Reprinted from Preprints Ninth Conferer on Probability and Statistics in Atmospheric Sciences, October 9-11, 1985, Virginia Beach, Virginia, American Meteorological Society.

A METHOD TO ESTIMATE THE CONTINUOUS RANKED PROBABILITY SCORE

David A. Unger
Techniques Development Laboratory
Office of Systems Development
National Weather Service, NOAA
Silver Spring, Maryland 20910

1. INTRODUCTION

An estimate of the uncertainty in a prediction can significantly increase its value. This is particularly true when the forecast accuracy varies significantly from prediction to prediction, as is the case for most meteorological forecasts. Certain synoptic situations are very difficult to forecast while others are quite easy to predict. For example, the forecast of elements near the expected position of a frontal boundary may be considerably less accurate than average. A statement of the forecaster's confidence in the prediction can alert the user to those forecasts which are not expected to perform as well as others. This can have a significant effect on decisions which are made on the basis of the information.

Uncertainty is often expressed in meteorology by probability forecasts. Probability forecasts express the chance of the occurrence of a specific event. The probability of measurable precipitation forecasts, for example, have been issued by the National Weather Service since 1965. A forecast of a continuous variable is usually expressed in probability form by specification of the probability that the variable will fall within a specified range of values. The range of assumable values for a continuous variable can be divided into a series of intervals, and the forecast probability that the observation will fall in each can be used to express the probability distribution of the variable.

Many scoring rules have been developed for the verification of probability forecasts. Probabilistic predictions in meteorology have almost exclusively been verified by quadratic probability scoring rules. The Ranked Probability Score (RPS) (Epstein, 1969) has been used for the verification of probability forecasts for a series of ordinal categories. The RPS can be used for forecasts of continuous variables which are expressed by the probability that the observation will occur in each of a series of ordered mutually exclusive, collectively exhaustive intervals.

The RPS verifies the probability statements for the intervals that are used to make the forecast. While this is advantageous when the intervals themselves have special significance to the user, for other applications it may be desirable to use a probability scoring rule which verifies a continuous probability distribution function. Such a score would depend only on the forecast distribution and the observation, and not on the intervals used to make the prediction.

A quadratic scoring rule for the verification of a continuous probability distribution function known as the Continuous Ranked Probability Score (CRPS) was developed in the 1970's (Brown, 1974; Matheson and Winkler, 1976). To the author's knowledge, the CRPS has rarely been applied. This may be because a continuous forecast distribution function is difficult to specify and disseminate. The score also requires integration to be performed on the forecast distribution, which also might have deterred its use in the past.

The vastly increased amount of computing power available today provides the opportunity to issue forecasts in greater detail than ever before. This makes it possible to specify and analyze, on a routine basis, more information relating to a forecast, such as the expression of a forecast probability distribution for a variable. Verification may be required not only to determine the accuracy of the forecasts themselves, but also to help determine the value of the probability information.

The CRPS will be examined here to demonstrate how it can be used in practical forecasting situations. The score will be partitioned into components which can help interpret the score. The CRPS will help evaluate forecasts of continuous variables in a continuous, rather than a discrete, framework. This can be valuable for the analysis of forecasts when a discrete representation is inconvenient.

A method to approximate the CRPS for probability forecasts for discrete intervals shall also be developed. This score is, to a large extent, independent of the intervals used to make the forecast, so that forecasts can be evaluated in terms of the continuous distributions represented, rather than for the specific intervals for which forecasts are made. This can help compare forecasts of the same variable made from different sets of intervals. The RPS cannot be used to compare forecasts made from different intervals.

2. VARIABLE DEFINITIONS

The following definitions shall be used throughout the paper. The variable to be forecasted, X, is represented by the set of numbers, x. X might be the 12-h forecast of temperature at a location, for example, and x be the range in °C. T represents the verifying observation. R(x) is the forecast cumulative probability distribution function ($R(x) = P(T \leq x)$), and r(x) is the associated density function. K is the median of the forecast distribution. For convenience of representation of the scores, the function, D(x), is defined such that: $D(x) = 0$ when $x < T$, and $D(x) = 1$ when $x \geq T$.

For discrete representation of the forecasts, x is divided into N intervals defined by the breakpoints x_i, $i = 1, 2, \ldots, N-1$. R_i represents the forecast probability $T \leq x_i$. The forecast probability that $x_{i-1} < T \leq x_i$ is represented by r_i, $r_1 = R_1$, and $r_N = 1-R_{N-1}$. x_k represents the lowest value of x_i for which $R_i \geq .5$; in other words, $x_{k-1} < K < x_k$. The discrete representation of D(x) is $D_i = D(x_i)$.

When the intervals into which x is divided extend well above and below the possible range of X, a value, x_L is defined to be the highest value for which both R_i and $D_i = 0$. The value x_M is defined as the lowest value for which both R_i and D_i are 1.0.

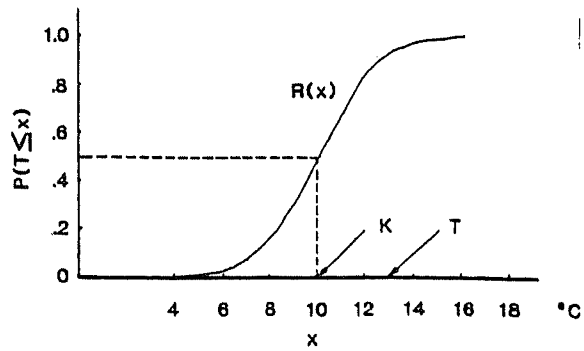
Table 1 summarizes these definitions.

Table 1. Summary of definitions used in this paper.

Variable	Meaning
X	The variable to be forecast
x	The domain of X
r(x)	The forecast probability density function
R(x)	$P(T \leq x)$
T	The verifying observation
K	The median of the forecast distribution: $R(K) = .5$
D(x)	0 for $x < T$ 1 for $x \geq T$.
R_i	$R(x_i)$
D_i	$D(x_i)$
r_i	$R_i - R_{i-1}$
N	The total number of intervals

The discussions presented here shall represent a single forecast event unless otherwise noted. The performance on a series of forecast events can be represented by a simple average of the scores presented.

An example to illustrate a probability forecast of a continuous variable is shown in Fig. 1. The forecast and observation are arbitrarily chosen for the purpose of illustration. The forecast distribution, R(x), might be a 12-h prediction for temperature at a particular location. R(x) is from a Gaussian distribution with a standard deviation of 2.0°C and a mean of



i	1	2	3	4	5	6	7
x_i (°C)	4	7	9	11	12	14	16
R_i	0	.07	.31	.69	.84	.98	1.0

K=10°C T=13°C

Figure 1. Example forecast probability distribution, R(x) for a hypothetical temperature forecast. See text for explanation.

10.0°C. The forecast has been expressed in discrete form at the specific points, x_i. The observation, T, can fall into any one of eight intervals, (N = 8). For this example, it was assumed to be 13.0°C.

3. THE RANKED PROBABILITY SCORE

The RPS is a probability scoring rule for prediction made for ordinal categories. Early references (e.g., Epstein, 1969) display the form of the score which is computed by the r_i's. A much simplified formulation was presented by Murphy (1971) in which the calculations are performed on the cumulative probabilities, R_i's.

In its scalar form, the RPS can be represented by:

$$RPS_N = \sum_{i=1}^{N-1} (R_i - D_i)^2, \qquad (1)$$

where the subscript, N, on the left side of the equation denotes the number of intervals from which the score was computed.

The RPS is a probability score which not only measures the accuracy of the probability statements, but also is sensitive to distance. A better score will be attained when the observed event occurs in a category which is close to the categories in which the forecast probabilities are concentrated.

The score has a negative orientation, which means that a lower value signifies a better forecast. The RPS takes on a minimum value of zero for a perfect forecast, and has a maximum possible value of N-1. It is a strictly proper measure of the probability forecasts--one in which the best score is obtained when the forecaster's stated probabilities conform to his/her true beliefs (Murphy, 1969). This is an extremely important quality for a probability scoring rule, because a probability forecast must ultimately be judged in terms of the correspondence between the forecast and observed probabilities. Great care must be exercised in the use of a scoring rule to ensure that the forecaster will be encouraged to state the frequencies that he/she actually believes will occur; otherwise, the scoring rule used will actually contribute to an inaccurate forecast.

Not only should a scoring rule be proper, but it also should reward forecasts which are highly certain. The scores for a reliable forecast which have a high degree of certainty (probabilities close to either 0 or 1) should yield better scores than those which have probabilities near .5. This feature will encourage the forecaster to forecast with as little uncertainty as possible. The RPS also has this characteristic.

The expected value of the RPS can be computed under the assumption that the forecast probabilities are correct. In that case, the observation will fall in the ith category with a frequency of r_i. If $RPS_N(i)$ represents the value of RPS_N when the observation occurs in the ith category, then,

$$E(RPS_N) = \sum_{i=1}^{N-1} r_i RPS_N(i) \qquad (2)$$

(Murphy and Daan, 1985). This measure shall be referred to as the expected score.

The RPS measures the accuracy for forecasts at the specific breakpoints for which forecasts are available. The score is highly dependent on the specific number and location of the breakpoints used. This is advantageous when the forecast at each breakpoint is considered equally as important as another, or, when the breakpoints themselves have significance in terms of economic impact.

On the other hand, two forecasts with identical distributions and verifying observations will have different RPS scores for a different set of intervals, even if the exterior breakpoints, x_1 and x_{N-1} (those that determine the range of x for which forecasts are made) are the same. Since the breakpoints are often arbitrarily chosen, frequently to represent easily remembered numbers, it would be advantageous, in some cases, to obtain a measure which represents the forecast in terms of a continuous probability distribution function rather than at specific points.

The RPS can be computed from the discrete probabilities shown in Fig. 1. For that example, the RPS was computed to be 1.28 and its expected value to be .65.

4. THE CONTINUOUS RANKED PROBABILITY SCORE

A quadratic probability scoring rule for continuous functions was introduced by Brown (1974). A more general form, independently developed and reported by Matheson and Winkler (1976), shall be used in this paper.

In its general form, the CRPS is represented by

$$CRPS = \int_{-\infty}^{\infty} (R(x) - D(x))^2 dG(x). \qquad (3)$$

G(x) is a function which can be used to weight the forecast more in certain ranges than in others. This weighting function will be discussed briefly in a later section.

The CRPS is negatively oriented with a minimum value of 0 for a perfect forecast, and no upper bound. Its units are the same as for G(x). Matheson and Winkler showed this score to be strictly proper for continuous distributions. That is, if the forecaster believes that the probability distribution of a variable is P(x), then the lowest (best) CRPS for that case will be obtained when the stated probability distribution, R(x), is everywhere equal to P(x).

For the current discussion, G(x) shall be selected to give a score which will evaluate a forecast made with absolute certainty by the absolute error in x. For this purpose, G(x) = x, and Eq. (3) can be stated as,

$$CRPS = \int_{-\infty}^{\infty} (R(x) - D(x))^2 dx. \qquad (4)$$

It is convenient to express a verification score in terms which are easy to comprehend. This not only helps in the interpretation of results, but also can help the forecaster construct a forecast which will provide a better score, and, by virtue of its strictly proper property, a better forecast. If a function, Q(x) is defined such that:

$$Q(x) = R(x) \text{ for } x \leq K$$
$$Q(x) = (1-R(x)) \text{ for } x > K; \; 0 \leq Q(x) \leq .5.$$

Then, Eq. (4) can be presented in expanded form as,

$$CRPS = \int_{-\infty}^{T} R^2(x)dx + \int_{T}^{\infty} (R(x)-1)^2 dx.$$

For T < K,

$$CRPS = \int_{-\infty}^{T} Q^2(x)dx + \int_{T}^{K} (1-R(x))^2 dx + \int_{K}^{\infty} (-Q(x))^2 dx$$
$$= \int_{-\infty}^{\infty} Q^2(x)dx + \int_{T}^{K} [-Q^2(x) + (1-R(x))^2] dx.$$

Recall:

$$R(x) = Q(x) \text{ for } T \leq x \leq K$$

so, with simplification:

$$CRPS = \int_{-\infty}^{\infty} Q^2(x)dx + \int_{T}^{K} dx - 2\int_{T}^{K} Q(x)dx.$$

The same procedure can be followed when $T \geq K$ if the direction of integration is reversed. The equations obtained for T < K and $T \geq K$ can be summarized as:

$$CRPS = \int_{-\infty}^{\infty} Q^2(x)dx + |K-T| - 2\left|\int_{T}^{K} Q(x)dx\right| .$$

or,

$$CRPS = S_0 + A - W. \quad (5)$$

The first component of the score,

$$S_0 = \int_{-\infty}^{\infty} Q^2(x)dx,$$

is a measure of the degree of uncertainty which is expressed in the forecast, and is independent of the verifying observation. $S_0 = 0$ for a forecast made with absolute certainty and increases as the degree of uncertainty increases. The forecaster is rewarded for expressing as little uncertainty as possible. The lowest CRPS that can be attained for any given forecast distribution occurs when T = K, where CRPS = S_0.

The term, A, is the absolute difference of the verifying observation from the median of the forecast distribution, $A = |K-T|$. This term measures the accuracy of the forecast independently of the uncertainty estimate.

The remaining term, $W = 2\left|\int_T^K Q(x)dx\right|$, is difficult to interpret, since it depends on the forecast probability only in the interval between K and T. It can be regarded as a measure of the forecasted likelihood of the departure of the verifying observation from the median. When departures are considered likely (the forecast probability that the observation will fall between T and K, is small), W is close to A. When the forecast distribution in the range between T and K indicates that the verifying observation is expected to occur much closer to the median than it actually occurred, W << A. A larger value for this term represents a better forecast. Note that $0 < W \leq A$, so that $(A-W) \geq 0$.

The CRPS penalizes the forecaster according to the amount of uncertainty expressed in the prediction, and by the distance of the verifying observation from the forecast median of the distribution. It rewards the forecaster when departures from the median occur in a region in which forecast distribution indicated it to be likely. The score's reward and penalties are proportioned so that the total score is strictly proper.

From Matheson and Winkler (1976), the expected score is

$$E(CRPS) = \int_{-\infty}^{\infty} R(x)\ (1-R(x))\ dx. \quad (6)$$

The CRPS for a series of forecasts can be combined to yield a measure of average forecast performance. If $CRPS_j$ represents the score on the jth occasion, and the overbar represents the average value, then,

$$\overline{CRPS} = \frac{1}{J}\sum_{j=1}^{J} CRPS_j = \overline{S}_0 + \overline{A} + \overline{W}, \text{ and}$$

$$\overline{E(CRPS)} = \frac{1}{J}\sum_{j=1}^{J} E(CRPS)_j.$$

5. APPROXIMATION OF THE CRPS

In practice, the CRPS will almost always be estimated by numerical integration. Most probability distribution functions are difficult to integrate analytically. Continuous probability distribution functions are also difficult for a forecaster to specify unless he/she is restricted to very simple ones. Even when the forecast is specified by a continuous distribution which can be integrated analytically, the CRPS still might have to be approximated by numerical integration if the reporting interval for the observation is large when compared to the forecast distribution. For example, for a temperature reported in whole degrees C, a verifying observation of 10°C might represent any value $9.5 \leq x < 10.5$°C. The CRPS then cannot be computed as a continuous function because T is not known precisely. It is unacceptable to assume that T occurs at some point within the interval, such as the midpoint, because this will encourage the forecaster to adjust the stated forecast probability distribution to conform to the assumption used to obtain T. The forecaster would, then, predict the probability distribution to be a step function, regardless of what he/she actually believes. This is a form of hedging.

Eq. (4) can be approximated from the values of D(x) and R(x) at discrete points by numerical integration. The trapezoidal rule shall be used here for integration. This rule states that, any function, f, can be approximated in the increment between points a and b, by

$$\int_a^b f(x)dx = (b-a)\ \frac{f(a) + f(b)}{2}.$$

For a continuous probability distribution function, and an infinite series of regularly spaced intervals which cover the entire possible range of X and r(x), the CRPS can be approximated by the score,

$$C = \sum_{i=L}^{M} \Delta x_i\ \frac{(R_{i-1} - D_{i-1})^2 + (R_i - D_i)^2}{2}$$

where

$$\Delta x_i = x_i - x_{i-1}.$$

The values of the subscripts, L and M are defined such that $(R(x) - D(x))^2 = 0$ for $x \leq x_L$ and $x \geq x_M$.

The terms in this equation can be sorted to form the equation,

$$C = \Delta x \sum_{i=L}^{M} (R_i - D_i)^2, \quad (7)$$

where $\Delta x = \Delta x_i$. Some rounding of the probabilities, R_i may be needed to limit the number of categories required for integration, i.e., set $R_i = 0$ or 1 when it becomes very close to those values.

Eq. (7), is identical to the RPS computed from the same categories except for the constant, Δx.

The integration need not be restricted to regular intervals, nor does it require that the complete distribution be forecasted. This indicates that the breakpoints for which forecasts were made need not cover the entire range of possible values of X. ((R_1-D_1) and $(R_{N-1}-D_{N-1})$ need not be 0.)

For a set of N intervals defined by the points x_i, i=1 ... N-1, the score can be written as:

$$C_N = \frac{\Delta x_2}{2}(R_1-D_1)^2 + \sum_{i=2}^{N-2}\frac{(\Delta x_i + \Delta x_{i+1})}{2}(R_i-D_i)^2 + \frac{\Delta x_{n-1}}{2}(R_{N-1}-D_{N-1})^2$$

or,

$$C_N = \frac{x_2 - x_1}{2}(R_1-D_1)^2 + \sum_{i=2}^{N-2}\frac{x_{i+1} - x_{i-1}}{2}(R_i-D_i)^2 + \frac{x_{N-1} - x_{N-2}}{2}(R_{N-1}-D_{N-1})^2. \quad (8)$$

It is very important to note that the variable's distribution is only verified in the interval $x_1 \leq x \leq x_{N-1}$.

Eq. (8) indicates that the C is no longer equivalent to the RPS for irregularly spaced intervals. C weights the total score according to interval width, which may vary in the range of X.

The score, C, can be shown to be a strictly proper measure of the stated probabilities by comparison to the RPS. Since the RPS is strictly proper, the minimum score will be obtained when the forecaster's stated probabilities, R_i's, conform to his/her beliefs, P_i's. Because the stated probability for any given breakpoint, x_i, only appears in the ith term in Eq. (1), that term must be at a minimum when $R_i = P_i$ as will any positive, finite multiple of that term. The sum of positive multiples from many terms will be a minimum when each $R_i = P_i$. Because C_N is equivalent to the RPS with each term weighted by a positive, finite value, it also is strictly proper.

The expected score for C can be computed by numerical integration of Eq. (6) to give:

$$E(C_N) = \frac{x_2 - x_1}{2}R_1(1-R_1) + \sum_{i=2}^{N-2}\frac{x_{i+1} - x_{i-1}}{2}R_i(1-R_i) + \frac{x_{N-1} - x_{N-2}}{2}R_{N-1}(1-R_{N-1}). \quad (9)$$

Note that the accuracy of the estimate of the CRPS depends on the interval widths, Δx_i's. Trapezoidal integration approximates a function between the intervals as if it varied linearly. The approximation will be accurate if the probability density distribution remains fairly constant throughout the interval, and the interval width is reasonably small in relation to the probability distribution. It should be emphasized that the actual forecast distribution is not measured within the intervals, so there is no reason for the forecaster to alter R(x) from his/her belief's between the breakpoints on the basis of the numerical integration rule used.

Numerical integration to estimate the CRPS from Eq. (5) should be performed with caution, since the method used to select the precise value of K may promote hedging of the probability forecasts.

The CRPS for the example forecast in Fig. 1, approximated from the discrete intervals presented, was 1.98°C with the expected score of 1.15°C.

6. EXAMPLES

Two examples shall be presented here; the first for a Gaussian forecast probability distribution, and the second for a climatological forecast of precipitation amounts for a series of breakpoints.

6.1 A Gaussian Forecast Distribution

Let r(x) be Gaussian with a standard deviation of σ and a mean of K. R(x) is the cumulative distribution function for the same forecast. The standard cumulative normal distribution, Φ(y), can be used by defining the variable, y, as

$$y = \frac{x-K}{\sigma},\ dx = \sigma dy,\ \Phi(y) = R(x).$$

Because Φ(y) is symmetric about K, the CRPS for a Gaussian distribution can be represented by:

$$\mathrm{CRPS} = 2\sigma\int_{-\infty}^{0}\Phi^2(y)dy + \left|\sigma\int_{(T-K)/\sigma}^{0}dy\right| - 2\sigma\int_{-|(T-K)/\sigma|}^{0}\Phi(y)dy.$$

This can be simplified to:

$$\mathrm{CRPS} = 2\sigma\int (y)^2 dy + |K-T| - 2\sigma\int_{-|(T-K)/\sigma|}^{0}\Phi(y)dy.$$

From Eq. (5), the expected score is

$$E(\mathrm{CRPS}) = 2\sigma\int_{-\infty}^{0}\Phi(y)\ (1-\Phi(y))dy.$$

These relationships show that E(CRPS) and S_0 for a Gaussian distributed forecast are linearly related to σ. This can greatly simplify estimation of CRPS for Gaussian forecast distributions.

The discrete approximation to the CRPS for a Gaussian forecast distribution can be computed by substitution of

$$R_i = \Phi\left(\frac{x_i-K}{\sigma}\right)$$

into Eq. (8). A similar substitution can be used with Eq. (9) to obtain the expected score.

Recall that the sample forecast from Fig. 1 was selected from a Gaussian distribution with σ = 2.0°C, and K = 10.0°C. T was 13.0°C. The CRPS for this distribution and observation was computed to a high degree of accuracy from Eq. (7) with an integration step, Δx, of .02°C. This yielded the CRPS of 1.97°C. The expected score was computed from numerical integration of Eq. (6) with the same Δx to obtain the value, E(CRPS) = 1.14°C. These values can be compared to those estimated by the discrete approximation (see section 5). The CRPS and E(CRPS) are slightly lower than the values estimated by the less precise approximation, although the difference is relatively small.

Fig. 2 shows the scores which would result for various values of σ as a function of (K-T) for a forecast of temperature reported in whole degrees C. The reported temperature is assumed to be rounded to the nearest degree. If T is reported to be 13°C, for example, then the actual value can be anywhere within the interval $12.5 \leq x < 13.5$°C. Thus, the breakpoints are at one degree intervals at the half degree. K is assumed to occur at the center of the interval.

Fig. 2 illustrates the properties of the score for Gaussian forecast distributions. The lowest value for any given σ always occurs when K-T = 0. The best score is obtained when the forecast is made with a high degree of certainty (low σ) and (K-T) is near zero. A larger σ produces a better score than one from a forecast with a smaller σ when $|K-T|$ is large.

6.2 A Precipitation Amount Forecast from Climatic Frequencies

Climatic frequencies for precipitation greater than or equal to specified amounts at Norfolk, Virginia for the 12-h period ending at 0000 GMT for the months of September through November are shown in Table 2. The values, originally in hundredths of inches, have been converted to millimeters. Rounding of the observations has been accounted for in the threshold

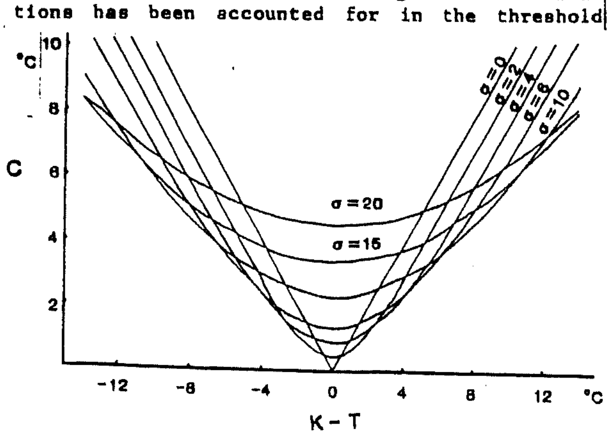


Figure 2. Value of C, approximated from one degree intervals, for a series of Gaussian forecast distributions with varying σ and K-T. Lines are labeled according to the σ in °C of the forecast distribution verified.

values, so, for example, the threshold for reported precipitation $\geq$ .01 actually includes all values: $x \geq .005$ in, or $\geq$ .1 mm. Columns 4 and 5 display the value of the RPS and CRPS, respectively, when the observation falls in the interval $x_i \leq T < x_{i-1}$. The natural breakpoint, x_1 = 0.0 mm has been added for the computations, F_1 = 1.0.

Note that the forecast probability for precipitation amount is $F(x) = P(T \geq x)$, and not $P(T \leq x)$ as used in previous discussions, is stated. For this definition, the CRPS can be shown to be:

$$CRPS = \int_{-\infty}^{\infty} [F(x) - (1-D(x))]^2 dx.$$

This was derived by defining a variable $y = -x$, so that $R(y) = F(x)$ and $D(y) = 1-D(x)$.

The RPS is computed from the probabilities, $f_i = P(x_{i-1} \leq T < x)$ and cumulative probabilities,

$$(1-F_i) = \sum_{k=1}^{i} f_k,$$

The RPS for this example is:

$$RPS_9 = \sum_{i=2}^{9} z_i$$

where

$$z_i = (1 - F_i - D_i)^2.$$

Note that $z_1 = 0$.

Since

$$z_i = [-(F(x_i) - (1-D(x_i)))]^2,$$

C can be expressed from Eq. (8), for these forecast categories, by

$$C_9 = 1.20\, z_2 + 3.05\, z_3 + 5.10\, z_4 + 9.55\, z_5 + 12.70\, z_6 + 12.70\, z_7 + 6.35\, z_8.$$

Table 2. Climatic frequencies, F_i, for 12-h precipitation amount greater than given thresholds, x_i, for the Fall (September through November) at Norfolk, Va. (Jorgensen et al., 1969). The RPS and CRPS for precipitation occurrence within categories, and the expected score are also shown.

			Scores for ($x_i \leq T < x_{i+1}$)	
i	x_i (mm)	F_i	RPS_9	C_9 (mm)
1	0.0	1.00	.04	.09
2	0.1	.17	.70	.88
3	2.4	.09	1.52	3.38
4	6.2	.06	2.40	7.87
5	12.6	.03	3.34	16.84
6	25.3	.01	4.32	29.29
7	38.0	.01	5.30	41.74
8	50.7	.00	6.30	48.08
	Expected score		0.33	1.27

These equations show that C weights the forecasts considerably more by the errors in the forecasts at the higher precipitation amounts than does the RPS.

7. USE OF THE WEIGHTING FUNCTION

The function, $G(x)$ in Eq. (3) can be used to emphasize the forecasts for ranges of x where the user requires a more accurate forecast. This function increases the amount that errors in some regions contribute to the total score. The CRPS with the general weighting function, $CRPS_{G(x)}$, analogous to Eq. (4) is:

$$CRPS_{G(x)} = \int_{-\infty}^{\infty} Q^2(x)dG(x) + |G(K) - G(T)| - |2 \int_{T}^{K} Q(x)dG(x)|.$$

The second term on the right side might be used to select $G(x)$ for a particular application since this term specifies the score which would result if the forecast were specified with no uncertainty. $G(x)$ can be selected to produce a probability scoring rule which is compatible with a particular categorical verification method.

For example, the Log Score (LS) was developed at the National Weather Service for the verification of ceiling height and visibility forecasts with the assumption that the hazard to aviation of a missed forecast is approximately proportional to the absolute value of the logarithm of the error (National Weather Service, 1982).

$$LS = |\log_{10}(\frac{K}{T})| = |\log_{10}(K) - \log_{10}(T)|$$

The scaling factor presented in the original reference shall be neglected for this discussion.

The selection of $G(x) = \log_{10}(x)$ will produce a form of CRPS to conform to the Log Score.

The discrete form of the $CRPS_{log10}$ for a series of probability forecasts can be represented by:

$$C(\log_{10})_N = .5\ [\log_{10}(\frac{x_2}{x_1})\ (R_1-D_1)^2 + \sum_{i=2}^{N-2} \log_{10}(\frac{x_{i+1}}{x_{i-1}})\ (R_i-D_i)^2 + \log_{10}(\frac{x_{N-1}}{x_{N-2}})\ (R_{N-1}-D_{N-1})^2].$$

Here $x_1 > 0$. This form can be derived by substitution of the variable, $y = \log_{10}(x)$ before integration of Eq. (3).

Notice that if x_i's are selected to be regularly spaced on a $\log_{10}$ scale, and cover the entire range of X, $((R_1-D_1)$ and $(R_{N-1}-D_{N-1}) = 0)$, then the RPS is equivalent to C with the weighting function, $G(x) = \log_{10}(x)$, since $\log_{10}(\frac{x_{i+1}}{x_{i-1}})$ would be constant. This suggests that the RPS effectively emphasizes different ranges of the forecast variable by the location of the breakpoints in a similar way that the CRPS emphasizes the forecasts by the $G(x)$.

8. CONCLUSION

The CRPS is a quadratic scoring rule for the verification of a continuous variable forecast specified by a continuous probability distribution function. The score requires integration of the distribution function over the range of the forecast variable.

The numerical integration can be used to produce an approximation of the CRPS when the forecasts are available only at discrete points, when the probability distribution function cannot be integrated analytically, or when the observations are not reported with sufficient precision to be considered continuous. This allows the score to be used for practical forecasting situations.

The discrete approximation to the CRPS, C, is similar in form to the RPS except that the probability forecasts are weighted to account for the distance separating the breakpoints. The RPS weights the forecasts from all categories equally. The RPS effectively emphasizes certain ranges in the predicted variable by the location of the breakpoints which define the categorical probability forecasts. The CRPS emphasizes the forecast in certain ranges by a weighting function.

The location and distance between breakpoints used to specify the forecast determine the accuracy of the approximation to the CRPS. C cannot be considered an accurate measure of the CRPS when the intervals are too large to satisfactorily represent the probability distribution function for the numerical integration. The score, however, will remain a strictly proper measure of the stated forecasts regardless of the intervals used to make the prediction.

The CRPS is an appropriate score to verify forecasts of continuous variables made in probability form. The score both measures the accuracy of the probability distribution and is also sensitive to distance in the continuous sense. The score's sensitivity to distance is determined by a weighting function. This function can be selected to produce a probability score which corresponds to certain categorical verification scores in use for continuous variables.

9. ACKNOWLEDGMENTS

I appreciate the help of Dr. Allan H. Murphy who informed me of the CRPS and provided some difficult-to-find literature on the score. I would also like to thank Mr. Chad E. Myers for drafting the figures.

10. REFERENCES

Brown, T. A., 1974: Admissible scoring systems for continuous distributions. Unpublished manuscript, P-5235, The Rand Corporation, Santa Monica, California, 22 pp.

Epstein, E. S., 1969: A scoring system for probability forecasts of ranked categories. J. Appl. Meteor., 8, 985-987.

Jorgensen, D. L., W. H. Klein, C. F. Roberts, 1969: Conditional probabilities of precipitation amounts in the conterminous United States. ESSA Technical Memorandum WBTM TDL 18, Environmental Science Services Administration, U.S. Department of Commerce, 89 pp.

Matheson, J. E., and R. L. Winkler, 1976: Scoring rules for continuous probability distributions. Management Science, 22, 1087-1095.

Murphy, A. H., 1969: On the Ranked Probability Score. J. Appl. Meteor., 8, 988-989.

______, 1971: A note on the Ranked Probability Score. J. Appl. Meteor., 10, 155-156.

______, and H. Daan, 1985: Forecast Evaluation. Probability, Statistics, and Decision Making in the Atmospheric Sciences. A. H. Murphy and R. W. Katz, editors. Westview Press, Boulder, Co., 545 pp.

National Weather Service, 1982: National Verification Plan. National Oceanic and Atmospheric Administration, U.S. Department of Commerce, 81 pp.