

## 基準空間の決め方とマハラノビス距離についてシミュレーションによる考察

## — ORリテラシーの普及事例（第4報） —

01102345 オーアールとく塾 権藤 元 GONDO Hajime

## 1. はじめに

正常なグループを基にしてマハラノビスの距離<sup>(1)(2)</sup>を求め、ある閾値を越えるときに異常発生を予知する方式は、計算処理はエクセルのシート上で手軽にでき、データを持ち正常と判断する技術力ある者には実用上かなり有効な方法でORリテラシーの一つとして取上げることは前回<sup>(3)</sup>述べた。しかし、基準空間・閾値の設定などは課題を持つ者が自分の知見を活かして自ら決める必要があり、それには基準空間の決め方によってマハラノビスの距離が変化する状況を知っていることが有効であり、そのためにシミュレーションによりそれらの状況を明らかにした。なお、前回発表時に座長より指摘のあったSN比による評価も行った。

## 2. 考察の対象

基準空間は正常なグループを対象に決めるものであり、この際、異常を若干含んでないか、逆に厳選し過ぎていないか、また、データ項目に基本的なデータを欠除していないか、無関係なデータ項目を多く含んでいないか、などなど基準区間を決めるにときにいろいろ悩むものである。これらを考察の対象とした。

## 3. シミュレーションの方法

相互に関連し相関のある測定データをシミュレーションするために、次のモデルを考えた。図表1に示すように真の要因（母要因と称する）としてA、B、Cの3つを用意し、この3つの母要因から2つの要因を選び、それと乱数を線形結合して測定しているデータ項目を得るモデルとした。線形結合のウエイトを3種類用意して計9項目の測定データ生成する。なお、A、B、Cおよび線形結合の対象となる9つの乱数はともに独立な正規乱数系列である。

次ぎにシミュレーションのケースについて述べる。

\* 基準空間：(1)正常なもの（純正基準空間とした）(2)異常と正常の境界レベルのデータを追加したもの（混じり基準空間とした）(3)要因ごとに外れに近い測定データを削除したもの（カット基準空間とした）(4)母要因Cのウエイトをゼロとしたもの（ABのみのみなし基準空間とした）

\* データ項目：(1)無関係項目を含まない9項目（基本項目とした）(2)基本項目に無関係の乱数を加え15項目（含無関係項目とした）

\* 評価対象：(1)SN比対象は母要因のマハラノビス距離1~4.5（0.5刻み）繰り返し数4 (2)正常グループは正規乱数40個

## 4. シミュレーションによる考察の結果

シミュレーションの結果は、図表2にデータの一部を、図表3に主なグラフを示す。これらから大胆に結果をまとめると以下のようなになる。

- \* 母要因を1つ欠いても影響は大きく、基本的なデータを欠くときは有効な結果は得られない。
- \* 正常グループの中に異常が混じると期待する結果は得られない。
- \* 正常グループを厳選するとマハラノビスの距離はやや過大になるが異常は明確になる。
- \* 無関係なデータ項目を含むと若干の効果が薄れるが、後で効果のないデータ項目の削除は可能であり、最初の段階ではあまり気にすることはない。

## 5. おわりに

マハラノビスの距離を用いた予測においては、課題を持つ者の知見を取り入れた基準空間の絞り込みを行うことが有効であることが知られている。今回のシミュレーションはこれをおおむね確認出来た。しかし、まだ残された検証事項もあり今後もこの考察は続けたい。なお、今回のシミュレーションの計算処理はエクセルで行っており、課題を持つ者が必要に応じて自らシミュレーションすることが望まれる。このような姿勢をORリテラシーとして普及させたい所存である。

ご意見をお待ちしている。 Eメール：hajime.gondo@nifty.com

参考文献

- (1) 奥野ほか、多変量解析法、日科技連、1971. 10、P279
- (2) 田口、品質工学の数理、日本規格協会、1999. 6、P136
- (3) 権藤、血栓症予測に用いたマハラノビス距離基準空間選定方法の確認シミュレーション—ORリテラシーの普及事例（第4報）—、OR学会秋季研究発表会予稿集、2000. 9

図表1 測定データ生成モデル

母要因	基準空間—データ																	
	含無関係項目												無関係データ					
	基本データ項目									無関係データ								
No.	A	B	C	AB1	AB2	AB3	BC1	BC2	BC3	CA1	CA2	CA3	R1	R2	R3	R4	R5	R6
1	2.8	0.7	-0.3	4.4	3.2	3.2	-0.6	-0.8	-0.5	2.3	2.0	0.9	0.4	-0.4	-0.6	-0.2	0.3	-1.8
2	-0.1	3.0	0.2	3.9	1.9	1.4	1.4	2.9	0.8	-0.7	-0.6	2.3	0.6	0.7	-0.2	0.6	2.1	0.1
...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...
100	0.6	1.4	-1.0	3.4	0.2	2.0	-0.8	0.4	0.1	-0.4	-2.0	-1.2	0.5	-0.7	0.8	-0.2	-1.3	1.2
ウ	母要因A			0.7	0.8	0.6	0	0	0	0.7	0.6	0.8	0	0	0	0	0	0
エ	母要因B			0.7	0.6	0.8	0.7	0.8	0.6	0	0	0	0	0	0	0	0	0
イ	母要因C			0	0	0	0.7	0.6	0.8	0.7	0.8	0.6	0	0	0	0	0	0
ト	乱数R			1	1	1	1	1	1	1	1	1	1	1	1	1	1	1

図表2 SN比評価対象データの一部

母要因	マハラノビス距離								SN比			
	2	2	2	2	4	4	4	4	基本	混じり	カット	ABのみ
基本データ項目	1.5	1.5	1.5	1.5	2.5	3	2.7	2.7	32.9	31.7	32.3	29.4
含無関係項目	1.3	1.5	1.5	1.7	2.2	2.4	2.2	2.3	29.6	28.3	30.5	27.0

図表3 マハラノビス距離の状況

