# 寄書　天気予報の検証と要因直交化の問題

鈴　木　栄　一*

本誌前号に学会3月例会の批評がある．そのうち筆者の2つの報告について，その後調べたこと，討論されたことをまとめてのべておきたい．

まず，情報量を天気予報に導入する試みは，天気予報も不確定な要素を含んだ情報であり，情報の価値を確率的に計量する情報のエントロピーでその検証ができると見たからで，3月例会では主に結果の検証をのべ，5月のO.R.学会で，技術と結果の検証が統一的にできることを指摘したのである．情報のエントロピーは

連続量の場合……

$$H(x_1, x_2\cdots, x_n) = -E[\log f(x_1, x_2, \cdots. x_n)]$$

不連続量の場合……

$$H(x_1, x_2, \cdots, x_n) = -\Sigma p(x_1, \cdots, x_n) \log p(x_1, \cdots, x_n)$$

である．ただし $f(x_1, x_2, \cdots, x_n)$ は確率密度函数，$p(x_1, x_2, \cdots, x_n)$ は $x_1, x_2, \cdots, x_n$ の同時生起確率とする．(Stieltjes 積分をつかって2つをまとめた形にかくこともできる。)

(a)　予報技術の検証

予報対象 $x_1$ の不確定な程度 $H(x_1)$ が，それに関与する予報要因 $x_2, x_3, \cdots, x_n$ を知ることにより $H_{x_2,\cdots,x_n}(x_1)$ まで減るとする．$H_{x_2,\cdots,x_n}(x_1)$ を条件づきエントロピーという．$x_1$ の分散 $\sigma_1^2$，$x_1$ と $x_2, \cdots, x_n$ の重相関を $R_{x_1\cdot x_2\cdots x_n}$ とすると，多変数正規ならば

$$H(x_1) = \log\sqrt{2\pi l}\,\sigma_1,$$

$$H_{x_2\cdots x_n}(x_1) = \log\sqrt{2\pi l}\,\sigma_1\sqrt{1-R^2}_{x_1\cdot x_2\cdots x_n}$$

となり，その差が予報技術の効果をあらわすとすれば

$$I = H(x_1) - H_{x_2\cdots x_n}(x_1) = -\log\sqrt{1-R^2}_{x_1\cdot x_2\cdots x_n}$$

となり，これを Базров（バグロフ）にならって条件つき相対エントロピーと名付けよう．つまり，多変数正規なら重相関が大きい程，よい予報技術で，予報対象 $x_1$ の不確定さの減り高 $I$ も大きい．これは多変数正規に限らずどんな分布でもいえる．

(b)　予報結果の検証

実況 $x_1$，予報 $x_2$ とすると，実況について何も知らないときの不確定さ $H(x_1)$ は予報 $x_2$ という情報を得ることにより $H_{x_2}(x_1)$ まで減る．その減り高

$$I(x_1, x_2) = H(x_1) - H_{x_2}(x_1)$$

についても同様の関係が成立つ．($R^2$ の代りに $x_1$ と $x_2$ の相関係数 $r^2$ が入る) $I(x_1, x_2) = I(x_2, x_1)$ が成り立ち，この条件つき相対エントロピー（批判の条件つき相関エントロピーはミスプリ）を伝達情報量ともいう．つまり，どの位正確な情報が伝達されたかの度合を示すもので，統計でよく知られた尤度 $\lambda$ をつかうと，

$$I(x_1, x_2) = (\log \lambda)/n \quad (n：資料総数)$$

なることが証明され，これから，分割表の検定をするときつかわれる $\chi^2$ と結びつくことがわかる**．つまり，$I$ は連続量のときも，不連続量のときも定義され，

不連続のとき，$I$ は $n\to\infty$（実際は$\geqq 0$）として $\chi^2$

連続正規のとき，$I$ は $-\log\sqrt{1-r^2}$

正規でないとき，$I$ は $-\log\sqrt{1-\eta^2}$（$\eta$は相関比）

という一般性をもつ．(Skill Score なども $I$ の特別な変形にすぎないことは一寸厄介だが証明できる)

情報の条件つきエントロピー，条件つき相対エントロピーといった情報量をつかうことにより，これまでいろいろいわれてきた検証の目安（予報と実況の相関とか，分割表で対応させたときの $\chi^2$ とか，Skill Score とか）が一まとめにできる．

つぎに要因の直交化とは図で示すと，つぎつぎと順序に回帰直線，回帰平面をつくることで Schmidt の直交化とよばれる．$x_2$ に対する $x_1$ の回帰直線からの残差 $\xi_1$，$x_2$，$x_3$ に対する回帰平面からの残差 $\xi_2$ を $x_1$，$\xi_1$ をつかってあらわすという工合にやるが，残差 $\xi_i$ は $x_1$ と $\xi_1\cdots\xi_{i-1}$ であらわされ次第に小さくなり，$x_1$ に対しまづ $x_2$ が関係し，その関係を通して $x_3$ に関係が及ぶという after effect（気象の持続性）の問題はこのようにして解明され，after effect の寄与の度合がどこまで及ぶかを説明できる．これを単なる相関でやると見かけ上の関係がでてきてわからない．3月例会で報告した例はその意味では確かに適切でなかった．その時1つの簡便な近似法として

$$(x_1 - \bar{x}_1) = \frac{\sigma_1}{\sigma_2} r_{12}(x_2 - \bar{x}_2) + \xi_1$$

**分割表独立性検定のための $\chi^2$ は $-2\log\lambda = -\frac{2}{n}\lambda$ が $n\to\infty$ のとき漸近的に $\chi^2$ になることから尤度比検定法の近似法として導びかれたものである．

* 気象研究所—1958年7月5日受理—

なる残差 $\xi_1$ を

$$\xi_1=\frac{\sigma_{\xi_1}}{\sigma_{x_4}}r(\xi_1, x_3)(x_3-\bar{x}_3)+\xi_2$$

$$\xi_2=\frac{\sigma_{\xi_2}}{\sigma_{x_3}}r(\xi_2, x_4)(x_4-\bar{x}_4)+\xi_3$$

…………

として逐次小さくしてゆくことを考えたが，精度に問題がある．1948年1月 500mb 高度の (135°E, 50°N) における値 $x_1$ を前日の値 $x_2$，前日の周辺の値 $x_3$，$x_4$*をつかうと，(after effect として適切でないが，ここでは単なる計算例としてみて頂きたい) 逐次方式では

$$(x_1-\bar{x}_1)=0.5126(x_2-709.16)+\xi_1$$

$$\xi_1=0.2834(x_3-658.39)+\xi_2$$

$$\xi_2=0.0953(x_4-802.81)+\xi_3$$

となり，結局 $x_1$ に対する関係は代入して

$$(x_1-\bar{x}_1)=0.5126(x_2-709.16)+0.2834(x_3-658.39)+0.0953(x_4-802.81)+\xi_3$$

となる．一方正確な式は $x_3$ までのとき，

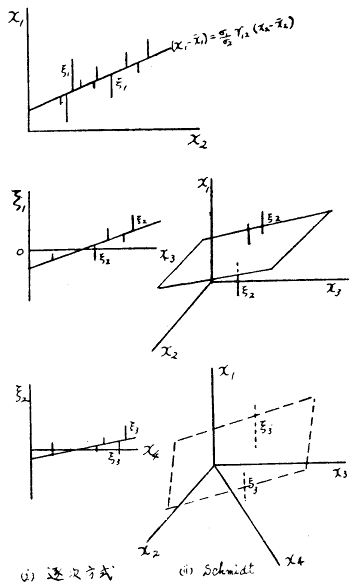
$$(x_1-\bar{x}_1)=0.538(x_2-709.16)+0.301(x_3-658.39)$$

$x_4$ までのとき

$$({}_1x-\bar{x}_1)=0.529(x_2-709.16)+0.298(x_3-658.39)+0.199(x_4-802.81)$$

で，逐次方式と大してちがわない．(ただし内部相関は $\gamma_{23}=0.3581$，$\gamma_{24}=0.2178$，$\gamma_{34}=0.2797$ で割合小さい) つまり，内部相関があまり大きくないと逐次方式でもかなり正確な式に近い．Schmidt の方式は正確な式を見方をかえて変形したもので，実質的には同じ内容と考えられる．

* すべて1万フィートをひいてある．




増山先生はこのような Schmidt の方式で after effet を研究されたが，気象の問題でも，どこまである現象の影響が及ぶかといった類の現象の統計的解析，つまり $x$ に対し $x_2$，$x_3$，…… といった寄与すべき変数の順序があらかじめ指定された現象の解析に直交化が有効につかわれる．

当日筆者が報告した例は確かに適切とはいえないわけだが，気象にもこの方法がつかえる面が相当あるのではないか．単なる相関が気象で広くつかわれているが，見かけ上の関係を追っている危険もあり，場合に応じて偏相関をとるとかして，なるべく現象にあった統計法をつかう工夫が必要なのではなかろうか．

〔書評〕 山本三郎著　富士山

朋文堂ガイドブック 28版　92頁　120円

富士山の研究案内書は多数あるので屋上屋を架することは甚だ困難である．それが今度山本君によって敢行されたことを喜びとしたい．従来の案内書には夏山の一般コースのみを示したものが総てであったが，ここでは冬山が主であり，然も総てが孫引をせず自分の足で書かれている．そしてここに特記したいことはその中に氷の状態の変化とか突風の来易い方向とかそういった気象条件が精しく書かれていることである．気象条件と切り離して登ることが不可能な冬山ではあるが，それを案内に織り込んだという点が全く新らしい試みだ．コースの図には主風向が書き加えられており，これは同君が一千個近い豆風船を飛ばせて調べたものである．夏山にしても古御岳流し右岸，七太郎尾根，主杖流し，大沢右岸，大沢左岸等従来見られなかったルートが加えられている．富士山の気候及び富士山の気象の二章は年間を通じての観天望気が体験として精密に説かれてあり，山の気象上からも貴重な文献である．写真も積雪期を多く取り入れて面白いものが多い．単なるガイドブックとして以上に読物としてもよくできた本である．この本は前述のように孫引が少しもなく全く独創的なものである．若し些か欲の深いことをいうならば余りに独創的過ぎて他の多くの人々に依る記録や遭難気象等がないことは残念である．

(1958.9.10)　(高層課　大井正一)