

稀な現象の確率予測にも使える適切な採点法 Proper scoring systems available for probability forecasts targeting rare phenomena

林 豊^{1*}
HAYASHI, Yutaka^{1*}

¹ 気象研究所
¹ Meteorological Research Institute

新たな手法を導入して予報を確実に改善できるためには、予報精度を適正に評価する採点法が存在し、新手法を導入した場合に出すであろう予報の採点結果（スコア）が従来の方法による予報のスコアを上回ることが必要条件である。「適正な」採点ルールは、予報者が自分の判断通りに予報を出すことを奨励する必要がある（Murphy and Epstein, 1967）。降水確率予報などの評価に用いられているブライアスコア（Brier, 1950）はこの数学的な条件を満たすが、カルバック・ライブラー情報量（Kullback and Leibler, 1951）などはこの条件を満たさない。「不適正な」採点ルールの下では、予報を細工した方が高いスコアを得られることがあるので、スコアを高める努力は、一般にはよりよい予報手法を目指す正直な行為とはいえない。このため、特に地震予知の分野で広く適用されている情報量等を尺度とした「不適正な」予報精度の測定や比較の意義の解釈には、慎重になるべきである。

ところで、ブライアスコアは予報値と実況値の平均自乗誤差の期待値で、現象の有と無に等しい重みを置いている。一方、稀な現象を対象の予報利用者にとっては、予報値が高い確率の場合や現象が実際に発生した場合の予報が重要な情報である。以下に、地震予測のように稀な現象を対象とした確率予報を適切に評価できる採点法を導出する。

予想値（予報者の真の信念に基づく確率）を p 、予報値を f 、実況値 i （現象が発生すれば 1、発生しなければ 0）、基準予報値（例えば統計値に基づく単純な予報値）を c 、スコアを $S_i(f,c)$ と定義する。適正な採点ルールでは、期待値 $E_s(x,c) = pS_1(x,c) + (1-p)S_0(x,c)$ (式 1) について、 $E_s(p,c) > E_s(x,c)$ for all $x \neq p$ (式 2) が必要である。また、稀な現象をより重視する場合でも、基準値から考えられる予報の難易度について公正な条件 $S_0(c,f) \equiv S_1(1-c,1-f)$ (式 3) と、難易度が高い予報ほど高いスコアが得られる条件 $\partial S_1 / \partial f|_{c=const} > 0$, $\partial S_1 / \partial c|_{c=const} < 0$ (式 4) の成立が必要である。

式 1,3 を満たす解は式 (5) を満たす。ここで A は任意の関数である。

$$S_1(f,c) = -(i-f)B'(f,c) + (i-x)B'(x,c) - B(f,c) + B(x,c), \quad A = d^2B/df^2 \quad (\text{式 5})$$

適当な境界値と A の関数形で式 5 を解き、 S_i が不等式 2 と 4 を満たせば、適正な採点ルールを満たし、かつ、都合のよい性質を持つ評価法を得られる。

例えば、基準値予報のスコア $S_i(c,c) \equiv 0$ 、完全な予報のスコア $S_i(i,c) \equiv 1$ の境界条件で解くと、 $A = -2$ で $S_i(f) = 1 - (i-f)^2$ (式 6) を得る。ここで、 $1 - S_i(f)$ はブライアスコアに一致する。

もう一例として、基準となる予報値でのスコア $S_i(c,c) \equiv 0$ 、完全な予報の期待スコア $E_{s,p} \equiv 1$ を境界条件として A が f の 0 次式の解を求めると、 $A = -2 / c(1-c)$ から $S_i(f,c) = \{(i-c)^2 - (i-f)^2\} / c(1-c)$ (式 7) を得られる。これも適正な採点ルールを満たす。 $c = 1/2$ の時、 $(1 - S_i(f))/4$ はブライアスコアであるから、式 7 は基準値予報の選択の任意性に関してブライアスコアを拡張した評価式と解釈してよい。

以上のように、基準予報値を参照した採点式にすることで、地震予測のように稀な現象を対象とした確率予報の適正な採点法の一般解と、特に、稀な現象により重みを置いて適正に評価できるいわゆる拡張ブライアスコアと呼ぶべき解を導いた。

大会では、式の導出の詳細と、これら採点法の具体的な適用に向けての課題も議論したい。なお、津波警報のように現象の有無を対象とする二値予報の一種を適切に評価する採点法は、効用理論から導出し議論済みである（林, 2014, JpGU）。

キーワード: 基準値予報, 地震予測, 拡張ブライアスコア, 情報利得, 確率予報, 適正な採点ルール

Keywords: benchmark forecast, earthquake prediction, extended Brier's score, information gain, probability forecast, proper scoring rule