MANAGEMENT SCIENCE
Vol. 22, No. 10, June, 1976
*Printed in U.S.A.*

# SCORING RULES FOR CONTINUOUS PROBABILITY DISTRIBUTIONS*

JAMES E. MATHESON† AND ROBERT L. WINKLER‡§**

Personal, or subjective, probabilities are used as inputs to many inferential and decision-making models, and various procedures have been developed for the elicitation of such probabilities. Included among these elicitation procedures are scoring rules, which involve the computation of a score based on the assessor's stated probabilities and on the event that actually occurs. The development of scoring rules has, in general, been restricted to the elicitation of discrete probability distributions. In this paper, families of scoring rules for the elicitation of continuous probability distributions are developed and discussed.

## 1. Introduction

The theory of personal, or subjective, probability, as developed by de Finetti [4] and Savage [14], plays an important role with respect to numerous inferential and decision-making models. In particular, Bayesian approaches to inferential problems often involve personal probabilities, and an important input to decision analysis models is a set of personal probabilities representing the judgments of a decision maker or of an expert consulted by a decision maker. Partly as a result of the increasing interest in Bayesian inference and decision analysis, the elicitation of personal probabilities has received a considerable amount of attention in recent years. Various methods have been developed to aid an individual in assessing (encoding) personal probabilities to be used in inferential and decision-making situations (e.g., see Winkler [19] and Spetzler and Stael von Holstein [16]). The methodology and experimental work regarding probability elicitation has ranged from investigations of details of the actual elicitation procedure (e.g., comparisons of different response modes) to studies of the relationship of the elicitation process to the broader framework of modeling in decision analysis.

Included among the elicitation procedures that have been studied are scoring rules, which encourage an assessor to reveal his "true" opinions and to make his stated probabilities correspond with his judgments. Scoring rules, which involve the computation of a score based on the assessor's stated probabilities and on the event that actually occurs, are useful in the evaluation of probability assessors as well as in the elicitation process itself. In terms of elicitation, the role of scoring rules is to encourage the assessor to make careful assessments and to be "honest," whereas in terms of evaluation, the role of scoring rules is to measure the "goodness" of the probabilities. For general discussions of scoring rules, see Winkler [20], Murphy and Winkler [12], Stael von Holstein [17], and Savage [15].

The development of scoring rules has, in general, been restricted to the elicitation of individual probabilities or discrete probability distributions. In many situations, of course, the variables of interest are discrete, or discrete approximations can be used. In this paper, we are concerned specifically with the development of scoring rules for continuous probability distributions. In Winkler [20], it is pointed out that scoring

* Processed by Professor Ronald A. Howard, Departmental Editor for Decision Analysis; received August 1974, revised May 1975. This paper has been with the authors 2 months for revision.

† Stanford Research Institute.

‡ Indiana University.

§ On leave and visiting the Department of Statistics, Stanford University, when some of the work reported here was conducted (January 1974 to August 1974).

** Supported in part by the National Science Foundation under Grants GA-31735 and GA-41232.

rules developed for discrete situations can be used to elicit continuous probability distributions through the use of randomly generated partitions that are not known to the assessor at the time of elicitation. The purpose of this paper is to develop classes of scoring rules based on the entire density function (or equivalently, the distribution function) rather than just on a set of probabilities determined from the density function via a partition. We generate an extremely rich set of scoring rules that includes previously developed rules (including discrete rules) as special cases and provides the experimenter with considerable flexibility in choosing a rule that is particularly appropriate for a given situation.

Families of scoring rules based on simple binary scoring rules are developed and compared with scoring rules for discrete situations in §2. In §3, families of scoring rules based on another type of payoff, or scoring, function are developed. §3 contains a brief summary and discussion.

## 2. The Generation of Scoring Rules for Continuous Distributions from Scoring Rules for Binary Situations

Consider the assessment of the probability of a single event $E$. We assume a subject assigns probability $p$ to the occurrence of the event, but when asked to reveal his probability assignment states a probability $r$ which might not be equal to $p$. A scoring rule $S(r)$ gives the subject a payoff

$$\begin{aligned} S(r) &= S_1(r) \quad \text{if the event occurs,} \\ &= S_2(r) \quad \text{if the event does not occur.} \end{aligned} \tag{1}$$

The subject's expected payoff for this binary situation is accordingly

$$E(S(r)) = pS_1(r) + (1-p)S_2(r), \tag{2}$$

and the scoring rule is defined as strictly proper if

$$E(S(p)) > E(S(r)) \quad \text{for } r \neq p. \tag{3}$$

The notion of scoring rules can be generalized quite easily to the assessment of any discrete probability distribution. Let $E_i$ represent the $i$th event (or $i$th value of a random variable), where $i \in I$ and $I$ is finite or countably infinite. Moreover, let $p_i$ and $r_i$ correspond to $p$ and $r$ in the binary situation, and suppose that the scoring rule $S(r_1, r_2, \ldots)$ gives the subject a payoff $S_j(r_1, r_2, \ldots)$ if $E_j$ occurs. Then

$$E(S(r_1, r_2, \ldots)) = \sum_{j \in I} p_j S_j(r_1, r_2, \ldots), \tag{4}$$

and $S$ is strictly proper if

$$E(S(p_1, p_2, \ldots)) > E(S(r_1, r_2, \ldots)) \quad \text{when } r_i \neq p_i \quad \text{for any } i \in I. \tag{5}$$

The literature regarding such rules is fairly extensive; several forms of strictly proper scoring rules have been developed (e.g., see the references given in §1). Three frequently-encountered examples are the quadratic, logarithmic, and spherical scoring rules, which are, respectively,

$$S_j(r_1, r_2, \ldots) = 2r_j - \sum_{i \in I} r_i^2, \tag{6}$$

$$S_j(r_1, r_2, \ldots) = \log r_j, \quad \text{and} \tag{7}$$

$$S_j(r_1, r_2, \ldots) = r_j \Big/ \Big( \sum_{i \in I} r_i^2 \Big)^{\frac{1}{2}}. \tag{8}$$

Scoring rules can be extended to the continuous case by limiting arguments. If $x$ is the revealed value of the variable of interest and $r(x)$ represents the density function assigned by the subject, continuous analogs of the quadratic, logarithmic, and spherical scoring rules are, respectively,

$$S(r(x)) = 2r(x) - \int_{-\infty}^{\infty} r^2(x)dx, \tag{9}$$

$$S(r(x)) = \log r(x), \quad \text{and} \tag{10}$$

$$S(r(x)) = r(x)\Big/\left(\int_{-\infty}^{\infty} r^2(x)dx\right)^{\frac{1}{2}} \tag{11}$$

(e.g., Buehler [2]). Rules such as these are strictly proper scoring rules for the continuous case, but they are not the only such rules. In the remainder of this section, we generate families of scoring rules for the continuous case without applying limiting arguments to scoring rules for discrete probability distributions. In particular, we work directly with binary scoring rules of the type represented in (1).

Consider the assessment of a probability distribution for a variable defined on the real line. We assume the subject assigns probability distribution function $F(x)$ to the variable, but when asked to reveal his probability assignment states $R(x)$. Let $u$ be an arbitrary real number we shall use to divide the real line into two intervals (see Figure 1), $I_1 = (-\infty, u]$ and $I_2 = (u, \infty)$. Let $E$ be the event that $x$ falls in $I_1$. Applying a scoring rule $S(r)$, as in (1), with the identification $p = F(u)$ and $r = R(u)$, we have

$$\begin{aligned} S(R(u)) &= S_1(R(u)) \quad \text{if } x \in I_1, \\ &= S_2(R(u)) \quad \text{if } x \in I_2, \quad \text{and} \end{aligned} \tag{12}$$

$$E(S(R(u))) = F(u)S_1(R(u)) + [1 - F(u)]S_2(R(u)). \tag{13}$$

(Note that $S$ is a function of $x$ as well as $R(u)$; throughout this section we suppress explicit mention of the dependence on $x$ in scoring-rule expressions in order to simplify the notation.)

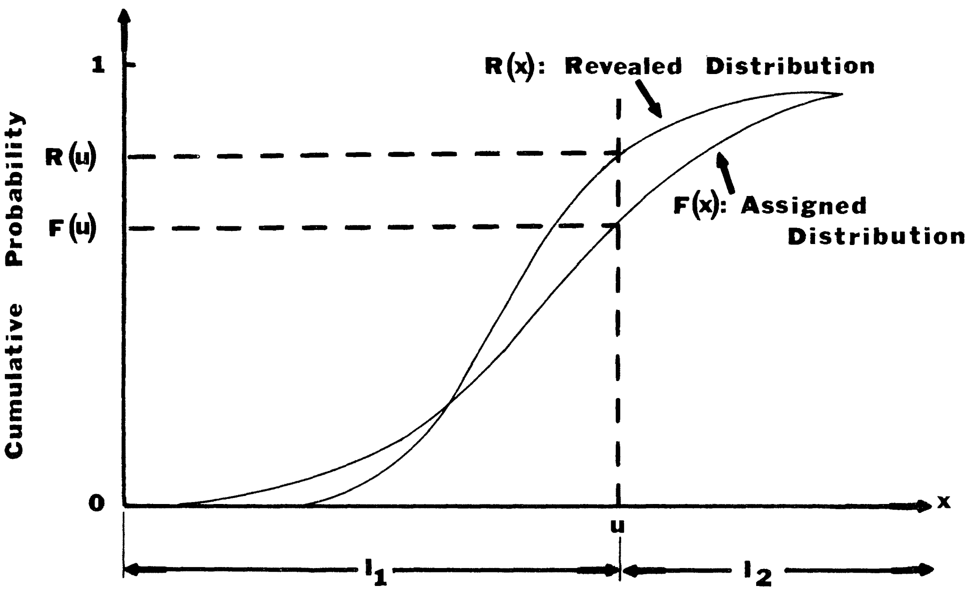


FIGURE 1. Generation of Probability-Oriented Scoring Rules.

If $S$ is strictly proper, then

$$E(S(F(u))) > E(S(R(u))) \quad \text{if } F(u) \neq R(u). \tag{14}$$

Thus, the subject will maximize the expected payoff by setting $R(u) = F(u)$. If the subject does not know the value of $u$, he should set $R(u) = F(u)$ for all possible values of $u$ to guarantee that regardless of the value of $u$, his expected payoff given $u$ is maximized. However, his payoff depends strongly on the arbitrarily selected value of $u$. To eliminate this dependence, we can simply integrate $S(R(u))$ over all $u$ and pay the subject this amount, which is

$$S^*(R(\cdot)) = \int_{-\infty}^{x} S_2(R(u))du + \int_{x}^{\infty} S_1(R(u))du. \tag{15}$$

The corresponding expected score is

$$E(S^*(R(\cdot))) = \int_{-\infty}^{\infty} E(S(R(u)))du. \tag{16}$$

(15) and (16) are in direct analogy with (12) and (13). (16) can be derived as the expectation of (15) with an interchange of order of integration. If $S$ is strictly proper, then $S^*$ is strictly proper, and the subject maximizes the expected payoff by setting $R(u) = F(u)$ for each $u$.

To increase the generality and usefulness of the above result, we assume that the experimenter selects a probability distribution function $G(u)$ for $u$. After a value of $x$ has been revealed, he pays the subject the expected score using this distribution. The expected score, given the revealed value $x$, is

$$S^{**}(R(\cdot)) = E_{u\,|\,x}(S(R(u))) = \int_{-\infty}^{x} S_2(R(u))dG(u) + \int_{x}^{\infty} S_1(R(u))dG(u), \tag{17}$$

and before $x$ is revealed the subject's expected score is

$$E(S^{**}(R(\cdot))) = \int_{-\infty}^{\infty} E(S(R(u)))dG(u). \tag{18}$$

Since $S$ is strictly proper, $S^{**}$ is also strictly proper, and the subject maximizes his expected payoff by setting $R(u) = F(u)$ for each $u$. Incidentally, note that the experimenter could simply generate a single value from $G(u)$ and use that value to reward the subject via (12). However, although the mathematical results are identical, it seems preferable to pay the expected score given by (17) instead of the score obtained from a single value generated from $G(u)$.

If we write (18) in density form,

$$E(S^{**}(R(\cdot))) = \int_{-\infty}^{+\infty} E(S(R(u)))g(u)du, \tag{19}$$

we see that $g(u)$ serves as a weighting function which should encourage the subject to pay more attention to his assessments where $g(u)$ is highest. Thus, if certain regions of values of the variable are of particular interest, the experimenter might make $g(u)$ higher in these regions than it is elsewhere. Savage [15, p. 798] notes, "Insofar as responding is hard work, the scoring rule should encourage the respondent to work hardest at what the questioner most wants to know." Of course, $g(u)$ could be a general weighting function [i.e., it is not necessary for $G(u)$ to be a probability distribution function], but this does not increase the generality of our results. Technically, $G(u)$ must be selected so that the integral of (19), which depends on both the scoring rule and the probability distributions, will exist. If the interval of definition is finite or the integrand is well behaved, $g(u)$ can be selected as uniform or "diffuse" to yield the earlier results of (15) and (16).

This process generates continuous scoring rules from each binary scoring rule. For instance, consider the quadratic scoring rule defined by

$$S_1(r) = -(1-r)^2 \quad \text{and} \tag{20}$$

$$S_2(r) = -r^2, \quad \text{with} \tag{21}$$

$$\begin{aligned} E(S(r)) &= -p(1-r)^2 - (1-p)r^2 \\ &= -(p-r)^2 - p(1-p). \end{aligned} \tag{22}$$

The generated continuous quadratic case defined by (17) and (18) is a payoff of[1]

$$S^{**}(R(\cdot)) = -\int_{-\infty}^{x} R^2(u)\,dG(u) - \int_{x}^{\infty}[1 - R(u)]^2\,dG(u) \tag{23}$$

and an expected score of

$$E(S^{**}(R(\cdot))) = -\int_{-\infty}^{+\infty}[F(u) - R(u)]^2\,dG(u) - \int_{-\infty}^{+\infty}F(u)[1 - F(u)]\,dG(u). \tag{24}$$

If the subject sets $R(u) = F(u)$, then his expected score is

$$E(S^{**}(F(\cdot))) = -\int_{-\infty}^{+\infty}F(u)[1 - F(u)]\,dG(u), \tag{25}$$

which is a measure of the dispersion in his true probability assignment. Note that (24) is the sum of two terms, the first rewarding honesty and the second rewarding expertise or sharpness. Although partitioning of the quadratic scoring rule and the resulting "attributes" measured by elements of various partitions have been studied (e.g., Sanders [13], Murphy and Epstein [11], Murphy [9], [10]), it appears that partitioning of the function representing the expected score has not been considered.

Although this work was motivated by the desire for better continuous scoring rules, the results are applicable for any probability distribution function $F(x)$. Thus, they are applicable to the discrete case. Moreover, the continuous case can be discretized by choosing $G(u)$ as a step function. For example, suppose that $G(u)$ is a step function with positive steps $g_1, g_2, \ldots, g_n$ at $u_1 < u_2 < \cdots < u_n$ and that $R(u_i) = R_i$ and $F(u_i) = F_i$ for $i = 1, 2, \ldots, n$. Then the quadratic binary scoring rule generates

$$S^{**}(R(\cdot)) = -\sum_{i=1}^{j-1} R_i^2 g_i - \sum_{i=j}^{n-1}(1 - R_i)^2 g_i \quad \text{if } x = u_j \quad \text{and} \tag{26}$$

$$E(S^{**}(R(\cdot))) = -\sum_{i=1}^{n-1}(F_i - R_i)^2 g_i - \sum_{i=1}^{n-1} F_i(1 - F_i) g_i. \tag{27}$$

It is interesting to note that when the continuous rules generated in this section are discretized by choosing $G(u)$ as a step function, the resulting rules are not equivalent to the standard discrete rules such as the quadratic, logarithmic, and spherical rules

[1] During the final preparation of this paper, we learned that in an unpublished paper, Brown [1] used a different approach to generate a rule that is apparently equivalent to the rule given by (23) with $dG(u)$ replaced by $du$. Setting $dG(u) = du$ implies that $g(u)$ is "diffuse" and provides a special member of the general family of continuous scoring rules generated from the binary quadratic scoring rule.

given in (6), (7), and (8), respectively. One important difference is the weighting function $g(u)$, which consists of $n$ positive weights $g_1, g_2, \ldots, g_n$ at $u_1, u_2, \ldots, u_n$ in the discrete case.

To remove the effect of the weights, let $g_i = n^{-1}$ for $i = 1, \ldots, n$. The rule generated in (26) with equal weights is not equivalent to the usual quadratic rule given in (6). Instead, it is equivalent to a scoring rule known as the ranked probability score (e.g., Epstein [3], Murphy [7], [8], Stael von Holstein [17]),

$$S_j(r_1, r_2, \ldots, r_n) = \frac{3}{2} - \frac{1}{2(n-1)} \sum_{i=1}^{n-1} \left[ \left( \sum_{j=1}^{i} r_j \right)^2 + \left( \sum_{j=i+1}^{n} r_j \right)^2 \right] - \frac{1}{n-1} \sum_{i=1}^{n} |i-j| r_i, \tag{28}$$

in the sense that it is a linear transformation of the ranked probability score. When $g_1 = g_2 = \cdots = g_n$, the rules given in (26) and (28) are both linear transformations of $-\sum_{i=1}^{n-1} R_i^2 + 2\sum_{i=j}^{n-1} R_i - (n-j)$.

The ranked probability score has properties quite different from those of the usual quadratic score. In particular, the ranked probability score assigns a higher score to $R$ than to $R'$ if $R'$ $(\neq R)$ is "more distant" from the true event $u_j$ according to the following definition of "distance" (Stael von Holstein [17]; see Murphy [7] for a different definition of "distance"):

$$R_i' \geqslant R_i \quad \text{for } i = 1, \ldots, j-1 \quad \text{and}$$

$$R_i' \leqslant R_i \quad \text{for } i = j, \ldots, n-1. \tag{29}$$

(Note that this definition only provides a partial ordering of distributions and does not provide any measure of distance.) Thus, the ranked probability score is said to be "sensitive to distance," and the procedures discussed in this section can be used to generate classes of scoring rules that are sensitive to distance for the continuous case. Continuous scoring rules such as the rules given in (6), (7), and (8) are sensitive to the probability density function at the precise point of the revealed value of the variable but not to the amount of probability mass nearby. In contrast, continuous rules generated by (17) are sensitive to the entire density function, not just to the density at a single value.

## 3. The Generation of Scoring Rules for Continuous Distributions from Payoff Functions Other than Binary Scoring Rules

The rules generated in §2 are based on binary scoring rules, which are defined on the probability space (the unit interval). Other rules for continuous distributions can be generated from different types of payoff functions. In this section, we consider families of scoring rules based on payoff functions that are defined on the space of values of the variable of interest (in general, the real line).

As in §2, we assume that the subject assigns probability distribution function $F(x)$ to the variable of interest but states $R(x)$ when asked to reveal his probability assignment. In order to treat the cases of discrete points and zero-probability intervals we shall define the inverse functions

$$F^{-1}(z) = \min_u \{ u \mid F(u) \geqslant z \} \quad \text{and} \tag{30}$$

$$R^{-1}(z) = \min_u \{ u \mid R(u) \geqslant z \} \tag{31}$$

for all $z \in (0, 1)$. The typical case is illustrated in Figure 2.

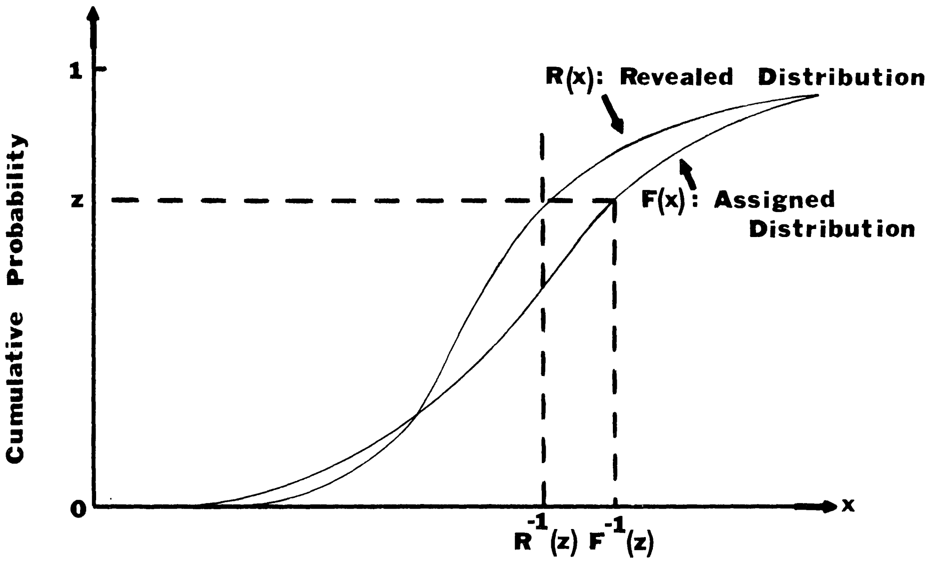


FIGURE 2. Generation of Value-Oriented Scoring Rules.

For any arbitrary $z \in [0, 1]$, let the subject receive a payoff according to the rule $T(R^{-1}(z))$. (Note that $T$ may be a function of $x$ and $z$ as well as $R^{-1}(z)$; as in §2, we suppress explicit mention of the dependence on $x$ and $z$ in scoring-rule expressions in order to simplify the notation.) If $T$ is strictly proper, then

$$E\big(T(F^{-1}(z))\big) > E\big(T(R^{-1}(z))\big) \quad \text{if } R^{-1}(z) \neq F^{-1}(z), \quad \text{where} \tag{32}$$

$$E\big(T(R^{-1}(z))\big) = \int_{-\infty}^{\infty} T(R^{-1}(z))\,dF(x). \tag{33}$$

For example, let $T$ represent a payoff function for a Bayesian point estimation problem under linear loss (e.g., see Winkler [21, pp. 397–405]). This situation is often called the "newsboy problem" because it can be expressed in the context of a newsboy who must order papers when he is uncertain about the demand for papers and any unsold papers are worthless to him. Suppose that each paper costs the newsboy $c$ and sells for $s$, the demand for papers is $x$, and $R^{-1}(z)$ is the number of papers bought by the newsboy as a function of $z = (s-c)/s$. Then the payoff function can be represented as follows:

$$\begin{aligned} T(R^{-1}(z)) &= (s-c)x - s(1-z)\big[R^{-1}(z) - x\big] \quad \text{if } x \leqslant R^{-1}(z), \\ &= (s-c)x - sz\big[x - R^{-1}(z)\big] \quad \text{if } x > R^{-1}(z). \end{aligned} \tag{34}$$

If the newsboy orders exactly the right number of papers [i.e., if $x = R^{-1}(z)$], then his payoff is $(s-c)x$. However, the payoff of $(s-c)x$ is reduced by $s(1-z) = c$ for each unsold paper if $x < R^{-1}(z)$ (demand overestimated) and is reduced by $sz = s - c$ for each demanded but unavailable paper if $x > R^{-1}(z)$ (demand underestimated). The newsboy's expected payoff,

$$\begin{aligned} E\big(T(R^{-1}(z))\big) = &\int_{-\infty}^{\infty}(s-c)x\,dF(x) - \int_{-\infty}^{R^{-1}(z)} s(1-z)\big[R^{-1}(z) - x\big]dF(x) \\ &- \int_{R^{-1}(z)}^{\infty} sz\big[x - R^{-1}(z)\big]dF(x), \end{aligned} \tag{35}$$

where $F(x)$ is the newsboy's probability distribution function for $x$, is maximized when $R^{-1}(z) = F^{-1}(z)$.

If the subject does not know the value of $z$, he should set $R(x) = F(x)$; however, the actual payoff depends strongly on the arbitrarily selected value of $z$. To eliminate this dependence, we integrate over all $z$ and pay the subject

$$T^*(R^{-1}(\cdot)) = \int_0^1 T(R^{-1}(z))dz. \tag{36}$$

The expected score is then

$$E\big(T^*(R^{-1}(\cdot))\big) = \int_{-\infty}^{\infty}\int_0^1 T(R^{-1}(z))dz dF(x) = \int_0^1 E\big(T(R^{-1}(z))\big)dz. \tag{37}$$

The integration over $z$ is analogous to the integration over $u$ in §2. If $T$ is strictly proper then $T^*$ is also strictly proper, and the subject maximizes his expected payoff by setting $R(x) = F(x)$.

We can now generalize the above result in a manner analogous to the generalization represented by (17) in §2. Assume that the experimenter selects a probability distribution function $H(z)$ for $z$. After a value of $x$ has been revealed, the subject is paid the expected score using $H$:

$$T^{**}(R^{-1}(\cdot)) = E_{z\,|\,x}\big(T(R^{-1}(z))\big) = \int_0^1 T(R^{-1}(z))dH(z). \tag{38}$$

Before $x$ is revealed, the subject's expected score is

$$E\big(T^{**}(R^{-1}(\cdot))\big) = \int_{-\infty}^{\infty}\int_0^1 T(R^{-1}(z))dH(z)dF(x) = \int_0^1 E\big(T(R^{-1}(z))\big)dH(z). \tag{39}$$

For example, the payoff generated by the scoring rule of (34) is

$$T^{**}(R^{-1}(\cdot)) = (s-c)x - \int_0^{R(x)} sz\big[x - R^{-1}(z)\big]dH(z)$$
$$- \int_{R(x)}^1 s(1-z)\big[R^{-1}(z) - x\big]dH(z). \tag{40}$$

If $T$ is strictly proper, the subject maximizes his expected payoff from (39) by setting $R(x) = F(x)$. Here $H$ is similar to $G$ (from §2) in that $dH(z)$ serves as a weighting function which should encourage the subject to pay more attention to his assessments where $dH(z)$ is highest. For example, if the experimenter is particularly concerned about the extreme tails of the distribution, he might select a $U$-shaped $dH(z)$; if the middle of the distribution is of interest, $dH(z)$ might be taken to be symmetric and unimodal with mode at $z = 0.5$. Of course, $dH(z)$ can simply be uniform, in which case (38) and (39) reduce to (36) and (37). If only certain fractiles are of interest, $H(z)$ can be chosen as a step function with positive steps $h_1, h_2, \ldots, h_m$ at $z_1 < z_2 < \cdots < z_m$.

## 4. Summary and Discussion

In this paper we have developed classes of scoring rules for continuous probability distributions. The rules developed in §2 are based on binary scoring rules, and each member of the family of scoring rules generated in this manner corresponds to a particular choice of $S$ (the binary scoring rule) and $G$ (the weighting function). The rules developed in §3 are based on a different type of payoff function, and each member of the family of scoring rules generated in §3 corresponds to a particular choice of $T$ (the payoff function) and $H$ (the weighting function). The two families of scoring rules are completely different, however. The difference relates to the fact that $S$ is defined on the probability space (the unit interval), whereas $T$ is defined on the space of values of the variable of interest (in general, the real line).

The families of scoring rules developed here are quite rich, and the choice of a single scoring rule for use in a particular situation has not been discussed. In practice, the choice of a single rule might be based primarily on convenience and on psychological considerations relating to the elicitation procedure. To the extent that scoring rules are used to encourage careful assessments and to the extent that the assessor perceives that he is being evaluated in terms of scores determined from scoring rules, such rules play an important role in the elicitation process and may have a significant effect on the resulting probabilities. For instance, experimental results suggest that different elicitation techniques may yield quite different results (e.g., see Jensen and Peterson [6], Tversky [18], and Hogarth [5]). Clearly such factors need to be investigated further; in particular, further study of the role of scoring rules in the elicitation process and the impact of scoring rules on assessed probabilities would be most valuable.

Of course, it must be remembered that scoring rules are just one aspect of the elicitation process. With respect to the elicitation of probabilities in general and the elicitation of continuous probability distributions in particular, further work is also needed regarding the other aspects of the elicitation process and the role of the elicitation process in the broader sphere of modeling inferential and decision-making problems.

## References

1. Brown, T. A., "Admissible Scoring Systems for Continuous Distributions," unpublished manuscript, P-5235, The Rand Corporation, Santa Monica, California, 1974.
2. Buehler, R. J., "Measuring Information and Uncertainty," in V. P. Godambe and D. A. Sprott, eds., *Foundations of Statistical Inference*, Holt, Rinehart and Winston, Toronto, 1971.
3. Epstein, E. S., "A Scoring System for Probability Forecasts of Ranked Categories," *Journal of Applied Meteorology*, 8, 1969, pp. 985–987.
4. de Finetti, B., "La prévision: ses lois logiques, ses sources subjectives," *Annales de l'Institut Henri Poincaré*, 7, 1937, pp. 1–68.
5. Hogarth, R., "Cognitive Processes and the Assessment of Subjective Probability Distributions," *Journal of the American Statistical Association*, 70, 1975, pp. 271–294.
6. Jensen, F. A. and Peterson, C. R., "Psychological Effects of Proper Scoring Rules," *Organizational Behavior and Human Performance*, 9, 1973, pp. 307–317.
7. Murphy, A. H., "The Ranked Probability Score and the Probability Score: A Comparison," *Monthly Weather Review*, 98, 1970, pp. 917–924.
8. ———, "A Note on the Ranked Probability Score," *Journal of Applied Meteorology*, 10, 1971, pp. 155–156.
9. ———, "Scalar and Vector Partitions of the Probability Score: Part I. Two-State Situation," *Journal of Applied Meteorology*, 11, 1972, pp. 273–282.
10. ———, "Scalar and Vector Partitions of the Probability Score: Part II. $N$-State Situation," *Journal of Applied Meteorology*, 11, 1972, pp. 1183–1192.
11. ——— and Epstein, E. S., "Verification of Probabilistic Predictions: A Brief Review," *Journal of Applied Meteorology*, 6, 1967, pp. 748–755.
12. ——— and Winkler, R. L., "Scoring Rules in Probability Assessment and Evaluation," *Acta Psychologica*, 34, 1970, pp. 273–286.
13. Sanders, F., "On Subjective Probability Forecasting," *Journal of Applied Meteorology*, 2, 1963, pp. 191–201.
14. Savage, L. J., *The Foundations of Statistics*, New York, Wiley, 1954.
15. ———, "Elicitation of Personal Probabilities and Expectations," *Journal of the American Statistical Association*, 66, 1971, pp. 783–801.
16. Spetzler, C. S. and Stael von Holstein, C.-A. S., "Probability Encoding in Decision Analysis," *Management Science*, 22, 1975, pp. 340–358.

17. Stael von Holstein, C.-A. S., *Assessment and Evaluation of Subjective Probability Distributions*, Stockholm, Economic Research Institute, Stockholm School of Economics, 1970.
18. Tversky, A., "Assessing Uncertainty," *Journal of the Royal Statistical Society* B, 36, 1974, pp. 148–159.
19. Winkler, R. L., "The Assessment of Prior Distributions in Bayesian Analysis," *Journal of the American Statistical Association*, 62, 1967, pp. 776–800.
20. ———, "The Quantification of Judgment: Some Methodological Suggestions," *Journal of the American Statistical Association*, 62, 1967, pp. 1105–1120.
21. ———, *An Introduction to Bayesian Inference and Decision*, Holt, Rinehart and Winston, New York, 1972.