境界復元に使う誤差分布の種類

さまざまな確率分布

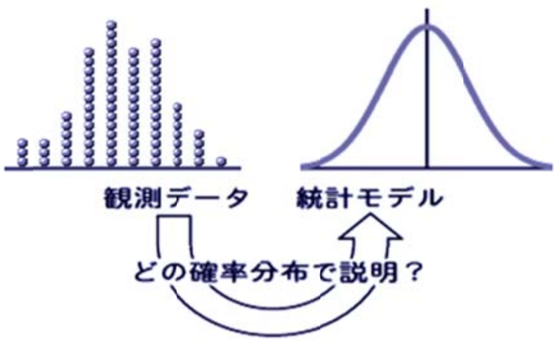
観測される現象は，確率的に変動するものが多いと考えられます。その観測されたデータを説明する統計モデルに，どの確率分布を使えばうまく説明できるでしょうか。

正規分布や二項分布など，確率分布の種類は数多く，いろいろなカタチ（分布形）があります。確率分布の当てはめを考えるには，そのカタチ（分布形）を知ることが重要です。各確率分布の母数（パラメータ）によってそのカタチ（分布形）が決まります。確率変数には離散型と連続型があり，その範囲もさまざまです。

確率変数のとる値が有限個，あるいは無限個であっても自然数で番号が付けられる場合は「離散型」といいます。離散型のグラフは，横軸の確率変数がその値になるときの確率を縦軸で表しています。

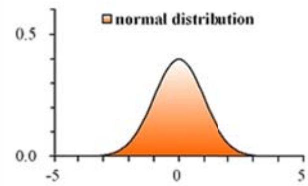
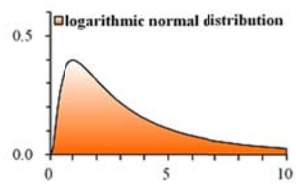
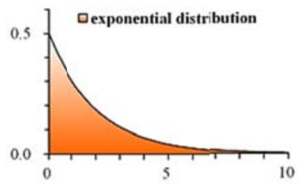
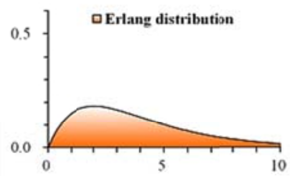
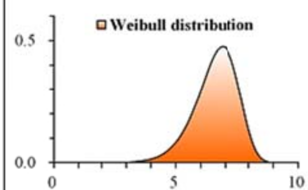
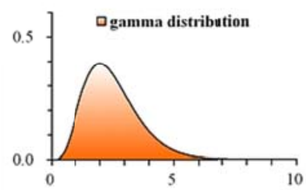
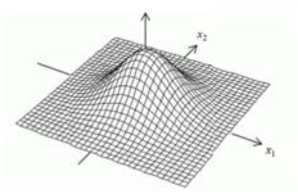
確率変数がある区間内の全ての実数を取り得る場合は「連続型」といいます。連続型のグラフは，横軸の確率変数が連続量なので，縦軸はその値での確率密度を表しており，区間内（横軸のある値とある値の間）を積分した面積がその確率に相当します。

離散型の代表格として正規分布があります，測量誤差は正規分布になりますので離散型の代表的な確率分布について，それらを比較・検討しやすいように母数（パラメータ）やグラフ等を一覧表にまとめたものです。



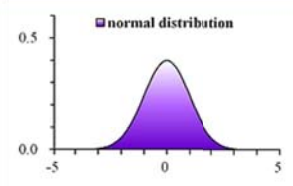
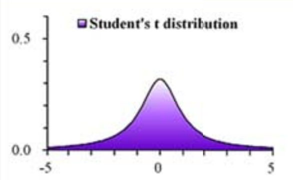
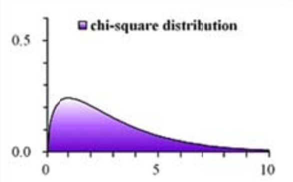
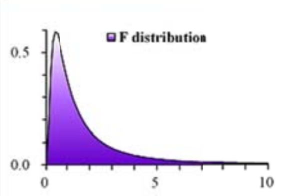
確率変数がある区間内の全ての実数を取り得る場合は「連続型」といいます。連続型のグラフは，横軸の確率変数が連続量なので，縦軸はその値での確率密度を表しており，区間内（横軸のある値とある値の間）を積分した面積がその確率に相当します。

連続型確率分布（Continuous probability distributions）

確率分布の名称	母数（パラメータ）	確率変数Xとその範囲（区間）	単峰形		単調形		その他の分布形	分布形（グラフ）の一例
			対称	歪み	減少	増加		
正規分布 （ガウス分布） N（μ，σ^2）	平均 μ， 分散 σ^2	X：実数， $-\infty < X < \infty$						
対数正規分布	平均 分散	X：実数， $0 < X < \infty$	−	●	−	−	−	
指数分布	ポアソン分布に従う事象が1回生じるまでの待ち時間の平均	X：待ち時間 $0 \leqq X < \infty$	−	−	●	−	−	
アーラン分布	ポアソン分布に従う事象が定めた回数生じるまでの待ち時間の平均， 定めた事象の回数	X：待ち時間 $0 \leqq X < \infty$	−	●	●	−	−	
ワイブル分布	形状パラメータ， 尺度パラメータ， 位置パラメータ	X：寿命時間 $0 \leqq X < \infty$	−	●	●	−	−	
ガンマ分布	形状パラメータ， 尺度パラメータ	X：寿命時間 $0 \leqq X < \infty$	−	●	●	−	−	
多変量正規分布	期待値ベクトル， 分散共分散行列	Xi：実数 $-\infty < X_i < \infty$	多次元型 （正規分布の多変量化）					

母数推定や仮説検定などの推測統計で用いられる確率分布

何を推定・検定するのかによって，それぞれの目的に応じた確率分布があります。推定・検定には多くの手法がありますが，ここでは，推定や検定でよく用いられる代表的な確率分布とその統計量の一例を次に示します。

確率分布の名称	母数（パラメータ）	確率変数Xとその範囲（区間）	推定・検定		分布形（グラフ）の一例
			統計量	用途の一例	
正規分布 N (0, 1)	平均 $\mu = 0$, 分散 $\sigma^2 = 1$ （母数は定数）	X :実数, $-\infty < X < \infty$	z 値	大標本 または 母分散が既知 のときの信頼区間の推定，平均の検定 統計量:(標本平均－検定する平均)を，√(分散／標本サイズ)で割ったもの	
t分布 （小標本の分布）	自由度	X :実数, $-\infty < X < \infty$	t 値	小標本 かつ 母分散が未知 のときの信頼区間の推定，平均の検定 統計量:(標本平均－検定する平均)を，√(分散／標本サイズ)で割ったもの	
カイ2乗分布 （誤差の2乗和の分布）	自由度	X :実数, $0 \leqq X < \infty$	χ^2 値	Pearsonのχ^2検定（適合度検定・独立性検定） 統計量:（観測度数－期待度数）の2乗を期待度数で除したものの総和	
F分布 （分散比の分布）	自由度1(分子), 自由度2(分母)	X :実数, $0 \leqq X < \infty$	F 値	例えば，要因の分散／誤差の分散，信号／ノイズ(S/N比)など，2種類の分散を比率で表したF値(F比)による検定 統計量: 分散$_1$ ／ 分散$_2$	

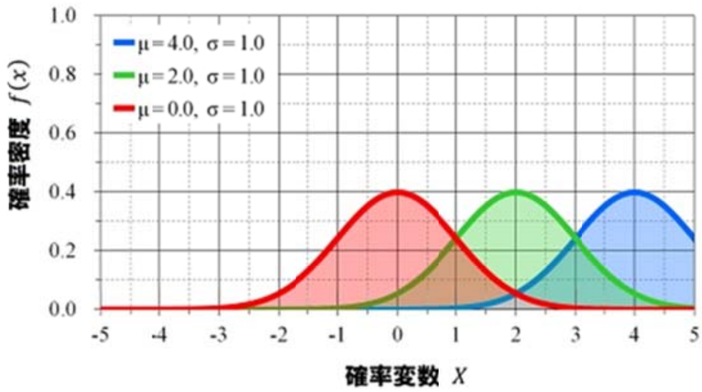
正規分布（normal distribution）

1．確率的な現象と正規分布

データをいくつかの階級に分けて度数分布表やヒストグラムを作成したとき，中心付近の度数が最も高くなり，そこから左右に同程度で度数が少なくなっていく形になることは多いと思います．測定誤差や社会現象あるいは自然現象の中に現れるバラツキは正規分布（normal distribution）に従うと見なせるものが多く，統計学の理論上も応用上も非常に重要で実用性の高い分布です．

正規分布は，自然現象や社会現象において広くあてはまる確率分布です． 19 世紀にガウス（Carl Friedrich Gauss 1777～1855）が観測誤差の研究から導いたことが有名であることからガウス分布（Gaussian distribution）と呼ばれることもありますが，それ以前にも各分野の

研究者によってこの確率分布が導出されており，また誤差の分布として基本的な性質を備えているために最も多くの確率事象に適用される分布であることから，normal （一般的：正規）な分布と呼ぶようになりました．



2．正規分布の形

正規分布は N（ μ ， σ 2 ）と表記します．これはカッコ内の2つの値， 平均 μ と 分散 σ 2 が決まれば正規分布が一意に定まることを意味しており，この平均 μ と 分散 σ 2 を母数(parameters)といいます．

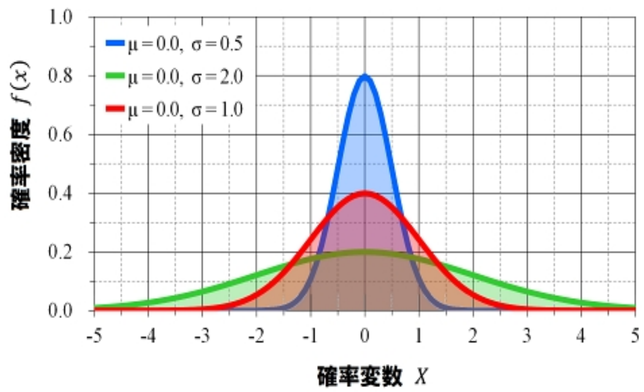
正規分布は平均 μ を中心として左右対称になった西洋の釣鐘と似た形状の曲線（ベルカーブ）の分布形を描きます．

平均 μ は分布形の中心的位置を表しているので，平均の違いは位置の違いとして表れます．また，分散 σ 2 （標準偏差 σ ）については，その値が大きくなるほど釣鐘型の曲線が横に伸びて裾野が広がる形になりますが，これは形が横に伸びただけで，正規分布の曲線の本質的な形状は，相対的に一定で決まった形をしています．

連続型の確率変数 X が正規分布 N（ μ ， σ 2 ）に従うとき，その確率密度関数 f（ x ）は

$$f(x) = \frac{1}{\sqrt{2\pi}\,\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

となります．確率変数 X は － ∞ ＜ x ＜ ＋ ∞ の範囲の実数をとります．この f（ x ）は x ＝ μ のときが最大値であり， x ＝ μ ± σ の点が変曲点となっています．

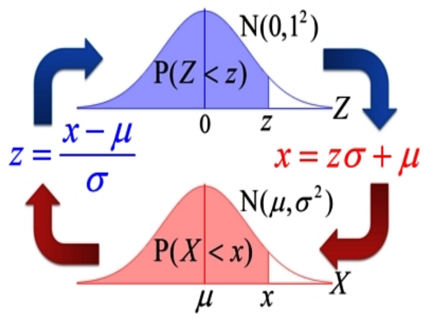


3．標準化変換で基準をつくる

正規分布に限らず，どのような分布であっても，平均 μ は確率変数 X の分布における位置を示す指標であり，標準偏差 σ は分布における確率変数 X のバラつきの尺度となります。 さまざまな分布が持っているこの平均と標準偏差の違いを，何らかの標準的な形に変換することができれば，さまざまな分布の姿を一定の基準で比較・検証することも可能となります。

そこで，位置の基準である平均を 0 ，尺度である標準偏差を 1 に変換することを考えます。確率変数 X を次式で変換すると，変換された確率変数 Z は，平均が 0 ，標準偏差が 1 の分布になります。

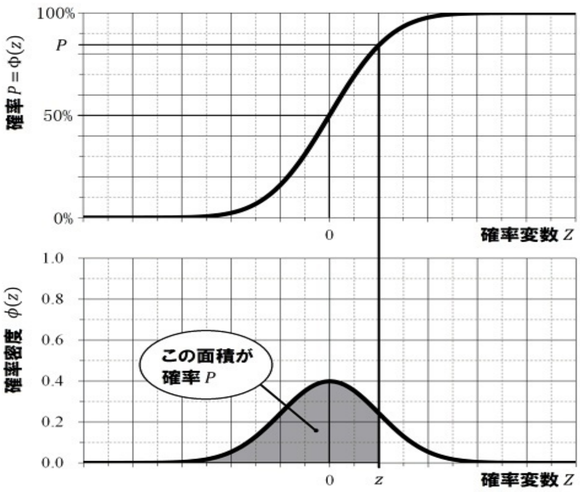
$$z=\frac{x-\mu}{\sigma}=\frac{\text{変数}-\text{平均}}{\text{標準偏差}}=\frac{X-\mathrm{E}(X)}{\sqrt{\mathrm{V}(X)}}$$



この変換を標準化変換(standardizing) といい，1 次式での変換（線形変換）となっていますので，その相対的な値（比率）は変化しません。この標準化変換された z は，標準化得点，z 得点 (z score) または z 値などと呼ばれます。

4. 基準となる標準正規分布

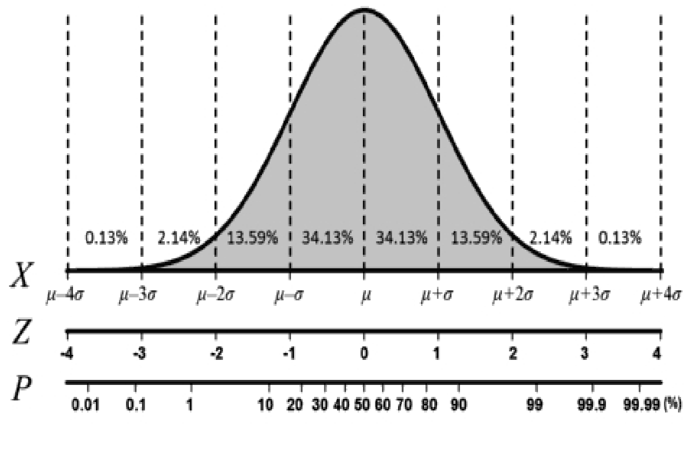
平均 μ ，標準偏差 σ の正規分布 N (μ ， σ 2) に従う確率変数 X を標準化変換した確率変数 Z は，平均が 0 ，標準偏差が 1 の正規分布 N (0, 1) に従うことになります。この N (0, 1) を特に標準正規分布 (standard normal distribution) といいます。



標準正規分布 N (0, 1) は， z = 0 の点を中心(平均)とした形です. したがって， z = 0 で確率は半々ですから P = 50% となります. 確率変数 X のある値 x を標準化変換した z の意味は, 元の一般的な正規分布の値 x が, 平均 μ から標準偏差 σ の z 倍だけ離れているということを示しています。 標準正規分布 N (0, 1) に従う確率変数 Z の確率密度関数を ，累積分布関数(下側確率)を と表すと次のような比較的簡単な式になります。

$$\phi(z) = \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}}$$

$$\Phi(z) = \int_{-\infty}^{z} \phi(y)\,dy$$



確率密度関数について，負の無限大から z までの範囲を積分計算したもの（面積）が累積分布関数（下側確率）です（図を参照）．この標準正規分布の累積分布関数には，正規分布の性質から次のような特徴があります。

累積分布関数の積分は解析的に計算できないので，コンピュータ等を用いて求めることになります。計算されたものを数表としてまとめた 正規分布表 があります．また，主な表計算ソフトには標準正規分布に関する関数（NORMSDIST, NORMSINV, ...）が用意されています。

$$\Phi(\infty) = 1\ ,\ \Phi(0) = 0.5$$
$$\Phi(-z) = 1 - \Phi(z)$$

の標準化変換によって，あらゆる正規分布は N（0, 1）の標準正規分布に帰着します。したがって，正規分布の計算は，標準化変換で z を求めて標準正規分布にて計算し，必要に応じて元の正規分布の変数 x に戻せばよいのです。

5．標準偏差で全体像がわかる

データを分析する上でとても有用な観点があります．それは，どのようなデータでも，分布の形が対称で単峰と見なせる場合には，「平均を中心」に「標準偏差を尺度」として見ることにより，大まかですがデータの全体像がわかるということです。

正規分布 N（ μ ， σ^2）において，平均 μ を中心に標準偏差 σ がプラスマイナス何個分の区間だと何パーセントの割合であるのかをまとめると右表のようになります．表中の数値をイメージできれば，標準偏差から全体像をつかみやすくなります。

区間の幅	$\mu\pm1\sigma$	$\mu\pm2\sigma$	$\mu\pm3\sigma$
区間に入る確率	68.26％	95.44％	99.74％
区間から外れる確率	31.74％ 約1/3	4.56％ 約1/20	0.26％ 約1/400

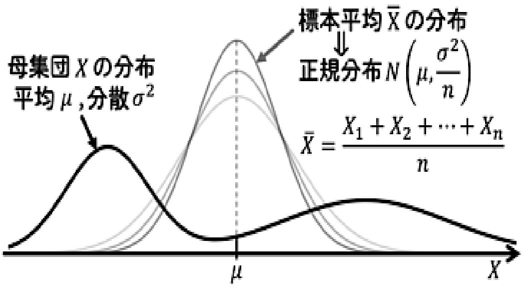
区間の幅	$\mu\pm1.96\sigma$	$\mu\pm2.58\sigma$	$\mu\pm2.81\sigma$
区間に入る確率	95.0％	99.0％	99.5％
区間から外れる確率	5.0％	1.0％	0.5％

区間の幅	$\mu\pm3.29\sigma$	$\mu\pm3.48\sigma$	$\mu\pm3.89\sigma$
区間に入る確率	99.9％	99.95％	99.99％
区間から外れる確率	0.1％	0.05％	0.01％

6．正規分布の理論的な位置づけ

正規分布に従わないどのような分布であっても，その標本平均の分布は標本が大きくなれば正規分布で近似できる性質（中心極限定理）があります。． また二項分布など確率分布の中にはその極限値が正規分布に近づくものも少なくありません（→確率シミュレーション）． 一方，推定や検定に用いられる主要な確率分布（標準正規分布，t 分布，カイ二乗分布，F 分布など）は正規分布を理論基盤としています。このように正規分布は推測統計の基礎となる最も重要な確率分布です。

「中心極限定理」が意味するもの



平均 μ ，分散 σ^2 をもつ 同一の確率分布に従う母集団から無作為抽出によって得られる n 個の標本平均はどのような分布になるでしょうか。．

n が大きくなるにつれて， X の分布は，平均 μ ，分散 σ^2／ n の正規分布に近づいていきます．これが 中心極限定理 といわれるものです。

すなわち，同一の母集団から n 個の標本平均をとると， n が大きくなればなるほど，その分布は 正規分布 N（ μ， σ^2／n ）に近づいていくことを示しています。また，標本の和(合計)をとった分布は 正規分布 N（ nμ， nσ^2 ）に近づきます。このように元の母集団の確率分布がどのような形であっても平均や和の分布は正規分布で近似できるようになることから，中心極限定理は正規分布の有用性を示す大きな根拠の一つとなっています。

【 例 】中心極限定理 ～ 一様分布の和 ～

一様分布(連続型)に従う範囲[0，1]の乱数（一様乱数）は，平均 μ ＝ 1／2 ，分散 σ^2 ＝ 1／12 です． ここで，この一様乱数 12 個分の合計値の分布について考えますと，同一の分布から得られる独立な確率変数の和ですので，平均と分散はそれぞれ 12 倍されて， μ ＝ 6 ，分散 σ^2 ＝ 1 の正規分布で近似できる分布となります。 ちなみに，この一様乱数 12 個の和から 6 を引いた値の分布は， μ ＝ 0 ，分散 σ^2 ＝ 1 の標準正規分布 N（0，1）で近似される分布になり，コンピュータシミュレーション等で正規分布に従う乱数を生成する方法の一つとしてこの定理を応用できます。

下図は n 個のサイコロを振って出た目の合計数の分布 を示したもので，その確率分布の形状を比較しやすいよう折れ線グラフで表現していま。1 個のサイコロは離散型の一様分布に従う範囲[1，6]の乱数で，各目が出る確率はそれぞれ 1／6 ，平均は 7／2 ，分散は 35／12 です．サイコロの数 n が大きくなると，和(合計数)の分布は正規分布に近づいていくことが

この図から読み取れます。

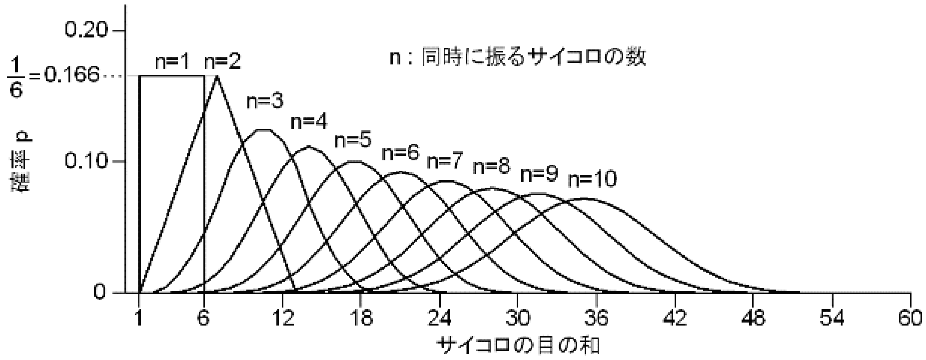
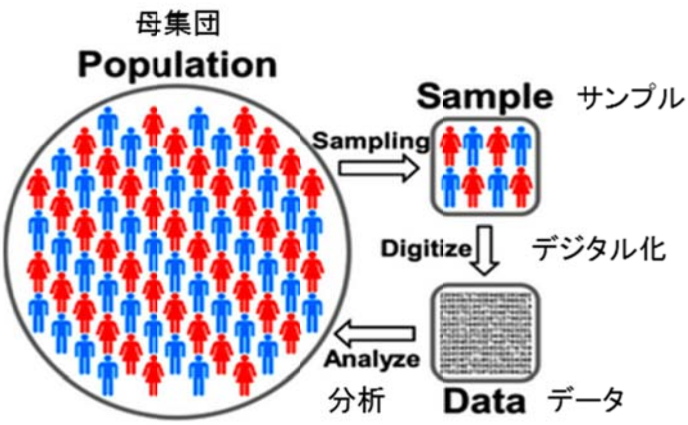


図　n 個のサイコロを振って出た目の合計数の分布

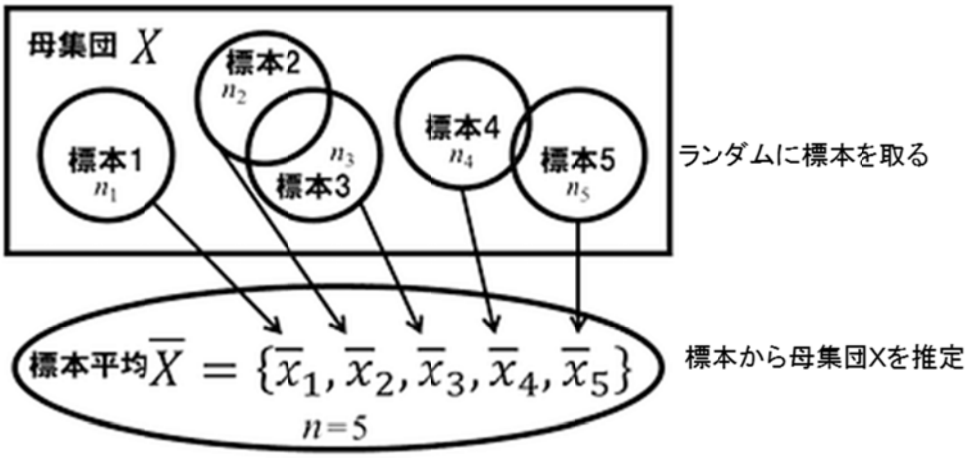
「信頼区間」が意味するもの 1. 母集団を標本から推測する データ分析において，その対象とする全ての要素が含まれる集合のことを「母集団 」といいます。母集団全体を分析できればよいのですが，規模が大きく，実施方法が難しいような場合には，母集団全体ではなく，その一部分からデータ分析をしなければならないこともあります．この母集団から抽出された一部分のことを「 標本（サンプル）」といいます。

実際には，母集団を標本から推測するという標本調査は，いろいろな場面で利用されています。しかし，一部分から推測しているので，どうしても母集団が持つ真の値との間に差が生じることは避けられません。では，その差とはどのようなものになるのでしょうか。



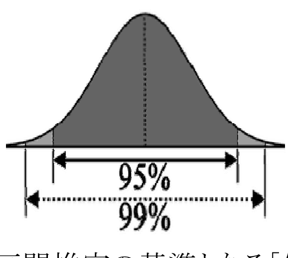
標本は偶然に

無作為抽出すなわちランダムに得られる標本から計算された統計量は，ランダムですので偶然に大きな値に偏ったり小さな値に偏ったりすることがあります。例えば，標本が今 10 個得られたとして，そこから母集団の平均を推定したい場合に標本平均を推定値としますが，その推定値が母集団の真の平均にどれだけ近いのかが問題となります。真の平均に極めて近い推定値になっているかも知れませんし，かなりズレタ値かも知れません。標本は偶然に得られたものですから・・・。



推定値に幅を持たせる

推定値を一点で決めてしまうことを点推定といいます。しかし，標本の統計量はバラつきますので，点推定の値だけではその推定値がどれだけズレているのかを知る余地がありません。そこで，推定値にある一定の幅を持たせることで，この幅の間隔を見れば推定値のズレの度合いを知ることができるようにします。このように幅を待たせる推定方法を区間推定といい，幅の間隔のことを「 信頼区間」といいます。



区間推定の基準となる「信頼度」

母集団から抽出した標本を使って区間推定する際に，その区間の幅を決めるための基準が必要です。信頼区間の幅を決める基準となるのが「 信頼度」と呼ばれるものになります。
信頼度として一般的によく用いられる基準は 90％，95% ， 99%の3種類ですが一般的には 95% ， 99%をつかいます。

　この信頼度という基準があれば，例えば図のように正規分布であれば平均を中心とした区間の幅が決まります。

信頼度 95%ならば，残りの 5%にはどういう意味があるのでしょうか。1 から信頼度を引いた値を「有意水準」といい，記号では　α　がよく用いられます。信頼度が 95%ならば有意水準は 5%となります。有意水準は危険率とも呼ばれるもので，いわば間違った答えを出してしまう割合です。 例えば，有意水準 5%というのは 5% ＝ 1／20 ですので，同じことを 20 回やったら 1 回ぐらいは間違った答えを出してしまうという程度を示しています。有意水準 1%なら 1% ＝ 1／100 ですので，100 回やれば 1 回ぐらいは間違うという程度になります。

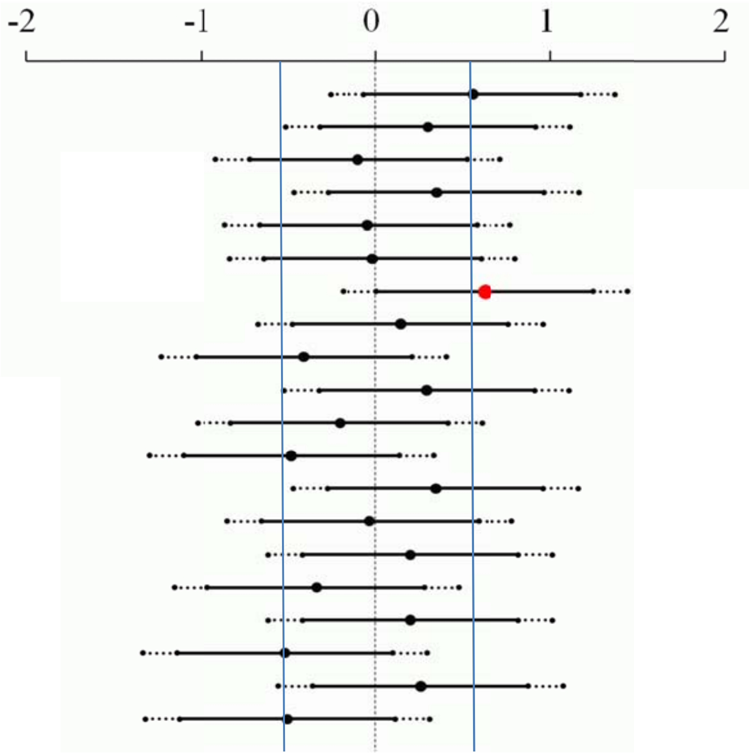
信頼区間が意味するものとは

次図をご覧ください．縦の点線は母集団の真の平均を表し，横棒 1 本で 1 回の標本抽出での区間推定の結果を示しています。20 本ある各横棒の中央の点が標本平均，実線が信頼度 95%の信頼区間，点線が信頼度 99%の信頼区間を表しています。信頼度が高いほど信頼区間は広い幅をとります。この図では同じ標本抽出を 20 回やったときに 1 回だけ 95%信頼区間内に真値が入っていない回（図中の赤印）があることがわかります。

このように，信頼区間とは，信頼度という基準で区間幅を推定したとき，同じ条件で標本抽出を何回か繰り返せば，区間内に真値が含まれる回数はこの程度だということを意味する幅を示しています．

シミュレーションで信頼区間をイメージする

信頼区間の考え方について説明してきましたが，百聞は一見にしかずで，目で見える形にすることで，より一層理解が深まると思います。本当にそうなるのか，実際にシミュレーションで確かめてみましょう。



例として，標準正規分布 N(0,1) に従う母集団から 10 個の標本を抽出して標本平均と信頼区間を推定し，実際に信頼区間に真値が入っているときと外れるときがどの程度あるのかを見ることにします。

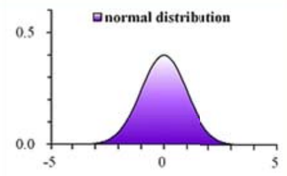
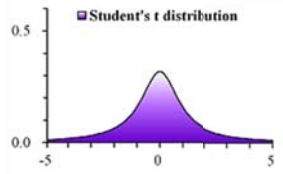
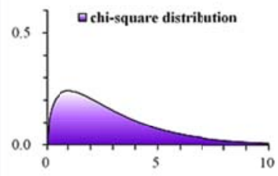
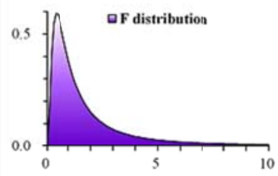
正規分布 $N(\mu,\ \sigma^2)$ に従う母集団から n 個の標本を得たとき，その標本平均の分布は正規分布 $N(\mu,\ \sigma^2/n)$に従うことがわかっています。したがって，シミュレーションは，標準正規分布 N(0,1) に従う乱数 Xi を 10 個発生させて，そこから次式によって計算される平均の推定値と信頼区間を表示し，母集団の平均である $\mu=0$ との関係を見ることになります。

- ・平均の推定値： Σ Xi ／10
- ・95％信頼区間： 平均 ± 1.96 ×（分散 1／10 の平方根）
- ・99％信頼区間： 平均 ± 2.58 ×（分散 1／10 の平方根）

これを何回か繰り返してみた結果の一部を右枠内の図に示します。図中の実線が信頼度 95%の信頼区間，点線が信頼度 99%の信頼区間を表し，20 回分をまとめて表示しています．およそ 20 回に 1 回は 95%信頼区間が真値から外れていることが分かります。このようなシミュレーション結果を視覚的に見ることで，信頼区間の意味がよく理解できます。

母数推定や仮説検定などの推測統計で用いられる確率分布

何を推定・検定するのかによって，それぞれの目的に応じた確率分布があります。推定・検定には多くの手法がありますが，推定や検定でよく土地家屋調査士等が使う四つの代表的な確率分布とその統計量の一例を次に示します。

確率分布の名称	母数（パラメータ）	確率変数Xとその範囲（区間）	推定・検定 統計量	推定・検定 用途の一例	分布形（グラフ）の一例
標準正規分布 N（0，1）	平均 $\mu = 0$, 分散 $\sigma^2 = 1$ （母数は定数）	X :実数， $-\infty < X < \infty$ [統計数値表]	z 値	大標本 または 母分散が既知 のときの信頼区間の推定，平均の検定 統計量:（標本平均−検定する平均）を，√（分散／標本サイズ）で割ったもの	
t分布 （小標本の分布）	自由度	X :実数， $-\infty < X < \infty$ [統計数値表]	t 値	小標本 かつ 母分散が未知 のときの信頼区間の推定，平均の検定 統計量:（標本平均−検定する平均）を，√（分散／標本サイズ）で割ったもの	
カイ2乗分布 （誤差の2乗和の分布）	自由度	X :実数， $0 \leqq X < \infty$ [統計数値表]	χ^2 値	Pearsonのχ^2検定（適合度検定・独立性検定） 統計量:（観測度数−期待度数）の2乗を期待度数で除したものの総和	
F分布 （分散比の分布）	自由度1（分子）， 自由度2（分母）	X :実数， $0 \leqq X < \infty$ [統計数値表]	F 値	例えば，要因の分散／誤差の分散，信号／ノイズ（S/N比）など，2種類の分散を比率で表したF値（F比）による検定 統計量: 分散1 ／ 分散2	

誤差は正規分布になる，これは数が相当量ある事が前提です，境界測量ではその数は少ないことが問題になります。

数が少ない場合に使うのがt分布です，t分布はt検定と平均値信頼区間の推定に使います，データ数が相当にある場合はχ^2分布を使います。t検定とχ^2二乗検定のデータ数の境は自由度（データ数－1）30以下ではt検定を31以上ではχ^2二乗検定を使うとされています。経験

上ではt検定では31個以上でも影響は出ません，χ^2二乗検定では20個以上でも影響は出ません，つまりt検定とχ^2二乗検定で差はないようです。

20個～30個の重なる部分は両方の結果を比較して判断すれば良いと思います。質の良いデータでは差がないはずです。

データの中に異常なデータがないかをt分布のt値を使って検定するのがt検定です，データが正規分布と判断出来るか否かをχ^2分布のχ^2値を使って検定するのがχ^2二乗検定です。

t分布のt値は平均値の信頼区間の計算にも使います。χ^2値は標準偏差の信頼限界の計算にも使います。

F分布のf値は二つの測量図を測量してそれぞれの標準偏差から二つの測量図データに有意差があるかないかの判断に使います。差があればどちらかを採用しない，差がなければ両方を同一データとして平均を使うなどと判断していきます，が測量図では有意差がなくても個別に検討しますので実用的な指標かどうかは疑問があります。そのような考え方もあると言おう程度の指標と覚えておけばいいでしょう。

土地家屋調査士がこれらの分布，検定について理解しておく必要があるかと言えば「正規分布については必要である」「他の3つについてはどのような時に使われるかを知っておく必要がある」です。「t検定」と「χ^2二乗検定」については理解しておく必要があります。

厳密網平均計算，TS機による測量が一般化した現在では測量成果に個人差はありません，土地家屋調査士が測量屋さんと言われる人たちとの違いを出すには最小二乗法と相まってこれらの統計知識が必要です。

データが正規分布に収まっていればそれらのデータを使って境界復元計算することを当たり前に考えなければなりません，限られたデータ数の中でどうやって判断するのか，ここが重要になります，統計でデータの解析をしたうえで最小二乗法を用いて計算をする，その上で法的な判断を加える訳ですが，注意したいのは**統計で許される範囲を超えて法的判断をしてはならないと言うことです。**

具体的には平均値は信頼区間の範囲を超えてはならない，精度は信頼限界を超えてはならない，この範囲を超えた位置に新たに**筆界を認定**してはならないと言うことです。

得てして，筆界確認が成されたことを理由に筆界位置を変えた地積測量図，地籍図を見ますがこれは行ってはならないことです。

（筆界を認定とは不動産登記法17条地図，土地台帳附属地図，地積測量図，地籍図，官民境界図等の公的資料から計算された位置と異なる位置を筆界とすること）

実際には学問的な統計判断を実務にどうやって反映させるかが課題なのですが。

次に分布について簡単に触れます，詳細は別途学習してください。

母集団と標本

無限のデータを母集団，抜き取ったデータを標本と言います，そもそも母集団とはどんなものなのかと言うことです，測量で例えればある条件下，もっと具体的には決められた作業基準（作業標準ともいう）の基で得られたデータを母集団といいます。

例えば法17条地図の作成にあたっては5W1H（When（いつ） Where（どこで） Who（誰が） What（何を） Why（なぜ）How（どのように））に添って作業基準を設定した上で得られたデータということです，ある法17条地図の境界点数が2万点あれば母集団は2万点です，この区域内の一区画（一筆）の測量を依頼され測った点が20個であれば標本を20個として考えます。

２万点のデータが正規分布なのだから20個のデータも正規分布に倣うと考える，今ある20個のデータは完全な正規分布になっていないけれどもデータ数が増えれば正規分布になると言うのが中心極限定理です。

明治の地租改正図も字単位に作成されましたので一字のデータが母集団という考え方です，地租改正図は市街地，農耕地（郷村地），山林原野の3つに区分けされて異なる基準で作成されましたので同じ字内でも分けて母集団が形成されていることに注意してください。

法17条地図の母集団を同じ作業基準で作成された全国のデータとする考え方もありますが広い意味ではそうなりますが5W1Hで考えれば一区域で考えるべきです。

では，地積測量図はどう考えるかです，地積測量図は一筆毎にしかありません，仮に Who（誰が）を条件にすればA土地家屋調査士が作成した他の地積測量図も含めて母集団とするのか，測量方法，測量機器が変われば当然母集団も変わります，ですから地積測量図には母集団は存在しないけれども架空の母集団が存在するとして，正規分布な母集団からの標本と見る地積測量図のデータも正規分布に倣うと考えるのです。

公差と精度（標準偏差）

公差とは最悪の場合，ここまでは許容できる範囲です，「平均的には」とか「通常は」という解釈ではありません。

ですから，公差ギリギリの範囲で合格するような測量は通常しません，公差とは言い方を変えれば誤差の大きさの限界値です，誤差とは被検査体（既成果）と検査値（実測成果）との差です。 誤差は正規分布になるとされています。

次の表は国土調査法施行令別表4の精度区分別の公差表です，これを使って公差と正規分布の関係について説明します。

筆界点の位置誤差についてですが甲2を除いた他は公差の3分の1が平均二乗誤差となっています，これで計算すると甲2の平均二乗誤差は6．7cm となりますが公差があって約3分の1＝7cm としたものと考えられます，したがって甲2の平均二乗誤差は細かく言えば6．7cmとなります。

位置誤差は2変量（x，yの誤差の分布）の分布ですから平均二乗誤差＝二変量の標準偏差となりますので一倍平均二乗誤差（一倍二変量標準偏差）の確率は0．683となります，三倍

平均二乗誤差（三倍二変量標準偏差）の確率は0．990です。

（國見利夫著，平成22改正「準則」地籍測量の18ページ，18～19行の記載「正規分布で表した誤差は，その約68%が平均二乗誤差の範囲内に，99%が平均二乗誤差の3倍までに収まるという性質を持っている。」との記載から筆界の位置誤差は二変量の確率を前提にしていると言える。）

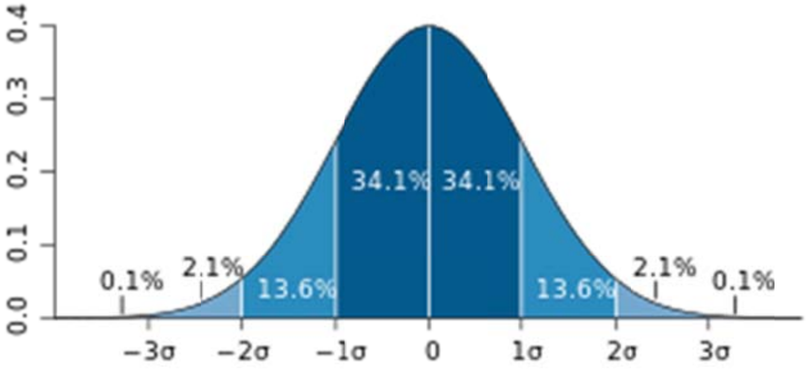
（書籍によっては

精度区分	筆界点の位置誤差		筆界点間の図上距離又は計算距離と直接測定による距離との差異の公差	地積測定の公差
	平均二乗誤差	公差		
甲一	2cm	6cm	$0.020m+0.003\sqrt{S}m+\alpha mm$	$(0.025+0.003\sqrt[4]{F})\sqrt{F}m^2$
甲二	7cm	20cm	$0.04m+0.01\sqrt{S}m+\alpha mm$	$(0.05+0.01\sqrt[4]{F})\sqrt{F}m^2$
甲三	15cm	45cm	$0.08m+0.02\sqrt{S}m+\alpha mm$	$(0.10+0.02\sqrt[4]{F})\sqrt{F}m^2$
乙一	25cm	75cm	$0.13m+0.04\sqrt{S}m+\alpha mm$	$(0.10+0.04\sqrt[4]{F})\sqrt{F}m^2$
乙二	50cm	150cm	$0.25m+0.07\sqrt{S}m+\alpha mm$	$(0.25+0.07\sqrt[4]{F})\sqrt{F}m^2$
乙三	100cm	300cm	$0.50m+0.14\sqrt{S}m+\alpha mm$	$(0.50+0.14\sqrt[4]{F})\sqrt{F}m^2$

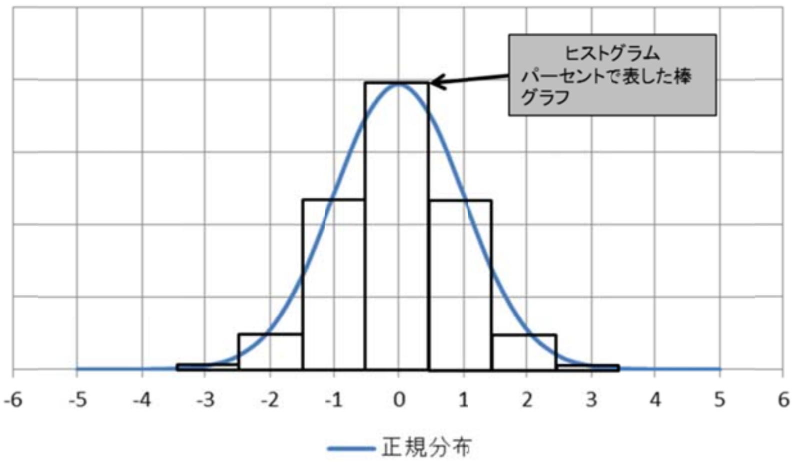
備考
一　精度区分とは、誤差の限度の区分をいい、その適用の基準は、国土交通大臣が定める。
二　筆界点の位置誤差とは、当該筆界点のこれを決定した与点に対する位置誤差をいう。
三　Sは、筆界点間の距離をメートル単位で示した数とする。
四　αは、図解法を用いる場合において、図解作業の級が、A級であるときは0.2に、その他であるときは0.3に当該地籍図の縮尺の分母の数を乗じて得た数とする。
図解作業のA級とは、図解法による与点のプロットの誤差が0.1ミリメートル以内である級をいう。
五　Fは、一筆地の地積を平方メートル単位で示した数とする。
六　mはメートル、cmはセンチメートル、mmはミリメートル、m2は平方メートルの略字とする。

国土調査法施行令別表第四

1．　正規分布



正規分布密度関数グラフ



ヒストグラム

級(幅＝標準偏差)から確率を求める方法

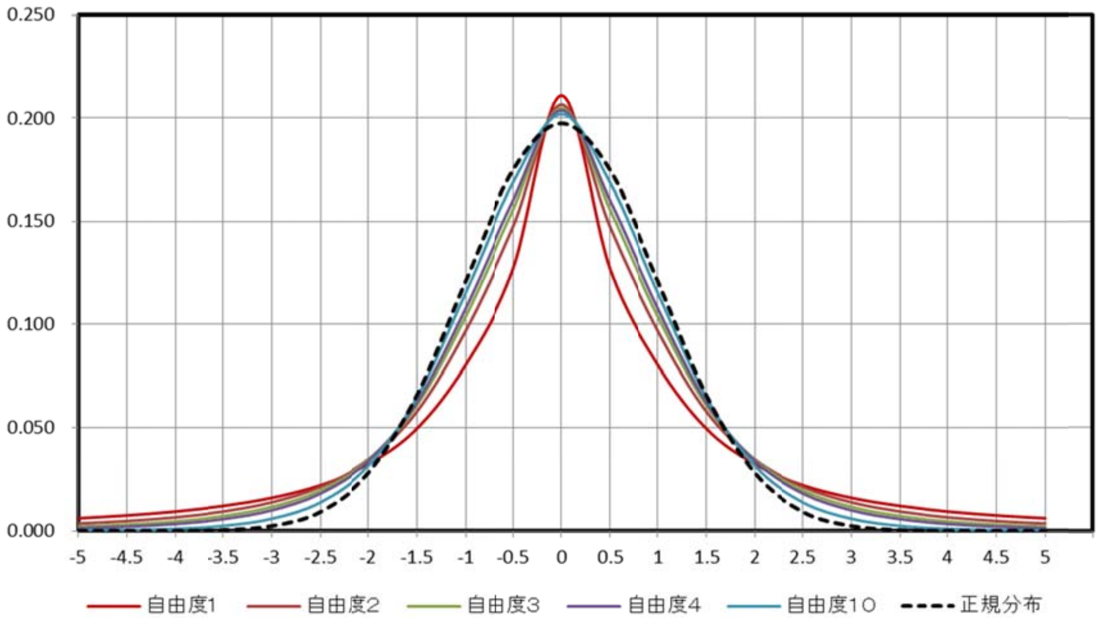
確率から級(幅＝標準偏差)を求める方法

t分布

次の図がt分布の自由度別と正規分布の密度関数グラフです,自由度が小さいと山の頂上が尖っており, 裾野が0から離れています。

自由度が増すにつれて正規分布（点線）に近づいていきます，∞になると正規分布と一致します，自由度31以上の場合は正規分布として考えχ^2分布で考えます。

土地家屋調査士が行う境界点の数は30点以下の場合が多いのでt分布を使ったデータ解析を理解しておく必要があります。

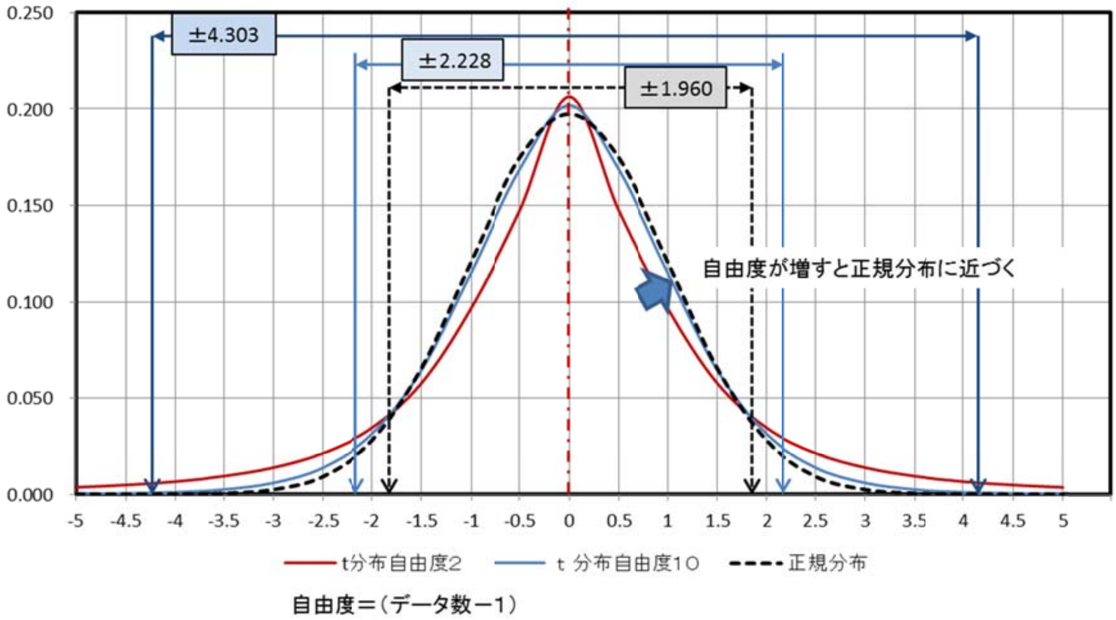


t分布密度関数　　点線は正規分布密度関数

次の図は確率95％の標準偏差の幅を自由度別に表した図です，自由度2では±4．303倍，自由度10では±2．228倍となります，正規分布では±1．960倍です。

ですから自由度（データ数－）が小さいほど同じ確率に対して標準偏差の幅が大きくなり，それだけその標準偏差への信頼性が落ちることになります。同じ標準偏差であってもデータ数が多いほど信頼性が高いと言うことです。

標準偏差の値＝t値です。



確率95%のσの幅

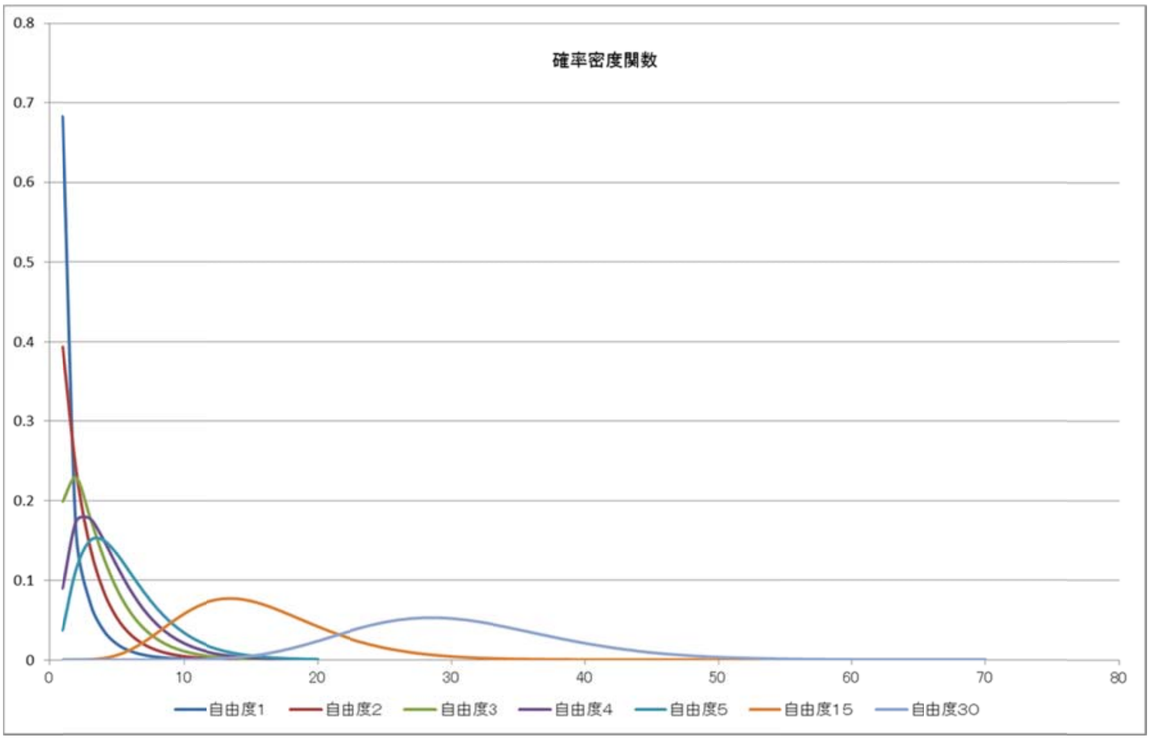
自由度（データ数ー1）からt値を求める方法

t 値はエクセル関数 TINV(有意水準,自由度)・・自由度＝データ数ー1
正規分布は NORM.INV((1-(1-確率)/2),平均, σ)・・σは標準偏差の幅 から求める。

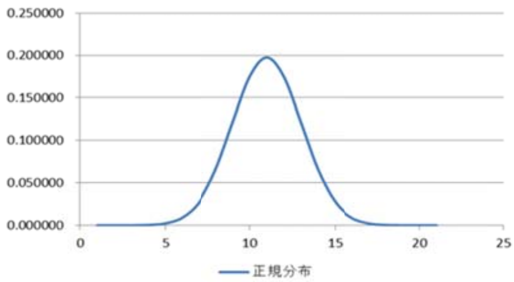
確率からt値を求める方法

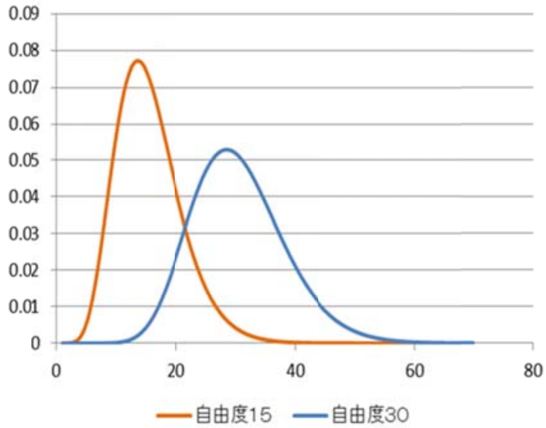
この表を密度関数でグラフにすると次の図になる。

χ^2(カイ二乗)分布

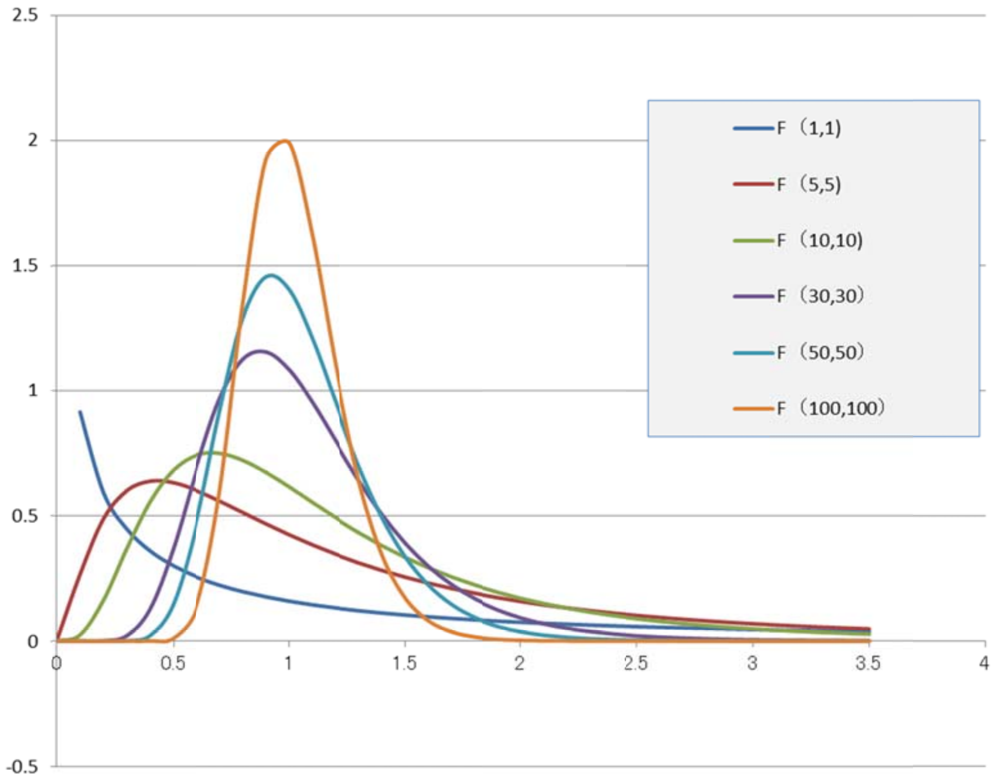


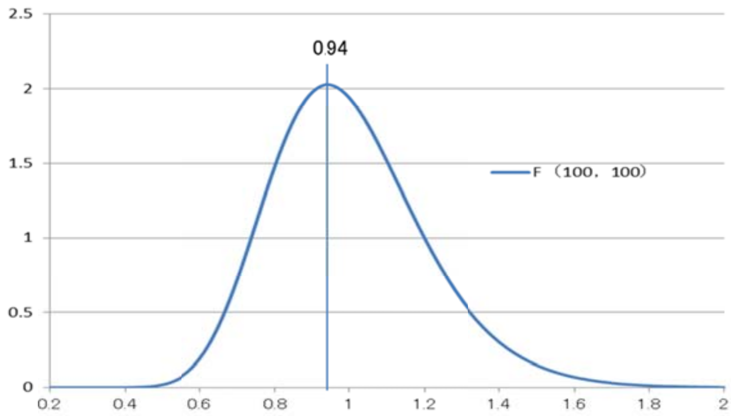
確率密度関数
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
0
10
20
30
40
50
60
70
80
自由度1
自由度2
自由度3
自由度4
自由度5
自由度15
自由度30


0.250000
0.200000
0.150000
0.100000
0.050000
0.000000
0
5
10
15
20
25
正規分布


0.09
0.08
0.07
0.06
0.05
0.04
0.03
0.02
0.01
0
0
20
40
60
80
自由度15
自由度30

F分布


2.5
2
1.5
1
0.5
0
-0.5
0
0.5
1
1.5
2
2.5
3
3.5
4
F（1,1)
F（5,5)
F（10,10)
F（30,30)
F（50,50)
F（100,100)



土地家屋調査士･測量士　小野孝治