DOCUMENT RESUME

ED 194 565 TM 800 694

AUTHOR van der Linden, Wim J.
TITLE Statistical Aspects of Optimal Treatment Assignment. Twente Educational Memorandum No. 18.
INSTITUTION Twente Univ. of Technology, Enschede (Netherlands).
PUB DATE Jun 80
NOTE 17p.: Paper presented at the European Meeting of the Psychometric Society (Groningen, Netherlands, June 18-20, 1980).

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS *Aptitude Treatment Interaction: *Bayesian Statistics: Decision Making: Elementary Secondary Education: Hypothesis Testing: *Individualized Instruction: *Mastery Tests: Predictor Variables: Regression (Statistics): *Research Design
IDENTIFIERS Criterion Variables: *Utility Theory

ABSTRACT

The issues of treatment assignment is ordinarily dealt with within the framework of testing aptitude treatment interaction (ATI) hypothesis. ATI research mostly uses linear regression techniques, and an ATI exists when the aptitude treatment (AT) regression lines cross each other within the relevant interval of the aptitude variable. Consistent with this approach is the use of the points of interaction of AT regression lines as treatment-assignment rule. The replacement of such rules by monotone, nonrandomized (Bayes) rules is proposed. Both continuous and dichotomous criteria for treatment success are considered. An example of the latter is evaluated using a mastery test. Solutions are given based on linear, normal ogive, and threshold utility functions. Some modifications of these functions are discussed which are believed to be more realistic in the context of individualized instruction, but for which no optimal monotone assignment rules are available yet. (Author/RL)

Statistical Aspects of Optimal Treatment Assignment

Wim J. van der Linden

Twente University of Technology

Enschede, The Netherlands

Paper presented at the European Meeting of the Psychometric Society, Groningen, The Netherlands, June 18-20, 1980.

Twente Educational Memorandum No. 18. Enschede, The Netherlands: Twente University of Technology, Department of Educational Technology, June,1980.

Abstract

The issue of treatment assignment is ordinarily dealt with within the framework of testing Aptitude Treatment Interaction hypotheses. ATI research mostly uses linear regression techniques, and an ATI exist when the AT regression lines cross each other within the relevant interval of the aptitude variable. Consistent with this approach is the use of the points of interaction of AT regression lines as treatment-assignment rule. In this paper, it is proposed to replace such rules by monotone, nonrandomized (Bayes) rules. Both continuous and dichotomous criteria for treatment success are considered. An example of the latter can be found in individualized instruction when learning success is evaluated using a mastery test. The solutions given in this paper are based on linear, normal ogive, and threshold utility functions. Finally, some modifications of these functions are discussed which are believed to be more realistic in the context of individualized instruction but for which no optimal monotone assignment rules are available yet.

Keywords: Aptitude Treatment Interaction, Decision Theory, Mastery Testing.

Statistical Aspects of Optimal Treatment Assignment

The problem we will consider in this paper is the following: Suppose we have a population of subjects and a number of different treatments to which subjects of this population are to be assigned. Furthermore, it is supposed that a criterion variable is available measuring the effect of the treatment. Finally, there is a predictor variable which can be used for predicting the criterion scores of the subjects for each of the treatments. The problem now consists of choosing a decision rule that assigns subjects to treatments on the basis of their predictor scores such that the assignment procedure is optimal in some sense.

An example of this problem can be found in individualized instructional systems. In these systems students are required to reach the same learning objectives but different instructional programs or treatments are available. Typically, the assignment of students to these treatments is based on their scores on an aptitude test administered previous to the instructional unit, while at the end of the unit a mastery test is administered to determine whether the student has reached the learning objectives and may proceed with the next unit or has to receive additional instruction. Individualized instruction has mainly been motivated by the view-point underlying Aptitude Treatment Interaction (ATI) research (Cronbach & Snow, 1977), namely that subjects can react differently to treatments and that treatments which are best on the average may therefore be worst in individual cases.

Other examples of situations to which the problem of this paper applies can be found, e.g., in psychotherapy, management sciences, medicine, and agricultural sciences.

Let X and Y denote the predictor and criterion variable measured before and after the treatment, respectively. Since educational and psychological measurement instruments are mostly tests, X and Y will be assumed to be integer-valued variables ranging from 0, ..., m and 0, ..., n, respectively. The possible treatments will be indexed by $j = 0, \ldots, t$. Furthermore, for each treatment j a probability function $\eta_j(x,y)$ will be adopted representing the relation between X and Y under treatment j. To select optimal decision rules, an evaluation of the decision outcomes or utility function is needed. For the present paper it is sufficiently general to consider the utility, U say, as a function of the criterion Y, which is allowed to assume a different shape for each treatment: $U = u_j(Y)$.

We shall first restrict the treatment-assignment problem to the case of two treatments. Moreover, it will be assumed that the optimal assignment rules we will be looking for can be found in the subclass of rules known as <u>monotone</u> rules (Ferguson, 1967, sect. 6.1). For the case of two treatments this means that the optimal rule has the form of a cutting score b on the predictor X such that for predictor scores $X < b$ one of the treatments and for $X \geq b$ the other treatment is assigned. The conditions which must be imposed on the utility and probability functions to arrive at optimal rules of a monotone form are discussed in van der Linden (1980a).

For each possible monotone assignment rule, the expected utility is given by

$$(1) \qquad B(b) = \sum_{x=0}^{b-1} \sum_{y=0}^{n} u_0(y)\eta_0(x,y) + \sum_{x=b}^{m} \sum_{y=0}^{n} u_1(y)\eta_1(x,y).$$

We shall use (1) as our optimization criterion and look for cutting scores for which (1) is maximal. In doing so, we may use the important fact that, although the bivariate distributions $\eta_j(x,y)$ can be expected to assume a different shape for each treatment, this does not apply to the marginal distributions of the predictor scores inasmuch as these are measured previous to the treatments. Thus, denoting the probability function of X by $\lambda(x)$, it holds that (1) may be optimized using the fact that

$$(2) \qquad \lambda_j(x) = \lambda(x)$$

for all values of j.

In this paper we will show some results for linear, normal ogive, and threshold utility functions. No derivations will be given; these can be found in van der Linden (1980a). The threshold utility function seems realistic when the criterion is a dichotomy, as is the case, for example, in individualized instruction when the criterion is a test used for mastery decisions (van der Linden, 1980b). We will also discuss other utility functions suited for the case of a dichotomous criterion, which are believed to be more realistic in the context of individualized instruction than this threshold utility function but for which no solutions are available yet.

Before proceeding, however, we observe that, although in part based on different assumptions, the approach taken in this paper comes close to the one chosen by Cronbach and Gleser (1965, Appendix 1) in their model for placement decisions. Another approach to the present problem has been used by Vijn (1980).

Linear Utility

For the case of two treatments (t=1) with utility and probability functions obeying the conditions leading to solutions in the subclass of monotone rules, it holds that the optimal cutting score, b^*, on the predictor is equal to the smallest value of x for which

$$(3) \qquad E_1[u_1(Y)|x] - E_0[u_0(Y)|x]$$

is positive. This solution, which involves a simple comparison between two conditional expected utilities, follows in a few steps from (1) and (2). It should be noted, however, that (3) is no closed-form solution and that further restrictions are required to arrive at such solutions.

We next suppose that the utility functions $u_j(Y)$ have a linear shape:

$$(4) \qquad u_j(Y) = f_jY + g_j.$$

The parameter g_j can, for instance, be a nonpositive constant representing the costs of treatment j. The relation between utility and criterion score is also determined by the (treatment-dependent) parameter f_j.

If the regression functions of Y on X may be assumed to be linear, it follows from (3) that for utility function (4) the treatment assignments are optimal for that value of b equal to

$$(5) \qquad b^* = \text{entier}\left(\frac{g_0 - g_1 + f_0\alpha_0 - f_1\alpha_1}{f_1\beta_1 - f_0\beta_0}\right) + 1,$$

where the entier function replaces the value of its argument by the greatest integer not greater than this value and α_j and β_j are the regression parameters given by $E_j(Y|x) = \alpha_j + \beta_j x$.

Note that in applications the regression parameters in (5) must be estimated and that this solution can therefore be unstable when $f_0\beta_0 \approx f_1\beta_1$.

Normal Ogive Utility

As an alternative to linear utility function (4), we next consider the following normal ogive utility function

$$(6) \qquad u_j(Y) = \Phi\left(\frac{Y - \mu_j}{\sigma_j}\right),$$

where Φ is the standard normal c.d.f. and the (treatment-dependent) parameters μ_j and σ_j govern the location and slope of (6). The use of cumulative distribution functions as utility function has been plead by Berhold (1973), Lindley (1976), and Novick and Lindley (1978). An attractive feature of (6) is that it can readily be combined with the model of a normal distribution for Y given X = x. Assuming such a model along with linear regression functions of Y on X and homoscedasticity, it can be shown that the following optimal assignment rule is obtained.

$$(7) \qquad b^* = \text{entier}\left[\frac{\varepsilon_0(\mu_1 - \alpha_1) - \varepsilon_1(\mu_0 - \alpha_0)}{\varepsilon_0\beta_1 - \varepsilon_1\beta_0}\right] + 1$$

with

$$\varepsilon_j \equiv [\mathrm{Var}_j(Y.X) + \sigma_j^2]^{1/2}$$

and

$$\mathrm{Var}_j(Y.X) \equiv \{1 - [\mathrm{Cor}_j(X,Y)]^2\}\mathrm{Var}_j(Y),$$

where $\mathrm{Cor}_j(X,Y)$ and $\mathrm{Var}_j(Y)$ are the linear correlation coefficient between X and Y and the variance of Y under treatment j, respectively.

Threshold Utility

Next, we suppose that criterion Y has a threshold value c so that $Y \geq c$ means success and $Y < c$ failure. An example of this arises in individualized instruction when the criterion is a mastery test with mastery score c.

A utility function fitting this situation can be the following threshold utility function with parameter c

$$(8) \qquad u_j(Y) = \begin{cases} w + a_j & \text{for } Y \geq c \\ v + a_j & \text{for } Y < c, \end{cases}$$

in which w and v represent the treatment-independent and a_j the treatment-dependent part (e.g., treatment costs) of the utility structure. For utility function (8), it appears that the optimal treatment-assignment rule is the smallest value of x for which

(9) $\Omega_0(c|x) = \Omega_1(c|x) + a_0 - a_1$

is positive, Ω_j being the c.d.f. of $Y|x$. It also appears that when $a_0 = a_1$ and the conditional distributions of Y given x are normal with linear regression functions and homoscedasticity, the optimal assignment rule is given by

(10) $$b^* = \text{entier}\left[\frac{(c - \alpha_1)[\text{Var}_0(Y.X)]^{1/2} - (c - \alpha_0)[\text{Var}_1(Y.X)]^{1/2}}{\beta_1[\text{Var}_0(Y.X)]^{1/2} - \beta_0[\text{Var}_1(Y.X.)]^{1/2}}\right] + 1$$

Note that this solution is a function not only of the regression parameters, α_j and β_j, and the "unexplained" variances, $\text{Var}_j(Y.X)$, but also of the threshold value c.

In all three utility functions considered so far, utility is a function of the observed criterion scores. In the event of unreliable criterion scores or criterion scores that are inefficient estimators of an underlying, latent parameter, it seems better to revise the utility functions and to define them as functions of the true criterion score T. As has been indicated in van der Linden (1980a), this does not change the solutions given above for the linear and normal ogive utility function but has consequences for (10). Not only must c be replaced by the true threshold value d on T but also $\text{Var}_j(Y.X)$ by $\text{Var}_j(T.X)$, while in the more general solution in (9) the c.d.f. of $Y|x$ must be replaced by the c.d.f. of $T|x$.

Treatment Assignment and Mastery Testing

As noted earlier, an example of a treatment-assignment problem with a dichotomized criterion can be found in individualized instruction when learning success is evaluated by the administration of a mastery test at the end of the instructional unit. Mastery testing has also been approached from a decision-theoretic point of view (e.g. Hambleton & Novick, 1973? for a review, see van der Linden, 1980b). In short, the mastery testing problem consists of two possible states obtained by dichotomizing the true-score scale, T, underlying the test by denoting a mastery score d. Students exceeding this score ($T \geq d$) are considered masters, the others ($T < d$) nonmasters. The problem is how to find a cutting score c on the observed-score variable, X, such that students are optimally classified as masters ($X \geq c$) and nonmasters ($X < c$).

It is interesting to note that this problem fits the treatment-assignment problem with an unreliable, dichotomized criterion, which suggests a further integration of treatment assignment and mastery testing. A fruitful approach seems to adapt the utility functions in use for mastery testing to the fact that in individualized instruction mastery decisions are preceded by treatment-assignment decisions and to optimize both decisions simultaneously.

Three examples of utility functions suited for this purpose will be shown. The first example is

$$
(11) \qquad u_j(T) = \begin{cases} u_{10} + a_j & \text{for } T \geq d,\ Y < c \\ u_{01} + a_j & \text{for } T < d,\ Y \geq c \\ a_j & \text{otherwise,} \end{cases}
$$

which is the threshold utility function in use in mastery testing (Hambleton & Novick, 1973; Huynh, 1976; Mellenbergh, Koppelaar, and van der Linden, 1977) extended with a treatment-dependent parameter a_j to represent, e.g., treatment costs.

The next example is an adaptation of the linear utility function introduced in mastery testing to remove the unrealistic discontinuity at $T = d$ in (11) (van der Linden & Mellenbergh, 1976):

$$
(12) \qquad u_j(T) = \begin{cases} b_{0j}(d - T) + a_{0j} & \text{for } Y < c \\ \\ b_{1j}(T - d) + a_{1j} & \text{for } Y \geq c. \end{cases}
$$

The adaption is that the parameters a_{0j}, a_{1j}, b_{0j}, and b_{1j} have been made treatment dependent to be able to account for possible differences in utility and costs between treatments.

The final example is an adaptation of the normal ogive utility function, which has been introduction in mastery testing by Novick and Lindley (1978) as an illustration of the use of cumulative distribution functions for representing utility structures:

$$(13) \qquad u_j(T) = \begin{cases} \Phi\left(\dfrac{\mu_{0j} - T}{\sigma_{0j}}\right) + a_{0j} & \text{for } Y < c \\[2em] \Phi\left(\dfrac{T - \mu_1}{\sigma_{1j}}\right) + a_{1j} & \text{for } Y \geq c \end{cases}$$

The parameters μ_{0j}, μ_{1j}, σ_{0j}, and σ_{1j} have, just as in the preceding example of the linear utility function, been made treatment dependent. Besides, the parameters a_{0j} and a_{1j} have been added to allow for treatment costs.

In principle, solutions for these three utility functions must be obtained by defining the expected utility using the trivariate distribution of (X,Y,T) and optimizing the resulting expression simultaneously to the cutting scores b on the predictor X and c on the criterion Y. Although it is believed that these can lead to an improvement on existing treatment-assignment and mastery decisions in individualized instruction, no closed-form solutions are available yet. This has to do with the fact that the optimization involved in this procedure is rather involved and that the conditions under which monotone solutions can be expected are not yet clear.

Concluding Remark

For a fuller discussion of the treatment-assignment problem in the first part of this paper, a generalization thereof to more than two treatments, a procedure for combining qualitative information with predictor scores to improve treatment assignment, and an "empirical" procedure to be used when the conditions leading to monotone rules are not met, we refer to van der Linden (1980b).

References

Berhold, M.H. The use of distribution functions to represent utility functions. Management Science, 1973, 19, 825-829.

Cronbach, L.J., & Gleser, G.C. Psychological tests and personnel decisions (2nd Ed.). Urbana, Ill.: University of Illinois Press, 1965.

Cronbach, L.J., & Snow, R.E. Aptitudes and instructional methods: A handbook for research on interactions. New York: Irvington Publishers, Inc., 1977.

Ferguson, T.S. Mathematical statistics: A decision theoretic approach. New York: Academic Press, 1967.

Hambleton, R.K., & Novick, M.R. Toward an integration of theory and method for criterion-referenced tests. Journal of Educational Measurement, 1973, 10, 159-170.

Huynh, H. Statistical considerations of mastery scores. Psychometrika, 1976, 41, 65-79.

Lindley, D.V. A class of utility functions. The Annals of Statistics, 1976, 4, 1-10.

Mellenbergh, G.J., Koppelaar, H., & van der Linden, W.J. Dichotomous decisions based on dichotomously scored items: A case study. Statistica Neerlandica, 1977, 31, 161-169.

Novick, M.R., & Lindley, D.V. The use of more realistic utility functions in educational applications. Journal of Educational Measurement, 1978, 15, 181-191.

van der Linden, W.J. Using aptitude measurements for the optimal assignment of subjects to treatments with and without mastery score. Submitted for publication, 1980. (a) van der Linden, W.J. Decision models for use with criterion-referenced testing. Applied Psychological Measurement, 1980, 4, in press. (b)

van der Linden, W.J., & Mellenbergh, G.J. Optimal cutting scores using a linear loss function. Applied Psychological Measurement, 1977, 1, 593-599.

Vijn, P. Prior information in linear models, Thesis, University of Groningen, 1980.

Acknowledgement

Gideon J. Mellenbergh for his commentary on the research reported in this paper. Paula Achterberg for typing the manuscript.

Author's Address

Wim J. van der Linden, Onderafdeling Toegepaste Onderwijskunde, Technische Hogeschool Twente, Postbus 217, 7500 AE Enschede, The Netherlands.