

境界(筆界)復元における混合分布, 層別, F検定とは

異なる要素の持ったデータが混在している状態を混合分布と言います, 混合分布が見つかった場合は層別, F検定と判断していきます, そのことによって最適な境界(筆界)復元に使うデータを分類することができます。ここでは混合分布, 層別, F検定と順番に説明していきます。

混合分布とは

混合分布とは, 例えばA工区とB工区と異なる時期の開発がなされたが, 一枚の境界(筆界)図に書かれている, A工区にある境界標を基準にB工区の境界標を見た時に一定方向に偏っている状態, 主に移動と回転が起きている状態を混合分布といいます。

下図においてA工区とB工区でA工区からB工区を見るとB工区が右に回転していることが判ります, 回転の他にズレも有りますので境界(筆界)復元ではこの現象を見つけて判断することが求められます。

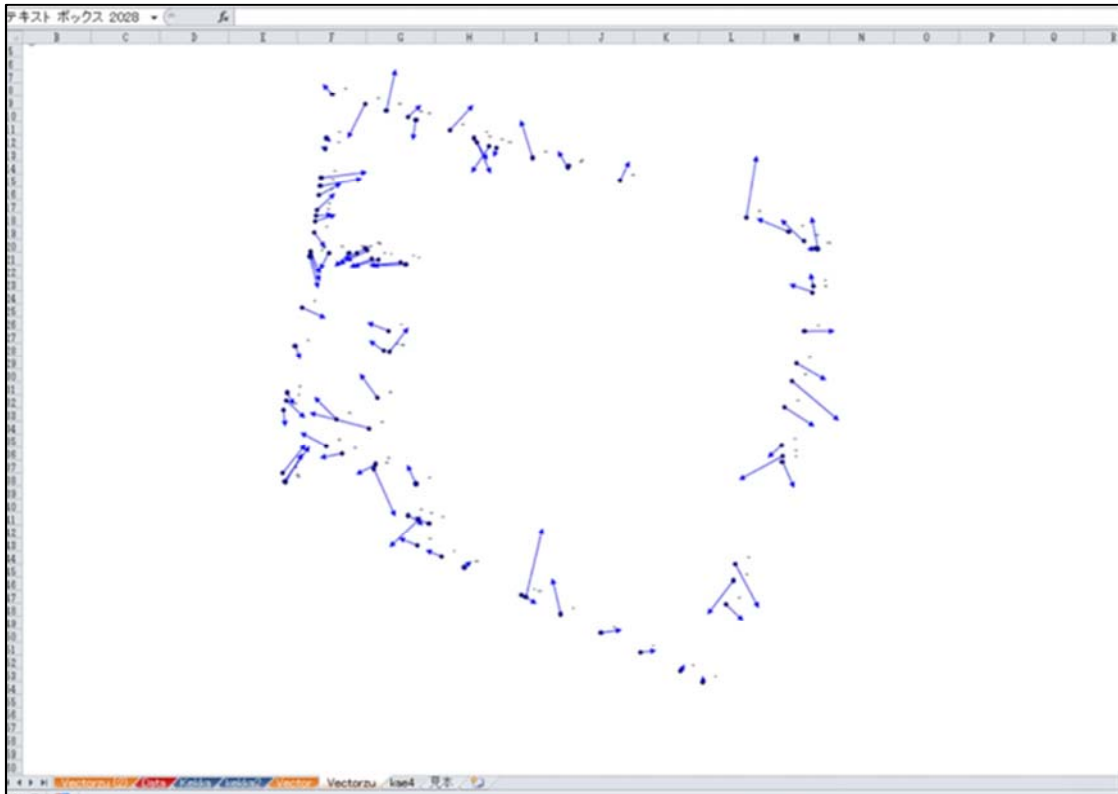
このような現象は大規模な宅地開発とか道路買収などで体験することがあります。一筆測量のような狭い範囲の測量ではなかなか体験することは有りませんが境界紛争などである程度広範囲に測量した場合に表れることがあります。このことが原因で境界紛争が起きていることもありますので, その時に混合分布に知識がないと誤った筆界の判断をすることがありますので重要知識です。

表れる現象

1. 図面值座標(以前の測量成果)と実測値座標(現在の測量成果)との関係をベクトル図にするとベクトル線に一定の方向を示す特徴が表れます。
2. 図面值座標(以前の測量成果)と実測値座標(現在の測量成果)の差を散布図にすると意外と簡単に混合分布を認識出来ます。

そのため普段からベクトル図, 散布図を作成する習慣があれば簡単に気づきます。

混合分布の無いベクトル図



この散布図はExcelで画いたものです。ベクトル線の方向がこれといった纏まりが無いのが特徴です、普通はこのような状態なのです。ベクトル線の向きが360度全方向に渡ってランダムになっている状態で誤差の三公理を当て嵌めて

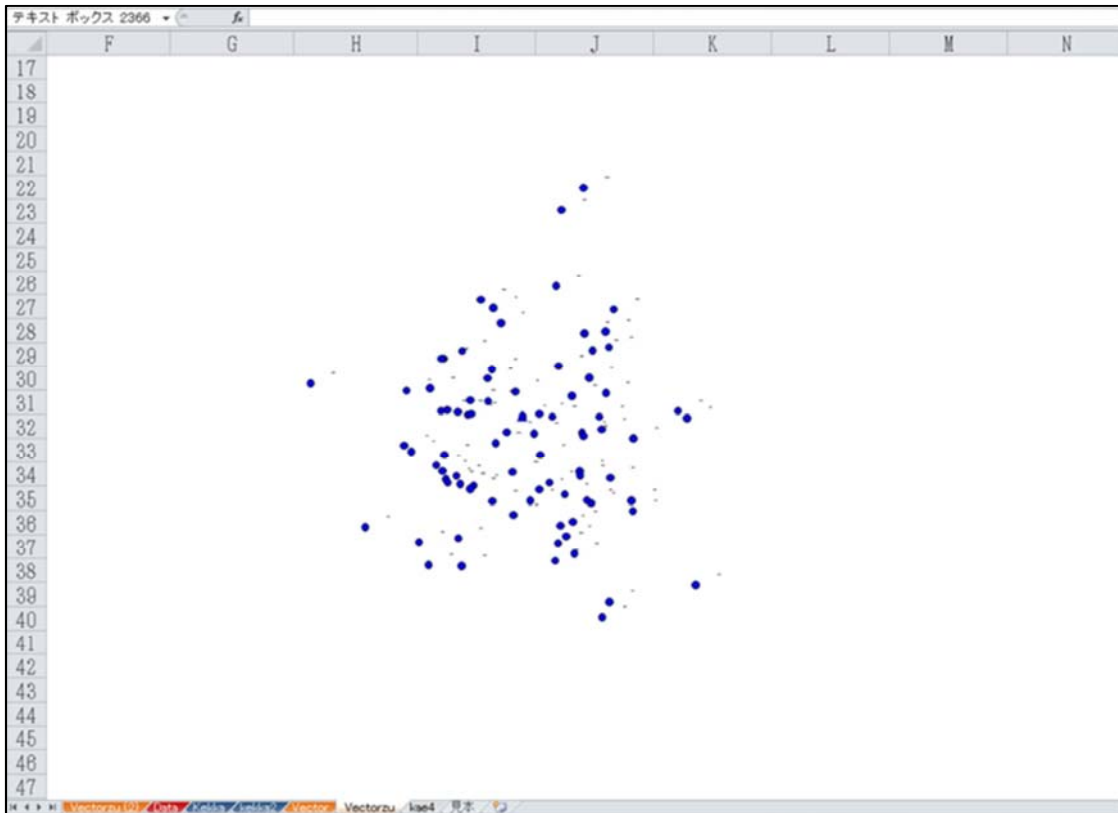
第一の公理:小さいベクトルは、大きいベクトルよりも多く起こる。

第二の公理:ベクトルは360度全方向に起こる。

第三の公理:大きいベクトルの発生する確率は非常に小さい。

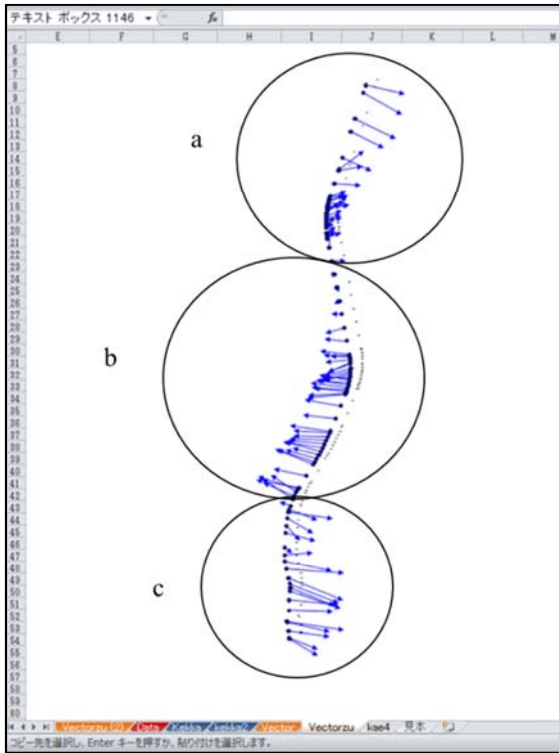
とすることになります。

混合分布の無い散布図



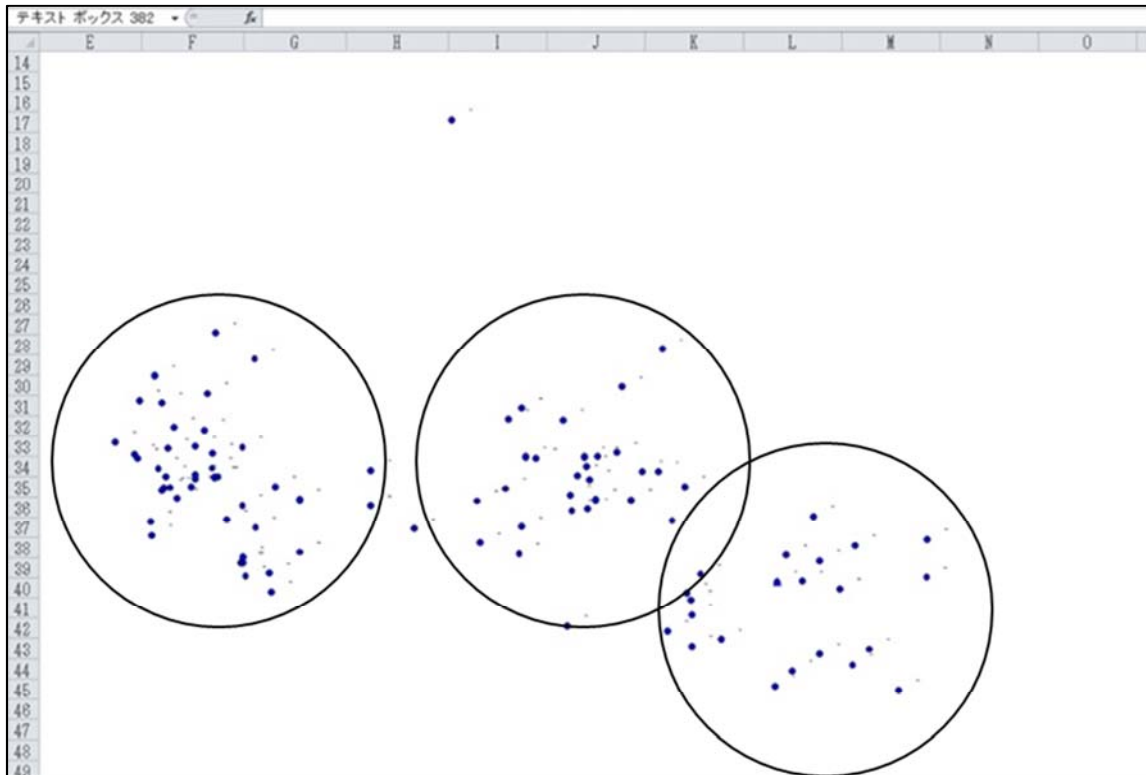
上図のデータを Excel で散布図に画いたものです, 分布の中心から外周に向かって点の配置が広がっているのが特徴です。

混合分布のあるベクトル図



実際の例をベクトル図にしたものです、○のつけた位置で三つ特徴あるベクトル線が確認できます。上はベクトル線が右を向いており、中は左、下は右を向いています、このようにハッキリとベクトル線の方が纏まってでる例は珍しいです。

混合分布のある散布図



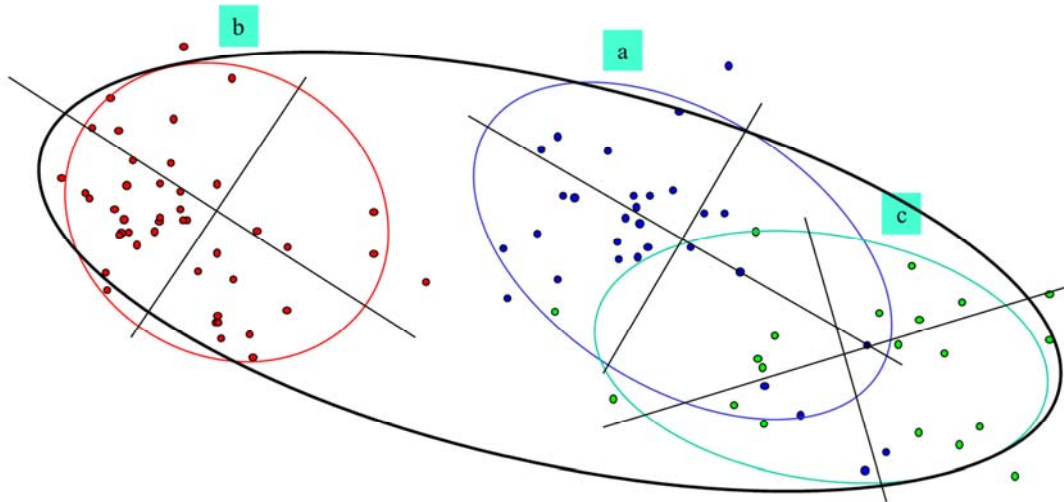
これは前のベクトル図のデータを Excel で散布図を画いたものです, ○のつけた位置で三つの纏まりが確認できます。このデータは精度の異なる分布が表れた例です。なかなかこのようにハッキリするのは珍しいのです。

その他にベクトル図, 散布図以外から表れる現象として

3. 座標軸が直交しているはずのデータでアフィン変換が優位とされたとき。

異なる座標変換の優位性は精度・標準偏差で判断せずにAICによって判断します, 座標軸が直交している場合はヘルマート変換が優位となります (AICがアフィン変換のAICより小さい)。座標軸が直交している状態とは基準点測定の精度が高い状態を言います。

上の図は次に示すようにa, b, c の三つに分類出来ます。



この図では全体の標準偏差ヘルマート0.040, アフィン0.032, ヘルマート AIC711, アフィン AIC670, でアフィン変換優位ですが a, b, c で分類すると

χ ² 検定・t検定 & ヘルマート変換									
件名	測地2011 市道成果と実測値比較								
係数a	0.999984	線の伸縮率	0.999984	平均二乗誤差	0.057				
係数b	0.000095	指定伸縮率	<input type="text" value="0.000095"/>	AIC	711				
移動量x	354.997	回転角	0° 0' 20"	標準偏差	0.040				
移動量y	-293.556	指定数	<input type="text" value="95"/>	点名セット	h				
指定数までリセット			準拠点をアフィンへ	空欄の実測値戻す	準拠点数				
					95				

χ ² 検定・t検定 & アフィン変換									
件名	測地2011 市道成果と実測値比較								
係数a	0.999968	移動量x	354.961	伸縮率x	0.999968	平均二乗誤差	0.045		
係数b	0.000132	移動量y	-294.514	伸縮率y	1.000964	AIC	670		
係数c	0.000208	回転角x	0° 0' 27"	回転角y	90° 0' 43"	標準偏差	0.032		
係数d	1.000964	指定数	<input type="text" value="95"/>	点名セット	h				
指定数までリセット			準拠点をヘルマートへ	空欄の実測値戻す	準拠点数	95			

aは標準偏差ヘルマート0.013, アフィン0.013, ヘルマート AIC171, アフィン AIC174, でアフィン変換優位です。

χ ² 検定・t検定 & ヘルマート変換									
件名	測地2011 市道成果と実測値比較								
係数a	1.000059	線の伸縮率	1.000059	平均二乗誤差	0.019				
係数b	0.000280	指定伸縮率	<input type="text" value="0.000280"/>	AIC	171				
移動量x	354.765	回転角	0° 0' 58"	標準偏差	0.013				
移動量y	-293.566	指定数	29	点名セット	h				
指定数までリセット			準拠点をアフィンへ	空欄の実測値戻す	準拠点数	31			

χ ² 検定・t検定 & アフィン変換							
件名	測地2011 市道成果と実測値比較						
係数a	1.000008	移動量x	354.509	伸縮率x	1.000009	平均二乗誤差	0.019
係数b	0.000558	移動量y	-293.462	伸縮率y	0.999942	AIC	174
係数c	0.000254			回転角x	0° 1' 55"	標準偏差	0.013
係数d	0.999942	指定数		回転角y	90° 0' 52"	点名セット	
指定数までリセット		準拠点をヘルマートへ		空欄の実測値戻す		準拠点数	31

bは標準偏差ヘルマート0.010, アフィン0.008, ヘルマートAIC211, アフィンAIC200, でアフィン変換優位です。

χ ² 検定・t検定 & ヘルマート変換								
件名	測地2011 市道成果と実測値比較							
係数a	1.000046	線の伸縮率	1.000046			平均二乗誤差	0.014	
係数b	0.000172	指定伸縮率	<input type="text" value="1.000"/>	<input type="radio" value="1.000"/>	<input checked="" type="radio" value="furee"/>	AIC	211	
移動量x	354.909	回転角	0° 0' 35"	<input type="text" value="0.0000031779"/>		標準偏差	0.010	
移動量y	-293.562	指定数			点名セット			
指定数までリセット		29	準拠点をアフィンへ		空欄の実測値戻す	h	準拠点数	44

χ ² 検定・t検定 & アフィン変換							
件名	測地2011 市道成果と実測値比較						
係数a	1.000017	移動量x	354.847	伸縮率x	1.000017	平均二乗誤差	0.012
係数b	0.000235	移動量y	-293.956	伸縮率y	1.000443	AIC	200
係数c	0.000300			回転角x	0° 0' 48"	標準偏差	0.008
係数d	1.000443	指定数		回転角y	90° 1' 2"	点名セット	
指定数までリセット		準拠点をヘルマートへ		空欄の実測値戻す		準拠点数	44

cは標準偏差ヘルマート0.018, アフィン0.017, ヘルマートAIC126, アフィンAIC127, でアフィン変換優位です。

χ ² 検定・t検定 & ヘルマート変換								
件名	測地2011 市道成果と実測値比較							
係数a	0.999975	線の伸縮率	0.999975			平均二乗誤差	0.026	
係数b	0.000002	指定伸縮率	<input type="text" value="1.000"/>	<input type="radio" value="1.000"/>	<input checked="" type="radio" value="furee"/>	AIC	126	
移動量x	355.097	回転角	0° 0' 0"	<input type="text" value="0.000105261"/>		標準偏差	0.018	
移動量y	-293.586	指定数			点名セット			
指定数までリセット		43	準拠点をアフィンへ		空欄の実測値戻す	h	準拠点数	20

χ ² 検定・t検定 & アフィン変換							
件名	測地2011 市道成果と実測値比較						
係数a	1.000003	移動量x	353.508	伸縮率x	1.000005	平均二乗誤差	0.024
係数b	0.001746	移動量y	-292.183	伸縮率y	0.998437	AIC	127
係数c	0.000022			回転角x	0° 6' 0"	標準偏差	0.017
係数d	0.998437	指定数		回転角y	90° 0' 5"	点名セット	
指定数までリセット		準拠点をヘルマートへ		空欄の実測値戻す		準拠点数	20

細かい点で言えば色々ありますが, 標準偏差は改善されましたが相変わらずアフィンが優位と図の歪みがあることとなります。

このデータの場合は三つに分類してそれぞれから準拠点選択を経てアフィン変換すれば最良の結果が得られると判断します。

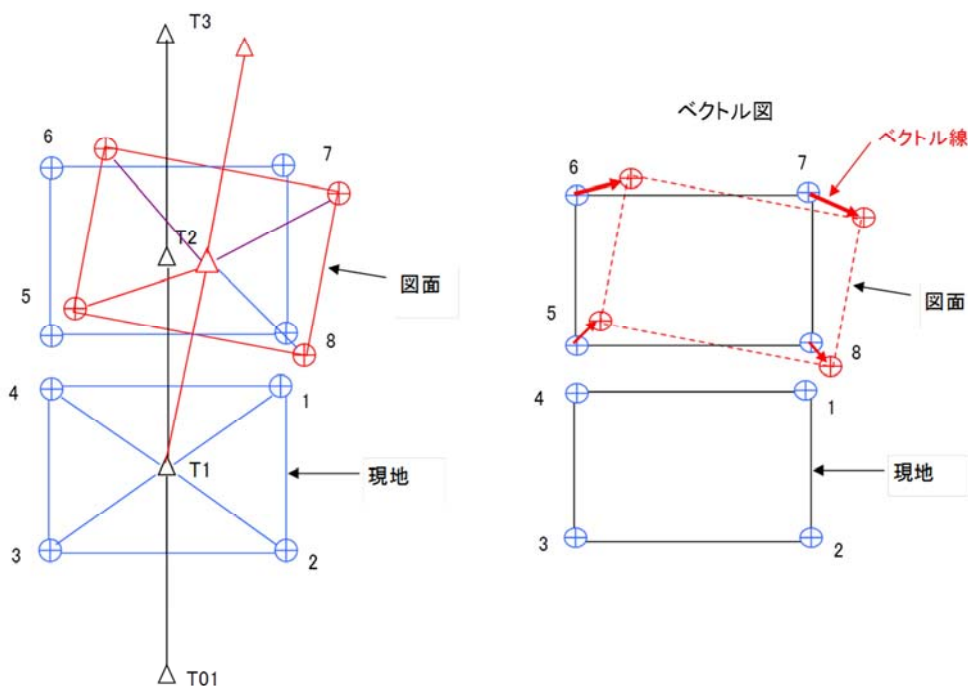
単独誤差によるベクトル線と混合によるベクトル線の違い

単独で起こる境界標の埋設誤差とか境界点の測量誤差によるベクトル線の表れ方と混合分

布に見られる纏まったベクトル線の表れ方には特徴があります。

次の図は混合分布に見られるベクトル線です，一定方向に向いている，特定の回転が見られるなどの特徴があります，これには左図の様な基準点のズレが原因にあるからと考えられます。その結果，右図のベクトル線のように回転して見える例です。

多角点のズレがある境界標の埋設誤差

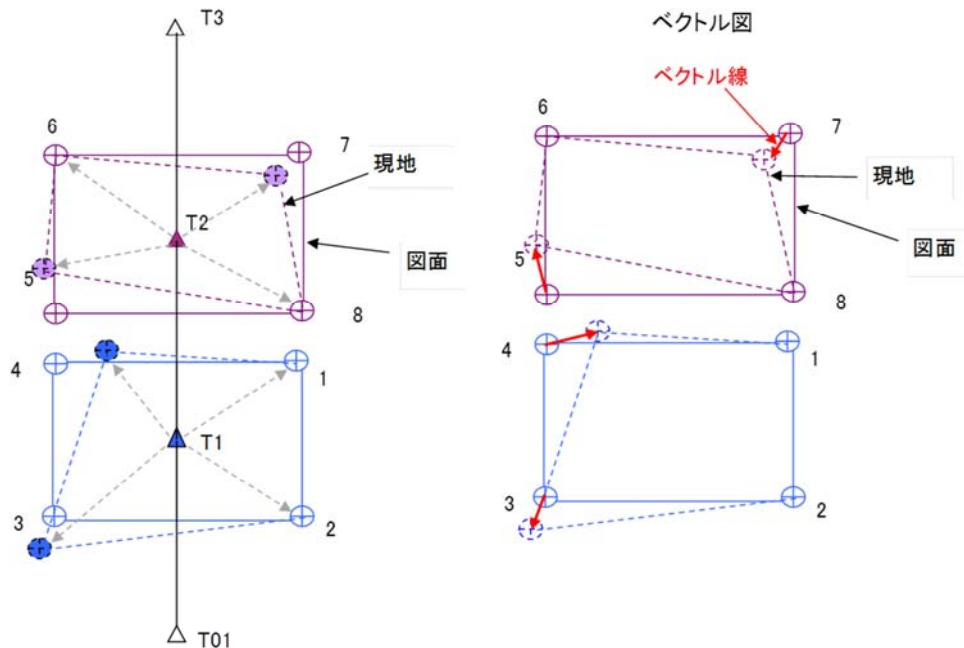


次の図は単独で起こった境界標の設置誤差とか測量誤差に起因するものです，ベクトル線に方向性が無く，ベクトル線の長さにも特徴がありません。

左図のように多角点の回転も見られず，右図のようにベクトル線に一定の傾向が見られないのが特徴です。

混合分布は一枚の図面に二つの分布が混在していますので，これを見極めることが重要です。

多角点のズレが無い境界標の埋設誤差



層別とは

混合分布のあるデータを要素別に分けることを層別と言います、データが明らかに異なる要素によって構成されている場合は問題なく分類できますので個別に境界(筆界)復元をすれば良いのです、難しいのがこの要素をどのように推測するかです。

よく使われるフレーズに「様々な現象を総合的に判断した結果……」と言うのがあります、この言葉はそれだけの能力がある者が言うことで軽々しく使う言葉ではありません、何がどのように総合的なか意外と説明は出来ないものなのです。

推測される要素が確定すれば、層別したグループ間で有意差があるか無いか、つまり異なるデータなのかそうでないのかをF検定を使い判断します。統計的に証明されている手法で判断することで適正化が図られます、この考え方が重要なのです。

層別の考えられる要素

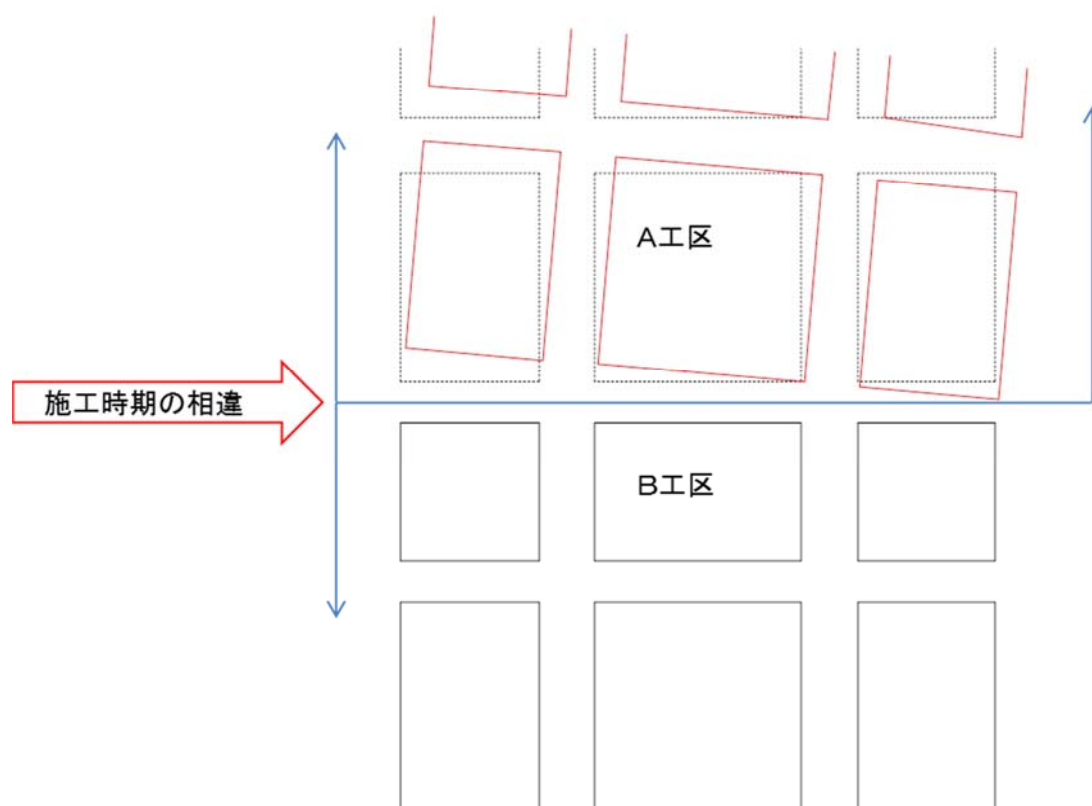
層別の注意点は必ず現地を調べて混合の起きている要素を推定することです、誤ったデータの分類がなされないためには必ず必要なことです。

多くの場合、多角点の展開の要素と境界標設置時期要素が問題になるからで現地の地形、利用状況、境界標の状態などを注意深く観察して要素絞り込みの道具としなければなりません。

1, 工区で回転

施工時期の違いによって工区単位で回転、移動が見られる例です。工区の境は幅の広い

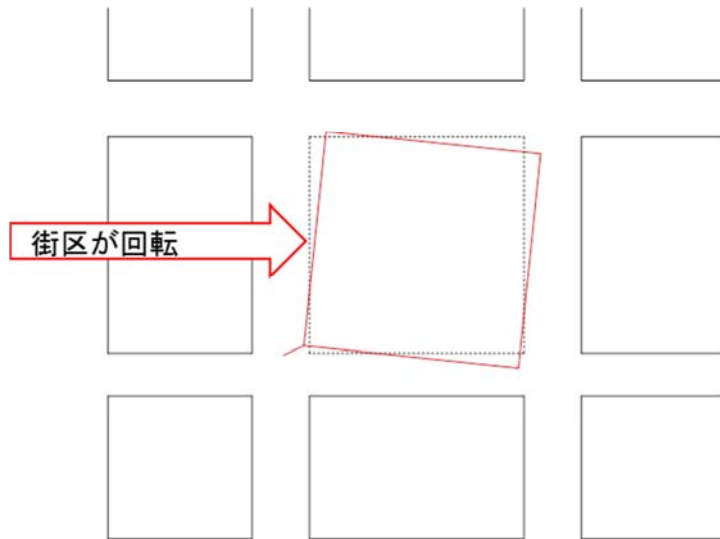
道路のことが多く、道路幅に余裕あり気がつかないことが多いです。



2, 区画整理において街区が回転

一つの街区が回転している例です、区画整理施行時に起きていたのか、その後の官民境界確定で起きたのかは回転している街区内にある境界標との位置関係から検証する必要があります。

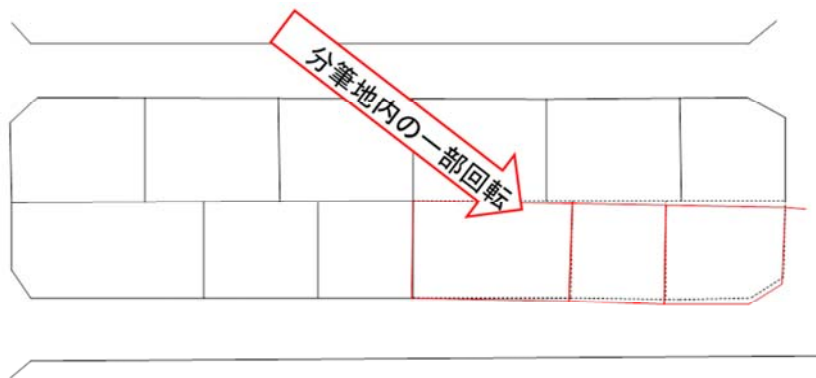
前者であれば基準点の測量ミスが考えられます、後者では移動、回転点有りとしたときの現地にある基準点又は境界標の選択ミスの可能性があります。



3, 土地の一部分に起きている回転

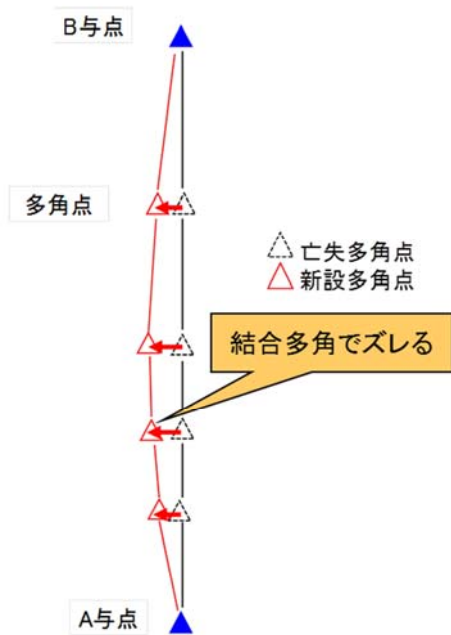
大規模分筆地でいくつかの土地がまとまって回転, 移動している例です, 上側と下側を区画する横の境界を挟んで上側が数メートル高いひな壇になっているような時に見られます。

一筆だけの測量では精度が多少悪いという認識で済みますが街区全体を測るとこの現象に気がつきます。

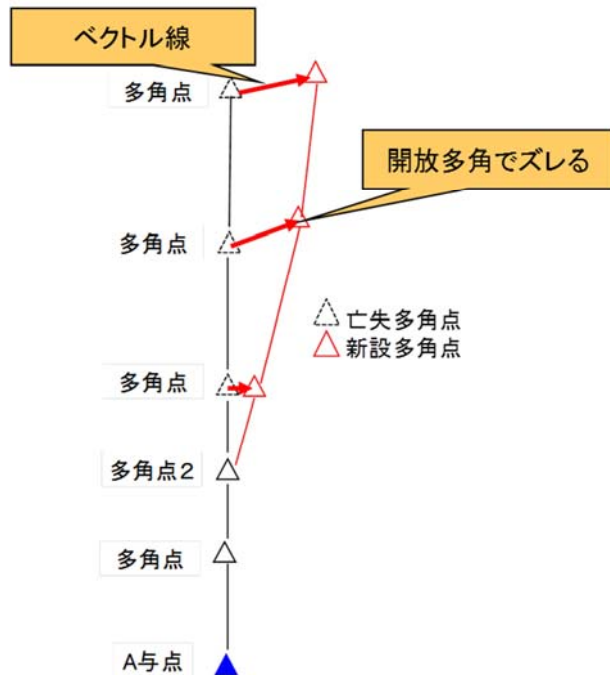


多くの場合, 混合分布の要素の元は多角点の誤測からきていると考えられます, 境界標の復元が原因の場合は旧成果の多角点が亡失している状態で新たな多角点を既存境界標に適した, 適正な多角点を展開出来なかったことにあります

多角点どこかで測角の誤測があった



開放で多角点2で測角の誤測があった



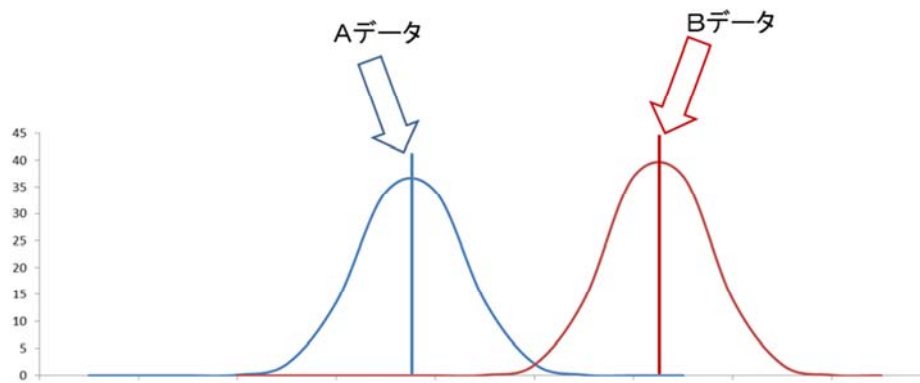
F検定とは

F 検定という検定方法があったのは以前から知っていましたが使い方が分からずにおりました。使うチャンスが少ない、使い方が確立されていないからその面の研究が遅れていたのではないのでしょうかと言うより誰も研究しなかったと言うのが正しい表現でしょう。

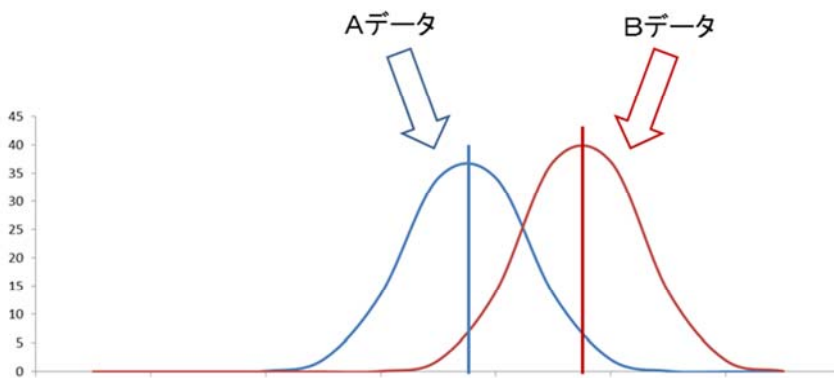
F検定の説明よれば「2つのデータ群の分散(標準偏差に二乗)の比であって、両群のデータとも正規分布に従う場合にはFはF分布に従う。これを用いて、Fの計算値が有意水準内に入るかどうかを検定するのがF検定です。」とされている。

一群のデータの中に要素の異なる複数のデータ郡が混在している場合にそのデータを分類します、この分類の方法は層別という考え方で出来ます。この分類されたデータに差があるかないかを判断します、境界(筆界)復元では差があれば別々に境界(筆界)復元をしなければなりません、差が無ければ一群のデータとして境界(筆界)復元に使えば良いわけです。

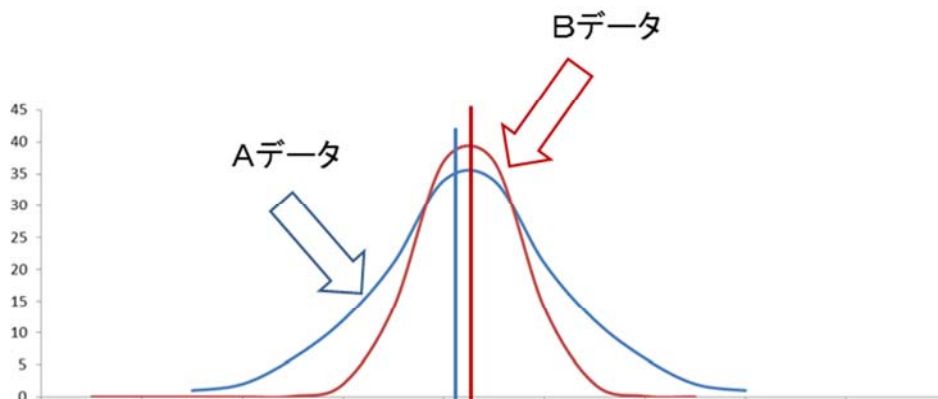
次の1図は較差の分布をAデータとBデータをグラフにしたものです、この二つのデータは同じものか違うものか、データのバラツキはほとんど同じですから精度・標準偏差は変わりませんが明らかに分布の中心が大きくズレていますので違うデータと言えます。



次の2図は、AデータとBデータに差が有るかどうかが、難しいところです。分布の中心が近づいてくるほどの判断が難しくなります。



次の3図は、AデータとBデータに差が有るかどうかが、ここまで重なれば同じデータと言えますがAデータの分布が図より潰れていたらどう判断するのかです。



二つのデータに有意差があるか無いかの判断をF検定で行うものです、F検定は分散比の検定ですから標準偏差が計算出来れば判断できます。

帰無仮説は「同じ作業手順、同じ精度の器機を使って測量した結果は同じ精度である」つまり「分散比<F境界値」「P>有意水準」なら同じ精度と判断します。

上図でAデータと B データの判断要素は山の潰れ具合 (F値)とデータのズレ具合 (P値)の両方を判断します, このような判断を「F検定」という方法で判断します。

エクセルによる F 検定

Excel で分析ツール Excel で分析ツールを使うには、最初にそのアドインを読み込む必要があります。

1. [Excel のファイル]をクリックし、[オプション] をクリックします。
2. [アドイン] をクリックし、[管理] ボックスの一覧で [Excel アドイン] を選びます。
[移動] をクリックします。
3. [有効なアドイン] ボックスの一覧で、[分析ツール] チェック ボックスをオンにし、[OK] をクリックします。

ヒント [有効なアドイン] の一覧に [分析ツール] が表示されない場合は、[参照] をクリックしてアドイン ファイルを検索します。

分析ツールが現在コンピューターにインストールされていないというメッセージが表示されたら、[はい] をクリックして分析ツールをインストールします。

分析ツールを読み込むと、[データ] タブの [分析] で [データ分析] を利用できるようになります。

エクセルによる一変量の F 検定

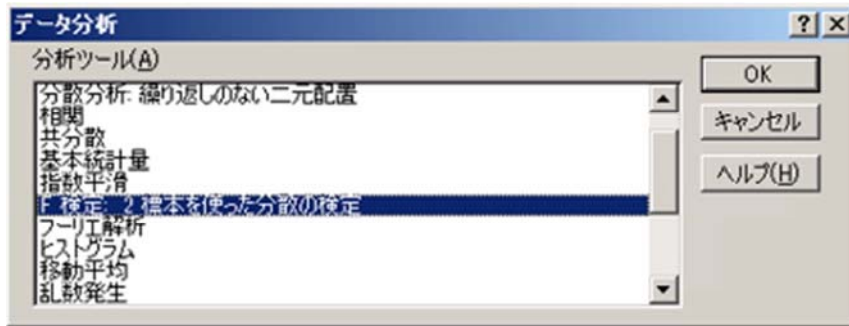
境界(筆界)図に基づいて10個所の図面辺長と実測値辺長の差についてA調査士の結果, B 調査士の結果が表の通りであった, A調査士と B 調査士に精度の差はあるかを調べて見ます。

	A	B
1	A調査士	B調査士
2	7.3	7.1
3	7.4	6.8
4	7	7.3
5	7.7	7.1
6	7.2	7.1
7	7.2	7.2
8	6.9	6.9
9	7.3	7.2
10	7.2	6.8
11	7.5	6.7

1. エクセル分析ツールで F 検定してみます。
[データ] タブの [分析] で [データ分析] 分析ツールを起動します。

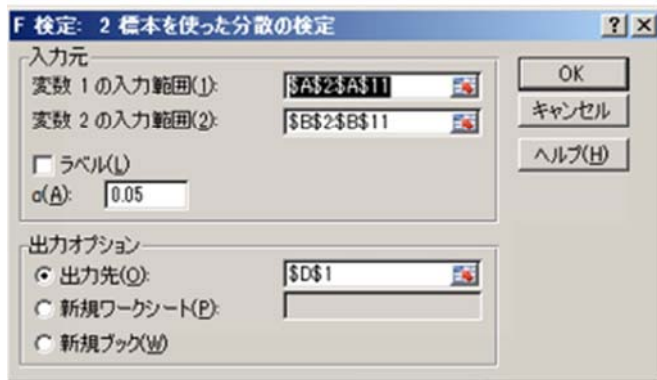


2. F 検定 2標本を使った分散の検定を選択します。



3. 必要な情報を入力します。片側検定のときは、有意水準をそのまま α に入れます。

変数1の入力範囲はエクセルのシートに書き込まれたセルをドラックします，変数2の入力範囲も同様に行います。 α (A)に有意水準を入力します，境界(筆界)の検定では0.05 でかまいません。出力先にチェックを入れ，最初のセルをクリックし，OK ボタンで「F 検定」の結果が次の表のように表示されます。



4. 以下のような結果を得ました。

「同じ作業手順，同じ精度の器機を使って測量した結果は同じ精度である」つまり「分散比 < F境界値」「P > 有意水準」なら同じ精度と判断します。

その逆，「分散比 > F境界値」「P < 有意水準」なら差があると判断します。

分散比(1.28) < F境界値(3.18), P(0.36) > 有意水準(0.05)ですから精度に差が無いと結論できます。

ここで，F境界値は分布の山の潰れ具合をPは山のズレ具合の指標と理解してください。

F-検定: 2 標本を使った分散の検定

	変数 1	変数 2
平均	7.27	7.02
分散	0.053444444	0.041777778
観測数	10	10
自由度	9	9
観測された分散	1.279255319	
P(F<=f) 片側	0.359843267	
F 境界値 片側	3.178893104	

「F 検定 2標本を使った分散の検定」は片側検定の結果が計算されますので両側検定の結果を知りたいときは P 値と F 値を Excel 関数「F.TEST(変数 1, 変数 2)」「F.INV(1-有意水準/2, 自由度 1, 自由度 2)」で次の表の通り求められます。

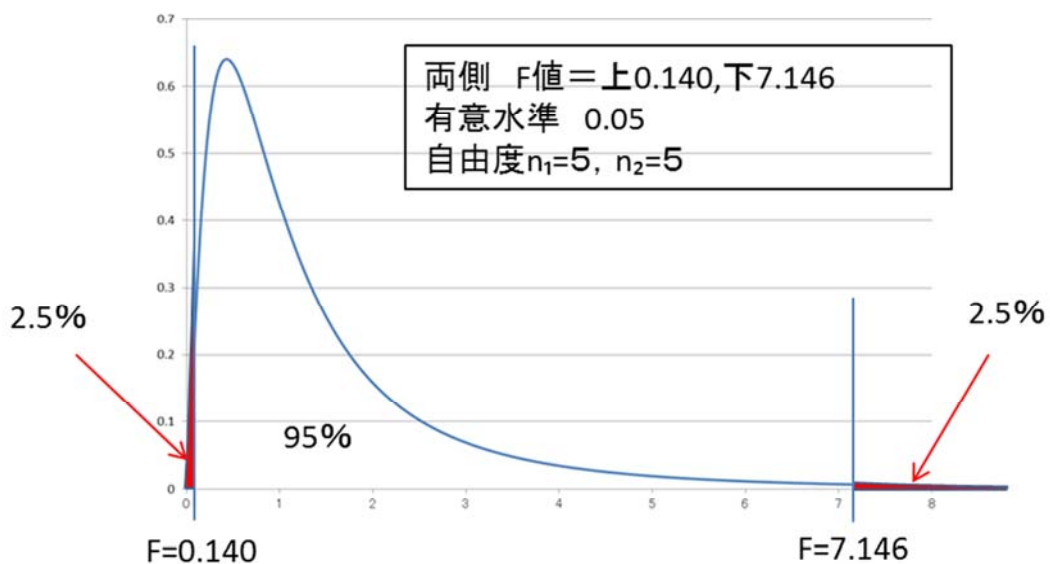
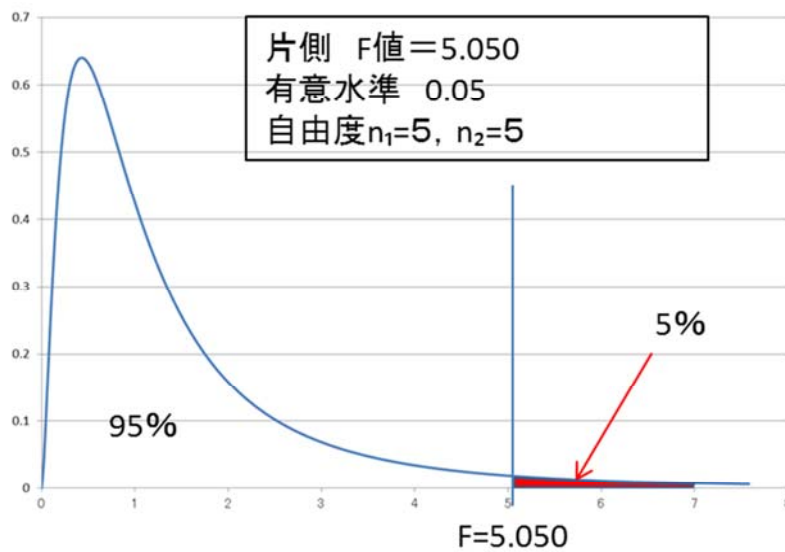
有意水準	0.050	
	変数A	変数B
平均	7.270	7.056
分散	0.053	0.033
観測数	10	9
自由度	9	8
観測された分散比	1.631	
P(F<=f) 片側	0.251	
P(F<=f) 両側	0.502	
F境界値(片側)	3.388	
F境界値(両側)	4.357	

F境界値(片側)「F. INV(1-有意水準, 自由度 1, 自由度 2)」は分析ツールの F 検定 2標本を使った分散の検定 と同じ数値です。

片側検定と両側検定のF値の違い

両側検定は調べたい対象の分母と分子が対等、例えば 2 群の分散比を見ているなら両側検定です。測量成果の場合、群(データの塊)の比較ですから両側検定を使いますので分析ツールの F 検定 2標本を使った分散の検定はそのままでは片側なのでそのままは使えませんので、上表のように計算表を作って計算します。

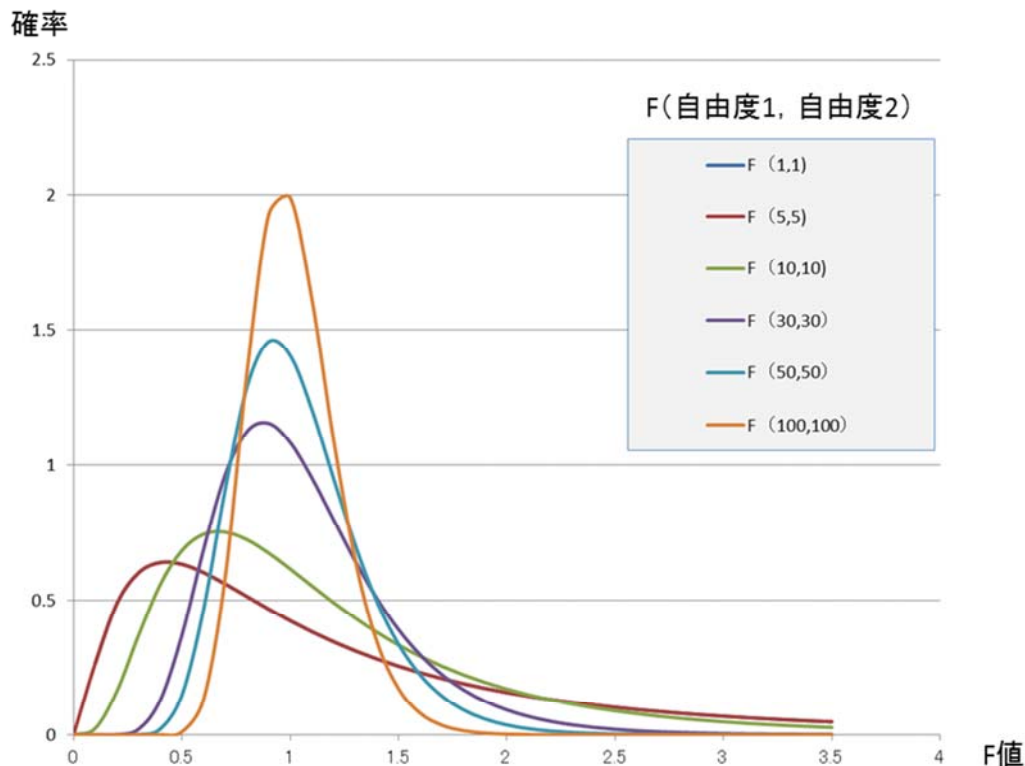
下図では同じ有意水準でも両側と片側では F 値が違って来ます, これを図で示しました。同じ 95%の確率でもこれだけの幅が違ってきます。(エクセルで計算しエクセルで作図したものです。)



自由度によって変わる曲線

自由度によって曲線の形状が変化します, 自由度1と自由度2は数が違ってもかまいません。

自由度が大きくなると正規分布曲線に近づきます, これを大雑把に言い換えると自由度, つまりデータが少ないほど判定は信頼性に欠けるという意味です, これは t 検定でも χ^2 二乗検定でも同じです。



(エクセルで計算しエクセルで作図したものです。)

エクセルによる二変量の F 検定

境界(筆界)復元は座標値で行います,これは二変量の検定になります,ここまでは一変量の検定でした,しかし二変量の検定そのものはありませんので二変量から指標とする一変量を取り出す必要があります。

標準偏差の計算と変数の一覧を取り出す必要が有ります,一変量として取り出した標準偏差の値と二変量標準偏差の値に差が出来るだけ生じないように配慮して行います,ここの部分の説明は省略します,メチャクチャ複雑なので。

F値=「F.INV(1-有意水準/2, 自由度 1, 自由度 2)」では自由度1はデータ×2-1 となります,これを使って F 値を求めます。

P値=「F.TEST(変数 1, 変数 2)」では変数1は連続したデータでなければ計算できません,そこで△xを1~n個並べて,△yを△xの次に連続して並べます。10個の座標値があれば20個の変数1が出来上がります。これを変数2でも行います。

変数1と変数2は一応座標変換した値ではあるのですがそのままは使えません,そこで変数1の中心を△x=0, △y=0において変数2を同量移動させた値を使います。(下図,ここがポイント)

これで,変数1,2が出来ましたのでExcelの関数P値=「F.TEST(変数 1, 変数 2)」で求めれば完了です。

混合分布のある境界(筆界)について

混合分布のある状況では筆界とは何かを自問, 自答しなければならず, ここが技術と制度の狭間にあって結論の難しいところです。要するに, 筆界は図面上で形成されるのか現地で形成されるのかということです。図面上で形成されたのであれば混合分布が確認された場合は区画単位で現地の境界標を設置し直すこととなりますが現地で形成されたのであれば図面を直すことになるからです。ただし図面を直すと言っても現状に境界標の位置を境界(筆界)とするとということではありませんので誤解の無いようにして下さい, このような状態が見つかるときは現地が作られてから相当の時間が経過してから見つかりますので権利関係と相まってどちらの方法を取るにしても大変なことなのです。

2016/12/02

土地家屋調査士・測量士 小野孝治