A Note on the Ranked Probability Score

ALLAN H. MURPHY

Dept. of Meteorology and Oceanography, The University of Michigan, Ann Arbor

4 August 1970 and 29 September 1970

The ranked probability score (RPS) (Epstein, 1969) is both strictly proper (Murphy, 1969) and sensitive to distance (Staël von Holstein, 1970), and, as a result, the RPS appears to be a particularly appropriate scoring rule for the evaluation of probability forecasts of ordered variables (Murphy, 1970). Recently, F. Sanders, in a personal communication, has described a scoring rule $T$ (say), where

$$T(r,d) = (1/K) \sum_{i=1}^{K} \left( \sum_{k=1}^{K} r_k - \sum_{k=1}^{K} d_k \right)^2,$$

which he believes also possesses these properties. The purposes of this note are to demonstrate: 1) that $T$ is equivalent, i.e., linearly related, to the RPS, which means that $T$ is both strictly proper and sensitive to distance; and 2) that, as a result, the RPS, as well as the probability score (PS) (Brier, 1950), are related to the difference between the forecast and the observed probability distributions, which provides an additional measure of support for the use of the RPS as an evaluation measure.

When class $j$ occurs, $T(r,d)$ becomes $T_j(r)$, where

$$T_j(r) = (1/K) \left[ \sum_{i=1}^{j-1} \left( \sum_{k=1}^{i} r_k \right)^2 + \sum_{i=j}^{K} \left( \sum_{k=1}^{i} r_k - 1 \right)^2 \right],$$

or

$$T_j(r) = (1/K) \left[ \sum_{i=1}^{j-1} \left( \sum_{k=1}^{i} r_k \right)^2 + \sum_{i=j}^{K-1} \left( \sum_{k=i+1}^{K} r_k \right)^2 \right].$$

Now,

$$\text{RPS}_j(r) = 1 - \left[ 1/(K-1) \right] \left[ \sum_{i=1}^{j-1} \left( \sum_{k=1}^{i} r_k \right)^2 + \sum_{i=j}^{K-1} \left( \sum_{k=i+1}^{K} r_k \right)^2 \right].$$

The vector $r = (r_1, \ldots, r_K)$, where $r_k \geq 0$ and $\sum_{k=1}^{K} r_k = 1$, represents the forecast and the vector $d = (d_1, \ldots, d_K)$, where $d_k$ equals one if class $k$ occurs and zero otherwise, represents the observation.
(Murphy, 1970, p. 924). Thus,
\[ T_i(x) = \left[ (K - 1)/K \right] \left[ 1 - \text{RPS}_i(x) \right]. \]
Since a scoring rule which is equivalent to a scoring rule which is strictly proper and sensitive to distance is itself strictly proper (Winkler and Murphy, 1968) and sensitive to distance (Staël von Holstein, 1970), \( T \) is both strictly proper and sensitive to distance.

The standard expression for the PS is
\[ \text{PS}(r, d) = \sum_{i=1}^{K} (r_i - d_i)^2. \]

Let \( r(x_i) \) and \( d(x_i) \) denote the (probability) mass functions [the discrete counterparts of (probability) density functions] for the forecast and the observation, respectively, where \( x_i \) represents the \( i \)th class of the variable of concern \( x \). Then,
\[ \text{PS}(r, d) = \sum_{i=1}^{K} \left[ r(x_i) - d(x_i) \right]^2. \]
Thus, the PS represents the sum of the squares of the differences between the forecast and the observed (probability) mass functions. Now, since \( T \) is equivalent to the RPS, the RPS can be expressed simply as
\[ \text{RPS}(r, d) = \sum_{i=1}^{K} \left( \sum_{k=1}^{i} r(x_k) - \sum_{k=1}^{i} d(x_k) \right)^2, \]
or, in terms of (probability) mass functions, as
\[ \text{RPS}(r, d) = \sum_{i=1}^{K} \left[ \left( \sum_{k=1}^{i} r(x_k) \right) - \left( \sum_{k=1}^{i} d(x_k) \right) \right]^2. \]

Note that the
\[ \sum_{k=1}^{i} r(x_k) \]
and the
\[ \sum_{k=1}^{i} d(x_k), \]
i = 1, \ldots, \( K \), describe the forecast and the observed cumulative (probability) mass functions, respectively [the discrete counterparts of (probability) distribution functions]. Let \( R(x_i) \), where
\[ R(x_i) = \sum_{k=1}^{i} r(x_k), \]
and \( D(x_i) \), where
\[ D(x_i) = \sum_{k=1}^{i} d(x_k), \]
denote these functions. Then,
\[ \text{RPS}(r, d) = \sum_{i=1}^{K} \left[ R(x_i) - D(x_i) \right]^2. \]

Thus, the RPS represents the sum of the squares of the differences between the forecast and the observed cumulative (probability) mass functions. Therefore, the RPS, as well as the PS, are concerned with the (square of the) differences between the forecast and the observed probability distributions. The relationship between the RPS and these probability distributions provides 1) a convenient means of comparing the RPS and the PS, and 2) an additional measure of support for the use of the RPS as an evaluation measure for forecasts of ordered variables.

REFERENCES