

## A Geometrical Framework for the Ranked Probability Score

ALLAN H. MURPHY<sup>1</sup>

*Advanced Study Program, National Center for Atmospheric Research,<sup>2</sup> Boulder, Colo. 80303*

AND

CARL-AXEL S. STAËL VON HOLSTEIN

*Stanford Research Institute, SRI-Scandinavia, Stockholm, Sweden*

(Manuscript received 18 June 1974; in revised form 17 October 1974)

### ABSTRACT

A geometrical framework for the representation of *cumulative* forecasts and observations is described. The ranked probability score is shown to be the square of the distance between the points in this framework which represent a cumulative forecast and the relevant cumulative observation. The relationship between this framework and the geometrical framework for the probability score is indicated.

### 1. Introduction

Probability forecasts and observations defined for a variable with  $N$  states can be depicted within the geometrical framework of a regular  $(N-1)$ -dimensional simplex (de Finetti, 1962, 1965; Epstein and Murphy, 1965). Within this framework the so-called Brier, or probability, score ( $PS$ ) (Brier, 1950), a measure of the "accuracy" of probability forecasts (Murphy and Winkler, 1970, 1971), is the square of the (euclidean) distance between the point in the simplex which represents the forecast of concern and the vertex of the simplex which represents the relevant observation. Such a simplex provides a framework for the evaluation of probability forecasts from a decision theoretic, as well as an inferential, point of view (Murphy, 1972).<sup>3</sup>

In contrast to the  $PS$ , the ranked probability score ( $RPS$ ), formulated by Epstein (1969), takes into account the order inherent in *ordered* variables.<sup>4</sup> Since

most meteorological variables with  $N$  states ( $N \geq 3$ ) are ordered variables, the  $RPS$  is expected to receive increasing attention and use in the future.<sup>5</sup> The  $RPS$  can be expressed in a simplified form in terms of *cumulative* forecasts and observations (Murphy, 1971). The purposes of this paper are to demonstrate that these cumulative forecasts and observations can be depicted within the framework of an irregular  $(N-1)$ -dimensional simplex, within which the  $RPS$  is the square of the distance between points representing individual cumulative forecasts and observations, and to describe the relationships between the geometrical frameworks for the  $RPS$  and the  $PS$ . We believe that these geometrical frameworks provide important insight into the nature of the evaluation process as well as into the relationship between the  $PS$  and the  $RPS$ .

The  $PS$  and the  $RPS$  are defined in Section 2, and the geometrical framework for the  $PS$  is briefly described in Section 3. In Section 4, the geometrical framework for the  $RPS$  is presented and the relationships between the frameworks for the  $RPS$  and the  $PS$  are described. For illustrative purposes, the three-state ( $N=3$ ) situation is examined in some detail in both Sections 3 and 4. Section 5 consists of a brief summary and conclusion.

<sup>1</sup> On leave and visiting the International Institute for Applied Systems Analysis, Laxenburg, Austria.

<sup>2</sup> The National Center for Atmospheric Research is sponsored by the National Science Foundation.

<sup>3</sup> Inferential evaluation relates to the "accuracy," "reliability," "skill," etc., of forecasts, while decision theoretic evaluation relates to their "value" to potential users or decision makers. These two aspects of evaluation have often been referred to as empirical (or scientific) and operational (or economic) evaluation, respectively (see Winkler and Murphy, 1968).

<sup>4</sup> Ordered variables are variables defined on at least an ordinal scale (e.g., temperature, precipitation amount, wind speed, cloud amount), while unordered variables are variables defined on at most a nominal scale (e.g., the "present weather" classes). For a discussion of some of the implications of the order inherent in ordered variables for the evaluation of probability forecasts, refer to Murphy (1970) and Staël von Holstein (1970).

<sup>5</sup> The  $RPS$  and the  $PS$  are equivalent in two-state ( $N=2$ ) situations. Until recently, meteorologists have been primarily, if not exclusively, concerned with the evaluation of forecasts of precipitation occurrence (which is a two-state situation). However, a number of experiments and operational programs involving probability forecasts of  $N$ -state ( $N \geq 3$ ) variables have been initiated in recent years (e.g., Klein and Glahn, 1974; Sanders, 1973; Staël von Holstein, 1971).

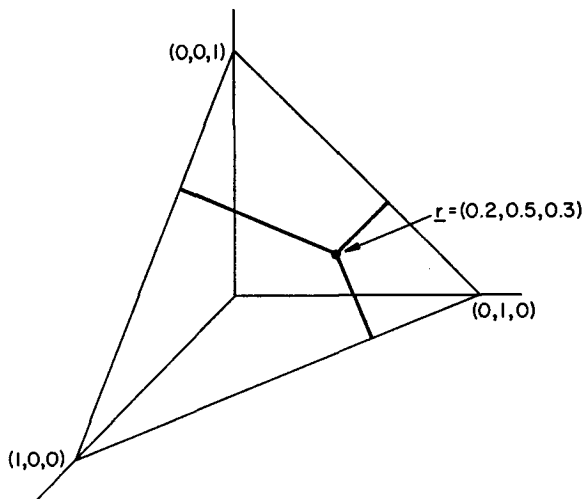


FIG. 1. The geometrical framework,  $\pi^3$ , for the representation of forecasts and observations in three-state situations.

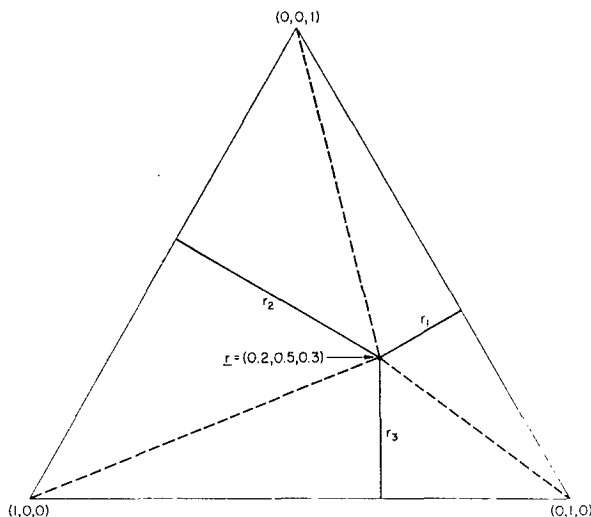


FIG. 2. The forecast  $r = (0.2, 0.5, 0.3)$  depicted within the framework of the equilateral triangle  $\pi^3$ .

2. The probability score and the ranked probability score

a. Forecasts and observations

In this paper, we assume that the range of values of the variable of concern has been divided into a set of  $N$  mutually exclusive and collectively exhaustive states  $\{s_1, \dots, s_N\}$ . A probability forecast is then a (row) vector  $r = (r_1, \dots, r_N)$ , where  $r_n$  is the probability assigned to state  $s_n$  ( $r_n \geq 0, \sum_n r_n = 1; n = 1, \dots, N$ ). An observation is a (row) vector  $d = (d_1, \dots, d_N)$ , where  $d_n$  equals one if state  $s_n$  obtains and zero otherwise ( $n = 1, \dots, N$ ). The  $N$  possible categorical forecasts are denoted by the vectors  $r_n$ , in which  $r_n = 1$  and  $r_m = 0$  for all  $m \neq n$  ( $m, n = 1, \dots, N$ ). Thus,  $d_n = r_n$ .

b. The PS

The PS for the forecast  $r$  and an observation  $d$  is  $PS(r, d)$ , where

$$PS(r, d) = (r - d)(r - d)' = \sum_{n=1}^N (r_n - d_n)^2. \quad (1)$$

If state  $s_j$  obtains (i.e., if  $d_j = 1$ ), then  $PS(r, d)$  becomes  $PS_j(r)$ , where

$$PS_j(r) = 1 - 2r_j + \sum_{n=1}^N r_n^2.$$

The range of the PS is the closed interval  $[0, 2]$  and the PS has a negative orientation (i.e., the smaller the score the better the score).

c. Cumulative forecasts and observations

We define a cumulative (probability) forecast as a

(row) vector  $R = (R_1, \dots, R_N)$ , where

$$R_n = \sum_{i=1}^n r_i \quad (n = 1, \dots, N), \quad (2)$$

and a cumulative observation as a (row) vector  $D = (D_1, \dots, D_N)$ , where

$$D_n = \sum_{i=1}^n d_i \quad (n = 1, \dots, N). \quad (3)$$

In this regard,  $R_N = 1$  for all cumulative forecasts and  $D_N = 1$  for all cumulative observations. Note that, since  $r_n = R_n - R_{n-1}$  ( $R_0 = 0$ ) and  $d_n = D_n - D_{n-1}$  ( $D_0 = 0$ ), a

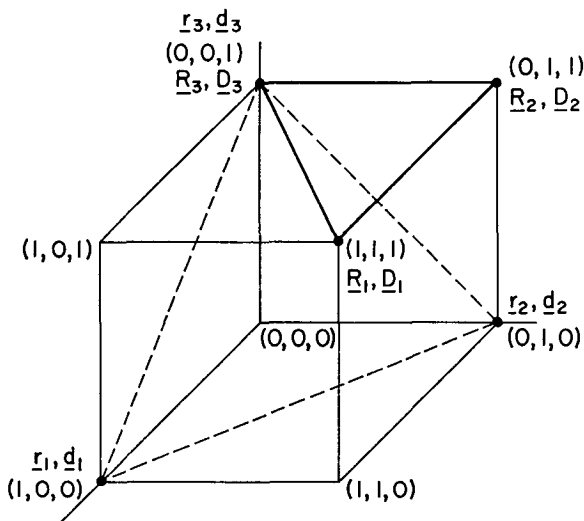


FIG. 3. The geometrical framework for the representation of cumulative forecasts and observations (heavy solid lines),  $\Pi^3$ , and of forecasts and observations (dashed lines),  $\pi^3$ .

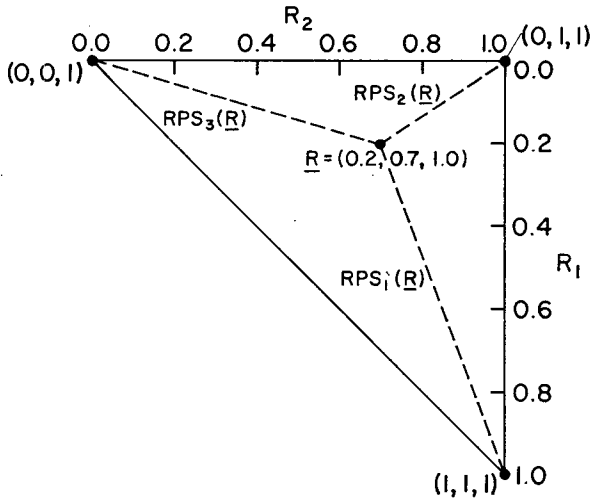


FIG. 4. The cumulative forecast  $\mathbf{R}=(0.2, 0.7, 1.0)$  depicted within the framework of the right, isosceles triangle  $\Pi^3$  in a two-dimensional Cartesian coordinate system.

one-to-one relationship exists between forecasts and cumulative forecasts (as well as between observations and cumulative observations).

We denote the  $N$  possible categorical forecasts by the vectors  $\mathbf{R}_n$  ( $n=1, \dots, N$ ), in which  $R_m=0$  for all  $m < n$  and  $R_m=1$  for all  $m \geq n$ . Similarly, we denote the  $N$  possible cumulative observations by the vectors  $\mathbf{D}_n$  ( $n=1, \dots, N$ ), in which  $D_m=0$  for all  $m < n$  and  $D_m=1$  for all  $m \geq n$ . Thus,  $\mathbf{D}_n = \mathbf{R}_n$  for all  $n$ .

d. The RPS

The RPS for the cumulative forecast  $\mathbf{R}$  and a cumulative observation  $\mathbf{D}$  is  $RPS(\mathbf{R}, \mathbf{D})$ , where

$$RPS(\mathbf{R}, \mathbf{D}) = (\mathbf{R} - \mathbf{D})(\mathbf{R} - \mathbf{D})' = \sum_{n=1}^N (R_n - D_n)^2 \quad (4)$$

(see Murphy, 1971). If state  $s_j$  obtains, then  $D_n=1$  for all  $n \geq j$  and  $RPS(\mathbf{R}, \mathbf{D})$  becomes  $RPS_j(\mathbf{R})$ , where

$$\begin{aligned} RPS_j(\mathbf{R}) &= \sum_{n=1}^{j-1} R_n^2 + \sum_{n=j}^N (R_n - 1)^2 \\ &= (N - j + 1) - 2\left(\sum_{n=j}^N R_n\right) + \sum_{n=1}^N R_n^2. \end{aligned}$$

The range of the RPS is the closed interval  $[0, N-1]$  and the RPS has a negative orientation.

3. The geometrical framework for the probability score

Denote the set of all possible forecasts by  $\pi^N$ , where

$$\pi^N = \{ (r_1, \dots, r_N) \mid r_n \geq 0, \sum_{n=1}^N r_n = 1; n = 1, \dots, N \}.$$

The set  $\pi^N$  represents a regular  $(N-1)$ -dimensional simplex. This set is a unit line segment,  $\pi^2$ , in two-state ( $N=2$ ) situations, an equilateral triangle,  $\pi^3$ , in three-state ( $N=3$ ) situations, and a regular tetrahedron,  $\pi^4$ , in four-state ( $N=4$ ) situations. The  $N$  vertices of the simplex  $\pi^N$  represent the  $N$  possible observations  $\mathbf{d}_n$  as well as the  $N$  possible categorical forecasts  $\mathbf{r}_n$  ( $n=1, \dots, N$ ).

The equilateral triangle  $\pi^3$  is shown in Fig. 1 in a three-dimensional cartesian coordinate system, together with the forecast  $\mathbf{r}=(0.2, 0.5, 0.3)$ . The distances between the point representing the forecast  $\mathbf{r}$  and the sides of the triangle, identified by heavy solid lines, are proportional to the three components of the forecast [the ratio between the distance and the probability is  $(3/2)^{1/2}$ ]. The triangle in Fig. 2 has a unit height and is an equivalent two-dimensional representation of  $\pi^3$ , in which the distances between the point representing

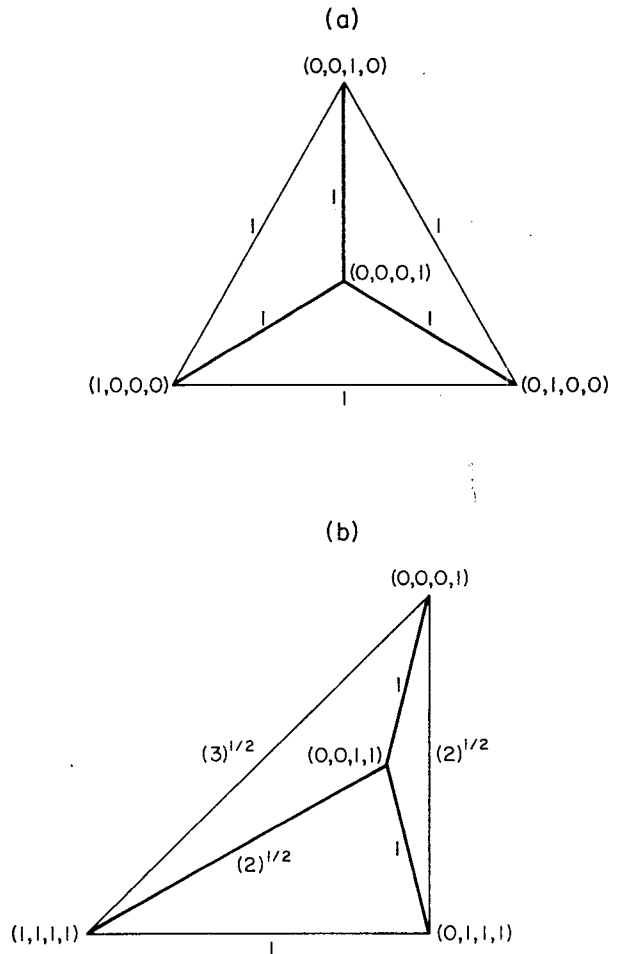


FIG. 5. (a) The regular tetrahedron  $\pi^4$ , the framework for the representation of forecasts and observations in four-state situations. (b) The irregular tetrahedron  $\Pi^4$ , the framework for the representation of cumulative forecasts and observations in four-state situations. The numbers adjacent to the edges of the faces of the tetrahedrons indicate the lengths of the respective edges.

the forecast and the sides of the triangle are equal to the three probabilities which constitute the forecast.

The (euclidean) distance between two points  $\mathbf{r} = (r_1, \dots, r_N)$  and  $\mathbf{r}' = (r'_1, \dots, r'_N)$  is  $D(\mathbf{r}, \mathbf{r}')$ , where

$$D(\mathbf{r}, \mathbf{r}') = \left[ \sum_{n=1}^N (r_n - r'_n)^2 \right]^{1/2}.$$

Thus, the distance between a forecast  $\mathbf{r} = (r_1, \dots, r_N)$  and the relevant observation  $\mathbf{d} = (d_1, \dots, d_N)$  is  $D(\mathbf{r}, \mathbf{d})$ , where

$$D(\mathbf{r}, \mathbf{d}) = \left[ \sum_{n=1}^N (r_n - d_n)^2 \right]^{1/2}. \tag{5}$$

Note, from (1) and (5), that

$$PS(\mathbf{r}, \mathbf{d}) = [D(\mathbf{r}, \mathbf{d})]^2. \tag{6}$$

That is, within this framework, the *PS* for an individual forecast is the square of the distance between the forecast and the relevant observation. The dashed lines in Fig. 2 represent the distances between the forecast  $\mathbf{r} = (0.2, 0.5, 0.3)$  and the three possible observations,  $\mathbf{d}_1 = (1, 0, 0)$ ,  $\mathbf{d}_2 = (0, 1, 0)$ , and  $\mathbf{d}_3 = (0, 0, 1)$ . The squares of these distances are  $PS_1(\mathbf{r})$ ,  $PS_2(\mathbf{r})$ , and  $PS_3(\mathbf{r})$ , respectively.

#### 4. A geometrical framework for the ranked probability score

Denote the set of all possible cumulative forecasts by  $\Pi^N$ . It follows from Eq. (2) that  $\Pi^N$  is an irregular  $(N-1)$ -dimensional simplex defined by

$$\Pi^N = \{ (R_1, \dots, R_N) \mid R_1 \geq 0, R_n \geq R_m \text{ for } 1 \leq m < n \leq N, R_N = 1 \}.$$

A cumulative forecast  $\mathbf{R} = (R_1, \dots, R_N)$  corresponds to a point in this simplex (and vice versa), and the  $N$  vertices of the simplex represent the  $N$  cumulative observations  $\mathbf{D}_n$  and the  $N$  categorical forecasts  $\mathbf{R}_n$  ( $n = 1, \dots, N$ ).

The simplex for three-state ( $N=3$ ) situations,  $\Pi^3$ , is depicted in a Cartesian coordinate system in Fig. 3 as a right, isosceles triangle (heavy solid lines) on the top face of the unit cube. This triangle is then a framework for the representation of cumulative forecasts and observations in three-state situations. The regular two-dimensional simplex, the equilateral triangle  $\pi^3$ , which represents the set of all forecasts  $\mathbf{r}$ , is shown in dashed lines. Since a one-to-one correspondence exists between a forecast  $\mathbf{r}$  and a cumulative forecast  $\mathbf{R}$ , a one-to-one correspondence exists between points in the two triangles.

The (euclidean) distance between a cumulative forecast  $\mathbf{R}$  and the relevant observation  $\mathbf{D}$  in this

framework is  $D(\mathbf{R}, \mathbf{D})$ , where

$$D(\mathbf{R}, \mathbf{D}) = \left[ \sum_{n=1}^N (R_n - D_n)^2 \right]^{1/2}. \tag{7}$$

Thus, from (4) and (7),

$$RPS(\mathbf{R}, \mathbf{D}) = [D(\mathbf{R}, \mathbf{D})]^2, \tag{8}$$

i.e., the *RPS* is the square of the distance between  $\mathbf{R}$  and  $\mathbf{D}$ . The cumulative forecast  $\mathbf{R} = (0.2, 0.7, 1.0)$ , corresponding to the forecast  $\mathbf{r} = (0.2, 0.5, 0.3)$ , is depicted within the framework of the triangle  $\Pi^3$  in a two-dimensional cartesian coordinate system in Fig. 4. As previously indicated, the vertices of this triangle represent the cumulative observations  $\mathbf{D}_n$  and the categorical cumulative forecasts  $\mathbf{R}_n$  ( $n = 1, 2, 3$ ). The dashed lines in Fig. 4 represent the distances between the cumulative forecast  $\mathbf{R}$  and the cumulative observations  $\mathbf{D}_1 = (1, 1, 1)$ ,  $\mathbf{D}_2 = (0, 1, 1)$ , and  $\mathbf{D}_3 = (0, 0, 1)$ . The squares of these distances are  $RPS_1(\mathbf{R})$ ,  $RPS_2(\mathbf{R})$ , and  $RPS_3(\mathbf{R})$ , respectively.

For illustrative and comparative purposes, the frameworks for the representation of forecasts and observations and cumulative forecasts and observations in four-state ( $N=4$ ) situations are depicted in Fig. 5. The former is a regular tetrahedron, while the latter is an irregular tetrahedron. The numbers adjacent to the edges of the faces of the tetrahedrons indicate the lengths of the respective edges. Note that the faces of the regular tetrahedron are equilateral triangles, while the faces of the irregular tetrahedron are right triangles.

#### 5. Summary and conclusion

In this paper, we have 1) described a geometrical framework, an irregular  $(N-1)$ -dimensional simplex, for the representation of cumulative forecasts and observations; 2) demonstrated that, within this framework, the ranked probability score (*RPS*) is the square of the (euclidean) distance between the point in the simplex which represents the forecast of concern and the vertex of the simplex which represents the relevant observation; and 3) identified the relationships between this framework for the *RPS* and the geometrical framework for the probability score (*PS*).

In a forthcoming paper we will show that the *RPS* and the *PS* are both members of a family of quadratic scoring rules. In addition to examining the nature and properties of the members of this family of scoring rules, the paper will consider their representation in geometric terms.

#### REFERENCES

Brier, G. W., 1950: Verification of forecasts expressed in terms of probability. *Mon. Wea. Rev.*, **78**, 1-3.  
 de Finetti, B., 1962: Does it make sense to speak of "good probability appraisers"? *The Scientist Speculates: An Anthology of*

- Partly-Baked Ideas* (I. J. Good, General Editor). New York, Basic Books, Inc., pp. 357-364.
- , 1965: Methods for discriminating levels of partial knowledge concerning a test item. *Brit. J. Math. and Stat. Psychol.*, **18**, 87-123.
- Epstein, E. S., 1969: A scoring system for probability forecasts of ranked categories. *J. Appl. Meteor.*, **8**, 985-987.
- , and A. H. Murphy, 1965: A note on the attributes of probabilistic predictions and the probability score. *J. Appl. Meteor.*, **4**, 297-299.
- Klein, W. H., and H. R. Glahn, 1974: Forecasting local weather by means of model output statistics. *Bull. Amer. Meteor. Soc.*, **55**, 1217-1227.
- Murphy, A. H., 1970: The ranked probability score and the probability score: a comparison. *Mon. Wea. Rev.*, **98**, 917-924.
- , 1971: A note on the ranked probability score. *J. Appl. Meteor.*, **10**, 155-156.
- , 1972: Ordinal relationships between measures of the "accuracy" and "value" of probability forecasts: preliminary results. *Tellus*, **24**, 531-542.
- , and R. L. Winkler, 1970: Scoring rules in probability assessment and evaluation. *Acta Psychol.*, **34**, 273-286.
- , and —, 1971: Forecasters and probability forecasts: some current problems. *Bull. Amer. Meteor. Soc.*, **52**, 239-247.
- Sanders, F., 1973: Skill in forecasting daily temperature and precipitation: some experimental results. *Bull. Amer. Meteor. Soc.*, **54**, 1171-1179.
- Staël von Holstein, C.-A. S., 1970: A family of strictly proper scoring rules which are sensitive to distance. *J. Appl. Meteor.*, **9**, 360-364.
- , 1971: An experiment in probabilistic weather forecasting. *J. Appl. Meteor.*, **10**, 635-645.
- Winkler, R. L., and A. H. Murphy, 1968: "Good" probability assessors. *J. Appl. Meteor.*, **7**, 751-758.