

法・平均二乗誤差について

【法・平均二乗誤差について】

国土調査法、及び不動産登記法に平均二乗誤差というのがあります、統計学にも平均二乗誤差(MSE)があります。

同じ平均二乗誤差という言葉ですが計算式が全く違います、ここで私が説明するのは国土調査法、及び不動産登記法に定められた平均二乗誤差についてです。誤解の無いように統計の平均二乗誤差と区別するため**法平均二乗誤差**として説明します。

法平均二乗誤差は二変量のデータになります、測量では座標軸を置いて、座標軸原点から測点の位置を x 、 y の二変量で表します、数百点から数万点といった大量の測量結果の良否を判定する場合、その中からランダムに抜き取り検査を行います。

この時の測量結果と検査結果の較差、 $\Delta x \cdot \Delta y$ から各種のデータを計算し、与えられた公差から全量の測量結果が良か不良かを判断するものです。

統計上 100%の良、0%の不良率という判断はありませんので決められた公差に対して〇〇%の不良率以下なら全体を良(合格)と判断するものです、これは一般的に行われている品質管理の手法と同じです。

この判断基準として法平均二乗誤差が決められています、法平均二乗誤差が何パーセントの不良率を想定して定められたものなのかが分かりません。それを調べるために二変量・良率計算方法(本ExcelBOOK)を考え計算した結果です。

その結果、法平均二乗誤差(国土調査法、及び不動産登記法に定められている平均二乗誤差)と良率(良品率、不良率)の関係はありません、つまり、法平均二乗誤差の値が幾らなら良率(良品率、不良率)は幾らに相当するという関係は無いということです。

これについて議論する前提として二変量の良率(良品率、不良率)を計算する必要があります、そこで私なりの知識の範囲から知恵を絞って二変量の分布と公差から良率(不良率)を計算する方法を考えました、このプログラムで計算した結果には 1000 分の 2%程度の計算誤差があるようです、もっと正確な方法があるはずですがこの程度しか思いつきませんでした。

【結論】

品質管理でつかわれるフォーシグマ(4 σ)法とシックスシグマ(6 σ)法について簡単に説明すれば、4 σ 法と6 σ 法は一変量の良品率を定めた基準でその会社独自の基準がなければこの基準を使っているはずです。

4 σ 法とはワンランク下の3 σ の一変量の確率99.73%(不良率0.027%)以上に良品率(不良率は以下)を管理するものです、この基準は高度成長期(1960年代)に取り入れられましたがその後の技術の高度化により6 σ 法が使われ、ツーランク下の4 σ の確率、良品率が99.9936658%以上になるように管理するものです。

ワンランク。ツーランク下の確率を使うのが公差中心と誤差中心のズレ、これをユラギと言いますがこのユラギを考慮するものです。

この考え方をそのまま適用すれば、検査を数値座標で実測値を平板から読み取った値を使った場合は4 σ 法で、検査、実測とも数値座標の場合は6 σ 法で判断するという考えが適切と考えられます。

座標データは二変量ですから二変量確率から4 σ で99.9665%以上(不良率では0.034%以下)となります。

良率・不良率を計算した場合の二変量確率から4 σ での良率 99.9665%以上(不良率では0.034%以下)で判断すれば良いこととなります。

実務上、正確な良率・不良率が得られたとしてもなんらかの判断に使われるものではないし、使われていませので“それがどうした!”と言われるでしょう、が少なくとも法平均二乗誤差

(国土調査法施行令別表4、統計の平均二乗誤差とは異なるので法をつけて区別した)よりはまともな判断資料になります。

法平均二乗誤差に関しては様々な資料があります、後尾にそれらの資料を紹介しました。これらの資料の式の説明は微妙に異なりますが説明に忠実に計算すると同じ計算結果になります。

したがってこの業界で使われている法平均二乗誤差は誤っていることになります。

【実際の計算例】

このデータは甲1地区、公差0.020mです。計算には下表のデータが必要です(平均二乗誤差mとデータ数nはいりません)。

法平均二乗誤差の公差は0.020です、結果は0.0187ですから法平均二乗誤差からの判定は“良”になります。

計算データ		saとdaenから	
平均二乗誤差m	0.0187		
誤差中心 $\angle X_{avg}$	0.0098		
誤差中心 $\angle Y_{avg}$	0.0140		
楕円回転角	103.9521	103°	57' 07"
相関係数 ρ	-0.1266		
二変量標準偏差 σ	0.0051		
データ数 n	26		

下表が良率・不良率の計算結果です。

良率計算		①を②へコピー	
①求良率の計	0.34326		
①315/45の率	0.00000	①を③へコピー	
②楕円角の率	0.11484	⑦良率	99.25725
③45/135の率	0.34326		

これを良率・不良率で計算すると良率99.25725%であり(不良率0.74275%、想定22倍の不良がある)、全体が不良と判断されます。

このデータは某法務局管内の筆界特定事件に法務局が選んだ筆界特定測量実施者が測量した結果から計算したものです。元のデータは法14条地図のデータで、地籍調査測量を2017年市は実施し、2020年に法14条地図に認定、その翌年2018年に筆界特定調査測量が法務局で実施したものです。

このままでは測量のデータを比較した場合の良、不良の判断は相対的な比較になるので地籍調査測量に問題があったのか筆界特定調査測量に問題があったのかは判断できません。通常は地籍調査測量の精度が高いのが普通です。

品質管理の場合は測定精度の高い検査機器で測って判断しますが現代の測量では法14条地図のデータ、筆界特定調査測量データとも同程度の高精度測量機器を使っていますのでどちらの精度が悪いという判断は難しくなります、相対的にみてどちらかの測量に問題があることはわかりますが座標データだけでは判りません。

ただし、その時の**基準点網図観測データ、境界点観測データ**があればある程度の解析は可能です、本例の場合はそれらのデータが入手できたので良率の低い原因を特定できました。このデータは筆界特定調査測量の基準点網図の組み方に問題があったのです。

”基準点測量の知識が無いか乏しい知識の者が計画し、測量した不適切な結果だったと言うことです“、その結果に基づいて筆界特定がなされてことになります。その結果誤った筆界特定がなされたことになります。この旨を主張したのですが裁判官、代理人弁護士に測量の観点から説明しても理解が得られなかったものです。

測量の専門家でない裁判官、代理人弁護士には単純に良率・不良率の説明ができれば状況は変わっていたかもしれませんが、この時はまだ二変量の良率・不良率を計算する手立てが得られず歯がゆい思いをした経験があります。

この時のことは当方の依頼人が依頼した弁護士が裁判官、相手の弁護士に同調して反論をしなかったことが今でもショックで残っている。これだけが原因ではないのだが。
この時の悔しさがこのプログラムを作る原動力になった。

【同じ法平均二乗誤差で良率・不良率が大きく異なる】

下表は法平均二乗誤差 0.020 (甲 1 の公差) のデータです。データは二つです、この二つのデータから良率・不良率を計算して見ると公差=、二変量標準偏差=、相関係数=、楕円回転角=、ユラギ ΔX_{avg} =、 ΔY_{avg} =を要素として計算すると良率 87.49192%・不良率 12.50808% と良率 98.51927%・不良率 1.48073% と大きく異なる。

このことから法平均二乗誤差は何を意味しているのか解らない。

法平均二乗誤差が同じなのに良率・不良率がこんなにも違うのか・・・驚きです。

平均二乗誤差 0.02009

計算データ	saとdaenから	
平均二乗誤差m	0.0201	
誤差中心 ΔX_{avg}	-0.0006	
誤差中心 ΔY_{avg}	0.0099	
楕円回転角	103.9521	103° 57' 07"
相関係数 ρ	-0.1266	
二変量標準偏差 σ	0.0128	
データ数 n	26	

平均二乗誤差 0.02004

計算データ	saとdaenから	
平均二乗誤差m	0.0200	
誤差中心 ΔX_{avg}	0.0108	
誤差中心 ΔY_{avg}	0.0150	
楕円回転角	103.9521	103° 57' 07"
相関係数 ρ	-0.1266	
二変量標準偏差 σ	0.0051	
データ数 n	26	

ΔX_{avg} ・ ΔY_{avg} と二変量標準偏差が異なる

(1)を(3)(4)へコピー) 良率計算			
①求良率の計	0.34455	①を③へコピー	
②315/45の率	0.00000	①を④へコピー	
③楕円角の率	0.34455	⑦ 良 率	87.49192
④45/135の率	0.04642	⑧ 不良率	12.50808

(1)を(3)(4)へコピー) 良率計算			
①求良率の計	0.55472	①を③へコピー	
②315/45の率	0.00000	①を④へコピー	
③楕円角の率	0.17835	⑦ 良 率	98.51927
④45/135の率	0.55472	⑧ 不良率	1.48073

20231215 作成
作成 小野孝治

【資料】

(1)まず、土地家屋調査士連合会が発行している「調査・測量実施要領」の附42頁の説明です。

(2) 平均 2 乗誤差

いくつかの筆界点をもっている誤差 (公差以下) の平均 2 乗誤差 (標準偏差) (m) を計算した値が、この数値以下でなければならないという意味で、次式で示される。

$$m = \pm \sqrt{\frac{[\delta^2]}{n-1}}$$

ただし、 δ : 残差, n : 筆界
点数

(3) 筆界点の位置誤差 (平均 2 乗誤差) の求め方

筆界点の位置は、一般に、1 点 1 観測で決める場合が多い。1 点 1 観測では、平均 2 乗誤差 (標準偏差) の算出はできないので、便法として、次のようにして求める。

(i) あるグループ (10~20 点) を定め、そのグループについて、それぞれの点が、どの程度の偏位でプロットされているかを測定する。

(要 領)

- ① 平板上にプロットされた筆界点の図上座標 (x, y) を図上から求める。
- ② 高性能のトランシットを用いて地上座標 (x_0, y_0) を計算から求める。
- ③ x_0, y_0 を最確値とみなし、

$$dx = x_0 - x$$

$$dy = y_0 - y$$

$$\delta = \sqrt{dx^2 + dy^2}$$

δ を残差とみなし、次式によって平均 2 乗誤差 (標準偏差) m を求める。

$$m = \pm \sqrt{\frac{[\delta^2]}{n-1}}$$

ただし、 $[\delta^2] = \delta_1^2 + \delta_2^2 + \delta_3^2 + \dots$

n : グループの点数

(2) 次は日本加除出版が発行している「14条地図活用マニュアル」の41頁です。

ここでいう標準偏差とは、JIS規格によれば、

標準偏差 = 分散の平方根

となっており、また、分散 (標本分散、不偏分散ともいう。) は、

$$\text{分散} = \frac{(x_i - x_m)^2 \text{の総和}}{n-1} \quad \dots\dots (2-1)$$

$i = 1 \sim n$

n : 試料 (標本) の個数

x_m : 試料 (標本) の平均値

なる式で表される。なお、標準偏差は分散の平方根の正の値のみを、平均 2 乗誤差は正負の値を指す。

(3) 次は建設大学校が使っている教材から編集して出版されている「測量のための最小二乗法」の14頁からです。

(3) 標準偏差(平均二乗誤差) 一連の観測誤差 $\Delta_1, \Delta_2, \dots, \Delta_n$ の 2 乗の和を誤差の数で割った平方根をその一連の観測の標準偏差(standard deviation) または平均二乗誤差(mean square error)という。

いま、標準偏差を m とすれば

$$m = \sqrt{\frac{\Delta_1^2 + \Delta_2^2 + \dots + \Delta_n^2}{n}} = \sqrt{\frac{[\Delta\Delta]}{n}} \quad (2.17)$$

ただし、 $[\Delta\Delta]$ は Δ の 2 乗の総和を示す。

(4) 次は日本加除出版発行の「地籍測量」18頁からです。

(3) 平均二乗誤差

平均二乗誤差(標準偏差ともいう。)は、測定値と真値との差の二乗の相加平均の正の平方根をとることにより求められるもので、測定値のバラツキ具合を数量的に表す。平均二乗誤差が小さいほど、その測定精度が高いという。精度区分に示された数値は、筆界点位置の平均二乗誤差の限度の目安を表す。地籍調査では、平均二乗誤差を次式で求める。

$$m = \sqrt{[\sigma^2]/(n-1)}$$

ただし、 σ : 較差

n : 筆界点数

これらの資料で平均二乗誤差(標準偏差)と言っている、平均二乗誤差は標準偏差ではありません。平均二乗誤差(法平均二乗誤差)で良、不良の判定は出来ないことは本文で説明済みです。