第 3 章　平均の区間推定

本章では，母集団が正規分布の場合の平均の推定値のばらつきの度合いから，真の平均値が存在する区間をモンテカルロシミュレーションで体験します。

3.1　正規分布における平均値の推定

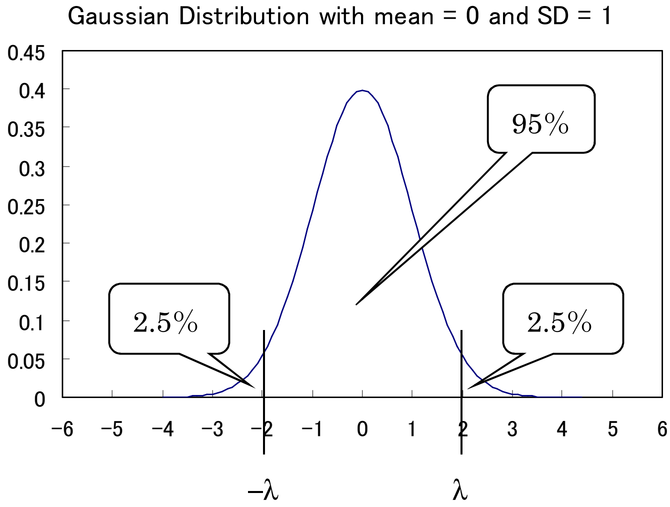


図 3.1　正規分布と確率

平均 0，標準偏差 1 の正規分布に従う乱数 X は確率 95%で，

$$-1.96 \leq \mathrm{X} \leq 1.96$$

の区間に存在します（図 3.1 参照）。第 2 章で示した標準偏差と確率の関係（式（2.3））から，一般に，平均μ，標準偏差σの正規分布に従う乱数は確率 P で

$$-\lambda \leq \frac{\mathrm{X}-\mu}{\sigma} \leq \lambda \qquad (3.1)$$

に存在することが分かります。確率 P とλの関係は，表 2.1 の通りです。たとえば，λ=1.65 ならば P＝90%です。

式（3.1）を書き換えてみます。式（3.1）にσをかけて，X を引くと

$$-\mathrm{X}-\lambda\sigma \leq -\mu \leq -\mathrm{X}+\lambda\sigma$$

となります。次に，－1 をかける（または不等号の左辺と右辺を入れ替える）と

$$X + \lambda\sigma \geq \mu \geq X - \lambda\sigma \tag{3.2}$$

となります。

λ=1.65 で P＝90%，σを既知，μを未知（つまり，知りたい量）として，式（3.2）を解釈してみます。X はいろいろな値をとる可能性がありますが，90%の確率で式（3.1）が成り立ちますから，式（3.2）も同じ確率で成り立ちます。つまり，式（3.1）より，X は 90%の確率で，区間μ ± 1.65σに存在し，この区間に X がない確率は高々10%になります。式（3.2）より，平均μは確率 90%で X ± 1.65σの区間に存在し，この区間にμがない確率は高々10%となります。この区間（X–1.65σ, X+1.65σ）を信頼区間といいます。

正規乱数で X を 100 個発生させてみます。すると，100 個の区間（式（3.2））が得られますが，λ=1.65 ならば，90 個の区間は平均値μを含み，10 個の区間はμを含まないことが想像できます。つまり，式（3.2）を平均の推定と考えると，90%の確率で正解が得られるわけです。すると，1 回の実験から X を求め，式（3.2）の推定を行った場合，μの存在する区間を 90%の確率で正しく推定できたと言ってもよいでしょう。

残念ながら，式（3.2）に基づく平均の推定は現実的ではありません。実験値 X が正規分布に従うと仮定すれば，λは任意に決められますから，問題はありません。しかし，標準偏差が既知σであるという仮定は現実的には問題があります。実験値の標準偏差が予め分かっていることはほとんどないからです。そこで，現実的な推定法を構築するために，σの代わりに，標準偏差の推定値 s を使って，式（3.1）を書き換えてみましょう。

3.2 t-分布における平均値の推定

式（3.1）を次のように書き換えます。

$$-t_\alpha \leq \frac{\mathrm{m} - \mu}{s/\sqrt{n}} \leq t_\alpha \tag{3.3}$$

ここで，m は，n 個の実験値 X_1，X_2，…，X_n からの平均の推定値

$$m = \frac{1}{n}\sum_{i=1}^{n} X_i \tag{3.4}$$

であり，s は，分散の推定値 s^2

$$s^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - m)^2 \tag{3.5}$$

の平方根です．式（3.1）における X の標準偏差をσとすると，n 個の X の平均 m の標準偏差は$\sigma/\sqrt{n}$ですから，式（3.3）は，式（3.1）の X とその標準偏差を，実験の平均値とその標準偏差の推定値で置き換えたことになります。式(3.3)で X の代わりに m を使う理由は，1 つの値 X だけからでは，s を推定できないからです（式（3.5）参照）。

ここで

$$\mathrm{t} = \frac{m-\mu}{s/\sqrt{n}} \tag{3.6}$$

と定義します。すると，t は実験ごとにランダムな値をとります。

(X – μ)/σ（式（3.1）参照）がランダムな値をとり，平均 0，標準偏差 1 の正規分布に従うことを考慮すると，t も正規分布に似た分布に従うと想像できます。もしそうならば，前項と同じ論理を使って，任意の信頼率（確率 P）をもって，平均μを区間で推定することができます。つまり，式（3.3）から，第 2 章と同じ論理を展開すると，その信頼区間は

$$m + t_\alpha \frac{s}{\sqrt{n}} \geq \mu \geq m - t_\alpha \frac{s}{\sqrt{n}} \tag{3.7}$$

となります[1]。t の分布は，自由度 n－1 の t-分布として知られています。つまり，P を指定すれば，対応する t_αが一意に定まります。正規分布の場合は，表 2.1 で確率 P とλの関係が分かります。同様に，P と t_αの関係が知られています（表 3.1 と 3.2)。

標準偏差が分かっている場合の平均の区間推定が正規分布を使い，標準偏差が分からな場合の平均の区間推定が t-分布を使います。実際に，式(2.3)はσを含んでいますが，式(3.7)は s を含んでいます。

3.3　t-分布のヒストグラム

正規乱数を使って，t-分布のヒストグラムを作ります。5 回のくり返し実験（n=5）から平均 m と標準偏差 s を推定し，式（3.6）から t を計算します。ただし，μ=5 とします。5 回のくり返し実験から 1 つの t 値が得られますので，この 5 回の実験を 1 シリーズの実験ということにします。ここでは，5000 シリーズの実験から，5000 個の t 値を計算し，そのヒストグラムを作ります。式（3.2）では 1 回の実験結果 X からμの信頼区間を推定できま

[1] 式（3.3）に$\mathrm{s}/\sqrt{\mathrm{n}}$を掛け，m を引くと，$-\mathrm{m} - \mathrm{t}_\alpha \mathrm{s}/\sqrt{\mathrm{n}} \leq -\mu \leq -\mathrm{m} + \mathrm{t}_\alpha \mathrm{s}/\sqrt{\mathrm{n}}$となりますから，－1 を掛ければ，式（3.7）が得られます。

すが，式（3.7）では，平均と標準偏差を使いますので，最低 2 回以上の実験の結果を使う必要があります。

図 3.2 に示されているように，行 1 に，「シリーズ」，「実験 1」，「実験 2」，「実験 3」，「実験 4」，「実験 5」，「平均」，「標準偏差」，「t」，「区間配列」，「度数」とタイプします。シリーズは，列 A に A2～A5001 まで 1 から順に番号を入力します[2]。列 B～F は実験値を入力します[3]。これは，平均値は 5，標準偏差は 1 の正規乱数です。平均，標準偏差，t の値は，各行で計算を行います（以下参照）。区間配列は－15～＋15 まで 0.2 間隔で入力します。度数は FREQUENCY 関数で計算します[4]。

	A	B	C	D	E	F	G	H	I	J	K
1	シリーズ	実験1	実験2	実験3	実験4	実験5	平均	標準偏差	t	区間配列	度数
2	1	4.113442	5.668044	5.592365	3.117855	4.50521	4.599366	1.068644	−0.83828	−15	0
3	2	5.448704	3.392077	5.703956	5.971727	4.571594	5.017611	1.049826	0.03751	−14.8	0
4	3	4.369265	6.134028	4.481638	4.79036	6.065139	5.168086	0.864545	0.434727	−14.6	0
5	4	5.45387	5.989419	6.269764	4.60457	5.817918	5.627108	0.643483	2.179101	−14.4	0
6	5	5.170457	6.426192	4.314352	4.564287	5.638124	5.222683	0.848867	0.586568	−14.2	0
7	6	5.123823	4.928068	5.820164	5.249305	4.332815	5.090835	0.538333	0.377289	−14	0
8	7	4.837764	6.337198	5.104881	5.277898	5.339201	5.379388	0.569655	1.48917	−13.8	0
9	8	2.985979	5.15658	3.457151	6.932103	4.16464	4.539291	1.567703	−0.65711	−13.6	0
10	9	5.11504	4.011579	5.422042	5.620686	4.828999	4.999669	0.629051	−0.00118	−13.4	0
11	10	4.21499	6.321466	3.89023	5.758632	5.600774	5.157218	1.049639	0.334916	−13.2	0
12	11	4.188904	4.666385	5.644891	5.074847	5.166345	4.948274	0.548852	−0.21073	−13	0
13	12	4.763223	3.388448	3.462163	4.540951	5.743821	4.379721	0.982137	−1.41217	−12.8	0
14	13	6.72563	5.32175	6.942599	4.38052	5.486336	5.771367	1.06069	1.626089	−12.6	0
15	14	3.676699	5.507544	5.015185	4.991012	6.461331	5.130354	1.007418	0.289325	−12.4	0
16	15	6.183271	4.761885	5.795148	6.713879	4.972878	5.685412	0.818269	1.872956	−12.2	0
17	16	4.545112	4.586889	5.740296	4.819831	6.863582	5.311142	0.99335	0.700371	−12	0
18	17	4.193783	4.094009	5.94796	5.660605	4.46416	4.8721	0.867637	−0.32961	−11.8	0
19	18	5.355615	4.81711	5.551463	3.844912	7.351917	5.384203	1.283716	0.669213	−11.6	0

図 3.2　データの入力

列 G の平均は次のように行います。セル G2 から G5001 まで選択し，数式ボックスに「=AVERAGE(B2:F2)」と入力し，Ctrl キーを押しながら Enter キーを押します[5]。列 H の標準偏差と列 I の t も同様に計算します。t の値は，式（3.6）を使います。

図 3.3 に 5000 個の t の値（列 I）の散布図を示します。幾つかの値は中心から 10 以上離れていますが，このヒストグラム（図 3.4）では，図を見やすくするため，これらは表示しません。図 3.4 のヒストグラムは，区間配列としては－10～＋10 の範囲で示してあります。図 3.4 から分かるように，t－分布は正規分布と似た形をしています。これらの相違については，次の項で調べます。

---

[2] 2.1 を参照してください。

[3] [データ]→分析[データ分析]→[乱数発生]で，［平均]を 5，[標準偏差]を 1，[出力先]を $B$2:$F$5001 と設定します。

[4] 項 1.4 を参照してください。

[5] Ctrl キー＋Enter キーは配列の設定を意味します。第 1 章付録の配列数式を参照してください。

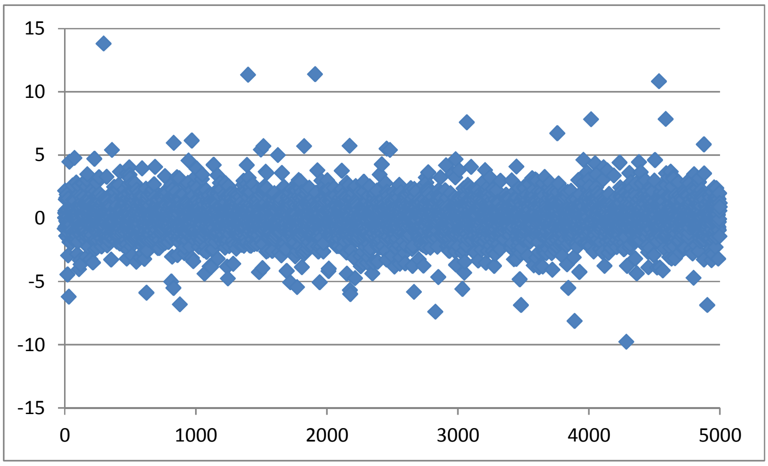


図 3.3　t の例

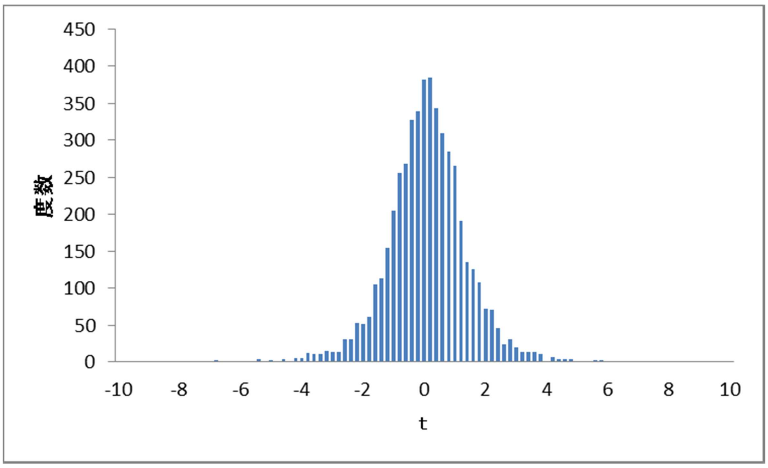


図 3.4　t のヒストグラム

3.4　t-分布の確率密度

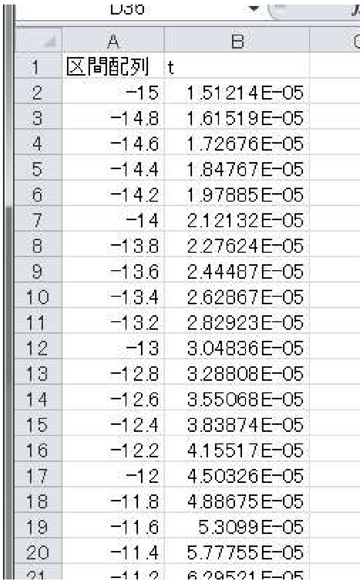
前項で得られた t-分布ヒストグラムが t-分布の確率密度関数と一致することを確認します。前項で設定した区間配列－15，－14.8，…と同じ区間（－15，15）と間隔（= 0.2）で関数 T.DIST を使って，t-分布を描きます。図 3.5 にあるように，セル A1 に「区間配列」，B1 に「t」と入力します。セル A2～A152 には，前項で設定した区間配列の値をコピーします。B2 を選択し，[数式]→[その他の関数]→[統計]→[T.DIST][6]を選択し，現れた[関数の引数]ダイアログボックスに次のように入力します（図 3.6）：

X :　　　　A2

自由度 :　　4

[6] Excel2007 の場合は，対応する関数は TDIST ですが，この関数では t-分布密度を簡単に得ることはできません。

関数形式：　false
この関数を B3～B152 までコピーします[7]。この結果を図示したのが図 3.7（左）です。前項で得られた t-分布ヒストグラム（図 3.7（右））と一致することが分かります。



	A	B
1	区間配列	t
2	−15	1.51214E-05
3	−14.8	1.61519E-05
4	−14.6	1.72676E-05
5	−14.4	1.84767E-05
6	−14.2	1.97885E-05
7	−14	2.12132E-05
8	−13.8	2.27624E-05
9	−13.6	2.44487E-05
10	−13.4	2.62867E-05
11	−13.2	2.82923E-05
12	−13	3.04836E-05
13	−12.8	3.28808E-05
14	−12.6	3.55068E-05
15	−12.4	3.83874E-05
16	−12.2	4.15517E-05
17	−12	4.50326E-05
18	−11.8	4.88675E-05
19	−11.6	5.3099E-05
20	−11.4	5.77755E-05

図 3.5　t-分布密度の計算

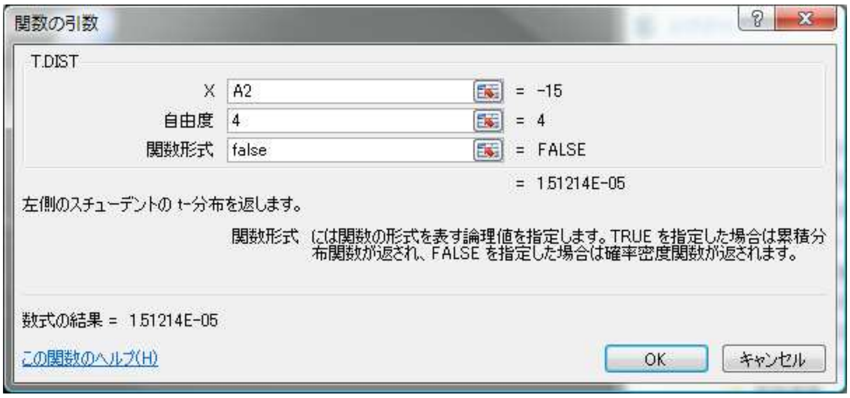


図 3.6　T.DIST の設定

　t-分布ヒストグラム（図 3.4）と t-分布確率密度（図 3.7（左））では Y-軸のスケールが異なります。ヒストグラムの Y-軸は度数ですが，確率密度は X-軸の単位長さ分の確率であり，定義により，確率密度の面積は 1 と定められています。一方，図 3.4 のヒストグラムでは，

[7] 普通にコピーした場合は，F9 を押して再計算させます。あるいは，最初にセル B2 から B152 を選択し，数式ボックスに=T.DIST(A2,4,FALSE)と入力し，Ctrl を押しながら，Enter を押します。

度数 1 を表す（仮想）ブロックは高さが 1 で，X-軸は 0.2 の間隔で設定しましたので，幅が 0.2 ですから，面積は 0.2 になります。すると，度数の合計は 5000 ですから，全体の面積は 1000（=5000×0.2）です。そこで，ヒストグラム全体の面積を 1 とするためには，図 3.4 の Y-軸のスケールに 1/1000 をかければ良いことになります。図 3.4 の最大値は約 400 ですが，図 3.7（右）は約 0.4 ですから，1/1000 の意味が分かると思います。

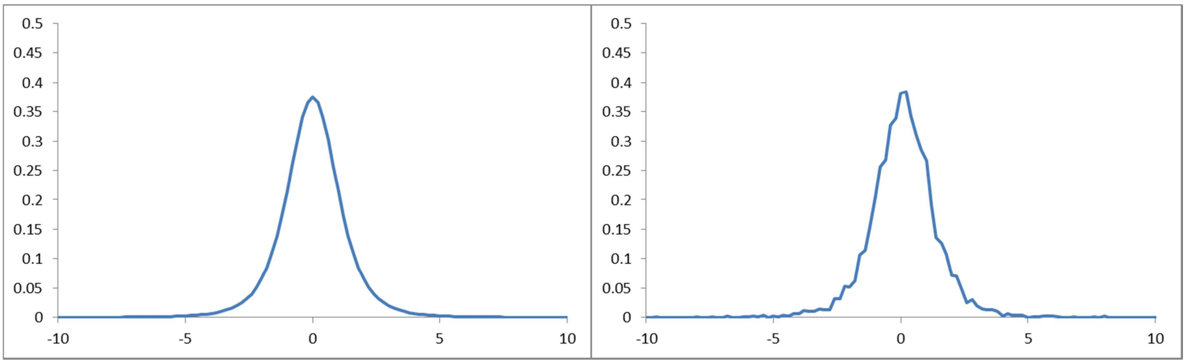


図 3.7　t-分布の確率密度（n=5）（左）と図 3.4 のヒストグラムからの確率密度（右）

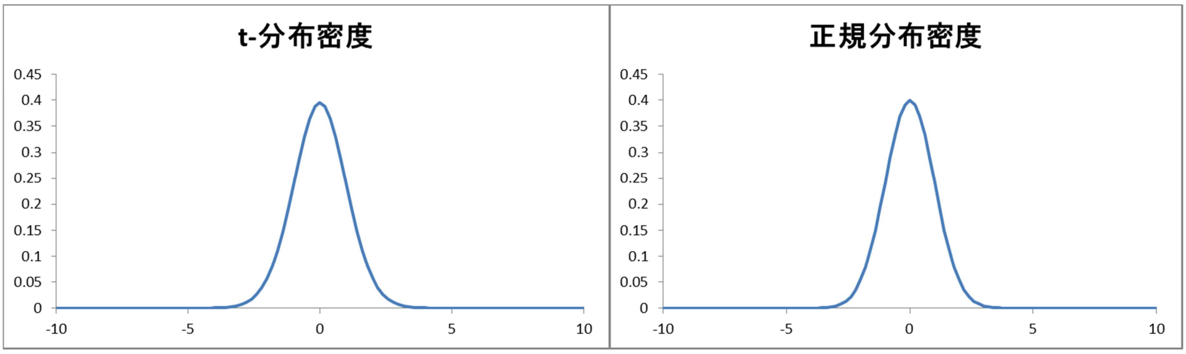


図 3.8　t-分布の確率密度（n=31）と正規分布密度[8]

3.5　t-分布の性質

t-分布は正規分布によく似た分布をしていますが，正規分布よりも少し裾が広がっています。図 3.8 の t-分布の確率密度と正規分布の確率密度の形が非常に似ていることから想像できるように，くり返し数 n が大きくなるほど，t-分布は正規分布に近い形になります。

正規分布では，変数 X の区間と確率の関係は需要です（表 2.1 と図 3.1 参照）。同様に，t-分布では，変数 t の区間と確率の関係が重要です。確率が 95%となる

$P\{|t| \le t_\alpha\} = 0.95$

t_αは，n=5 では 2.776，n=31 では 2.042 です。正規分布では，この値は 1.96 ですから，n

[8] NORMDIST 関数を使って描きます。

の増大とともに，t 分布は正規分布に近づくことが分かります。

表 3.1　t-分布の確率と区間（n=5）

t_α	2.132	2.776	4.604
$P\{\lvert t \rvert \le t_\alpha\}$	0.900	0.950	0.990

表 3.2　t-分布の確率と区間（n=31）

t_α	1.697	2.042	2.750
$P\{\lvert t \rvert \le t_\alpha\}$	0.900	0.950	0.990

式（3.7）を見ると，実験から推定した標準偏差 s に$t_\alpha/\sqrt{n}$を掛けた値が，信頼区間の幅の 1/2 であることが分かります。信頼区間の中心は平均 m であり，m はくり返し数 n が大きくなるに従って，真の値μに近づいて行きますから，信頼区間の幅はしだいに狭くなるはずです。つまり，$t_\alpha/\sqrt{n}$は，n が大きくなるに従って，小さくなります。

3.6　平均値の区間推定

95%信頼区間の 95%の意味をモンテカルロシミュレーションで示します。95%は，たとえば，5 回の繰り返し実験（1 シリーズ）を 100 シリーズ行ない，100 個の信頼区間を得た場合，真の平均を含んでいる信頼区間は 95 個あるという意味です。

平均μ=5，標準偏差σ=1 の正規乱数表を使って，真の平均値μの 95%信頼区間と 90%信頼区間を 100 個作り，比較します。図 3.9 にあるように，信頼区間を棒グラフとして表示します。グラフとしては，2 つの棒が縦に重なっている[積み上げ縦棒]を使い，下の棒を非表示とします。 下の棒の長さは，信頼区間の下限

$$\mathrm{m} - t_\alpha \frac{s}{\sqrt{n}} \tag{3.8}$$

であり，上の棒の長さは信頼区間の幅

$$2t_\alpha \frac{s}{\sqrt{n}} \tag{3.9}$$

です。

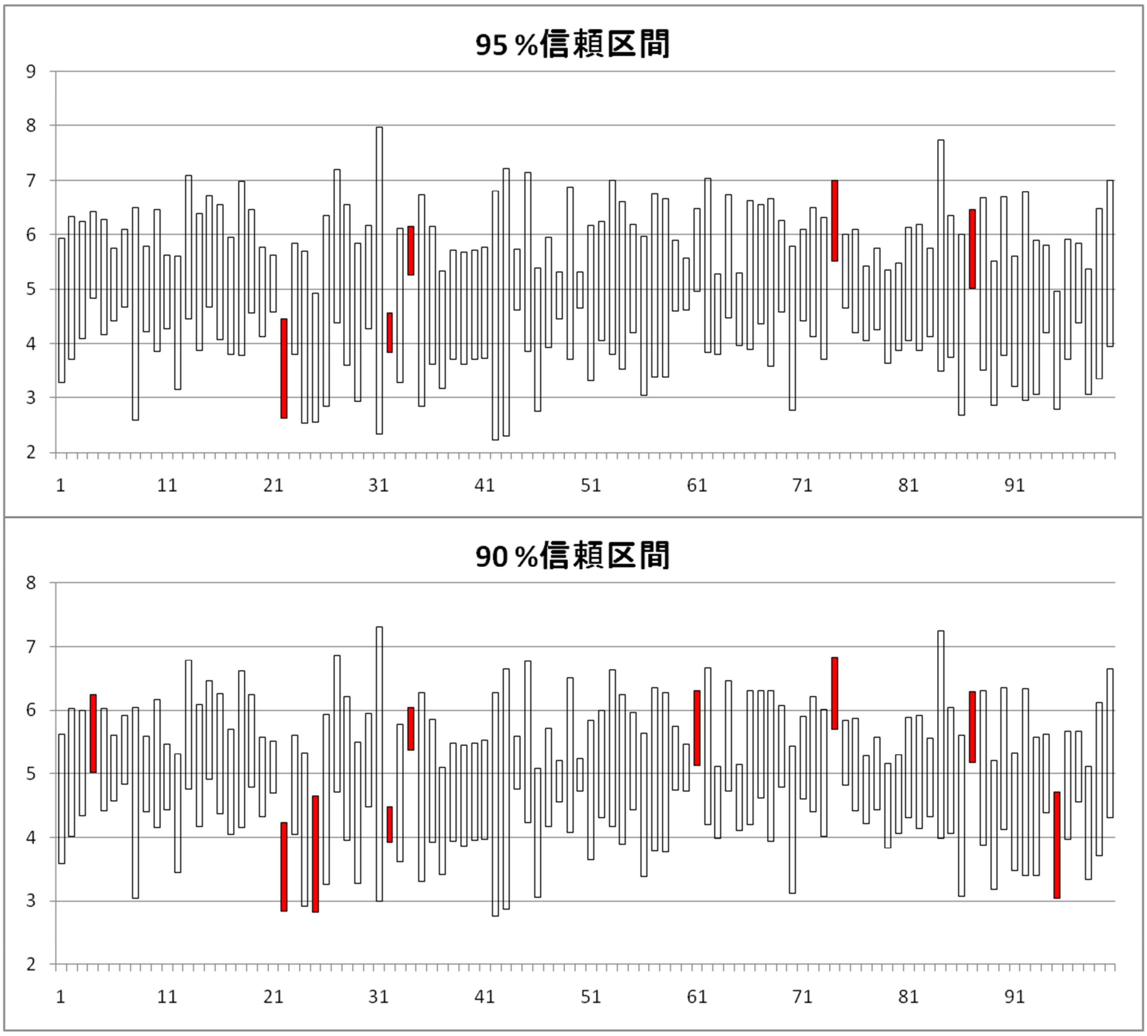


図 3.9　95%信頼区間と 90%信頼区間のグラフ表示

	A	B	C	D	E	F	G	H	I	J	K	L
1	セット	実験1	実験2	実験3	実験4	実験5	平均	標準偏差	95%CI下	95%CI幅	90%CI下	90%CI幅
2	1	4.113442	5.668044	5.592365	3.117855	4.505121	4.599366	1.068644	3.272641	2.653449	3.580426	2.037879
3	2	5.448704	3.392077	5.703956	5.971727	4.571594	5.017611	1.049826	3.71425	2.606724	4.016615	2.001994
4	3	4.369265	6.134028	4.481638	4.79036	6.065139	5.168086	0.864545	4.094752	2.146669	4.343753	1.648667
5	4	5.45387	5.989419	6.269764	4.60457	5.817918	5.627108	0.643483	4.828223	1.597771	5.013555	1.227106
6	5	5.170457	6.426192	4.314352	4.564287	5.638124	5.222683	0.848867	4.168812	2.107742	4.413298	1.61877
7	6	5.123823	4.928068	5.820164	5.249305	4.332815	5.090835	0.538333	4.422493	1.336683	4.577541	1.026588
8	7	4.837764	6.337198	5.104881	5.277898	5.339201	5.379388	0.569655	4.672161	1.414456	4.836229	1.086318
9	8	2.985979	5.15658	3.457151	6.932103	4.16464	4.539291	1.567703	2.592984	3.892614	3.044505	2.989572
10	9	5.11504	4.011579	5.422042	5.620686	4.828999	4.999669	0.629051	4.218701	1.561937	4.399876	1.199586
11	10	4.21499	6.321466	3.89023	5.758632	5.600774	5.157218	1.049639	3.854089	2.606259	4.1564	2.001637
12	11	4.188904	4.666385	5.644891	5.074847	5.166345	4.948274	0.548852	4.266873	1.362803	4.42495	1.046649
13	12	4.763223	3.388448	3.462163	4.540951	5.743821	4.379721	0.982137	3.160395	2.438652	3.443265	1.872913
14	13	6.72563	5.32175	6.942599	4.38052	5.486336	5.771367	1.06069	4.454517	2.633699	4.760011	2.022711
15	14	3.676699	5.507544	5.015185	4.991012	6.461331	5.130354	1.007418	3.879642	2.501424	4.169793	1.921123
16	15	6.183271	4.761885	5.795148	6.713879	4.972878	5.685412	0.818269	4.669529	2.031766	4.905202	1.56042

図 3.10　100 個の 95%信頼区間の計算

図 3.2 において t 値を計算するときに使ったデータをそのまま使います。実験 1~5，平均と標準偏差をそれぞれ 100 個含むデータをコピーし, 新しいシートに貼り付けます(図 3.10 参照)。行 1 の I1~L1 まで，「95%IC 下」,「95%IC 幅」,「90%IC 下」,「90%IC 幅」とタイプします。

式（3.8）をセル I2 に入力します。実際には,「=G2–2.776*H2/2.236」とタイプします。G2 は平均の推定値が収容されているセル，2.776 は 95%信頼区間を表す t 値（表 3.1），H2 は標準偏差推定値が収容されているセル，2.236 は$\sqrt{5}$です。セル J2 には，式（3.9）を使い,「=2*2.776*H2/2.236」とタイプします。次に，I2 と J2 の数式を，I3~J101 までコピー[9]すれば 95%信頼区間の計算は終了です。

95%信頼区間の下限と幅を示す配列 I2~J101 を選択し，[挿入]→[縦棒]→[積み上げ縦棒]の順で進めば，図 3.9 の元となるグラフが現れます。後は，付録にあるように体裁を整えれば，図 3.9 が得られます。

90%信頼区間も同様な方法で計算できます。ただし，t 値としては 2.132 を使います（表 3.1 参照)。

100 シリーズの実験から得られた 100 個の信頼区間のグラフ（図 3.9）を見ると，信頼区間の位置と幅がシリーズ実験毎に大きく異なることが分かります。95%信頼区間では，5 個の信頼区間の外に真の平均があり, 90%信頼区間では, 9 個の信頼区間から真の平均が外れています（赤で示してあります)。95%と 90%の違いは，ここにあります。つまり，真の平均μを推定する信頼区間は，実験から得られた平均値と標準偏差を含むため，実験毎に異なります．しかし，同じくり返し実験のシリーズを数多く行なうと，95%または 90%の確率で，真の平均値が信頼区間に含まれることになります。

式（3.7）が示すように，信頼区間の中央は平均推定値 m です。一方，信頼区間の幅は，$2t_\alpha\frac{s}{\sqrt{n}}$であり，同じ回数のくり返し実験では，$t_\alpha$ と n は定数ですから，この幅のバラツキは実験値 s だけに依ります。では，くり返し実験から求まる標準偏差 s は，実験毎にどのくらいバラツクのでしょうか．この答えが χ^2-分布です（4 章参照)。

3.7　信頼区間と宝くじ

宝くじは，江戸時代に流行した富くじの一種です。富くじは，販売された多数の富札の中から抽選により賞金の当たる賭博です。現在では，富札は宝くじ券と呼ばれています。2010 年現在，当選金額が約一億円のある宝くじの当選確率は約 600 万分の 1，つまり，0.000017%ですから，当たる確率はほとんどゼロで，この数字は実感がわきません。しかし，四等 1 万円の当選確率は 610 分 (=0.16%) くらいですから，ずっと身近に感じられます。千三つと言われている大うそつきでも，1000 回に 3 回は本当のことを言うのですから，性善説に立脚すれば，この人物の善意を期待できるくらいの確率です。

[9] I2 と J2 を同時に選択し，Ctrl キーと C キーを同時に押してから離した後，I3 をクリックした後，Shift キーを押しながら J101 をクリックし，Enter キーを押します。

ある人が，当選確率が 610 分の 1 の宝くじ券を 1 枚持っていると仮定します。1 枚の券は，当たる場合と当たらない場合の 2 つの場合しかないわけですから，610 分の 1 の確率はどのように解釈すればよいのでしょうか。やはり，もし宝くじ券を 610 枚持っていれば，そのうち 1 枚は当たると考えることが妥当でしょう。

この状況は, 図 3.9 の信頼区間と似ています。科学実験は, 仮説を証明するために行われます。通常，科学では，1 シリーズの実験から 1 つの信頼区間を得て，実験結果を評価します。複数回のシリーズ実験を行うことはしません。つまり，1 枚の宝くじ券を買うことに相当します。

1 回のシリーズ実験から得られた 95%信頼区間が 95%の確率で真の平均を含んでいると仮定しても, 実際に自分が得た信頼区間が真の平均を含んでいるかどうかは実験者には分からないわけです。宝くじならば，後日抽選会の日に，自分の持っている宝くじ券の当落は明らかになりますが，信頼区間の場合は抽選会はありませんから，その真偽が明らかになることはありません。

科学における仮説が 95%の確率で正しいとする結論は科学では満足できるものではありません。全く別な実験でこの仮説を証明する，関連した知識を総合して仮説を傍証するなどの手段が必要です。つまり，違った形式の抽選会を行うわけです。

3.8　系統誤差と有意差検定

自分が新しい測定法を開発したことを想像します。この測定法が信頼できることを確認するために第一にすることは，量が分かっている標準物質を測定して，測定値が正しい測定値 μ に一致することを確認することです。測定には必ず誤差が伴いますが，多くのくり返し測定から得られた測定値の平均は μ に近いと期待できるでしょう。残念ながら，現実にはそうならない場合があります。たとえば，測定値（たとえば電圧）が測定対象物質の量に比例していない，常に余分な電圧が加わっているなどの原因から誤差が生じます。この場合，誤差の原因は常に同じですから，誤差はランダムではなく，一定の値を取ります。この誤差を系統誤差といいます。系統誤差があるかどうかを調べるのが，この項の主題です。

系統誤差とは違い，測定ごとにランダムに現れる誤差を偶然誤差といいます。ある測定系に系統誤差はなく，偶然誤差だけがあれば，無限界の測定値の平均は真の値 μ に一致します。以下で，系統誤差と偶然誤差を含んでいる測定値から，偶然誤差の存在を調べます。

測定値は，一般に，

測定値　＝　真の値 μ　＋　系統誤差　＋　偶然誤差

と表されます。

試験に標準物質を使う場合，今問題にしている測定法が取るべき真の値 μ を我々は知っていると仮定します。そこで，測定値と真の値の差

測定値　－　真の値 μ　＝　系統誤差　＋　偶然誤差

から，系統誤差を調べることにします。しかし，偶然誤差が大きい測定系では，偶然誤差が小さい測定系に比べて，測定値が大きくばらつきますから，系統誤差を調べるためには，どちらの系でも偶然誤差を同じスケールで測る工夫が必要です。そこで，測定値は正規分布に従うと仮定し，上の式の左辺を測定値の標準偏差で割ることにします：

$$\frac{\text{測定値} - \text{真の値}\ \mu}{\text{測定値の標準偏差}}$$

1 個の測定値と μ の差で，測定法の系統誤差を調べるのはあまりにも荒っぽいですし，測定値の標準偏差も求まりませんから, n 回のくり返し測定から得られる n 個の測定値の平均で上の式を書き換えてみます：

$$\frac{\text{平均} - \text{真の値}\ \mu}{\text{平均の標準偏差}}$$

平均の標準偏差は，（測定値の標準偏差）÷$\sqrt{n}$ですから，この式は

$$\frac{\text{測定値の平均} - \text{真の値}\ \mu}{\text{測定値の標準偏差}/\sqrt{n}} \tag{3.10}$$

となります。

式(3.10)と式(3.3)を比べると, 式(3.10)は t-分布に従い, 式(3.10)の値は, $-t_\alpha$と$+t_\alpha$の間に決まった確率で存在することが分かります。ただし，測定値は正規分布に従い，系統誤差はないとする過程が必要です。

3.9　帰無仮説と有意差検定

式（3.6）は t の定義です。ある実験系でくり返し測定を行えば，平均と標準偏差は計算できますが，真の値 μ は実験からは分かりませんから，t は計算できません。しかし，何らかの根拠により，その実験系の μ の値を決定できれば t は計算できます。μ を定めるための根拠を帰無仮説といいます。すると，t の値を使って，帰無仮説の真偽を調べることができます。

一般に，式（3.3）より，t がある確率で存在する範囲は

$$t_\alpha \geq |t| \tag{3.11}$$

です。帰無仮説が正しければ，実験から得た t 値は高い確率（たとえば 95%）で式（3.11）の範囲にあります。しかし，帰無仮説が正しい場合でも，低い確率（たとえば 5%）で t 値は範囲外にあります。もし帰無仮説が間違っていれば，高い確率で t 値は範囲（3.11）外にあるでしょうか？残念ながら，現実にはこれは分かりません。帰無仮説が正しくない場合でも，仮定したμが真の値と無視できるくらいしか違わなければ，t 値が範囲（3.11）外にある確率は依然として低いでしょう。しかし，μと真の値の差が十分に大きければ，t 値が範囲（3.11）外にある確率は高いと考えられます。すると，次のような判定方法が考えられます。問題の実験系の t が，式（3.11）範囲内にあれば帰無仮説は棄却されませんが，範囲外ならば帰無仮説は棄却されます。

分析の例を挙げて，有意差検定と帰無仮説を考察します。ある分析法により，水銀 5.00%を含む標準物質について下の表の分析結果が得られたとします。この分析法の系統誤差の有無を調べます。

	実験 1	実験 2	実験 3	実験 4	実験 5
分析結果（%）	4.11	5.67	5.59	3.12	4.51

測定に用いた標準物質の水銀濃度は 5%と分かっていますので，式（3.6）の$\mu=5$と仮定します。この仮定が帰無仮説であり，系統誤差はないとする仮定でもあります。この分析結果の平均は 4.60，標準偏差は 1.07 ですから，t は

$$t = (4.60 - 5)/(1.07/\sqrt{5}) = -0.84$$

となります。n=5（自由度=4）の場合，t が 95%の確率（危険率 5%）で存在する範囲はt_α=2.78 です。この場合，t は 2.78 を超えませんから，帰無仮説は棄却されません。つまり，系統誤差はないと考えてよいことになります。

この項では，系の平均 m と帰無仮説μとの差が有意であるかどうかを検定しました。有意差検定という名称はここから来ています。上の例では，5%の水準で有意とは言えませんでした。この有意水準 5%は，危険率ともいいます。また，P=0.05 と書くこともあります。

3.10　対になったデータの t 検定

前項では，帰無仮説が否定されない方が都合がよい例を示しました。この項では，帰無仮説を否定したい例を示します。

A，B 二つの睡眠薬の優劣を調べる仮想実験を考察します。下の表は，それぞれの睡眠薬によって伸びた睡眠時間とその差を示しています。睡眠薬の効果の差（行 3）の平均 m が，0 と有意に差があるかを調べることにします。つまり，帰無仮説は$\mu=0$であり，二つの睡眠薬には差がないとする仮説です。B－A（行 3）を見ると，m は正であることが分かりま

すから，B の方が，A よりも効果が高いと思えます。この直感を統計的に検証します。

A の効果(時間)	5.4	3.3	5.7	5.9	4.5	4.3	6.1	4.4	4.7	6.1
B の効果(時間)	6.9	7.4	7.7	6.1	7.3	6.6	7.9	5.8	6	7.1
B−A	1.5	4.1	2	0.2	2.8	2.3	1.8	1.4	1.3	1

差 B−A の平均は 1.84，標準偏差は 1.07 ですから，t は

$$t = 1.84/(1.07/\sqrt{10}) = 5.45$$

となります。n=10（自由度=9）の場合，t が 95%の確率（危険率 5%）で存在する範囲はt_α=2.26 です。この場合，t は 5.45 ですから，危険率 5%で，B は A よりも効果が高いと判定することになります。

3.11　付録

信頼区間のグラフ表示

95%信頼区間の下限と幅の積み上げ縦棒のグラフは，3.6 の操作では，図 3.11 のようになるはずです。この項では，図 3.9 のように体裁を整える方法を説明します。

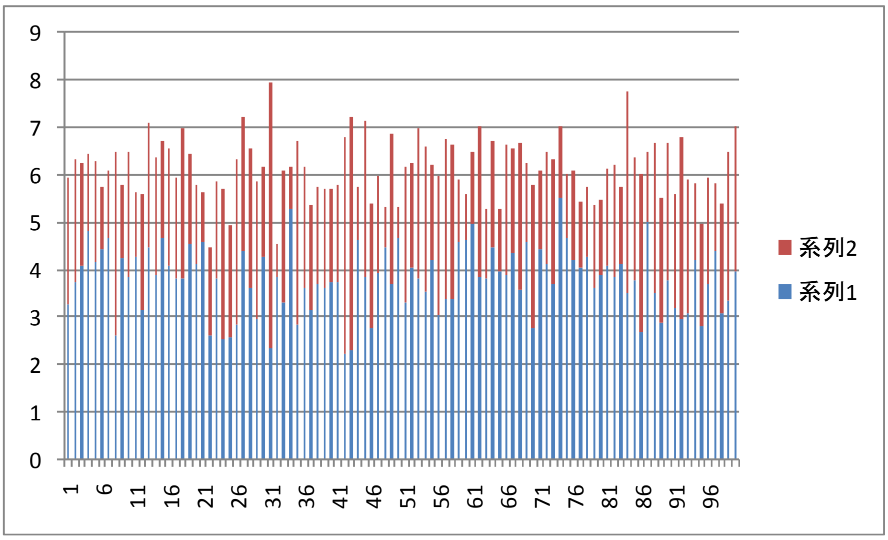


図 3.11　95%信頼区間の下限と幅の積み上げ縦棒のグラフ

凡例を削除し，図全体（グラフエリア）を横に広げます。X-軸の目盛の上を右クリックし，現れたポップアップメニューから[軸の書式設定（F）...]を選択し，[軸の書式設定]ダイアログボックスの[間隔の単位（S）：]に 10 と入力します。[閉じる]ボタンを押します。

積み上げ縦棒の下の棒の上で右クリックし，現れたポップアップメニューから[データ系列の書式設定（F）...]を選択し，[データ系列の書式設定]ダイアログボックスの[塗りつぶし]から，[塗りつぶしなし（N）]のラジオボタンをクリックし，[閉じる]ボタンを押します。下の棒が消えたはずです。

上の棒の上を右クリックし，現れたポップアップメニューから[データ系列の書式設定（F）...]を選択し，[データ系列の書式設定]ダイアログボックスの[塗りつぶし]から[塗りつぶしなし（N）]ラジオボタンをクリックし，[枠線の色]から[線（単色）（S）]ラジオボタンをクリックします。次に，[系列のオプション]から[要素の間隔（W）]スライダーを「80%」に設定します。[閉じる]ボタンを押します。

真の平均を含んでいない信頼区間を塗りつぶすために，そのような棒の上を左クリック[10]します。目的の棒だけが選択されたことを確認した後，その棒上で右クリックし，現れたポップアップメニューから[データ要素の書式設定（F）...]を選択し，[データ要素の書式設定]ダイアログボックスの[塗りつぶし]から[塗りつぶし（単色）（S）]ラジオボタンをクリックし，[色（C）：]から適当な色を選択します。[閉じる]ボタンを押します。この操作を目的とするすべての信頼区間の棒に対して行います。

グラフにタイトルを付けるために，まず，グラフの端をクリックし，グラフ全体を選択します。選択状態は，[デザイン]，[レイアウト]と[書式]の 3 つのタブが新たに現れることで確認できます。[レイアウト]→ラベル[グラフタイトル]→[グラフの上]の順で進み，グラフタイトルとして「95%信頼区間」とタイプします。

以上の操作で，図 3.9 のような体裁のグラフが得られたはずです。

[10] 全ての棒を選択するときは右クリックを使い，特定の棒を選択するときは左クリックを使います。現れるポップアップメニューとダイアログボックスには「データ系列」と「データ要素」の違いがあります。