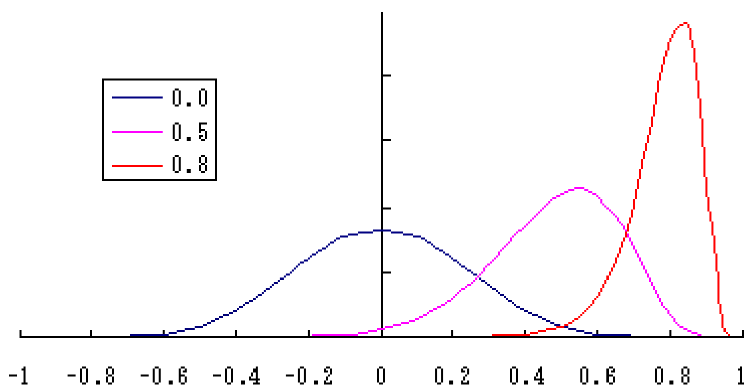
第2単元で2つの連続変数の関係の強さを表わす尺度として相関係数について説明した.

サンプルデータから計算した相関係数も，平均や標準偏差と同様に誤差を含む．すなわち，同じ母集団から取り出した別のサンプルから計算した相関係数は別の値になる.

(1) 相関係数の分布

母相関係数 ρ がそれぞれ $0.0, 0.5, 0.8$ の母集団から抽出した $n = 20$ のサンプルの相関係数 r の理論分布を表示4.5に示す.

表示4.5: 相関係数の分布



このグラフから分かるように，相関係数の分布は $\rho = 0$ のときは左右対称であるが，$\rho \neq 0$ のとき非対象である．したがって，そのまま正規分布で近似することはできない.

相関係数 r を

$$z = \frac{1}{2} \ln \left(\frac{1+r}{1-r} \right) \tag{4.2}$$

という式で変換した z は 正規分布

$$\mathrm{N}\left(\frac{1}{2}\ln\left(\frac{1+\rho}{1-\rho}\right),\frac{1}{n-3}\right)$$

で近似できる．この変換を フィッシャーの z-変換 と呼ぶ．

r から z，z から r を求める関数が

```
z = FISHER(r)
r = FISHERINV(z)
```

である．

これらの関数を使って，$n = 20,\ r = 0.81$ というデータが得られたとき，$\rho = 0.5$ の仮説検定と ρ の区間推定をしてみる．

Excel で計算する過程を表示4.6に示す．

表示4.6: 相関係数の検定と区間推定

	M	N	O	P	Q	R	S	T	U
3	相関係数の仮説検定								
4	ρ	n	r	z(ρ)	z(r)	z	下側確率	上側確率	両側確率
5	0.5	20	0.81	0.5493	1.1270	2.3820	0.009	0.991	0.017
6	P5:	=FISHER(M5)				S5:	=1-NORMSDIST($R5)		
7	Q5:	=FISHER(O5)				T5:	=1-S5		
8	R5:	=(Q5-P5)*SQRT(N5-3)				U5:	=2*MIN(S5:T5)		
9									
10	相関係数の区間推定								
11	n	r	α(両側)	z(r)	半幅	z(下側)	z(上側)	下側	上側
12	20	0.81	0.05	1.1270	0.475	0.6517	1.6024	0.573	0.922
13	P12:	=FISHER(N12)							
14	Q12:	=NORMSINV(1-(O12/2))/SQRT(M12-3)							
15	R12:	=P12-Q12				T12:	=FISHERINV(R12)		
16	S12:	=P12+Q12				U12:	=FISHERINV(S12)		

まず，$\rho = 0.5,\ r = 0.81$ のz 変換値を求める．

$$z(\rho) = z(0.5) = 0.5493$$

$$z(r) = z(0.81) = 1.1279$$

$z(\rho)$ と $z(r)$ の差を z の標準偏差 $1/\sqrt{n-3} = 1/\sqrt{20-3} = 0.2425$ で割って

$$z = \frac{1.1279 - 0.5493}{0.2425} = 2.3820$$

を計算する．帰無仮説が正しいとき，z は標準正規分布に従う．標準正規分布で z 以下と以上の確率を計算すると $0.009,\ 0.991$ となる．片側検定の場合は，対立仮説の側の確率が α よりも小さければ有意となる．対立仮説が $r > \rho$ であれば，有意差がある．

両側検定の場合のp値は，下側確率と上側確率の小さい方の2倍となる．

サンプルから計算された相関係数 r から母相関係数 ρ の区間推定も同じ考えで求められる．

r のz変換値を求め，それに，信頼区間の半幅を加減する．$\alpha = 0.05$ のとき，半幅は

$$\text{半幅} = \frac{1.96}{\sqrt{n-3}} = \frac{1.96}{\sqrt{18}} = 0.475$$

となる．z変換値の信頼区間 $(0.6517, 1.6024)$ を元の相関係数に戻すと母相関係数の信頼区間 $(0.573, 0.922)$ が得られる．

n 組のデータから相関係数 r が求められたとき，母相関係数 ρ の信頼区間の概数を読み取ることのできるグラフを表示4.7に示す．

横軸に r を取り，n に近い曲線の縦軸の値を読み取る．

例えば，$n = 100$ のデータから $r = 0.4$ が求められたとき，$n = 50, 200$ の2本の曲線の間を取って，$0.2 < \rho < 0.55$ という近似信頼区間を求めることができる．

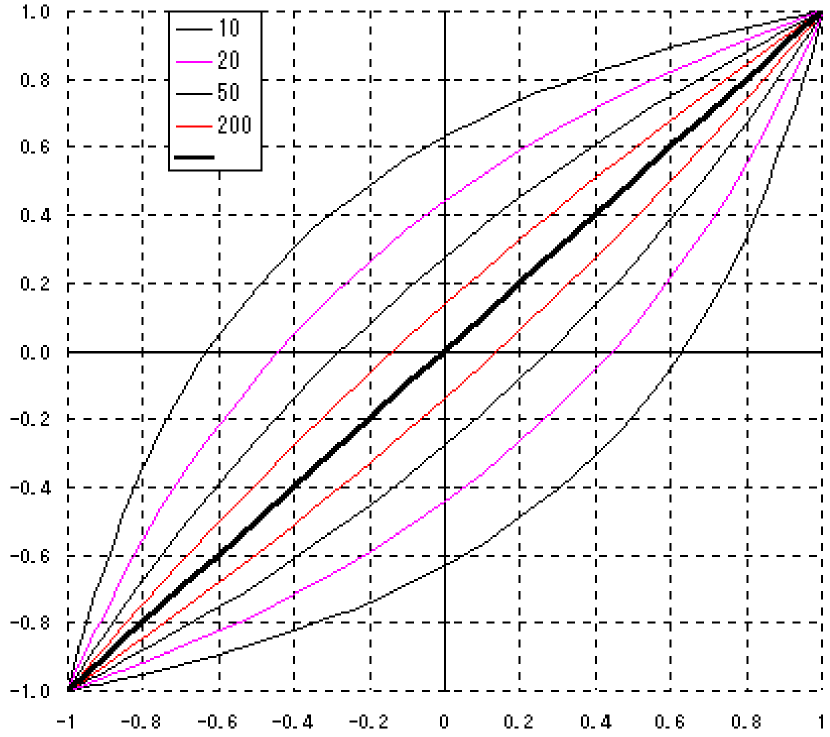
表示4.6に $n = 100,\ r = 0.4$ を入力すると，信頼区間として，$0.221, 0.553$ が得られる．グラフから求めた近似値にかなり近い．

■　**演習 14**　相関係数が $r = 0.4$ であるとき，母相関係数を $0.3 \leq \rho \leq 0.5$ 程度の信頼区間で推定するためには，どの位の n が必要かを考えよ．

ヒント　表示4.7で，横軸 r が 0.4 のとき，縦軸の $0.3, 0.5$ 付近を通る曲線を探す．

また，表示4.6の「区間推定」の計算表に，$r = 0.4$ と，上の方法で求めた n を入力し，右の下側，上側の値を見る．

表示4.7: 相関係数の信頼区間



期待する幅と離れているときは，n を増減させて，期待する幅の得られる n を探索する．

ゴールシークが使える人は，ゴールシークを使って n を求めて見よ．　■

本日のまとめ

例えば，標本相関係数の値が 0.5 のとき，相関があるといって良いのか，もし相関があるとしたら母相関係数の値を区間推定したいという要望があるであろう．今日は，それに対する答えを与えた．

最後に示すグラフは，母相関係数の信頼区間の概数を知るために役立つであろう．

5　演習解答

5.1　第1章　母平均に関する推定と検定

演習1(p.4)

第1**単元の演習**12**のデータから，平均値と標準偏差を計算し，偏差値** z **を計算する．**

$|z|$ **が** 1.960 **より大きいときは 太字の赤，**1.645 **より大きいときは 斜体の青で表示した．** 11/1, 11/2 **の** z **は** –2.87, –1.99 **で太字となり，**10/5 **の** z **は** 1.91 **で斜体となった．**

Excel **ファイル「第**4**単元演習」，シート「演習**1**」参照．**

演習2(p.8)

分析ツールを使って，100**行**7**列の正規乱数** N(130, 10^2) **をシート「演習**2**」の** `B4:H103` **に生成した．それから，**145**以上の個数を求めると** 57**個であった．**

表示5.1: **正規乱数のチェック**

	J	K	L
4	n	700	=COUNT(B4:H103)
5	145以上個数(x)	57	=K4-FREQUENCY(B4:H103,145)
6	割合 p=x/n	0.081	=K5/K4
7	pの期待値	0.067	=1-NORMSDIST((145-130)/10)
8	xの期待値	46.8	=K4*K7
9	xの分散	43.6	=K4*K7*(1-K7)
10	xの標準偏差	6.6	=SQRT(K9)
11	xの棄却限界（下限）	33.8	=K8-1.96*K10
12	（上限）	59.7	=K8+1.96*K10

正規分布 N(130, 10^2) **が** 145 **以上となる確率は** 0.0668 **である．** $\pi = 0.0668$, $n = 700$ **の**2**項分布で，** x **の期待値と標準偏差を計算すると** 46.8, 6.6 **が得られる．**

これから，x の棄却限界値は 33.8，59.7 となる．

これらの計算過程を表示5.1 に示す．

シミュレーションで得られた $x = 57$ はこの範囲内に入っている．ちなみに，テキスト表示1.3 のデータから x が145以上の個数を求めると（表示1.4参照）38 個である．この値も上の範囲内に入っている．

演習 3(p.9)

テキストの z を計算する式の分子の138 を135 に修正すると，$z = 1.323$ となる．標準正規分布で 1.323 以上の確率は 0.093 である．これは，0.05 よりも大きいから，血圧が正常時よりも 上昇したとはいえない．

表示5.2: $\overline{x} \geq 135$ の確率

期待値	130
標準偏差	10
n	7
標準誤差	3.780
xの平均	135
z	1.323
上側確率	0.093

演習 4(p.16)

表示5.3 の左上に x を入力する．さらに，帰無仮説の値 μ, σ と α（両側）に 1.54, 0.08 と 0.05 を入力する．

$z = -2.174$ で，両側p値が 0.030 となり，帰無仮説は棄却される．

95% 信頼区間は $1.435 < \mu < 1.535$ となる．この範囲内には，帰無仮説の値 1.54 は含まれない．

演習 5(p.21)

表示5.3 の右半分に x をコピーする．さらに，帰無仮説の値 μ と α（両側）に 1.54 と 0.05 を入力する．

標準偏差が 0.084 と求められ，t 値は–2.065 となる．両側p 値は 0.069 で，0.05 よりも大きく，帰無仮説は棄却されない．

95% 信頼区間は $1.425 < \mu < 1.545$ となる．この範囲内には，帰無仮説の値 1.54 が含まれる．

表示5.3: 平均値の区間推定

σ既知	（z検定）
	x
	データ省略
個数	10
平均	1.49
H0: μ	**1.54**
σ	**0.08**
母平均の検定	
標準誤差	0.025
z	-2.174
α（両側）	**0.05**
z(α)	1.960
棄却（下限）	1.490
（上限）	1.590
p値（下側）	0.015
（上側）	0.985
（両側）	0.030
母平均の区間推定	
下限	1.435
上限	1.535

σ未知	（t検定）
	x
個数	10
平均	1.49
H0: μ	**1.54**
標準偏差の推定	
平方和	0.064
自由度(f)	9
平均平方	0.007
標準偏差	0.084
母平均の検定	
標準誤差	0.027
t	-2.065
α（両側）	**0.05**
t(f, α)	2.262
棄却（下限）	1.480
（上限）	1.600
p値（下側）	0.034
（上側）	0.966
（両側）	0.069
母平均の区間推定	
下限	1.425
上限	1.545

5.2 第2章 母平均の差に関する推定と検定

演習6(p.29)

ヒントに従って，表示2.1 の計算表をコピーし，行数を調整してデータを入力する．$\mu_1 - \mu_2 = 0.0$, $\sigma = 10$, α(両側) $= 0.05$ を入力すると，表示5.4 の左半分が得られる．

$z = 0.972$ は1未満であって，p値を見るまでもなく，有意差はない．

平均値の差の信頼区間は $-4.573 \leq \mu_1 - \mu_2 \leq 13.573$ で，この区間内には 0 が含まれる．

演習7(p.34)

表示2.2 について，演習6と同様の手順で解析すると，表示5.4 の右半分が得られる．

表示5.4: 平均値の差の検定

σ既知	（z検定） データ省略		
個数(n)	14	7	差
平均	128.50	124.00	4.50
H0: μ1-μ2	**0.0**		
σ	**10**	**10**	
母平均の差の検定			
差の標準誤差	4.629		
z	0.972		
α（両側）	**0.05**		
z(α)	1.960		
棄却（下限）	-9.073		
（上限）	9.073		
p値（下側）	0.835		
（上側）	0.165		
（両側）	0.331		
母平均の差の区間推定			
下限	-4.573		
上限	13.573		

σ未知	（t検定） データ省略		
個数(n)	14	7	差
平均	128.50	124.00	4.50
H0: μ1-μ2	**0.0**		
標準偏差の推定			
平方和	1131.50	296.00	1427.50
自由度(f)	13	6	19
平均平方	87.038	49.333	75.132
標準偏差			8.668
母平均の差の検定			
差の標準誤差	4.012		
t	1.122		
α（両側）	**0.05**		
t(α)	2.093		
棄却（下限）	-8.398		
（上限）	8.398		
p値（下側）	0.862		
（上側）	0.138		
（両側）	0.276		
母平均の差の区間推定			
下限	-3.898		
上限	12.898		

演習6 と同様に，有意差は認められず，平均の差の信頼限界には 0 が含まれる．

演習 8(p.36)

(i)　表示2.2 をコピーし，B の8,9,10 番目のデータとして，47.1，48.0，46.7 を追加した．両群の平均平方は共に 0.39 となった．

等分散としての残差の自由度は18 であるのに対して，ウェルチの検定の自由度も 18.0 （小数点以下を詳しく表示すると，17.99985）となり，その違いはごく僅かである．

(ii)　B のデータから平均値に近い 47.3, 47.6, 47.7 を除く．B 群の平均平方は 0.56 と A 群の平均平方の 0.39 よりも大きくなった．ウェルチの検定の自由度も 11.4 と，等分散としての残差の自由度15 よりもかなり小さくなった．

前とは逆に，B のデータから平均値から遠い 48.2, 48.4, 46.7 を除く．群の平均平方は 0.23 と小さくなった．ウェルチの検定の自由度は 14.8 と，15 からそれほど低下しない．

表示5.5: ウェルチの検定の自由度

	A	B			A	B			A	B		
	(t検定)			(Welch)	(t検定)			(Welch)	(t検定)			(Welch)
	A	B			A	B			A	B		
1	49.6	47.3			49.6	*47.3			49.6	47.3		
2	48.6	48.2			48.6	48.2			48.6	*48.2		
3	48.3	47.6			48.3	*47.6			48.3	47.6		
4	47.9	48.4			47.9	48.4			47.9	*48.4		
5	48.9	46.6			48.9	46.6			48.9	46.6		
6	48.4	47.7			48.4	*47.7			48.4	47.7		
7	48.1	47.0			48.1	47.0			48.1	47.0		
8	47.5	47.1			47.5	47.1			47.5	47.1		
9	48.5	48.0			48.5	48.0			48.5	48.0		
10	47.6	46.7			47.6	46.7			47.6	*46.7		
個数(n)	10	10	差		10	7	差		10	7	差	
平均	48.34	47.46	0.88		48.34	47.43	0.91		48.34	47.33	1.01	
H0: μ1-μ2	0.0				0.0				0.0			
標準偏差の推定												
平方和	3.50	3.48	6.99		3.50	3.37	6.88		3.50	1.35	4.86	
自由度(f)	9	9	18	18.0	9	6	15	11.4	9	6	15	14.8
平均平方	0.39	0.39	0.39		0.39	0.56	0.46		0.39	0.23	0.32	
標準偏差			0.62				0.68				0.57	

ウェルチの検定の自由度 f' は，一般に通常の t 検定の自由度 f よりも小さくなる．

2つの群の平均平方の差が大きくなると，自由度の差 $f - f'$ が大きくなる．特に，n の小さい群の平均平方が大きいとき，この傾向が顕著である．

演習9(p.39)

通常の t 検定を実行すると，平均値の差の標準誤差が 7.132 となり，対応のある t 検定での値 1.881 の約4倍になる．そのために，t 値は 0.617 と小さくなり，p 値は 0.545 と0.05 よりも大きくなるので，有意差は認められない．

表示5.6: 対応のあるｔ検定と通常のｔ検定の比較

	差	A	B		
データ	6	555	549		
	12	533	521		
	1	541	540		
	9	545	536		
	10	566	556		
	-6	524	530		
	5	567	562		
	-4	515	519		
	3	535	532		
	8	533	525		
	対応あり		対応なし		
個数(n)	10	10	10		差
平均	4.4	541.4	537.0		4.400
平方和	318.4	2640.4	1938.0	4578.4	
自由度(f)	9	9	9	18	
平均平方(V)	35.4			254.4	
標準偏差	5.9			15.9	
平均値の差の標準誤差	1.881	7.132		13.928	
t	2.339	0.617		0.316	
p値(両側)	0.044	0.545		0.756	

１組目のデータから100を引いて，455, 449 とする．対応のあるｔ検定の結果は全く変化しない．それに対して通常のｔ検定の結果は表示5.6 の右下の黒枠にように変化する．品物間のバラツキの影響を受け，バラツキが大きくなると，差の検出力が低下する．

5.3　第３章　バラツキに関する推定と検定

演習 10(p.61)

表示3.5 をコピーし，演習問題に無関係な行を削除する[1]．$n = 30$, $S = 1240$, α(両側) $= 0.1$ を入力すると，表示5.7 が得られる．

演習 11(p.69)

表示3.11 の計算表をコピーし，データ，α（両側）を入力すると表示5.8 が得られる（データは省略）．

[1] 求めたい値をクリックし,「ワークシート分析」で参照元を順次たぐることにより，必要なセルを知ることができる．

表示5.7: 標準偏差の信頼区間

	例2
個数	30
標準偏差の推定	
平方和	1240
自由度(f)	29
母標準偏差の検定	
α（両側）	**0.1**
χ^2(1-α/2)	17.708
χ^2(α/2)	42.557
母標準偏差の区間推定	
下限	5.398
上限	8.368

表示5.8: 分散比の検定

個数(n)	5	6
標準偏差の推定		
平方和	0.308	0.873
自由度(f)	4	5
平均平方	0.077	0.175
標準偏差	0.277	0.418
	(F検定)	
母標準偏差の差の検定		
F	0.441	
α（両側）	**0.05**	
F(1-α/2)	0.107	
F(α/2)	7.388	
p値(下側)	0.224	
p値(上側)	0.776	
p値(両側)	0.448	
母標準偏差の比の区間推定		
下限	0.244	
上限	2.032	

F=0.4408 は 0.107〜7.388 の範囲内であるから，有意差は認められない．

また，両側のp値は 0.448 で，0.05 よりも大きいので，上と同じ結論が得られる．

また，母標準偏差の比の信頼区間は $0.244 \leq \dfrac{\sigma_A}{\sigma_B} \leq 2.032$ となる．この区間には 1.0 が含まれている．

5.4　第4章　補遺

演習12(p.77)

各自試みよ．　解答は省略．

演習13(p.77)

カイ2乗の表で，χ2 に 60，f に 50 を入力すると，上側確率が 0.1572 となる．

自由度が50のカイ2乗分布の期待値は $f=50$，標準偏差は $\sqrt{2f}=\sqrt{100}=10$ である．正規分布の表で，x, μ, σ に 60, 50, 10 を入力すると，上側確率が 0.1587 となる．正確な値 0.1572 にかなり近いことが分かる．

標準正規分布の表で，$z=(60-50)/10=1.0$ を入力しても同じ値が得られる．

表示5.9: カイ2乗分布の正規近似

正規分布					
x	μ	σ	下側確率	両側確率	上側確率
60	50	10	0.8413	0.317311	**0.1587**
標準正規分布					
z			下側確率	両側確率	上側確率
1.000			0.8413	**0.3173**	**0.1587**
カイ2乗分布					
x2	f		下側確率	両側確率	上側確率
60.0000	50		0.8428	0.3145	**0.1572**

演習 14(p.85)

まず，表示4.7 より，横軸の相関係数 $r = 0.4$ であるときの縦軸の母相関係数が $0.3, 0.5$ 付近を通る曲線を探す．$n = 200$ の曲線が該当する．

表示4.6 に $n = 200$, $r = 0.4$, α(両側) $= 0.05$ を入力する．信頼区間は $0.277 \le \rho \le 0.510$ となる．$n = 200$ では不十分であることが分かる．

下側の信頼限界を 0.30 にするために必要な n をゴールシークで求める.「数式入力セル」に `H5`,「目標値」に 0.3,「変化させるセル」に `A5` を入力して 「OK」をクリックすると，$n = 297$ が得られる．上側の信頼限界を 0.50 にするために必要な n を求めると，$n = 244$ が得られる．

相関係数は左右が非対称の分布をするので，どちら側で計算するかによって異なる結果が得られる．そこで，上限と下限の差が 0.20 になるn を求めることにする．下限，上限の右のセルに差を計算し，差が 0.20 となる n をゴールシークで求めると，$n = 272$ が得られる．

表示5.10: 相関係数を希望する精度で推定するために必要な n

	A	B	C	D	E	F	G	H	I	J
3	n	r	α(両側)	z(r)	半幅	z(下側)	z(上側)	下側	上側	
4	200	0.4	0.05	0.4236	0.140	0.2840	0.5633	**0.277**	**0.510**	
5	**297**	0.4	0.05	0.4236	0.114	0.3094	0.5379	0.300	0.491	
6	**244**	0.4	0.05	0.4236	0.126	0.2973	0.5500	0.289	0.500	
7	**272**	0.4	0.05	0.4236	0.120	0.3041	0.5431	0.295	0.495	0.200