ED 164 548 TM 006 662

AUTHOR Levine, Michael V.
TITLE Item-Item Curves and Consistent Mental Test Parameter Estimates.
INSTITUTION Educational Testing Service, Princeton, N.J.
REPORT NO ETS-RB-76-36
PUB DATE Dec 76
NOTE 24p.

EDRS PRICE MF-$0.83 HC-$1.67 Plus Postage.
DESCRIPTORS *Item Analysis; Mathematical Applications; Mathematical Models; *Psychometrics; Statistical Analysis; *Test Construction; Test Interpretation; Test Reliability
IDENTIFIERS *Item Characteristic Curve Theory; *Latent Trait Theory

ABSTRACT

It is shown that empirical mental test P - P plots are approximately equal to theoretical item-item curves, at least for long tests administered to many people. This result is important because it leads to (1) a distribution free method for estimating points on item-item curves; (2) a general method for defining estimates of item parameters; and (3) the only parameter estimation method known to be consistent for some of the most important aptitude test models. (Author/CTM)

RB-76-36

RESEARCH BULLETIN

ITEM-ITEM CURVES AND CONSISTENT MENTAL TEST PARAMETER ESTIMATES

Michael V. Levine

This Bulletin is a draft for interoffice circulation. Corrections and suggestions for revision are solicited. The Bulletin should not be cited as a reference without the specific permission of the author. It is automatically superseded upon formal publication of the material.

Educational Testing Service
Princeton, New Jersey
December 1976

Item-Item Curves and Consistent Mental Test Parameter Estimates[1]

Abstract

It is shown that empirical mental test P - P plots are approximately equal to theoretical item-item curves, at least for long tests administered to many people. This result is important because it leads to (1) a distribution free method for estimating points on item-item curves, (2) a general method for defining estimates of item parameters and (3) the only parameter estimation method known to be consistent for some of the most important aptitude test models.

I. Background and Motivation

The "answer sheets" for multiple choice aptitude tests are commonly considered to be generated by a two stage process. First, an examinee with some (unobserved) ability θ is sampled. Second, the various items are answered independently so that the conditional probability of a randomly selected person with ability θ having a specified pattern of right and wrong answers on the first n items of the test can be expressed in the form

$$\prod_{i=1}^{n} P_i(\theta)^{u_i}[1 - P_i(\theta)]^{1-u_i}$$

where $u_i = 1$ or 0 according to whether item i is correctly answered or not and $P_i(\theta)$ is the conditional probability of passing the i-th item.

The conditional passing probabilities or "item characteristic functions" P_i are generally assumed to have one of three forms

(1) (Rasch model) $P_i(\theta) = P(\theta - b_i)$

(2) (Two parameter model) $P_i(\theta) = P(a_i\theta - b_i)$

(3) (Guessing model) $P_i(\theta) = c_i + (1 - c_i)P(a_i\theta - b_i)$.

P is usually specified to be the logistic function

$$(1 + e^{-t})^{-1}$$

or the normal ogive

$$\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{t} e^{-x^2/2}\, dx \quad .$$

Our results will be applicable to a much wider class of functions P and P_i.

For a discussion of these models and basic parameter estimation results, see Birnbaum (1968).

Throughout this paper we assume that the distribution of ability is unknown. For comments on this point see section II.4.1.

Our basic result provides a mathematical justification for a distribution-free method for estimating points on item-item curves, a method which is currently being used [Levine and Saxe, 1976]. (Item-item curves are defined below. They play an essential role in the application of functional equations and group theoretical methods to psychometric problems [Levine, 1970, 1972, 1975; Levine and Saxe, 1976].)

Our results also have implications for parameter estimation, especially with equation (2) and (3). The extra generality of equations (2) and (3) seems to be needed for such important applications as the design of optimal tests, computer supported individualized testing and the detection of bias in tests. However, the estimation of item parameters (a_i, b_i, c_i) is considerably more complicated for (2) and (3) than for (1). In fact, it is not known whether any of the parameter estimation procedures that are now used (that is, used in applications in which the ability distribution is not specified) are consistent.

Our results can be used to show that a simple, common-sense way to estimate parameters (for (2) as well as (1)) is consistent. The results may be applicable to (3) as well, but an elementary step in the verification for (3) has not been proved.

It is unlikely that the estimation technique suggested in this paper will be applied in its present form as it uses data inefficiently. However, the technique seems intrinsically interesting as an application of P - P plots [Gnanadesikan and Wilk, 1968]. Furthermore, it seems important to show that at least one consistent estimation procedure is available for the general models.

II. The Model and an Informal Discussion of the Basic Result

Since the result is complicated, we first present the central finding informally, in outline form, together with some explanatory comments.

II.1 Model and Notation

We assume that the "abilities" or Θ 's are obtained by sampling from a population with a continuous density denoted by f . The i -th item score random variable u_i has the interpretation: $u_i = 1$ if item i is answered correctly and zero otherwise. P_i is called the i -th item characteristic function.

We assume that for each n , each vector of zeros and ones $< v_1, v_2, \ldots, v_n >$ and each interval of numbers T ,

$$\mathrm{Prob}\{u_1 = v_1,\ u_2 = v_2,\ \ldots\ u_n = v_n \text{ and } \Theta \in T\}$$

$$= \int_T \prod_{i=1}^{n} P_i(t)^{v_i}[1 - P_i(t)]^{1-v_i} f(t)\,dt \quad .$$

This model will be used throughout the paper.

II.2 Basic Result

We assume that each P_i is strictly increasing and continuous, that f is continuous and that these functions satisfy technical conditions detailed in section III.

Consider any two items, say item 1 and 2. The item-item curve for these items is the point set $C = C(1,2)$ given by

$$C = \{ \langle x,y \rangle \in \mathbb{R}^2 \mid \text{for some } t \,,\ x = P_1(t) \text{ and } y = P_2(t) \} .$$

Let d be the usual distance from a point $\langle x,y \rangle$ in the plane to the curve C, i.e.,

$$d(\langle x,y \rangle) = \inf_t \sqrt{[x - P_1(t)]^2 + [y - P_2(t)]^2} .$$

Denote by z_n the average

$$z_n = \frac{1}{n} \sum_{i=1}^{n} u_i .$$

This is simply the usual average score or proportion correct score on the first n items.

Assume ζ is a proportion in the range of each P_i. Then it is possible to select a sequence of positive numbers s_n decreasing to zero such that the sequence of points $\langle x_n, y_n \rangle$ in the unit square with co-ordinates given by the conditional expectations

$$x_n = \mathcal{E}(u_1 \mid |z_n - \zeta| < s_n)$$

$$y_n = \mathcal{E}(u_2 \mid |z_n - \zeta| < s_n)$$

is eventually close to C in the sense that

$$d(\langle x_n, y_n \rangle) \longrightarrow 0 .$$

II.3 Significance of Basic Result

The basic result implies that if a large sample of people is administered an n-item test, the point in the plane defined by the sample conditional proportions

$\hat{x}_n$ = proportion passing item 1 among those with proportion correct within s_n of ζ

$\hat{y}_n$ = proportion passing item 2 among those with proportion correct within s_n of ζ

will be very likely to be close to a point on the curve C. For $\langle \hat{x}_n, \hat{y}_n \rangle$ can be made arbitrarily close to $\langle x_n, y_n \rangle$ by taking a sufficiently large sample. And $\langle x_n, y_n \rangle$ can be made arbitraily close to a point on the item-item curve C by taking sufficiently large n.

These curves can be used to obtain consistency results. For suppose (2) is correct and P is, say, the logistic function $y = (1 + e^{-x})^{-1}$ with continuous inverse

$$P^{-1}(y) = -\log[\frac{1}{y} - 1] .$$

Then the transformation

$$\langle x,y \rangle \longrightarrow \langle P^{-1}(x), P^{-1}(y) \rangle$$

carries C to the straight line

$$a_1 x - b_1 = a_2 y - b_2$$

and the identifiable parameters (e.g., the slope a_1/a_2 or intercept $(b_2 - b_1)/a_2$) can be computed by Cramer's rule from any two distinct

points on C. Since the formulas of Cramer's rule are continuous, a sequence of points converging to two distinct points on C could be used to define a sequence of estimates converging to parameter values.

The case on hand is somewhat more complicated. The technical assumption made in section III will imply that if two different proportions ζ and ζ' are used to define two different sequences $<\hat{x}_n, \hat{y}_n>$ and $<\hat{x}'_n, \hat{y}'_n>$, then both will eventually be arbitrarily close to C. Generally neither sequences will converge (see section II.4.3). But it can be shown that corresponding points in the sequence will be "separated," i.e., for some positive ϵ, for all n

$$(\hat{x}_n - \hat{x}'_n)^2 + (\hat{y}_n - \hat{y}'_n)^2 > \epsilon \quad .$$

This means that the line connecting (transformed) corresponding points will converge to the transformed item-item curve line, and Cramer's rule could be used to define a convergent sequence of parameter estimates.

(We consider only two ζ's here to keep the reasoning as simple as possible. In practice it would be preferable to use many ζ's and estimate parameters by fitting a line to transformed $<\hat{x}_n, \hat{y}_n>$'s.)

The key points of the above scheme for estimating parameters for equation (2) are: (i) There is at most one item-item curve passing through two distinct points in plane and (ii) a continuous formula (Cramer's rule) is available for expressing item parameters as functions

of points in the plane. The scheme, thus, clearly could be used for the model in which P is the normal ogive, or any other such function appearing in (2).

One would like to apply this estimation scheme to the commonly used model in which P is logistic and each P_i satisfies (3). But unfortunately, I have not yet succeeded in verifying that for some n there is at most one item-item curve consistent with (3) and the logistic assumption passing through n distinct points in the plane. This elementary step and the exhibiting of a continuous inversion formula presently blocks the application of the basic result to the most commonly used model, the logistic guessing model.

The item-item curves $< P_i, P_j >$ are interesting in their own right since they play a central role in the use of group theoretical and Fourier methods to analyze test data [Levine and Saxe, 1976]. The basic result supports a distribution free method for estimating points on these curves which is currently being used. (The method is to compute the $< \hat{x}_n, \hat{y}_n >$'s for many ζ's and then compute the best fitting monotonic function to the $< \hat{x}_n, \hat{y}_n >$'s.)

II.4 Comments

1. If f is assumed to be normal, then estimation problems can be greatly simplified. However large scale studies have clearly disconfirmed normal and other guesses about the form of f. Consequently we choose to make only plausible regularity assumptions about the form of f.
2. Our method circumvents the complications of simultaneously introducing new unknown "item parameters" by increasing test length and unknown abilities Θ by increasing sample size.

3. Since proportion correct is a fraction, z_n takes only finitely many different values. Consequently we condition upon z_n being in an interval containing ζ rather than $z_n = \zeta$.

4. Convergence of proportion correct z_n to ζ does <u>not</u> imply convergence of estimated ability. In fact since blocks of easy and hard items may be interspersed even the sequence,

$$\mathcal{E}(\theta \mid z_{2n} = 1/2)$$

in which proportion correct is constant can oscillate indefinitely. <u>A fortiriori</u>, convergence of ζ_n to ζ cannot imply convergence of either

$$\mathcal{E}(u_1 \mid z_n = \zeta_n)$$

or

$$\mathcal{E}(u_2 \mid z_n = \zeta_n) \quad .$$

Thus neither

$$x_n = \mathcal{E}(u_1 \mid \ |z_n - \zeta| < s_n)$$

nor

$$y_n = \mathcal{E}(u_2 \mid \ |z_n - \zeta| < s_n)$$

need converge. All that can be proved without further assumptions is that $< x_n, y_n >$ is eventually close to the <u>curve</u> $< P_1, P_2 >$.

5. Our results are fully distribution free. The basic result remains valid even if none of the formulas (1), (2) or (3) are valid for any P.

III. Technical Details

III.1

In accord with current practice we will generally use

$$(4) \qquad \mathcal{E}\left(u_1 \,\middle|\, \left|\frac{1}{n-1}\sum_{i=2}^{n} u_i - \zeta\right| \le s_n\right)$$

rather than

$$(5) \qquad \mathcal{E}\left(u_1 \,\middle|\, \left|\frac{1}{n}\sum_{i=1}^{n} u_i - \zeta\right| \le s_n\right) .$$

That is, we condition on the average score on the test without the item being studied rather than on the first n items. In applications (4) is preferred to (5) because for short tests the quantities

$$\mathcal{E}\left(u_j \,\middle|\, \left|\frac{1}{n}\sum_{1}^{n} u_i - \zeta\right| \le s_n\right)$$

obtained for different items j are markedly interdependent [Lord and Novick, 1968, Theorem 16.4.1] Asymptotically (4) and (5) are the same, but some of the proofs using (4) are simpler.

III.2

It is well known that "true score" or proportion correct on a very long test is very nearly equal to a function of ability. We will use a more specific result, namely that under general conditions, if $s_n \longrightarrow 0$ and $ns_n^3 \longrightarrow \infty$, then for a given proportion ζ there is a constant k such that

(6) $$\text{Prob}\{ks_n \leq \text{ability} - \Theta_n \leq ks_n \mid |\text{proportion correct} - \zeta| \leq s_n\}$$

$$\longrightarrow 1$$

where Θ_n is the unique solution to the equation

$$\zeta = \frac{1}{n} \sum_{i=1}^{n} P_i(\Theta_n) \quad .$$

(The uniqueness of Θ_n will follow from the assumptions to be made about the P_i 's.)

This result will imply the key part of the basic result. For denoting

$$z_n^* = \text{average score on the first } n \text{ items of the test, except for the first item}$$

$$= \frac{1}{n-1} \sum_{2}^{n} u_i \quad .$$

and

$$x_n = \mathcal{E}(u_1 \mid |z_n^* - \zeta| \leq s_n) \quad ,$$

we have

$$x_n = \text{Prob}\{u_1 = 1 \;\&\; |\Theta - \Theta_n^*| \leq ks_n \mid |z_n^* - \zeta| \leq s_n\}$$

$$+ \text{Prob}\{u_1 = 1 \;\&\; |\Theta - \Theta_n^*| > ks_n \mid |z_n^* - \zeta| \leq s_n\}$$

where Θ = the ability random variable and

$$\Theta_n^* \text{ satisfies } \sum_{i=2}^{n} P_i(\Theta_n^*) = \zeta \quad .$$

Formula (6) implies that the second term tends to zero. The first term

$$\text{Prob}\{u_1 = 1 \;\&\; |\Theta - \Theta_n^*| \leq ks_n \,\big|\, |z_n^* - \zeta| \leq s_n\}$$

equals

$$\frac{\displaystyle\int_{\Theta_n^*-ks_n}^{\Theta_n^*+ks_n} P_1(\theta)\text{Prob}\{|z_n^* - \zeta| \leq s_n \,\big|\, \Theta = \theta\}\, f(\theta)d\theta}{\text{Prob}\{|z_n^* - \zeta| \leq s_n\}} \quad .$$

But since P_1 will be assumed to be continuous, this will be very nearly

$$P_1(\Theta_n^*)\text{Prob}\{|\Theta - \Theta_n^*| \leq ks_n \,\big|\, |z_n^* - \zeta| \leq s_n\}$$

which by (6) is asymptotically equal to

$$P_1(\Theta_n^*) \quad .$$

Our assumptions about the P_i 's will imply that $\Theta_n^* - \Theta_n \longrightarrow 0$. Thus $x_n - P_1(\Theta_n)$ will tend to zero. Similarly, if

$$y_n = \mathcal{E}(u_2 \,\big|\, |\frac{1}{n-1}(u_1 + u_3 + \dots u_n) - \zeta| \leq s_n)$$

then $y_n - P_2(\Theta_n)$ also tends to zero, and the distance between $< x_n, y_n >$

and curve $< P_1(\theta), P_2(\theta) >$ will tend to zero. Thus to establish the result considered in II.2 it is necessary only to more precisely formulate and prove (6).

III.3

In this section inequality (6) is stated more completely and proved.

Notation and Assumptions

f = "ability" density, assumed to be continuous.

$[a,b]$ a fixed interval with $a < b$.

$k = \min_{\theta \in [a,b]} f(\theta)$, assumed to be positive.

P_i = "item characteristic function," assumed to be continuous, strictly increasing, differentiable.

$\alpha = \inf_i \inf_{\theta \in [a,b]} P_i'(\theta)$, assumed to be positive.

$\beta = \sup_i \sup_{\theta \in [a,b]} P_i'(\theta)$, assumed to be finite.

ζ = a constant in the range of each P_i.

θ_n = the unique solution to $\zeta = \frac{1}{n} \sum_{i=1}^{n} P_i(\theta)$.

Θ = "ability" random variable. Has density f.

u_i, $i \geq 1$ = "item score" random variable, satisfies

$$\text{Prob}\{u_i = 1 \mid \Theta = \theta\} = P_i(\theta) .$$

z_n = "proportion correct" random variable, satisfies

$$z_n = \frac{1}{n} \sum_1^n u_i \quad .$$

ε = a positive constant. We assume for all n

$$a + \varepsilon < \theta_n < b - \varepsilon \quad .$$

<u>Asymptotic relation between proportion correct and ability</u>:

<u>If</u> $s_n \longrightarrow 0$ <u>and</u> $ns_n^3 \longrightarrow \infty$, <u>then</u>

$$\text{Prob}\{|\Theta - \theta_n| \geq 2s_n/\alpha \mid |z_n - \zeta| \leq s_n\} \longrightarrow 0 \quad .$$

<u>Proof</u>: Let $\mu_n(\theta)$ denote the conditional expectation

$$\mathcal{E}(z_n | \Theta = \theta) = \frac{1}{n} \sum_{i=1}^{n} P_i(\theta) \quad .$$

Clearly $\mu_n(\theta)$ is continuous, differentiable, strictly increasing and satisfies $\alpha < \mu_n'(\theta) < \beta$ for all n and $\theta \in [a,b]$.

The conditional variance of z_n satisfies

$$\text{Var}(z_n | \Theta = \theta) = \frac{1}{n^2} \sum_1^n P_i(\theta)[1 - P_i(\theta)] \leq \frac{1}{4n} \quad .$$

Using this in Chebyshev's inequality gives

(7) $$\text{Prob}\{|z_n - \mu_n(\Theta)| \geq \delta\} \leq \frac{1}{4n}\frac{1}{\delta^2} \quad .$$

From the triangle inequality we have

$$|\mu_n(\theta) - z_n| \geq |\mu_n(\theta) - \zeta| - |z_n - \zeta| \quad .$$

Consequently

$$\text{(8)} \qquad \text{Prob}\{|z_n - \zeta| \leq s \,|\, \Theta = \theta\} \leq \text{Prob}\{|\mu_n(\theta) - z_n| \geq |\mu_n(\theta) - \zeta| - s \,|\, \Theta = \theta\} \; .$$

For $\theta \in (a,b)$ and $0 < s < \alpha|\theta_n - \theta|$ it follows from (8), the mean value theorem and Chebyshev's inequality that

$$\text{(9)} \qquad \text{Prob}\{|z_n - \zeta| \leq s \,|\, \Theta = \theta\} \leq \frac{1}{4n}\,[\alpha|\theta_n - \theta| - s]^{-2} \quad .$$

For $s > 0$, put $U_n = (a,b) \cap [\theta_n - 2s/\alpha, \theta_n + 2s/\alpha]^c$. Then $\theta \in U_n$ implies $0 < s < 2s < \alpha|\theta - \theta_n|$, so (9) can be used to derive

$$\text{(10)} \qquad \text{Prob}\{|z_n - \zeta| \leq s \;\&\; \Theta \in U_n\}$$

$$= \int_{U_n} \text{Prob}\{|z_n - \zeta| \leq s \,|\, \Theta = \theta\} f(\theta)d\theta$$

$$= \frac{1}{4n}\int_{U_n} [\alpha|\theta_n - \theta| - s]^{-2} f(\theta)d\theta$$

$$\leq \frac{1}{4ns^2} \quad .$$

For $\theta \in (a, b)^c$ and $0 < s < \alpha\epsilon/2$

(11) $\quad \text{Prob}\{|z_n - \zeta| \leq s \mid \Theta = \theta\}$

$$\leq \text{Prob}\{|z_n - \mu_n(\theta)| \geq |\mu_n(\theta) - \zeta| - s \mid \Theta = \theta\} \quad ,$$

from (8)

$$\leq \text{Prob}\{|z_n - \mu_n(\theta)| \geq \min_{x=a,b} |\mu_n(x) - \zeta| - s \mid \Theta = \theta\} \quad ,$$

since μ_n is monotonic

$$\leq \frac{1}{4n} \cdot \frac{1}{(\alpha\epsilon - s)^2} \quad ,$$

from Chebyshev's inequality and the mean value theorem,

$$\leq \frac{1}{4n} \frac{1}{s^2} \quad ,$$

since $\alpha\epsilon > 2s$.

Combining these results gives

for $s < \alpha\epsilon/2$, $t = 2s/\alpha$

(12) $\quad \text{Prob}\{|z_n - \zeta| \leq s \ \& \ |\theta_n - \Theta| \geq t\}$

$$= \int_{U_n \cap (|\theta_n - \theta| \geq t)} \text{Prob}\{|z_n - \zeta| \leq s \mid \Theta = \theta\} f(\theta) d\theta$$

$$\leq \int_{U_n^c \cap (|\theta_n - \theta| \geq t)} \text{Prob}\{|z_n - \zeta| \leq s \,|\, \Theta = \theta\} f(\theta) d\theta$$

$$\leq \frac{1}{4ns^2} \int_{[U_n \cup (a,b)^c] \cap (|\theta_n - \theta| \geq t)} f(\theta) d\theta \leq \frac{1}{4ns^2} .$$

Finally, we obtain a bound on $\text{Prob}\{|z_n - \zeta| \leq s\}$ as follows: For $s < 2\beta\epsilon$, $|\theta - \theta_n| < s/2\beta$ implies $\theta \in (a,b)$ and

$$\text{Prob}\{|z_n - \zeta| \leq s\}$$

$$\geq \text{Prob}\{|z_n - \zeta| \leq s \;\&\; |\theta_n - \theta| < s/2\beta\}$$

$$= \int_{\theta_n - s/2\beta}^{\theta_n + s/2\beta} \text{Prob}\{|z_n - \zeta| \leq s \,|\, \Theta = \theta\} f(\theta) d\theta$$

$$\geq \int_{\theta_n \pm s/2\beta} \text{Prob}\{|z_n - \mu_n(\theta)| + |\mu_n(\theta) - \zeta| \leq s \,|\, \Theta = \theta\} f(\theta) d\theta$$

$$\geq \int_{\theta_n \pm s/2\beta} \text{Prob}\{|z_n - \mu_n(\theta)| \leq s - \beta|\theta_n - \theta| \,|\, \Theta = \theta\} f(\theta) d\theta$$

$$\geq \int_{\theta_n \pm s/2\beta} \text{Prob}\{|z_n - \mu_n(\theta)| \leq s/2 \,|\, \Theta = \theta\} f(\theta) d\theta$$

$$\geq \left(1 - \frac{1}{4n} \cdot \frac{4}{s^2}\right) \int_{\theta_n \pm s/2\beta} f(\theta) d\theta$$

$$\geq \left(1 - \frac{1}{ns^2}\right) \min_{a \leq \theta \leq b} f(\theta) \cdot s/\beta$$

$$= \left(\frac{ns^2 - 1}{ns^2}\right) s \cdot k/\beta \quad .$$

Combining this result with (12) gives for $s < \alpha\varepsilon/2,\ 2\beta/\varepsilon$

$$\text{Prob}\{|\theta - \theta_n| \geq 2s/\alpha \mid |z_n - \zeta| \leq s\}$$

$$\leq \frac{1}{s(ns^2 - 1)} \frac{\beta}{4k}$$

which tends to zero as $s \longrightarrow 0$ and $ns^3 \longrightarrow \infty$, as was to be proved.

References

Birnbaum, A. Some latent trait models and their use in inferring an examinee's ability. In F. M. Lord and M. R. Novick, Statistical theories of mental test scores. Reading, Mass.: Addison-Wesley, 1968. Chapters 17-20.

Gnanadesikan, R. and Wilk, M. B. Probability plotting methods for the analysis of data. Biometrika, 1968, 55, 1-18.

Levine, M. V. Transformations that render curves parallel. Journal of Mathematical Psychology, 1970, 7, 410-443.

Levine, M. V. Transforming curves into curves with the same shape. Journal of Mathematical Psychology, 1972, 9, 1-16.

Levine, M. V. Additive measurement with short segments of curves. Journal of Mathematical Psychology, 1975, 12, 212-224.

Levine, M. V. and Saxe, D. The use of periodic functions to measure the difficulty of aptitude test items. Research Bulletin 76-17. Princeton, N.J.: Educational Testing Service, 1976.

Lord, F. M. and Novick, M. R. Statistical theories of mental test scores. Reading, Mass.: Addison-Wesley, 1968.

Footnote

[1]I am indebted to Joseph B. Kruskal, Cheryl Reed, Donald Rubin and Marilyn Shaw for useful comments on an earlier version of this paper.