OPTIMUM SELECTION OF DIFFICULTY LEVELS FOR TEST ITEMS*

Charles A. Graver

Technical Report No. 35

University of Minnesota
Minneapolis, Minnesota

*This research was supported in part by the National Science Foundation under Grant Number G-25164.

ACKNOWLEDGEMENTS

I would like to thank Dr. Charles H. Kraft who suggested the original form of the model and problems of interest to be looked at in relation to the model, and who spent many hours of consultation with the author during the preparation of this paper. Thanks also to Dr. Robert J. Buehler, Dr. Bernard W. Lindgren and Dr. Harold Ruben who freely gave of their time for consultation. Acknowledgement is also made to the National Science Foundation who provided partial financial support for this paper under Grant G-25164.

TABLE OF CONTENTS

SUMMARY

This paper has been written from the point of view that a person's ability, in a given area, can only be measured with respect to the ability, in that area, of some other specified set of individuals. This is in contrast to the classical point of view, in item analysis, that takes as its starting point the assumption of some absolute ability. For example, Sitgreaves, [8], assumes the existence of some underlying ability Y randomly spread throughout the population with a Standard Normal Distribution. The view is taken that Y cannot be observed or measured directly and therefore, to help the experimenter, certain more observable mental abilities $X_1,\dots,X_k$ are studied. The latter are related to Y by conditional distributions which are usually given as normal. Decisions about the value of Y are then accomplished by making decisions about the value of $X_1,\dots,X_k$. The latter is accomplished by selecting items, questions, in some optimum way and observing the individual's responses to them.

In this paper a model is developed that takes an existing set of items and indexes them on some given, well-defined population. Each item is indexed by the probability of selecting an individual from the population that cannot answer the item correctly. Thus, altering the set of items will not change the indexing on those items that remain from the old set. But altering the indexing population will, in general, alter the indexes on the items. The index for a particular item is referred to as the difficulty level of that item. Thus the difficulty levels of the items depend on the indexing population.

In the model, it is assumed that the items are nested. This means that if an individual, from the indexing population, cannot answer a certain item correctly then he cannot answer any harder item correctly. In terms of difficulty levels this means that if he cannot answer an item, with a certain difficulty level correctly, then he cannot answer correctly any item with a higher difficulty level. This assumption is certainly a limitation on the model that hopefully can be weakened at a future date.

A random variable, called ability level, is then defined over the indexing population by the difficulty level of the hardest item that an individual can answer correctly. It is shown that ability level over the indexing population has a uniform [0, 1] distribution. This result stems from the indexing procedure, the assumption of nested items, and the definitions of difficulty and ability levels.

Alternative populations are then introduced. Ability level over an alternative population is then defined, with respect to the already indexed items, by the difficulty level of the hardest item that an individual can answer correctly. The distribution of ability level over an alternative population is of course not in general uniform as the items have been indexed on a different population.

In the remaining sections of the paper various questions concerning the distribution of ability level are examined. For concreteness let us assume that we are interested in spelling ability. A set of items, words, have been indexed on all 4th-6th grade school children in a given city. It is known that the ability level of 4th graders X' has a distribution, with respect to the indexed items, given by $F_{X'}$. Suppose that a year has passed and we assume that the indexing and distribution for 4th graders would still be valid for the now existing 4th-6th grade children. Suppose we are interested in estimating the ability level of a specific 4th grade student. In section II rules are given for optimally choosing the estimators and the difficulty levels to obtain a good test, where the concept of minimum expected loss is used as the criterion for a good test. Some examples for the distribution of X' are taken and the resulting optimum difficulty levels for a two item test are given below.

Distribution of X'	Optimum Difficulty Levels		Loss Function
Beta with parameters (1/2, 1)	.247	.603	$(Y-X')^2$
Beta with parameters (1/2, 1)	.209	.566	$\lvert Y-X' \rvert$
Uniform [0, 1]	.333	.666	either of the above

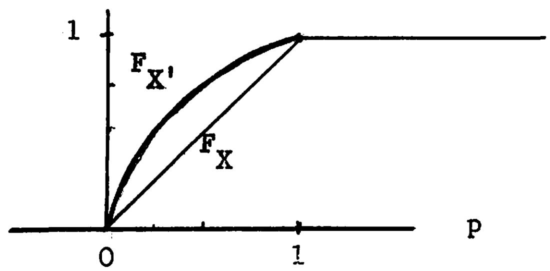
Y in the above table refers to the estimators.

Several results can be seen from the above table. First the optimum difficulty levels in general depend not only on the distribution of X' but also on the loss funtion used. Second the optimum difficulty levels when X' is uniform [0, 1] are the 33-1/3 and the 66-2/3 percentiles of the distribution of X'. However this is not the case when X' is distributed Beta (1/2, 1) as the 33-1/3 and the 66-2/3 percentiles of that distribution are .111 and .444. Finally, for the case when X' has a Beta (1/2, 1) distribution, a distribution has been choosen that is uniformly worse than the uniform [0, 1] distribution, which is the distribution of ability level for 4th-6th graders, X. That is for any item I_p, $P[X \leq I_p] \leq P[X' \leq I_p]$ or in other words

P[a 4th grader answers I_p incorrectly].

$\geq$ P[a 4th-6th grader answers I_p incorrectly].

Graphically we would have:



Since this is the case one would intuitively expect that easier questions ought to be asked to the 4th grader than to the 4th-6th grader. From the table one can see that intuition is right in this example.

In section III the problem of ranking the ability levels of two individuals from the same population, say two 4th graders, is examined. The criterion used for choosing optimum difficulty levels is that of maximizing the probability of making a correct ranking. It turns out that for an n item test one should use the $(k/(n+1))$ 100 percentiles of the distribution of X'. This is in contrast to the results of section II as the optimum difficulty levels for ranking, using the same examples that were used above, would be

Distribution of X'	Optimum Difficulty Levels n = 2	
Beta with parameters (1/2, 1)	.111	.444
Uniform [0, 1]	.333	.666

So the type of question that one is interested in answering effects the selection of the optimum difficulty levels.

The classification of an individual into one of two known populations is the problem considered in section IV. For example, let us assume that we are interested in classifying a 4th grader into one of the following:

1. Those 4th graders that can successfully complete the normal spelling program.
2. Those 4th graders that will need additional help.

It is assumed that the distribution of ability level, with respect to the items indexed on the 4th through 6th graders, is known and given by $F_{X'}$ and $F_{X''}$ for populations 1 and 2 respectively. Both the Most Powerful Test approach, using the Neymann-Pearson Lemma, and the Minimax approach, with equal losses assumed for the two types of misclassification, are presented. The results for two specific choices for the distribution of X' and X" are summarized below:

Approach	Distribution of X'	Distribution of X"	Significance Level of the Test	Optimum Difficulty Levels
Most Powerful (Significance	Beta (1/2,1)	Beta (1/3, 1)	.05	.0025
Level Specified)	Beta (6, 5)	Beta (4, 3)	.05	.2119, .7881
Minimax (Significance	Beta (1/2, 1)	Beta (1/3, 1)	.43	.185
Level Determined)	Beta (6, 5)	Beta (4, 3)	.4421	.37725, .62275

As can be seen from the table the distribution of X' and X" effect not only the optimum difficulty levels but also the number of test items required. Of course the choice of approach also effects the optimum difficulty levels.

In the final section attention is turned to testing whether or not the distribution of ability level over an alternative population, say present 4th-6th graders, is significantly different from the distribution of ability level over the indexing population, the past 4th-6th graders, where the former distribution is now assumed to be unknown. The Chi-Square Test is used in this section as the problem turns out to be a goodness of fit test against the uniform [0, 1] distribution. The criterion used for picking the difficulty levels is to minimize the deviation from the uniform [0, 1] distribution of distribution functions not differentiated from the uniform [0, 1] distribution by the Chi-Square Test. It turns out that for an n item test, the item difficulties should be chosen to be equal to $k/(n+1)$ for $k = 1, 2, \ldots, n$. These items are the $(k/(n+1))$ 100 percentiles of the uniform [0, 1] distribution.

Thus a number of questions concerning the distribution of ability level have been looked at in this paper. The list is by no means exhaustive and possible generalizations of problems looked at can be easily thought of. It is hoped, however, that the problems examined are extensive enough to show the potential uses of this model.

I. The Model

<u>General Assumptions</u>

In order to avoid any assumption about an underlying random variable or trait, either about its existence or its distribution, a model has been chosen that uses the performances of a group of individuals as a scale. The general situation is that there exists a population of individuals $\mathcal{S}$ and a set of items (questions) $\mathcal{I}$. The set of individuals from $\mathcal{S}$ that cannot answer item I from $\mathcal{I}$ correctly will be denoted by [I]. P[I] will then denote the probability of picking an individual from $\mathcal{S}$ who cannot answer item I correctly. It is assumed that $\mathcal{S}$ and $\mathcal{I}$ are interrelated by the following two assumptions hereafter referred to as the properties of the model.

1. For any real number p such that $0 \leq p \leq 1$, there exists a unique item, from $\mathcal{I}$, called I_p such that $P[I_p] = p$.
2. For any two items I_{p_1} and I_{p_2} from $\mathcal{I}$, such that $0 \leq p_1 \leq p_2 \leq 1$, we have $[I_{p_1}] \subset [I_{p_2}]$.

It becomes immediately clear, from property 1 of the model, that $\mathcal{S}$ and $\mathcal{I}$ must be very large, in fact uncountably infinite. So the situation described by the model is hypothetical. But this appears to be no more serious a restriction than assuming a continuous distribution on some underlying random variable as this restriction would also force the set of individuals to be uncountably infinite. Furthermore an approximation of the model may be set up for the practical case in which a finite population S and a finite set of items I are used. Index each item in I by using the relative frequency of incorrect responses of the individuals in S to it. Make sure that property 2 of the model is satisfied by I in relation to S. Then assume that S is a subset of a larger population $\mathcal{S}$ and I is a subset of a larger set of items $\mathcal{I}$ where $\mathcal{S}$ and $\mathcal{I}$ satisfy properties 1 and 2 of the model.

It becomes convenient, for later discussion, to refine $\mathcal{S}$ in such a way that $[I_1] = \mathcal{S}$. This involves nothing more than throwing out of $\mathcal{S}$ the set of individuals that can answer I_1 correctly. Call this set $[I_1]^c$. Since $P[I_1] = 1$ it follows that $P[I_1]^c = 0$. Because of this, the exclusion of the set $[I_1]^c$ from $\mathcal{S}$ will not effect any of the probabilities on sets $[I_p]$. Hence without loss of generality we may consider $\mathcal{S}$ and $[I_1]$ to be equal and it will be this interpretation that will be used throughout the remainder of the paper.

For any item I_p in $\mathcal{I}$, p will be called the difficulty level of the item or simply the item difficulty. It should be noted that the index p on item I_p will always refer to a probability based on $\mathcal{S}$ through property 1 of the model.

The Underlying Probability Space

By property 1 of the model, a probability measure P has been assigned to subsets of $\mathcal{S}$ called $[I_p]$, for $0 \leq p \leq 1$. These subsets have a certain ordering as defined by property 2 of the model. It is important to show that there exists a probability space that satisfies properties 1 and 2 of the model. That is it must be shown that there exists an $(\Omega, \mathcal{A}, P)$ such that

1. Ω is a set.

2. $\mathcal{A}$ is a σ-field of subsets of Ω containing Ω and $[I_p]$ for $0 \leq p \leq 1$.

and 3. P is a countably additive probability measure defined for all sets in $\mathcal{A}$ and such that $P[I_p] = p$ for all $0 \leq p \leq 1$.

To accomplish this it becomes necessary to examine another group of subsets of $\mathcal{S}$. For any p_1 and p_2 such that $0 \leq p_2 \leq p_1 \leq 1$, let $[I_{p_1} - I_{p_2}]$ be the set of individuals from $\mathcal{S}$ that can answer I_{p_2} correctly but cannot answer I_{p_1} correctly. If P is to be a probability measure it must satisfy the following:

If $B \subset A$ then $P(A-B) = P(A) - P(B)$. By property 2 of the model $[I_{p_2}] \subset [I_{p_1}]$ for $0 \leq p_2 \leq p_1 \leq 1$ and by definition $[I_{p_1} - I_{p_2}]$ $= [I_{p_1}] - [I_{p_2}]$. Therefore $P[I_{p_1} - I_{p_2}] = P[I_{p_1}] - P[I_{p_2}] = p_1 - p_2$.

Claim: The collection $\mathcal{C}$ of subsets of $\mathcal{S}$ defined by $\mathcal{C} = \{[I_p],[I_{p_1} - I_{p_2}];\ 0 \leq p \leq 1,\ 0 \leq p_2 \leq p_1 \leq 1\}$ forms a semi-ring.

By definition, ([3], page 17), "a collection of sets $\mathfrak{S}$ is called a semi-ring if it satisfies the following conditions:

1. $\mathfrak{S}$ contains the empty set $\emptyset$.

2. If $A, B \in \mathfrak{S}$, then $A \cap B \in \mathfrak{S}$.

   (ε stands for "is a member of")

3. If A and $A_1 \subset A$ are both elements of $\mathfrak{S}$, then $A = \bigcup_{k=1}^{n} A_k$ where the sets A_k are pairwise disjoint elements of $\mathfrak{S}$, and the first of the sets A_k is the given set A_1".

It is apparent that $\mathcal{C}$ satisfies these conditions and hence is a semi-ring. The proof of this appears in the appendix. Note also that $\mathcal{C}$ contains a unit, namely $\mathcal{J} = [I_1]$, as any member of $\mathcal{C}$ is contained in $\mathcal{J}$.

It must now be shown that P on $\mathcal{C}$ is a countably additive probability measure. To establish this, examine the following 1-1 correspondence between subsets of [0, 1] and the class of sets $\mathcal{C}$.

1. Let $[I_p]$ correspond to $[0, p]$ for $0 \leq p \leq 1$.

2. Let $[I_{p_1} - I_{p_2}]$ correspond to $(p_2, p_1]$ for $0 \leq p_2 \leq p_1 \leq 1$.

Let the probabilities of sets $[0, p]$ and $(p_2, p_1]$ be equal to the probability of the corresponding sets in $\mathcal{C}$. It is easy to see that P assigns to the sets $[0, p]$ and $(p_2, p_1]$ the Lebesgue Measure of the sets. Since Lebesgue Measure on [0, 1] is a countably additive probability measure, it must be a countably additive probability measure on the collection of sets $\{[0, p], (p_2, p_1]; 0 \leq p \leq 1, 0 \leq p_2 \leq p_1 \leq 1\}$. Therefore, by the 1-1 correspondence, P is a countably additive probability measure on $\mathcal{C}$.

The existence of a probability space is now easily obtained by appealing to the Lebesgue Extension Theorem, ([3], paragraph 38), which essentially states that any countably additive probability measure on a semi-ring with unit can be extended uniquely to a countably additive probability measure defined on the minimal σ-field containing the semi-ring. Furthermore this σ-field will have the same unit that the semi-ring has. Therefore $(\mathcal{J}, \sigma\{\mathcal{C}\}, P)$ is the probability space that satisfies properties 1 and 2 of the model where $\sigma\{\mathcal{C}\}$ is the minimal σ-field containing $\mathcal{C}$ and P, the unique extension of the P defined on $\mathcal{C}$, is a countably additive probability measure defined on $\sigma\{\mathcal{C}\}$.

An Interesting Random Variable

To obtain some measure of the ability, with respect to $\mathcal{J}$, for each member of $\mathcal{S}$ the following random variable will now be introduced. For each $s \varepsilon \mathcal{S}$ let

$$X(s) = \inf \{p: 0 \leq p \leq 1 \text{ and } s \text{ cannot answer } I_p \text{ correctly}\}.$$

It should be noted that by property 2 of the model, this is essentially equivalent to defining X(s) by the supremum of the difficulty levels of the items that s can answer correctly. By defining X(s) in this way the following theorem follows.

Theorem 1. X(s) is a random variable that is uniformly distributed on [0, 1].

Proof: The proof will be given in two parts.

1. X(s) is a random variable.

X(s) is a real valued function that maps $\mathcal{S}$ onto [0, 1]. Hence it must only be shown that X(s) is Borel Measurable. This can be established by showing that for any p_1 and p_2, such that $0 < p_2 \leq p_1 < 1$, $X^{-1}[p_2, p_1] \varepsilon \sigma\{\mathcal{C}\}$, for the set of closed sub-intervals of [0, 1] generate the class of Borel subsets of [0, 1].

Now for any $s \varepsilon \mathcal{S}$ with X(s) = p we have, by property 2 of the model and the definition of X(s), that

s cannot answer I_{p_1} correctly for $p_1 > p$ and

s can answer I_{p_2} correctly for $p_2 < p$.

Hence for any $[p_2, p_1]$, with $0 < p_2 \leq p_1 < 1$, it follows that

$$X^{-1}[p_2, p_1] = \{s\varepsilon\mathcal{S}: p_2 \leq X(s) \leq p_1\}$$

$$= \bigcap_{n=m}^{\infty} \left\{ s\varepsilon\mathcal{S}: \begin{array}{l} \text{s cannot answer } I_{p_1} + 1/n \text{ correctly and} \\ \text{s can answer } I_{p_2} - 1/n \text{ correctly} \end{array} \right\}$$

where $m > \max\left(\frac{1}{1-p_1}, \frac{1}{p_2}\right)$ which implies that

$p_1 + 1/m < 1$ and $p_2 - 1/m > 0$

$$= \bigcap_{n=m}^{\infty} [I_{p_1} + 1/n - I_{p_2} - 1/n] \varepsilon \sigma\{\mathcal{C}\}$$

since for each $n \geq m, [I_{p_1} + 1/n - I_{p_2} - 1/n] \varepsilon \mathcal{C} \epsilon \sigma\{\mathcal{C}\}$ and $\sigma\{\mathcal{C}\}$ is a σ-field. Hence $X(s)$ is Borel Measurable and therefore a random variable.

2. X is uniformly distributed on $[0, 1]$.

By definition $F_X(z) = P(X \leq z)$

Let $0 \leq z < 1$.

$$\begin{aligned} P(X \leq z) &= P[X \leq z | I_z \text{ is answered incorrectly}]\, P[I_z] \\ &+ P[X < z | I_z \text{ is answered correctly}]\, P[I_z]^c \\ &+ P[X = z | I_z \text{ is answered correctly}]\, P[I_z]^c \end{aligned}$$

where $[I_z]^c = \{s\varepsilon\mathcal{S}:$ s can answer I_z correctly$\}$.

Now by definition of X and property 2 of the model it follows that

$$P[X \leq z | I_z \text{ is answered incorrectly}] = 1 \text{ and}$$

$$P[X < z | I_z \text{ is answered correctly}] = 0$$

Also, by property 1, we have $P[I_z] = z$.

Furthermore, by definition of X and property 2 of the model, we have

$$\{s \varepsilon \mathcal{S}: X(s) = z | s \text{ can answer } I_z \text{ correctly}\}$$

$$= \left\{ s \varepsilon \mathcal{S}: \begin{array}{l} s \text{ can answer } I_z \text{ correctly and} \\ s \text{ cannot answer } I_p \text{ correctly for any } p > z \end{array} \right\}$$

$$= \bigcap_{n=m}^{\infty} [I_{z + 1/n} - I_z] \quad \text{where} \quad m < \frac{1}{1-z} \text{ which implies that } z + 1/m < 1.$$

Since $[I_{z + 1/n} - I_z] \supset [I_{z + 1/(n + 1)} - I_z]$ for all $n \geq m$ it follows that

$$P[I_z] \cdot P[X = z | I_z \text{ is answered correctly}] = P[\bigcap_{n=m}^{\infty} [I_{z + 1/n} - I_z]]$$

$$= \lim_{\substack{n \to \infty \\ n \geq m}} P[I_{z + 1/n} - I_z] = \lim_{\substack{n \to \infty \\ n \geq m}} (z + 1/n - z) = 0$$

Therefore for $0 \leq z < 1$

$$F_X(z) = P(X \leq z) = z.$$

Since, from the definition of X and property 1 of the model, it follows that

$$F_X(z) = 0 \quad \text{for } z < 0 \quad \text{and}$$

$$F_X(z) = 1 \quad \text{for } z \geq 1$$

we have that

$$F_X(z) = \begin{cases} 0 & \text{if } z < 0 \\ z & \text{if } 0 \leq z < 1 \\ 1 & \text{if } z \geq 1 \end{cases}$$

But this is the distribution function of the uniform [0, 1] distribution. Hence **X** has the uniform [0, 1] distribution.

It should be noted that the validity of theorem 1 stems from properties 1 and 2 of the model and the definition of **X**. It does not depend on any assumption about the existence of some underlying random variable but rather is a consequence of the model. Hereafter the value of the random variable **X**(s) will be referred to as the ability level of s.

Alternative Populations

The main interest of this paper is not in examining members of $\mathcal{S}$ but in examining members of some new population $\mathcal{S}'$. It would of course be nice if one could consider any alternative population $\mathcal{S}'$. The problems that arise are quite different for different types of alternative populations and considering all possible alternative populations would lead to very few conclusions. The class of alternative populations to be examined by this paper will therefore be limited by the following assumptions.

1. Property 2 of the model also holds for $\mathcal{S}'$. Therefore for any p_1 and p_2, with $0 \leq p_2 \leq p_1 \leq 1$,

   $\{s' \epsilon \mathcal{S}' :$ s' cannot answer I_{p_2} correctly$\}$

   $\subset \{s' \epsilon \mathcal{S}' :$ s' cannot answer I_{p_1} correctly$\}$.

2. **X'** is a continuous random variable where

   $\mathbf{X}'(s') = \inf\{p: 0 \leq p \leq 1$ and s' cannot answer I_p correctly$\}$

   for all $s' \epsilon \mathcal{S}'$. The density function of **X'**, $f_{\mathbf{X}'}$, is such that

$$f_{X'}(X) \text{ is } \begin{cases} > 0 & \text{if } x \varepsilon (0, 1) \\ = 0 & \text{otherwise} \end{cases}$$

These assumptions still leave a wide class of alternative populations to choose from. For example X' can be any Beta random variable and still fulfill constraint 2.

The value of the random variable X'(s') will be referred to as the ability level of s'. In the following sections various questions concerning the distribution of ability level will be examined.

II. Estimation of Ability Level

General Framework

A common question of interest would be to estimate $X'(s')$, that is the ability level of s', for any $s' \varepsilon \mathcal{S}'$ when the distribution of X' is known. The only information about the value of $X'(s')$ available is the responses of s' to a set of items selected from $\mathcal{S}$. For any $s' \varepsilon \mathcal{S}'$ the outcome of any test, consisting of n unique items $(I_{p_1}, I_{p_2}, \ldots, I_{p_n})$ with $0 < p_1 < p_2 < \ldots < p_n < 1$, can be described by one of the following n + 1 vectors:

$$V_0 = (0,0,\ldots,0)$$

$$V_1 = (1,0,\ldots,0)$$

$$\vdots$$

$$V_k = (1,1,\ldots,1,0,\ldots,0)$$

↑kth component

$$\vdots$$

$$V_n = (1,1,\ldots,1)$$

A one in the jth component of a vector stands for s' answering I_{p_j} correctly while a zero in the jth component stands for s' answering I_{p_j} incorrectly. This outcome description follows from taking $0 < p_1 < p_2 < \ldots < p_n < 1$ and property 2 of the model which is population invariant by assumption. (The ordering of the item difficulties was suggested by Sitgreaves [8], chapter 2, page 33).

Some justification might be needed for choosing the test $(I_{p_1}, \ldots, I_{p_n})$ such that $0 < p_1 < p_2 < \ldots < p_n < 1$. The test consists of n unique items and hence, by property 1 of the model, $p_i \neq p_j$ if $i \neq j$. Since the subscripts are unique, they may be ordered without loss of generality. Whether or not the items are asked in this order is immaterial to the evaluation of the test. It has been assumed that $f_{X'}(x) = 0$ if $x \notin (0, 1)$. Hence

$$P[s' \varepsilon \mathcal{S}': s' \text{ cannot answer } I_1 \text{ correctly}] = 1 \text{ and}$$
$$P[s' \varepsilon \mathcal{S}': s' \text{ cannot answer } I_0 \text{ correctly}] = 0 .$$

Therefore items I_0 and I_1 contain no information. Hence more information may be obtained from an n item test by picking p_1 and p_n such that $0 < p_1$ and $p_n < 1$. Thus there is no loss of generality in assuming that $0 < p_1 < p_2 < \ldots < p_n < 1$ and making this assumption provides a simple description of the possible outcomes of the test, i.e. $V_0, V_1, \ldots, V_n$.

For each outcome of the test V_k, $k=0,1,\ldots,n$, let y_k, a real number, be the estimate of X'. That is for a given individual $s' \varepsilon \mathcal{S}'$ answering the n-item test $(I_{p_1}, I_{p_2}, \ldots, I_{p_n})$ with vector V_k, y_k will be used as the estimate of X'(s'). A random variable Y may now be defined which takes on the values $y_0, y_1, \ldots, y_n$.

Before ability levels can be estimated it becomes necessary to obtain the joint probability distribution of X' and Y. Since X' is a continuous random variable and Y is a discrete random variable, the usual terminology of mass function or density function will not apply. Hence the joint probability that will be worked with in this paper will be called a probability element and will be denoted by

$g_{X',Y}(x,y_k)$. The value of the probability element is given by

$g_{X',Y}(x,y_k) = P(Y = y_k | X' = x)\, f_{X'}(x)$. Since $f_{X'}(x)$ is assumed to be known, specifying $P(Y = y_k | X' = x)$ for $k=0,1,\ldots,n$ and $0 \leq x \leq 1$ will completely determine $g_{X',Y}(x,y_k)$. For notational convenience let $p_0 = 0$ and $p_{n+1} = 1$. Then by property 2 of the model and the definition of X', it follows that, for any $k = 0,1,\ldots,n$ and any $s' \varepsilon \mathcal{S}'$, if $p_k < X'(s') < p_{k+1}$ then s' will answer the test $(I_{p_1}, I_{p_2},\ldots,I_{p_n})$ with vector V_k. Hence for $k = 0,1,\ldots,n$

$$P(Y=y_k | X'=x) = \begin{cases} 1 & \text{if } p_k < x < p_{k+1} \\ ? & \text{if } x = p_k \text{ or } x = p_{k+1} \\ 0 & \text{otherwise} \end{cases}$$

It should be noted that this conditional probability is not completely defined due to the fact that X' was defined as an infinum. This raises no problems. Since X' has been assumed to be continuous, it follows that $P(X' = p_k) = 0$ for $k = 0,1,\ldots,n$. Since one is not concerned with what happens on a set of probability zero, the conditional probabilities will arbitrarily be defined as follows. For any $k = 0,1,\ldots,n$

$$P(Y = y_k | X' = p_k) = 0 \quad \text{and}$$

$$P(Y = y_k | X' = p_{k+1}) = 1.$$

With this convention we have

$$P(Y = y_k | X' = x) = \begin{cases} 1 & \text{if } p_k < x \leq p_{k+1} \\ 0 & \text{otherwise} \end{cases}$$

and hence

$$g_{X',Y}(x,y_k) = \begin{cases} f_{X'}(x) & \text{if } p_k < x \leq p_{k+1} \\ 0 & \text{otherwise} \end{cases}$$

The problem of estimating $X'(s')$ for any $s' \varepsilon \mathcal{S}'$ with an n-item test now reduces to picking the item difficulties $p_1, p_2, \ldots, p_n$ and the estimators $y_0, y_1, \ldots, y_n$. These parameters may be picked in many different ways and therefore some criterion is needed for telling when a test defined by one set of parameters is better than or worse than another test defined by another set of parameters. The concept of loss will be used for this purpose. It will be assumed that the loss for estimating $X'(s')$ by y_k is given by either $(y_k - X'(s'))^2$ or $|y_k - X'(s')|$. The two functions

1. $(Y-X')^2$ and
2. $|Y-X'|$

will be referred to as loss functions. A good test will be defined as one that minimizes the expected value of the loss function under study.

Solutions for Good Tests

In this sub-section good tests will be found for the two loss functions mentioned above.

1. Theorem 2. A good test using the loss function $(Y-X')^2$, i.e. a test that minimizes $E(Y-X')^2$ among the class of tests that consist of n unique items $(I_{p_1}, I_{p_2}, \ldots, I_{p_n})$ from $\mathcal{S}$ with $0 < p_1 < p_2 < \ldots < p_n < 1$, is obtained if and only if the item difficulties and the estimators are chosen such that the following equations will be satisfied:

$$p_k = \frac{y_k + y_{k-1}}{2} \qquad \text{for } k = 1,2,\ldots,n$$

$$y_k = E(X' \mid p_k < X' \leq p_{k+1}) \qquad \text{for } k = 0,1,\ldots,n$$

where $p_0 = 0$ and $p_{n+1} = 1$.

The following lemma will be needed in the proof of this theorem.

Lemma: To minimize $E(Y-\hat{X})^2$ it is necessary that $y_{k-1} \leq p_k < y_k$ for $k = 1,2,\ldots,n$. For proof of the lemma see the appendix.

Proof of Theorem 2:

From the joint probability element $g_{X',Y}(x,y_k)$ it follows that

$$E(Y-X')^2 = \sum_{k=0}^{n} \int_{p_k}^{p_{k+1}} (y_k - x)^2 f_{X'}(x)\, dx .$$

For $k = 1,2,\ldots,n$

$$\frac{\partial E(Y-X')^2}{\partial p_k} = (y_{k-1} - p_k)^2 f_{X'}(p_k) - (y_k - p_k)^2 f_{X'}(p_k).$$

Since $f_{X'}(x) > 0$ for $x \varepsilon (0, 1)$ and $0 < p_k < 1$ for $k = 1,2,\ldots,n$ it follows that $f_{X'}(p_k) > 0$ for $k = 1,2,\ldots,n$. Therefore

$$\frac{\partial E(Y-X')^2}{\partial p_k} = 0 \qquad \text{if and only if}$$

$$(y_{k-1} - p_k)^2 = (y_k - p_k)^2$$

or $$y_{k-1}^2 - 2y_{k-1}\, p_k = y_k^2 - 2y_k p_k$$

or $$2p_k(y_k - y_{k-1}) = y_k^2 - y_{k-1}^2$$

But by the lemma $y_{k-1} \leq p_k < y_k$ for $E(Y-X')^2$ to be minimized.

Hence a necessary condition for $E(Y-X')^2$ to be minimized is that

$2p_k(y_k - y_{k-1}) = y_k^2 - y_{k-1}^2$ and $y_k - y_{k-1} \neq 0$ which implies that $p_k = \frac{y_k + y_{k-1}}{2}$ for $k = 1,2,\ldots,n$. Now for $k = 0,1,2,\ldots,n$

$$\frac{\partial E(Y-X')^2}{\partial y_k} = \int_{p_k}^{p_{k+1}} 2(y_k - x)\, f_{X'}(x)\, dx .$$

Therefore $\frac{\partial E(Y-X')^2}{\partial Y_k} = 0$ if and only if

$$y_k \int_{p_k}^{p_{k+1}} f_{X'}(x)dx = \int_{p_k}^{p_{k+1}} x f_{X'}(x)dx$$

or $$y_k = \frac{\int_{p_k}^{p_{k+1}} x f_{X'}(x)dx}{\int_{p_k}^{p_{k+1}} f_{X'}(x)dx} = E[X' \mid p_k < X' \leq p_{k+1}]$$

as $\int_{p_k}^{p_{k+1}} f_{X'}(x)dx > 0$ for $k = 0,1,2,\ldots,n$ for

$0 = p_0 < p_1 < p_2 < \ldots < p_n < p_{n+1} = 1$ and $f_{X'}(x) > 0$ for $x \in (0, 1)$.

Hence it has been established that picking the item difficulties and estimators that satisfy the equations stated in Theorem 2 is a necessary condition for minimizing $E(Y-X')^2$ and hence for obtaining a good test. The proof of the converse is rather long and uncomplicated and has therefore been relegated to the appendix.

2. Theorem 3: A good test using the loss function $|Y-X'|$, i.e.

a test that minimizes $E|Y-X'|$ among the class of tests that consist of n unique items $(I_{p_1}, I_{p_2}, \ldots, I_{p_n})$ from $\mathcal{J}$ with $0 < p_1 < p_2 < \ldots < p_n < 1$, is obtained if and only if the item difficulties and the estimators are chosen such that the following equations are satisfied:

$$p_k = \frac{y_k + y_{k-1}}{2} \qquad \text{for } k = 1,2,\ldots,n$$

$$F_{X'}(y_k) = \frac{F_{X'}(p_{k+1}) + F_{X'}(p_k)}{2} \qquad \text{for } k = 0,1,2,\ldots,n$$

where $p_0 = 0.$ and $p_{n+1} = 1.$

It should be noted that y_k in the second set of equations is really the conditional median of X' given that $p_k < X' \leq p_{k+1}$.

Proof of Theorem 3.

From the joint probability element $g_{X'Y}(x,y_k)$ it follows that

$$E|Y-X'| = \sum_{k=0}^{n} \int_{p_k}^{p_{k+1}} |Y_k - X| f_{X'}(x)dx$$

$$= \sum_{k=0}^{n} \left[\int_{p_k}^{y_k} (y_k - x)f_{X'}(x)dx + \int_{y_k}^{p_{k+1}} (x - y_k)f_{X'}(x)dx \right].$$

For $k = 1,2,\ldots,n$

$$\frac{\partial E|Y-X'|}{\partial p_k} = (p_k - y_{k-1})f_{X'}(p_k) - (y_k - p_k)f_{X'}(p_k).$$

Since $f_{X'}(x) > 0$ for $x \in (0, 1)$ and $0 < p_k < 1$ for $k = 1,2,\ldots,n$ it follows that $f_{X'}(p_k) > 0$ for $k = 1,2,\ldots,n$. Therefore

$$\frac{\partial E|Y-X'|}{\partial p_k} = 0 \text{ if and only if } p_k - y_{k-1} = y_k - p_k$$

or $$p_k = \frac{y_k + y_{k-1}}{2}$$

Hence a necessary condition for $E|Y-X'|$ to be minimized is that

$$p_k = \frac{y_k + y_{k-1}}{2} \quad \text{for } k = 1,2,\ldots,n.$$

Now for $k = 0,1,2,\ldots,n$

$$\frac{\partial E|Y-X'|}{\partial Y_k} = \int_{p_k}^{y_k} f_{X'}(x)dx - \int_{y_k}^{p_{k+1}} f_{X'}(x)dx.$$

Hence $\frac{\partial E|Y-X'|}{\partial Y_k} = 0$ if any only if

$$\int_{p_k}^{y_k} f_{X'}(x)dx = \int_{y_k}^{p_{k+1}} f_{X'}(x)dx \quad \text{or} \quad F_{X'}(y_k) - F_{X'}(p_k) = F_{X'}(p_{k+1}) - F_{X'}(y_k)$$

or $$F_{X'}(y_k) = \frac{F_{X'}(p_{k+1}) + F_{X'}(p_k)}{2}$$

Hence it has been established that picking the item difficulties and estimators that satisfy the equations stated in Theorem 3 is a necessary condition for minimizing $E|Y-X'|$, that is for obtaining a good test. The converse, being rather long and uncomplicated, has been relegated to the appendix.

<u>Examples Using Theorem 2 and Theorem 3</u>

1. Suppose X' is uniformly distributed on [0, 1]. Then

$$F_{X'}(x) = \begin{cases} 0 & \text{if } x \leq 0 \\ x & \text{if } 0 < x < 1 \\ 1 & \text{if } x \geq 1 \end{cases}$$

and $$f_{X'}(x) = \begin{cases} 0 & \text{if } x \leq 0 \text{ or } x \geq 1 \\ 1 & \text{if } 0 < x < 1 \end{cases}$$

a. Using the loss function $(Y-X')^2$ and Theorem 2, to obtain a good test the following system of equations must be solved.

$$\text{(i)} \quad p_k = \frac{y_k + y_{k-1}}{2} \quad \text{for } k = 1,2,\ldots,n$$

$$\text{(ii)} \quad y_k = E(X' \mid p_k < X' \leq p_{k+1}) \quad \text{for } k = 0,1,\ldots,n$$

subject to $0 = p_0 < p_1 < \ldots < p_n < p_{n+1} = 1$. In this case (ii) reduces to

$$\text{(ii)}' \quad y_k = \frac{p_k + p_{k+1}}{2} \quad \text{for } k = 0,1,\ldots,n \; .$$

The following unique solution is obtained for the set of equations in (i) and (ii)'.

$$y_k = \frac{2k+1}{2(n+1)} \qquad \text{for } k = 0,1,2,\ldots,n$$

$$\text{and} \quad p_k = \frac{2k}{2(n+1)} = \frac{k}{n+1} \quad \text{for } k = 1,2,\ldots,n$$

Hence $P_k = k/(n+1)$ which is the $(k/(n+1))100$ percentile of the distribution of X'. This result is unfortunately not true in general.

b. Using the loss function $|Y-X'|$ and Theorem 3, to obtain a good test the following system of equations must be solved.

$$\text{(i)} \quad p_k = \frac{y_k + y_{k-1}}{2} \qquad \text{for } k = 1,2,\ldots,n$$

$$\text{(ii)} \quad F(Y_k) = \frac{F(p_k) + F(p_{k+1})}{2} \qquad \text{for } k = 0,1,\ldots,n$$

subject to $0 = p_0 < p_1 < \ldots < p_n < p_{n+1} = 0$. In this case (ii) reduces to

$$\text{(ii)}' \quad y_k = \frac{p_k + p_{k+1}}{2} \qquad \text{for } k = 0,1,\ldots,n$$

This is the same set of equations that were obtained in finding a good test, using the loss function $(Y-X')^2$. Therefore in the case when X' is uniform on [0, 1], if the difficulty levels are chosen such that $p_k = k/(n+1)$

for $k = 1,\ldots,n$ and if the estimators are chosen such that $y_k = \frac{2k+1}{2(n+1)}$ for $k = 0,1,\ldots,n$, the test defined by these parameters is a good test using either the loss function $(Y-X')^2$ or the loss function $|Y-X'|$. This of course is not true in general.

It should be noted that this is the criteria for estimating X when individuals are taken from $\mathcal{I}$, the indexing population in the model, as Theorem 1 established that X has a uniform distribution on $[0, 1]$.

2. Suppose X' has the Beta distribution with parameters $(\alpha, 1)$, $\alpha > 0$. Then $f_{X'}(x) = \frac{1}{\beta(\alpha,1)} x^{\alpha-1} (1-x)^{1-1} = \frac{\Gamma(\alpha+1)}{\Gamma(\alpha)} x^{\alpha-1} = \alpha x^{\alpha-1}$ for $0 < x < 1$, and $f_{X'}(x) = 0$ for $x \leq 0$ or $x \geq 1$.

a. Using the loss function $(Y-X')^2$ and Theorem 2, to obtain a good test the following system of equations must be solved.

$$\text{(i)} \qquad p_k = \frac{y_k + y_{k-1}}{2} \qquad \text{for } k = 1,2,\ldots,n$$

$$\text{(ii)} \qquad y_k = E(X' | p_k < X' \leq p_{k+1}) = \frac{\int_{p_k}^{p_{k+1}} \alpha x^{\alpha} dx}{\int_{p_k}^{p_{k+1}} \alpha x^{\alpha-1} dx} = \frac{\alpha}{\alpha+1} \left[\frac{p_{k+1}^{\alpha+1} - p_k^{\alpha+1}}{p_{k+1}^{\alpha} - p_k^{\alpha}} \right] \qquad \text{for } k = 0,1,\ldots,n$$

subject to $0 = p_0 < p_1 < \ldots < p_n < p_{n+1} = 1$. Unhappily this system of equations cannot be easily solved. It should be noted however that

$$y_k = \frac{\int_{p_k}^{p_{k+1}} \alpha x^{\alpha} dx}{\int_{p_k}^{p_{k+1}} \alpha x^{\alpha-1} dx} = \frac{\alpha}{\alpha+1} \frac{F_{X''}(p_{k+1}) - F_{X''}(p_k)}{F_{X'}(p_{k+1}) - F_{X'}(p_k)}$$

where X'' has a Beta distribution with parameters $(\alpha + 1, 1)$. Therefore solutions can be found through the use of Tables of the Incomplete Beta Function [6].

In order to see what types of values should be given to the parameters a specific example will be taken. Suppose $n = 2$ and $\alpha = 1/2$. The solutions obtained are given in the following table. The calculations can be found in the appendix.

y_0	p_1	y_1	p_2	y_2
.082	.247	.412	.603	.793

b. To obtain a good test using the loss function $|Y-X'|$ and Theorem 3, the following system of equations must be solved.

(i) $p_k = \dfrac{y_k + y_{k-1}}{2}$ for $k = 1,2,\ldots,n$

(ii) $F_{X'}(y_k) = \dfrac{F_{X'}(p_{k+1}) + F_{X'}(p_k)}{2}$

or $y_k^{\alpha} = \dfrac{p_{k+1}^{\alpha} + p_k^{\alpha}}{2}$ for $k = 0,1,\ldots,n$.

As in part "a", this system of equations cannot be easily solved in general. It should be pointed out that, as in "a", solutions can be found by using Tables for the Incomplete Beta Function [6].

Using the same example as in "a", that is $n = 2$ and $\alpha = 1/2$, the following solution is obtained. (For calculations see the appendix).

y_0	p_1	y_1	p_2	y_2
.052	.209	.366	.566	.767

Note that the solutions are different for the two loss functions which is to be expected. Note also that the solutions for p_1 and p_2 are not the $(k/(n+1))100$ percentiles of the distribution of X', as was the case for example 1, for in this example the $(k/(n+1))100$ percentiles are given by .111 and .444.

Before going on to another problem, one final observation might be worthwhile. By choosing $\alpha = 1/2$ for the density of X' if follows that $F_{X'}(x) \geq F_X(x)$ for all real x, where X(s) is the ability level of an individual s from the indexing population. This implies that for any c, $0 < c < 1$, $P[X' \leq c] \geq P[X \leq c]$. Therefore for any item I_p from $\mathcal{A}$ the probability of picking an individual from $\mathcal{S}'$ that cannot answer I_p correctly is greater than or equal to the probability of picking an individual from $\mathcal{S}$ that cannot answer I_p correctly. In other words the distribution of ability level over $\mathcal{S}'$ is uniformly worse than the distribution of ability level over $\mathcal{S}$. In light of this observation it would be expected that the item difficulties for the good test designed to estimate X' would be less than or equal to the corresponding item difficulties for the good test designed to estimate X. Our intuition is borne out in the example, as one can easily see by the following table.

Random Variable Estimated	y_0	p_1	y_1	p_2	y_2	Loss Function
X'	.082	.247	.412	.603	.793	$(Y-X')^2$
X	.166	.333	.5	.666	.833	either
X'	.052	.209	.366	.566	.767	$\|Y-X'\|$

III Ranking of Ability Level of Two Individuals

Another common question of interest would be to rank the ability level of 2 individuals that were chosen at random from a given population $\mathcal{S}'$ for which the distribution of ability level is known. Assume that n distinct items may be chosen from $\mathcal{I}$. The problem then is to select the difficulty levels of the n items, that is choose $p_1, p_2, \ldots, p_n$ with $0 < p_1 < \ldots < p_n < 1$ in such a way as to maximize the probability of ranking correctly the ability level of the two individuals.

Let $F_{X'}$ be the distribution function of X' over the population $\mathcal{S}'$. Let $X_1' = X'(s_1)$ and $X_2' = X'(s_2)$ where s_1 and s_2 are chosen at random from $\mathcal{S}'$. It will be assumed that $\mathcal{S}'$ is large enough that the drawing of s_1 and s_2 from $\mathcal{S}'$ is essentially done with replacement and hence X_1' and X_2' are independent and identically distributed with the same distribution as X' that is $F_{X_1'}(x) = F_{X_2'}(x) = F_{X'}(x)$ for all real x.

Suppose that from the list of items $(I_{p_1}, \ldots, I_{p_n})$ there exists at least one item say I_{p_k} $k = 1,2,\ldots,n$ such that s_1 answers I_{p_k} incorrectly and s_2 answers I_{p_k} correctly. Then from property 2 of the model and the definition of X', $X_1' \leq X_2'$ and strict inequality would hold unless $X_1' = X_2' = p_k$. But X' is a continuous random variable and hence X_1' and X_2' are continuous, and therefore $P[X_1' = p_k] = P[X_2' = p_k] = 0$ for $k = 1,2,\ldots,n$. Hence this last situation may be ignored and it can be asserted with probability 1 that if s_1 answers I_{p_k} incorrectly and s_2 answers I_{p_k} correctly for at least one $k = 1,2,\ldots,n$ then $X_1' < X_2'$.

The ability level of s_1 and s_2 can then be ranked correctly, with probability one, if and only if there exists at least one item I_{p_k} in the test consisting of items $(I_{p_1}, I_{p_2}, \ldots, I_{p_n})$ from $\mathcal{I}$ such that s_1 answers I_{p_k}

correctly and s_2 answers I_{p_k} incorrectly or s_1 answers I_{p_k} incorrectly and s_2 answers I_{p_k} correctly. Call this event E. Then the problem is to pick $p_1, p_2, \ldots, p_n$ in such a way that P(E) will be maximized.

Now $P(E) = 1 - P(E^c)$ and maximizing the P(E) is therefore equivalent to minimizing $P(E^c)$, where E^c is the event that for all items I_{p_k} $k = 1,2,\ldots,n$, s_1 and s_2 answer I_{p_k} in the same way, i.e. they both answer I_{p_k} correctly or they both answer I_{p_k} incorrectly. Let $C_k = \{p_k < X_1', X_2' < p_{k+1}\}$ for $k = 0,1,\ldots,n$ where $p_0 = 0$ and $p_{n+1} = 1$. Then $E^c = \bigcup_{k=0}^{n} C_k$. For if one of the C_k occur then s_1 and s_2 will answer all items I_{p_k} in the same way. If one of the C_k does not occur then there will exist an item I_{p_k} from the list of items such that with probability one $X_1' < p_k < X_2'$ or $X_2' < p_k < X_1'$ and in such a case s_1 and s_2 will answer I_{p_k} differently by property 2 of the model and the definition of X'. Since $C_0, C_1, C_2, \ldots, C_n$ are obviously mutually disjoint it follows that $P(E^c) = P(\bigcup_{k=0}^{n} C_k) = \sum_{k=0}^{n} P(C_k)$. Since X_1' and X_2' are independent and identically distributed with the same distribution as X' it follows that:

$$P(C_k) = P(p_k < X_1', X_2' < p_{k+1}) = P(p_k < X_1' < p_{k+1})P(p_k < X_2' < p_{k+1})$$

$$= [P(p_k < X' < p_{k+1})]^2 = [F_{X'}(p_{k+1}) - F_{X'}(p_k)]^2$$

Hence $P(E^c) = \sum_{k=0}^{n} [F_{X'}(p_{k+1}) - F_{X'}(p_k)]^2$

Therefore the task becomes one of selecting $p_1, p_2, \ldots, p_n$ with $0 = p_0 < p_1 < \ldots < p_n < p_{n+1} = 1$ in such a way that $\sum_{k=0}^{n} [F_{X'}(p_{k+1}) - F_{X'}(p_k)]^2$ is minimized.

Theorem 4. To maximize the probability of E, i.e. of ranking correctly with probability 1 the ability levels of two individuals chosen at random from $\mathcal{S}'$, by using an n item test $(I_{p_1},\ldots,I_{p_n})$ such that $0 < p_1 < p_2 < \ldots < p_n < 1$, choose the item difficulties $p_1,\ldots,p_n$ in such a way that

p_k = 100(k/(n+1)) percentile of the distribution of X', that is $F_{X'}(p_k) = k/(n+1)$, $k = 1,2,\ldots,n$.

Proof: As was stated above maximizing P(E) is equivalent to minimizing $\sum_{k=0}^{n} [F_{X'}(p_{k+1}) - F_{X'}(p_k)]^2$ where $p_0 = 0$ and $p_{n+1} = 1$.

Let $a_k = F_{X'}(p_{k+1}) - F_{X'}(p_k)$. Now $\sum_{k=0}^{n} a_k = \sum_{k=0}^{n} F_{X'}(p_{k+1}) - F_{X'}(p_k) = F_{X'}(1) - F_{X'}(0) = 1$. Hence the problem reduces to minimizing $\sum_{k=0}^{n} a_k^2$ subject to $\sum_{k=0}^{n} a_k = 1$. The Cauchy--Schwarz inequality, ([1], page 6) states that if $a_0,a_1,\ldots,a_n$ and $b_0,b_1,\ldots,b_n$ are arbitrary real numbers we have

$$\left(\sum_{k=0}^{n} a_k b_k\right)^2 \leq \left(\sum_{k=0}^{n} a_k^2\right)\left(\sum_{k=0}^{n} b_k^2\right).$$

Letting $b_0 = b_1 = \ldots = b_n = 1$ it follows for the above case that

$$1 = \left(\sum_{k=0}^{n} a_k\right)^2 \leq \left(\sum_{k=0}^{n} a_k^2\right)(n+1)$$

$$\text{or } \sum_{k=0}^{n} a_k^2 \geq 1/(n+1).$$

This lower bound is achieved when $a_0 = a_1 = \ldots = a_n = 1/(n+1)$.

Hence $\sum_{k=0}^{n} (F_{X'}(p_{k+1}) - F_{X'}(p_k))^2$ is minimized if $F_{X'}(p_{k+1}) - F_{X'}(p_k) = 1/(n+1)$ for $k = 0,1,\ldots,n$. This implies that

$$F_{X'}(p_k) = k/(n+1) \text{ for } k = 1,2,\ldots,n,$$

as $F_{X'}(p_0) = F_{X'}(0) = 0$ and $F_{X'}(p_{n+1}) = F_{X'}(1) = 1$.

Therefore Theorem 4 is established.

In the above proof it has also been shown that

$P(E^c) = \sum_{k=0}^{n} (F_{X'}(p_{k+1}) - F_{X'}(p_k))^2 = 1/(n+1)$ when the item difficulties are picked optimally. This implies that $P(E) = 1 - 1/(n+1)$ when the items are optimally chosen. Thus, when items are picked optimally, the more items that one is allowed to use, the greater the probability of making a correct ranking.

Two more observations are worthwhile making at this time. First, unlike the problem of estimating ability level, the optimum solution for ranking the ability level of two individuals is attained by choosing the item difficulties p_k, $k = 1,2,\ldots,n$, to be the $(k/(n+1))$ 100 percentiles of the distribution of X'. Second, the optimum choice of difficulty levels when ranking the ability of two individuals is different, in general, than the optimum choice of difficulty levels when estimating ability level. This can be seen from the example of X' having a Beta (1/2, 1) distribution that has been worked out in Section II. Hence the type of question that one is interested in effects the optimum selection of difficulty levels.

IV Choice of Populations For an Individual

The Problem

Another possible problem that might arise is the classification of an individual as having been chosen from one of two possible populations, when the distributions of ability level over the two populations are known. Assume that an individual s has been chosen. It is known that s was taken from either $\mathcal{J}'$ or $\mathcal{J}''$. In practice, for example, $\mathcal{J}'$ might refer to those individuals that are capable of succeeding in college and $\mathcal{J}''$ those individuals that are not capable of succeeding in college. The problem is to decide which population s is a member of by selecting items from $\mathcal{J}$ and observing the responses of s to them.

Let X' and X" be the random variables, defined in section I, for $\mathcal{J}'$ and $\mathcal{J}''$ respectively and let $F_{X'}$ and $F_{X''}$ be the respective distribution functions. Let the ability level of s be a random variable Y that has the same distribution as X' or X" depending on whether s is from $\mathcal{J}'$ or $\mathcal{J}''$. Let the null hypothesis H_0 and the alternative hypothesis H_1 be

$H_0 : s \in \mathcal{J}'$

$H_1 : s \in \mathcal{J}''$.

So the problem is to decide which of the hypotheses, H_0 or H_1, is the true hypothesis by observing the responses of s to selected items from $\mathcal{J}$.

Most Powerful Test for Fixed Significance Level

There are two types of misclassification possible. They are

Type I : reject H_0 when H_0 is true

Type II: accept H_0 when H_1 is true.

If one is mainly interested in controlling the Type I misclassification, a "most powerful test" would be most appropriate. For letting

$$\alpha = P[\text{Type I misclassification}] = P_{H_0}(\text{Reject } H_0),$$

$$\beta = P[\text{Type II misclassification}] = P_{H_1}(\text{Accept } H_0)$$

and choosing a fixed α size say α_0, a "most powerful test," of significance level α_0, is defined ([5], page 238) as one that minimizes the size of β among all tests whose α size is no larger than α_0.

Now since $F_{X'}$ and $F_{X''}$ are completely specified, H_0 and H_1 are simple hypotheses. Therefore if the ability level of s were observable, i.e. $Y(s) = y$, where $0 \leq y \leq 1$, then by the Neyman-Pearson Lemma ([5],page 238) a most powerful test, of significance level α, would be given by the critical (rejection of H_0) region defined by

$$\frac{f_{X'}(y)}{f_{X''}(y)} < \text{Constant} = c$$

where c depends on α. The critical region will reduce to a region defined in terms of the ability level Y, with the bounds depending on c and hence α.

Though Y is not observable from the responses of s to a finite number of items from $\mathcal{S}$, this obstacle is of little consequence, since for any critical region, defined by a finite number of boundaries, items may be chosen from $\mathcal{S}$ such that, with probability one, it can be decided whether or not the ability level of s, Y(s), falls in the critical region. For example suppose that the critical region given by the Neyman-Pearson Lemma for a given X', X" and α turned out to be

$$Y < c' \quad \text{where } 0 < c' < 1 .$$

Now if $s \in \mathcal{S}'$ then $P[Y = c'] = P[X' = c'] = 0$ as X' is a continuous random variable. Similarly, if $s \in \mathcal{S}''$ then $P[Y = c'] = P[X'' = c'] = 0$ as X" is a continuous random variable. Hence $P[Y = c'] = 0$. Now assume that $s \in \mathcal{S}'$. If s answers $I_{c'}$ incorrectly then $Y(s) = X'(s) < c'$ with probability 1.

If s answers $I_{c'}$ correctly then, with probability 1, $Y(s) = X'(s) > c'$. These two statements follow from property 2 of the model and $P[Y = c'] = 0$. The same conclusions follow if it is assumed that $s \varepsilon \mathcal{S}''$. Therefore with probability one if s answers

1. $I_{c'}$ incorrectly, then $Y(s) < c'$ and therefore H_0 is rejected or equivalently the hypothesis that $s \varepsilon \mathcal{S}''$ is accepted;

2. $I_{c'}$ correctly, then $Y(s) > c'$ and therefore H_0 is accepted or equivalently the hypothesis that $s \varepsilon \mathcal{S}'$ is accepted.

Some more concrete examples might prove helpful in demonstrating the above.

1. Let X' have a Beta distribution with parameters $(\gamma,1)$ and X" have a Beta distribution with parameters $(\delta,1)$ where $\delta \neq \gamma$ and $\delta, \gamma > 0$.

Then $f_{X'}(y) = \gamma y^{\gamma-1} \qquad 0 < y < 1$

and $f_{X''}(y) = \delta y^{\delta-1} \qquad 0 < y < 1$

A most powerful test is given by the critical region

$$\frac{\gamma y^{\gamma-1}}{\delta y^{\delta-1}} < c \qquad 0 < y < 1$$

which implies that

$$y^{\gamma-\delta} < c' \qquad 0 < y < 1$$

or equivalently

a) $Y < c_1$ if $\gamma > \delta \qquad 0 < c_1 < 1$

b) $Y > c_2$ if $\delta > \gamma \qquad 0 < c_2 < 1$

In the case of $\gamma > \delta$, c_1 is found by specifying a significance level α. Once specified it follows that

$$\alpha = P_{H_0}[Y < c_1] = P[X' < c_1] = F_{X'}(c_1) = c_1^{\gamma}$$

therefore $\qquad c_1 = \alpha^{1/\gamma}$.

Now choose I_{c_1} from $\mathcal{S}$. If s answers

a. I_{c_1} incorrectly, then $Y < c_1$ with probability 1 and the hypothesis that $s \varepsilon \mathcal{S}''$ is accepted.

b. I_{c_1} correctly, then $Y > c_1$ with probability 1 and the hypothesis that $s \varepsilon \mathcal{S}'$ is accepted.

For the case in which $\gamma = 1/2$, $\delta < \gamma$ and $\alpha = .05$, $c_1 = (.05)^2 = .0025$. Choose $I_{.0025}$ from $\mathcal{S}$ and record the response of s to this question. Then follow the rules given in "a" and "b" above to find which hypothesis to accept at the .05 significance level.

Obvious analogous results can be worked out for the case when $\delta > \gamma$.

2. As an example of a case when two items must be used, i.e. a two tail test instead of a one tail test, let X' have a Beta distribution with parameters (γ_1, δ_1) and X" have a Beta distribution with parameters (γ_2, δ_2) where

$$\gamma_1 > \gamma_2 \text{ and } \delta_1 > \delta_2 .$$

Then $$f_{X'}(y) = \frac{y^{\gamma_1 - 1}(1-y)^{\delta_1 - 1}}{\beta(\gamma_1, \delta_1)} \qquad 0 < y < 1$$

and $$f_{X''}(y) = \frac{y^{\gamma_2 - 1}(1-y)^{\delta_2 - 1}}{\beta(\gamma_2, \delta_2)} \qquad 0 < y < 1$$

Then a most powerful test is given by

$$\frac{\beta(\gamma_2, \delta_2)\, y^{\gamma_1 - 1}\,(1-y)^{\delta_1 - 1}}{\beta(\gamma_1, \delta_1)\, y^{\gamma_2 - 1}\,(1-y)^{\delta_2 - 1}} < c \qquad 0 < y < 1$$

or equivalently

$$y^{\gamma_1 - \gamma_2}\,(1-y)^{\delta_1 - \delta_2} < c' \qquad 0 < y < 1 .$$

Let $\gamma_1 - \gamma_2 = \gamma$ and $\delta_1 - \delta_2 = \delta$. Then $y^{\gamma}(1-y)^{\delta} < c'$ where $0 < y < 1$ and $\delta, \gamma > 0$.

To see what type of critical region the Neyman-Pearson Lemma produces, examine

$$\frac{d \log y^{\gamma}(1-y)^{\delta}}{dy} .$$

$$\frac{d \log y^{\gamma}(1-y)^{\delta}}{dy} = \frac{\gamma}{y} - \frac{\delta}{1-y} = \frac{\gamma - (\gamma+\delta)y}{y(1-y)}$$

Therefore,

$$\frac{d \log y^{\gamma}(1-y)^{\delta}}{dy} \begin{cases} > 0 & y < \frac{\gamma}{\gamma+\delta} \\ = 0 & y = \frac{\gamma}{\gamma+\delta} \\ < 0 & y > \frac{\gamma}{\gamma+\delta} \end{cases}$$

and $0 < \frac{\gamma}{\gamma+\delta} < 1$ as $\gamma, \delta > 0$.

Hence $y^{\gamma}(1-y)^{\delta}$ has a maximum at $y = \frac{\gamma}{\gamma+\delta}$, is monotonically increasing for $y < \frac{\gamma}{\gamma+\delta}$, and is monotonically decreasing for $y > \frac{\gamma}{\gamma+\delta}$. Therefore the type of critical region is specified by $Y < c_1$ or $Y > c_2$ where $0 < c_1 < c_2 < 1$. Moreover c_1 and c_2 are related by

$$(*) \quad c_1^{\gamma}(1-c_1)^{\delta} = c_2^{\gamma}(1-c_2)^{\delta}$$

as c_1 and c_2 are the roots of $y^{\gamma}(1-y)^{\delta} - c' = 0$. Solving (*) for c_2 in terms of c_1 it will follow that $c_2 = g(c_1)$. Now choosing a significance level α, one finds the value of c_1 and hence c_2 by

$$\alpha = P_{H_0}(Y < c_1 \text{ or } Y > c_2) = P(X' < c_1 \text{ or } X' > c_2)$$

$$= P(X' < c_1) + P(X' > c_2) = 1 + F_{X'}(c_1) - F_{X'}(g(c_1)).$$

So the c_1 that satisfies

$\alpha = 1 + F_{X'}(c_1) - F_{X'}(g(c_1))$ will be the boundary we are looking for. c_1 may be found by using Tables of the Incomplete Beta Functions [6] for the Beta distribution with parameters (γ_1, δ_1).

After finding c_1 and hence c_2, select items I_{c_1} and I_{c_2} from $\mathcal{J}$ and carry out the test. If s answers

a. I_{c_1} correctly and I_{c_2} incorrectly then, with probability 1, $c_1 < Y < c_2$ and hence H_0 is accepted, that is accept the hypothesis that $s \epsilon \mathcal{S}'$.

b. I_{c_1} incorrectly or I_{c_2} correctly

then, with probability one, $Y < c_1$ or $Y > c_2$ and therefore H_0 is rejected, that is accept the hypothesis that $s \epsilon \mathcal{S}''$.

For the case when $\alpha = .05$, $\gamma_1 = 6$, $\gamma_2 = 4$, $\delta_1 = 5$, and $\delta_2 = 3$ it follows that $\gamma = 2$ and $\delta = 2$. Therefore it is obvious that $c_2 = 1 - c_1$ for by (*) $c_2^2(1-c_2)^2 = c_1^2(1-c_1)^2$. Now by use of Tables for the Incomplete Beta Function [6], and linear interpolation it follows that $c_1 = .2119$ and $c_2 = .7881$. For then

$$1 + F_{X'}(c_1) - F_{X'}(c_2) = 1 + .0086 - .9586 = .05 = \alpha.$$

For calculations see the appendix. Therefore choose items $I_{.2119}$ and $I_{.7881}$ from $\mathcal{S}$ and record the responses of s to these questions. Then follow the rules given in "a" and "b" above to find which hypothesis to accept at the .05 significance level.

Minimax Approach

So far attention has been confined to controlling the size of α. There are a number of procedures that might be used, which take into account both the size of α and β. To a limited extent the concept of a "most powerful test" controls the size of both α and β, but this is accomplished from the point of view of fixing α and then minimizing β. No actual control of the size of β is contained in the test.

Another type of procedure is that of the minimax approach. If one assumes that the losses for making the two types of misclassification are equal, say to c, then for any test with a given α and β, the risk associated with H_0 being the true hypothesis is $c\alpha$ and the risk associated with H_1 being the true hypothesis is $c\beta$. The minimax approach is to take the test that minimizes the maximum risk. In the present situation, the minimax solution is therefore given by the test that minimizes $\max(\alpha,\beta)$ where the minimization is taken over all possible tests.

Theorem 5: The minimax solution for the problem of classification, where the losses for the two types of misclassification are equal, is obtained by taking the test defined by the Neyman-Pearson Lemma, that is

$$\frac{f_{X'}(y)}{f_{X''}(y)} < \text{constant} = c ,$$

choosing c such that $\alpha_0 = \beta_0$ where $\alpha_0 = P_{H_0}(\text{Reject } H_0)$ and $\beta_0 = P_{H_1}(\text{Accept } H_0)$.

Proof: The proof will be given in two parts.

1. There exists no test with $(\alpha,\beta) \neq (\alpha_0,\beta_0)$ such that $\max(\alpha,\beta) < \max(\alpha_0,\beta_0)$.

Since the problem consists of two simple hypotheses, the test given by Theorem 5 is most powerful as it is defined by the Neyman-Pearson Lemma. Therefore for any test with $(\alpha,\beta) \neq (\alpha_0,\beta_0)$ it follows from the definition of a "most powerful test," that

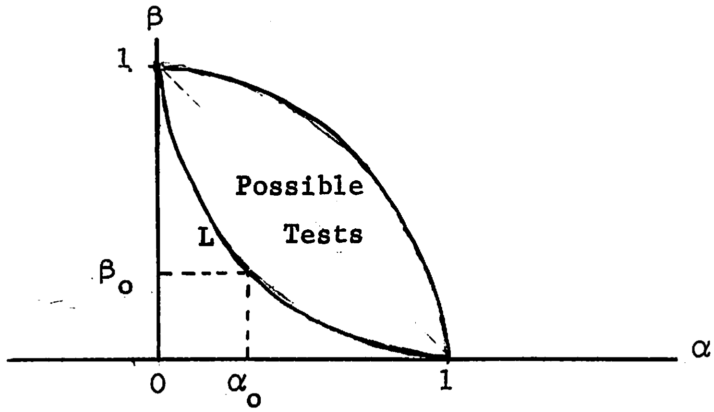
a. if $\alpha < \alpha_0$ then $\beta \geq \beta_0$ which implies that $\beta \geq \beta_0 = \alpha_0 > \alpha$ and hence $\max(\alpha,\beta) = \beta \geq \beta_0 = \max(\alpha_0,\beta_0)$.

b. if $\beta < \beta_0$ then $\alpha > \alpha_0$ which implies that $\alpha > \alpha_0 = \beta_0 > \beta$ and hence $\max(\alpha,\beta) = \alpha > \alpha_0 = \max(\alpha_0,\beta_0)$.

Since it is obvious that any test with either $\alpha > \alpha_0$ or $\beta > \beta_0$ will have $\max(\alpha,\beta) > \max(\alpha_0,\beta_0)$, the conclusion follows.

2. The minimax solution is unique.

The question of uniqueness arises from part "a" in the proof as it could only be concluded that for $\alpha < \alpha_0$, $\max(\alpha,\beta) \geq \max(\alpha_0,\beta_0)$. In all other cases the inequality was strict. However, it has been shown ([4], pages 67-68) that the set of possible tests forms a closed convex set that is symetric around the line $\alpha = 1-\beta$, contains the points (1,0) and (0,1), and is such that $0 \leq \alpha \leq 1$ and $0 \leq \beta \leq 1$.



The points on the line L represent the most powerful tests for the different α levels. It follows therefore, from the nature of the set of possible tests, that if $\alpha < \alpha_0$ then $\beta > \beta_0$ and therefore $\max(\alpha,\beta) = \beta > \beta_0 = \max(\alpha_0,\beta_0)$. Hence the test in theorem 5 is the unique minimax solution.

To demonstrate the above approach and its effect on the boundaries of the critical region, the previous examples of this section will again be examined.

1. X' has a Beta distribution with parameters $(\gamma,1)$ and X" has a Beta distribution with parameters $(\delta,1)$ where $\delta \neq \gamma$ and $\delta, \gamma > 0$.

For the case of $\gamma > \delta$, it has been shown that the critical region, given by the Neyman-Pearson Lemma, has the form $Y < c_1$. For the "most

powerful test" approach, α was specified and c_1 was found by $\alpha = c_1^\gamma$.
For the minimax solution α is not specified beyond $\alpha = \beta$. In this case

$$\beta = P_{H_1}[\text{Accept } H_0] = P_{H_1}[Y \geq c_1] = 1 - P_{H_1}[Y < c_1]$$

$$= 1 - P[X'' < c_1] = 1 - F_{X''}(c_1) = 1 - c_1^\delta .$$

Therefore c_1 is found by solving

$$c_1^\gamma = 1 - c_1^\delta .$$

For the case where $\gamma = 1/2$ and $\delta = 1/3$, $c_1 = .185$ and then $\alpha = \beta = .43$.

2. X' has a Beta distribution with parameters (γ_1, δ_1) and X" has a Beta distribution with parameters (γ_2, δ_2) where $\gamma_1 > \gamma_2$ and $\delta_1 > \delta_2$.

It has been shown that the critical region, given by the Neyman-Pearson Lemma, has the form $Y < c_1$ or $Y > c_2$, $0 < c_1 < c_2 < 1$, where $c_2 = g(c_1)$. For the "most powerful test" approach, α was specified and c_1 was found, through the use of Tables of The Incomplete Beta Function [6] by $\alpha = 1 + F_{X'}(c_1) - F_{X'}(g(c_1))$. For the minimax solution, α is not specified beyond $\alpha = \beta$. In this case we have

$$\beta = P_{H_1}(\text{Accept } H_0) = P_{H_1}(c_1 \leq Y \leq c_2) = F_{X''}(g(c_1)) - F_{X''}(c_1).$$

Then c_1 is found by use of the tables and

$$F_{X''}(g(c_1)) - F_{X''}(c_1) = 1 + F_{X'}(c_1) - F_{X'}(g(c_1)) .$$

For the case in which $\gamma_1 = 6$, $\delta_1 = 5$, $\gamma_2 = 4$ and $\delta_2 = 3$, it was found that $c_2 = 1 - c_1$. The minimax solution is then given by $c_1 = .37725$ and $c_2 = .62275$. For these boundaries $\alpha = \beta = .4421$. The solution was obtained through linear interpolation. Calculations may be found in the appendix.

A summary of the solutions obtained in this section is given below.

Approach	Example	Parameters	α	c_1	c_2
Most Powerful Test	1	$\gamma = \frac{1}{2} \quad \gamma > \delta$	.05	.0025	—
(α Specified)	2	$\gamma_1 = 6,\ \delta_1 = 5,\ \gamma_2 = 4,\ \delta_2 = 3$	.05	.2119	.7881
Minimax	1	$\gamma = \frac{1}{2} \quad \delta = \frac{1}{3}$	.43	.185	—
(α Determined)	2	$\gamma_1 = 6,\ \delta_1 = 5,\ \gamma_2 = 4,\ \delta_2 = 3$	.4421	.37725	.62275

From the table it can easily be seen that the distribution of X' and X" not only effects the values of the item difficulties, but also the number of items to be selected. Of course, the approach used also has an effect on the item difficulties.

Before going on to another type of problem, some additional points ought to be mentioned. In choosing the losses of a Type I and Type II misclassification to be equal, one is essentially assuming the same interest in controlling the two types of possible error. If the losses are assumed to be unequal, then one essentially weights the degree of control desired on the two types of misclassification. In this situation, the minimax approach can still be used but minimization is taken with respect to $\max(c_1\alpha, c_2\beta)$, instead of $\max(\alpha,\beta)$, where c_1 is the loss associated with making a Type I misclassification and c_2 is the loss associated with making a Type II misclassification.

Finally, the two approaches used in this section are not the only approaches to the problem. For instance one could use a Bayes procedure. So the approaches used in this section are by no means exhaustive.

V Comparison of the Distribution of Ability Level

The final question, concerning this model, that has been looked at is one of comparing an alternative population $\mathcal{J}'$ with the indexing population $\mathcal{J}$. An example of the usefulness of this type of question is to test the effect of geography on ability levels. For example the indexing population, $\mathcal{J}$, might be an entire city and the alternative population, $\mathcal{J}'$, the east side of that city. One might want to find out if the distribution of ability level is different over $\mathcal{J}'$ than $\mathcal{J}$.

Let X' be the random variable on $\mathcal{J}'$ with distribution function $F_{X'}$, which is assumed unknown. The problem is to pick n items from $\mathcal{J}$, in some optimum way, for the purpose of testing

$$H_0: \; F_{X'} = F_X \quad \text{against}$$

$$H_1: \; F_{X'} \neq F_X$$

where, by theorem 1, X is distributed uniformly on [0,1]. So the problem involves a goodness of fit test for $F_{X'}$ against the uniform [0,1] distribution and a natural test would be the "Chi-Square Test."

Choose any n items from $\mathcal{J}$, $(I_{p_1}, \ldots, I_{p_n})$, such that $0 < p_1 < p_2 < \ldots < p_n < 1$. Now select any individual s' from $\mathcal{J}'$. As has previously been discussed, s' will answer the test with one of the following n+1 vectors.

$$V_0 = (0,0,\ldots,0)$$

$$V_1 = (1,0,\ldots,0)$$

$$\vdots$$

$$V_n = (1,1,\ldots,1)$$

As before, the k^{th} component of the vector has a one if s' answers I_{p_k} correctly and a zero if s' answers I_{p_k} incorrectly. It has been previously shown that if s' answers the test with vector V_k, $k = 0,1,\ldots,n$, then, with probability one, $p_k < X'(s) < p_{k+1}$ where, as before, $p_0 = 0$ and $p_{n+1} = 1$. So the chosen items determine the class intervals for the Chi-Square Test.

In order to pick the items in some optimal way, a criterion for selecting optimum class intervals must be established. For any set of class intervals $(0,p_1],\ldots,(p_n,1]$, or equivalently for any set of items $(I_{p_1},\ldots,I_{p_n})$ such that $0 < p_1 < \ldots < p_n < 1$, there exists a class of distribution functions $\mathfrak{F}_{(p_1,\ldots,p_n)}$, such that for any $F \varepsilon \mathfrak{F}_{(p_1,\ldots,p_n)}$:

1. $F(x) \neq F_X(x) = x$ for $0 < x < 1$ (i.e. $F(x)$ is not uniform $[0,1]$).
2. $F(p_k) = F_X(p_k) = p_k$ for $k = 0,1,\ldots,n+1$ where $p_0 = 0$ and $p_{n+1} = 1$.

The class $\mathfrak{F}_{(p_1,\ldots,p_n)}$ then consists of all distributions that are not differentiated from the uniform $[0,1]$ distribution by the Chi-Square Test with class intervals $(0,p_1],\ldots,(p_n,1]$. This is easily seen from the following:

1. For any $F \varepsilon \mathfrak{F}_{(p_1,\ldots,p_n)}$ and any $k = 0,1,\ldots,n$,

   $F(p_{k+1}) - F(p_k) = p_{k+1} - p_k = P[p_k < X < p_{k+1}]$ where X is uniform $[0,1]$. Hence F is not differentiable from the uniform $[0,1]$ distribution by the Chi-Square Test with the given intervals.

2. For any $F \notin \mathfrak{F}_{(p_1,\ldots,p_n)}$, there exists a $k = 0,1,\ldots,n$ such that

$$F(p_{k+1}) - F(p_k) \neq p_{k+1} - p_k = P(p_k < X < p_{k+1}) \text{ where } X \text{ is}$$

uniform [0,1]. Hence F is differentiable from the uniform [0,1] distribution by the Chi-Square Test with the given intervals.

Based on this idea of non-differentiable alternative distribution functions, Wald, [10], suggested a criterion for picking optimum class intervals. Let

$$\Delta_{(p_1,\ldots,p_n)} = \sup_{\substack{0 < x < 1 \\ F \in \mathfrak{F}_{(p_1,\ldots,p_n)}}} |F(x)-x|$$

that is $\Delta_{(p_1,\ldots,p_n)}$ is the supremum of the deviation of the non-differentiable alternative distribution functions from the uniform [0,1] distribution function. Wald suggested that minimizing $\Delta_{(p_1,\ldots,p_n)}$, with respect to $(p_1,\ldots,p_n)$, would serve as a useful criterion for optimum selection of class intervals. This would force the class of non-differentiable alternative distribution functions to be made up of distribution functions that do not deviate from the uniform [0,1] distribution function too drastically.

Wald has shown that selecting $p_k = k/(n+1)$ for $k = 0,1,2,\ldots,n+1$ will minimize $\Delta_{(p_1,\ldots,p_n)}$. To see this let $p_k = k/(n+1)$ for $k = 0,1,\ldots,n+1$. Then for any $F \in \mathfrak{F}_{(p_1,\ldots,p_n)}$ and $k = 0,1,\ldots,n$,

$$|F(x)-x| \leqq p_{k+1} - p_k = \frac{k+1}{n+1} - \frac{k}{n+1} = \frac{1}{n+1} \quad \text{for } p_k \leqq x \leqq p_{k+1}$$

as $F(p_{k+1}) = p_{k+1}$ and $F(p_k) = p_k$ for $F \in \mathfrak{F}_{(p_1,\ldots,p_n)}$. Therefore

$\Delta_{(p_1,\ldots,p_n)} \leq 1/(n+1)$. Now take any other choice of $(p_1,\ldots,p_n)$ say $(p_1',\ldots,p_n')$. Since $\sum_{k=0}^{n} p_{k+1}' - p_k' = 1$ there must exist a k_0 such that $p_{k_0+1}' - p_{k_0}' = 1/(n+1) + \epsilon$ where $\epsilon > 0$.

$$\text{Define } F(x) = \begin{cases} x & \text{for } 0 \leq x \leq p_{k_0}' \text{ and } p_{k_0+1}' < x \leq 1 \\ p_{k_0+1}' & \text{for } p_{k_0}' < x \leq p_{k_0+1}' \\ 0 & \text{for } x < 0 \\ 1 & \text{for } x > 1 \end{cases}$$

Clearly $F \varepsilon \mathcal{F}_{(p_1',\ldots,p_n')}$. Now look at $x = p_{k_0}' + \delta$ where $0 < \delta < \epsilon$. Then

$$|F(p_{k_0}' + \delta) - (p_{k_0}' + \delta)| = p_{k_0+1}' - p_{k_0}' - \delta$$

$$= 1/(n+1) + \epsilon - \delta > 1/(n+1) \text{ as } \delta < \epsilon .$$

Therefore $\Delta_{(p_1',\ldots,p_n')} > 1/(n+1)$. Hence choosing items $(I_{p_1},\ldots,I_{p_n})$ such that

$p_k = k/(n+1)$ for $k = 1,\ldots,n$ satisfies Wald's criterion for optimum interval selection.

To carry out the test, select a sample of size m from $\mathcal{S}'$. Under the null hypothesis the expected number of individuals s' having $X'(s') \varepsilon (p_k, p_{k+1}]$, for $k = 0,1,\ldots,n$, is given by

$$mP_{H_0}(p_k < X' \leq p_{k+1}) = mP(p_k < X \leq p_{k+1})$$

$$= m\left(\frac{k+1}{n+1} - \frac{k}{n+1}\right) = \frac{m}{n+1} .$$

The test statistic is then calculated by $\chi^2 = \sum_{k=0}^{n} \frac{\left(f_k - \frac{m}{n+1}\right)^2}{m/(n+1)}$

where f_k = number of individuals from the sample answering the test with vector V_k.

It is known that, for large m, under the null hypothesis, χ^2 has approximately a chi-square distribution with n-1 degrees of freedom. Now for a given significance level $\alpha = P_{H_0}(\text{reject } H_0)$ find the $(1-\alpha)100$ percentile, $c_{1-\alpha}$, from chi-square tables. The critical region for the test having significance level α is then given by $\chi^2 > c_{1-\alpha}$. In other words if

1. $\chi^2 > c_{1-\alpha}$, reject H_0 i.e. accept $F_{X'} \neq F_X$

2. $\chi^2 \leq c_{1-\alpha}$, accept H_0 i.e. accept $F_{X'} = F_X$

One final remark should be made at this time. The optimum difficulty levels obtained in this section rest on the adoption of the Chi-Square Test and on the assumption that Wald's criterion is a good criterion for picking optimum class intervals. Using a different test or a different criterion for optimum class intervals will undoubtedly lead to different optimum difficulty levels.

APPENDIX

Proofs and calculations pertaining to but not contained in the body of the paper are presented here under the appropriate sections.

The Model

$\mathcal{C}$ is a semi-ring where $\mathcal{C} = \{[I_p], [I_{p_1} - I_{p_2}]; 0 \leq p \leq 1, 0 \leq p_2 \leq p_1 \leq 1\}$

Proof: It must be shown that

1. $\emptyset \varepsilon \mathcal{C}$

2. If A, B $\varepsilon \mathcal{C}$ then $A \cap B \varepsilon \mathcal{C}$.

3. If $A_1 \subset A$ and $A_1, A \varepsilon \mathcal{C}$ then there exists sets $A_2, A_3, \ldots, A_n \varepsilon \mathcal{C}$

such that

$$A_i \cap A_j = \emptyset \text{ if } i \neq j \text{ for } i,j = 1,2,\ldots,n$$

and

$$A = \bigcup_{i=1}^{n} A_i$$

1. Choose any p, $0 \leq p \leq 1$. Then $[I_p - I_p] \varepsilon \mathcal{C}$ by definition. But,

$$[I_p - I_p] = \left\{ s \varepsilon \mathcal{S} : \begin{array}{l} s \text{ cannot answer } I_p \text{ correctly and} \\ s \text{ can answer } I_p \text{ correctly} \end{array} \right\}$$

$$= \emptyset$$

Therefore $\emptyset \varepsilon \mathcal{C}$.

2. The proof of "2" will be broken up into the three possible choices for A and B such that $A, B \varepsilon \mathcal{C}$.

a. $A = [I_{p_1}]$ and $B = [I_{p_2}]$ where $0 \leq p_1 \leq 1$ and $0 \leq p_2 \leq 1$.

Then by property 2 of the model

$$A \cap B = \begin{cases} [I_{p_1}] & \text{if } p_1 \leq p_2 \\ [I_{p_2}] & \text{if } p_2 \leq p_1 \end{cases}$$

Therefore $A \cap B \varepsilon \mathcal{C}$.

b. $A = [I_p]$ and $B = [I_{p_1} - I_{p_2}]$ where $0 \leq p \leq 1$ and $0 \leq p_2 \leq p_1 \leq 1$.

Here there are three possible situations to consider

(i) If $0 \leq p_2 \leq p \leq p_1 \leq 1$ then by definition of the sets

$$[I_p] = [I_{p_2}] \cup [I_p - I_{p_2}]$$

and

$$[I_{p_1} - I_{p_2}] = [I_{p_1} - I_p] \cup [I_p - I_{p_2}]$$

where the unions are disjoint. Now $[I_{p_2}] \cap [I_{p_1} - I_{p_2}] = \emptyset$ and $[I_{p_1} - I_p] \subset [I_{p_1} - I_{p_2}]$ by definition of the sets and property 2 of the model. Hence

$$[I_{p_1} - I_p] \cap [I_{p_2}] = \emptyset$$

Therefore

$$[I_p] \cap [I_{p_1} - I_{p_2}] = [I_p - I_{p_2}]$$

(ii) If $0 \leq p_2 \leq p_1 \leq p \leq 1$ then by definition of the sets and property 2 of the model $[I_{p_1} - I_{p_2}] \subset [I_p]$ which implies that

$$[I_p] \cap [I_{p_1} - I_{p_2}] = [I_{p_1} - I_{p_2}]$$

(iii) If $0 \leq p \leq p_2 \leq p_1 \leq 1$ then

$[I_p] \subset [I_{p_2}]$ and

$[I_{p_2}] \cap [I_{p_1} - I_{p_2}] = \emptyset$ by definition of the sets and property 2 of the model. Hence $[I_p] \cap [I_{p_1} - I_{p_2}] = \emptyset$.

Therefore

$$A \cap B = \begin{cases} [I_p - I_{p_2}] & \text{if } 0 \leq p_2 \leq p \leq p_1 \leq 1 \\ [I_{p_1} - I_{p_2}] & \text{if } 0 \leq p_2 \leq p_1 \leq p \leq 1 \\ \emptyset & \text{if } 0 \leq p \leq p_2 \leq p_1 \leq 1 \end{cases}$$

Hence $A \cap B \in \mathcal{C}$.

c. $A = [I_{p_1} - I_{p_2}]$ and $B = [I_{p_3} - I_{p_4}]$ where $0 \leq p_2 \leq p_1 \leq 1$ and $0 \leq p_4 \leq p_3 \leq 1$. Here there are six possible situations to consider.

(i) If $0 \leq p_4 \leq p_2 \leq p_1 \leq p_3 \leq 1$ then by definition of the sets and property 2 of the model

$[I_{p_1} - I_{p_2}] \subset [I_{p_3} - I_{p_4}]$ which implies that

$[I_{p_1} - I_{p_2}] \cap [I_{p_3} - I_{p_4}] = [I_{p_1} - I_{p_2}]$

(ii) If $0 \leq p_4 \leq p_2 \leq p_3 \leq p_1 \leq 1$ then by the definition of the sets we have

$[I_{p_1} - I_{p_2}] = [I_{p_1} - I_{p_3}] \cup [I_{p_3} - I_{p_2}]$ and

$[I_{p_3} - I_{p_4}] = [I_{p_3} - I_{p_2}] \cup [I_{p_2} - I_{p_4}]$

where the unions are disjoint. Now by the definition of the sets and property 2 of the model

$[I_{p_1} - I_{p_3}] \subset [I_{p_1} - I_{p_2}]$ and

$[I_{p_2} - I_{p_4}] \cap [I_{p_1} - I_{p_2}] = \emptyset$

Hence $[I_{p_2} - I_{p_4}] \cap [I_{p_1} - I_{p_3}] = \emptyset$

Therefore

$$[I_{p_1} - I_{p_2}] \cap [I_{p_3} - I_{p_4}] = [I_{p_3} - I_{p_2}]$$

(iii) If $0 \leq p_4 \leq p_3 \leq p_2 \leq p_1 \leq 1$ then by definition of the sets and property 2 of the model we have

$$[I_{p_1} - I_{p_2}] \subset [I_{p_1} - I_{p_3}]$$

and

$$[I_{p_3} - I_{p_4}] \cap [I_{p_1} - I_{p_3}] = \emptyset$$

Therefore

$$[I_{p_1} - I_{p_2}] \cap [I_{p_3} - I_{p_4}] = \emptyset$$

For the other cases, that is

(iv) $0 \leq p_2 \leq p_4 \leq p_3 \leq p_1 \leq 1$

(v) $0 \leq p_2 \leq p_4 \leq p_1 \leq p_3 \leq 1$

(vi) $0 \leq p_2 \leq p_1 \leq p_4 \leq p_3 \leq 1$

interchange the roles of I_{p_4} with I_{p_2} and I_{p_3} with I_{p_1} in cases (i), (ii) and (iii) respectively

$$\text{Then } A \cap B = \begin{cases} [I_{p_1} - I_{p_2}] & \text{if } 0 \leq p_4 \leq p_2 \leq p_1 \leq p_3 \leq 1 \\ [I_{p_3} - I_{p_2}] & \text{if } 0 \leq p_4 \leq p_2 \leq p_3 \leq p_1 \leq 1 \\ \emptyset & \text{if } 0 \leq p_4 \leq p_3 \leq p_2 \leq p_1 \leq 1 \\ [I_{p_3} - I_{p_4}] & \text{if } 0 \leq p_2 \leq p_4 \leq p_3 \leq p_1 \leq 1 \\ [I_{p_1} - I_{p_4}] & \text{if } 0 \leq p_2 \leq p_4 \leq p_1 \leq p_3 \leq 1 \\ \emptyset & \text{if } 0 \leq p_2 \leq p_1 \leq p_4 \leq p_3 \leq 1 \end{cases}$$

Hence $A \cap B \in \mathcal{C}$.

Hence "2" holds for $\mathcal{C}$.

3. The proof of "3" will be broken up into the three possible choices for A and A_1 such that $A_1 \subset A$ and $A, A_1 \varepsilon \mathcal{C}$

a. $A_1 = [I_{p_1}]$ and $A = [I_p]$ where $0 \leqq p_1 \leqq p \leqq 1$.

Let $A_2 = [I_p - I_{p_1}]$. Then $A = A_1 \cup A_2$, $A_1 \cap A_2 = \emptyset$ and $A_2 \varepsilon \mathcal{C}$ by definition of the sets.

b. $A_1 = [I_{p_1} - I_{p_2}]$ and $A = [I_p]$ where $0 \leqq p_2 \leqq p_1 \leqq p \leqq 1$.

Let $A_2 = [I_{p_2}]$ and

$A_3 = [I_p - I_{p_1}]$.

Then $A_i \cap A_j = \emptyset$ if $i \neq j$ for $i,j = 1,2,3$, $A = A_1 \cup A_2 \cup A_3$ and $A_2, A_3 \varepsilon \mathcal{C}$ by definition of the sets.

c. $A_1 = [I_{p_3} - I_{p_4}]$ and $A = [I_{p_1} - I_{p_2}]$ where $0 \leqq p_2 \leqq p_4 \leqq p_3 \leqq p_1 \leqq 1$.

Let $A_2 = [I_{p_4} - I_{p_2}]$ and

$A_3 = [I_{p_1} - I_{p_3}]$

Then by definition of the sets if follows that $A_i \cap A_j = \emptyset$ if $i \neq j$ for $i,j = 1,2,3$, $A = A_1 \cup A_2 \cup A_3$ and $A_2, A_3 \varepsilon \mathcal{C}$.

Hence "3" holds for $\mathcal{C}$.

Since $\mathcal{C}$ satisfies conditions "1","2" and "3", $\mathcal{C}$ is a semi-ring.

Estimation of Ability Level

1. Lemma: In order that $E(Y-X')^2$ be minimized, it is necessary that $y_{k-1} \leqq p_k < y_k$ for $k = 1,2,\ldots,n$.

Proof:

a. $y_{k-1} \leq p_k$ for $k = 1,2,\ldots,n$. Assume $y_{k-1} > p_k$ for some $k = 1,2,\ldots,n$. Then for that k

$$(y_{k-1} - x)^2 > (p_k - x)^2 \quad \text{for all} \quad x \varepsilon (p_{k-1}, p_k] .$$

Therefore since $f_{X'}(x) > 0$ for $x \varepsilon (p_{k-1}, p_k)$ and $p_{k-1} < p_k$ it follows that

$$\int_{p_{k-1}}^{p_k} (y_{k-1} - x)^2 f_{X'}(x)dx > \int_{p_{k-1}}^{p_k} (p_k - x)^2 f_{X'}(x)dx$$

which implies that

$$E(Y-X')^2 = \sum_{j=0}^{n} \int_{p_j}^{p_{j+1}} (y_j - x)^2 f_{X'}(x)dx$$

$$> \sum_{j=0}^{n} \int_{p_j}^{p_{j+1}} (y_j' - x)^2 f_{X'}(x)dx$$

where $y_j' = y_j$ if $j \neq k - 1$ and $y_{k-1}' = p_k$.

Therefore $y_0, y_1, \ldots, y_n; p_1, \ldots, p_n$ cannot possibly minimize $E(Y-X')^2$ if $Y_{k-1} > p_k$ for some $k = 1,2,\ldots,n$ which implies that $y_{k-1} \leq p_k$ for $k = 1,2,\ldots,n$.

b. $y_k \geq p_k$ for $k = 1,2,\ldots,n$. Assume $y_k < p_k$ for some $k = 1,2,\ldots,n$. Then for that k $(y_k - x)^2 > (p_k - x)^2$ for all $x \varepsilon (p_k, p_{k+1}]$. Therefore since $f_{X'}(x) > 0$ for $x \varepsilon (p_k, p_{k+1})$ and $p_k < p_{k+1}$ it follows that

$$\int_{p_k}^{p_{k+1}} (y_k - x)^2 f_{X'}(x)dx > \int_{p_k}^{p_{k+1}} (p_k - x)^2 f_{X'}(x)dx$$

which implies that

$$E(Y-X')^2 = \sum_{j=0}^{n} \int_{p_j}^{p_{j+1}} (y_j - x)^2 f_{X'}(x)dx$$

$$> \sum_{j=0}^{n} \int_{p_j}^{p_{j+1}} (y_j' - x)^2 f_{X'}(x)dx$$

where $y_j' = y_j$ if $j \neq k$ and $y_k' = p_k$ if $j = k$.

Therefore $y_0, y_1, \ldots, y_n; p_1, \ldots, p_n$ cannot possibly minimize $E(Y-X')^2$ if $y_k < p_k$ for some $k = 1,2,\ldots,n$ which implies that $y_k \geq p_k$ for all $k = 1,2,\ldots,n$.

c. $y_k \neq p_k$ for $k = 1,2,\ldots,n$.

Assume $y_k = p_k$ for some $k = 1,2,\ldots,n$. For that k let

$$\delta_k = \int_{p_k}^{p_{k+1}} f_{X'}(x)dx \quad \text{for } k = 1,2,\ldots,n.$$ Then $0 < \delta_k$ as $f_{X'}(x) > 0$

for $x \in (p_k, p_{k+1})$ and $p_k < p_{k+1}$.

Let $\epsilon_k > 0$.

Then

$$\int_{p_k}^{p_{k+1}} (y_k - x)^2 f_{X'}(x)dx = \int_{p_k}^{p_{k+1}} (p_k - x)^2 f_{X'}(x)dx$$

$$= \int_{p_k}^{p_{k+1}} (p_k + \epsilon_k - x)^2 f_{X'}(x)\,dx - (2\epsilon_k p_k + \epsilon_k^2) \int_{p_k}^{p_{k+1}} f_{X'}(x)\,dx$$

$$+ 2\epsilon_k \int_{p_k}^{p_{k+1}} x\, f_{X'}(x)\,dx.$$

Claim: $\int_{p_k}^{p_{k+1}} x\, f_{X'}(x)\,dx = \bar{p}_k \int_{p_k}^{p_{k+1}} f_{X'}(x)\,dx$ for some $\bar{p}_k$ such that $p_k < \bar{p}_k \leq p_{k+1}$.

If this claim is true it will follow that

$$\int_{p_k}^{p_{k+1}} (y_k - x)^2 f_{X'}(x)\,dx = \int_{p_k}^{p_{k+1}} (p_k + \epsilon_k - x)^2 f_{X'}(x)\,dx$$

$$- (2\epsilon_k p_k + \epsilon_k^2)\delta_k + 2\epsilon_k \bar{p}_k \delta_k$$

$$= \int_{p_k}^{p_{k+1}} (p_k + \epsilon_k - x)^2 f_{X'}(x)\,dx + 2\epsilon_k \delta_k (\bar{p}_k - p_k - \frac{\epsilon_k}{2}).$$

Now choose ϵ_k^* such that $0 < \epsilon_k^* < 2(\bar{p}_k - p_k)$ which is possible since $\bar{p}_k > p_k$ by claim. Then $(\bar{p}_k - p_k - \frac{\epsilon_k^*}{2}) > 0$.
Hence

$$\int_{p_k}^{p_{k+1}} (y_k - x)^2 f_{X'}(x)\,dx = \int_{p_k}^{p_{k+1}} (p_k + \epsilon_k^* - x)^2 f_{X'}(x)\,dx$$

$$+ \; 2\epsilon_k^* \delta_k (\bar{p}_k - p_k - \frac{\epsilon_k^*}{2})$$

$$> \int_{p_k}^{p_{k+1}} (p_k + \epsilon_k^* - x)^2 \, f_{X'}(x)\,dx \qquad \text{as} \quad \delta_k > 0, \; \epsilon_k^* > 0 \quad \text{and} \quad (\bar{p}_k - p_k - \frac{\epsilon_k^*}{2}) > 0.$$

This implies that

$$E(Y - X')^2 = \sum_{j=0}^{n} \int_{p_j}^{p_{j+1}} (y_j - x)^2 \, f_{X'}(x)\,dx > \sum_{j=0}^{n} \int_{p_j}^{p_{j+1}} (y_j' - x)^2 \, f_{X'}(x)\,dx$$

where $y_j' = y_j$ if $j \neq k$

$y_k' = p_k + \epsilon_k^*$.

Therefore $y_0, y_1, \ldots, y_n;\; p_1, \ldots, p_n$ cannot possibly minimize $E(Y - X')^2$ if $y_k = p_k$ for some $k = 1,2,\ldots,n$ and hence $y_k \neq p_k$ for $k = 1,2,\ldots,n$.

Proof of Claim:

It must be shown that for $k = 1,2,\ldots,n$

$$\int_{p_k}^{p_{k+1}} x \, f_{X'}(x)\,dx = \bar{p}_k \int_{p_k}^{p_{k+1}} f_{X'}(x)\,dx \qquad \text{where } p_k < \bar{p}_k \leq p_{k+1}.$$

Writing $\displaystyle\int_{p_k}^{p_{k+1}} x \, f_{X'}(x)\,dx = \int_{p_k}^{p_{k+1}} x \, dF_{X'}(x)$

it follows from the mean value theorem ([7], Thm. 6.30, page 107) that

$$\int_{p_k}^{p_{k+1}} x \, dF_{X'}(x) = \bar{p}_k [F_{X'}(p_{k+1}) - F_{X'}(p_k)]$$

$$= \bar{p}_k \int_{p_k}^{p_{k+1}} dF_{X'}(x) \qquad \text{for some } \bar{p}_k \in [p_k, p_{k+1}].$$

But $$\int_{p_k}^{p_{k+1}} dF_{X'}(x) = \int_{p_k}^{p_{k+1}} f_{X'}(x)dx.$$

Therefore $$\int_{p_k}^{p_{k+1}} x \, f_{X'}(x)dx = \bar{p}_k \int_{p_k}^{p_{k+1}} f_{X'}(x)dx \quad \text{for some } \bar{p}_k \in [p_k, p_{k+1}].$$

To show that $p_k \neq \bar{p}_k$ assume that $p_k = \bar{p}_k$.

Then $$\int_{p_k}^{p_{k+1}} x \, f_{X'}(x)dx = p_k \int_{p_k}^{p_{k+1}} f_{X'}(x)dx \qquad \text{or}$$

$$\int_{p_k}^{p_{k+1}} (x - p_k) \, f_{X'}(x)dx = 0.$$

Viewing the above as a definite Lebesgue integral it follows that $(x - p_k) \, f_{X'}(x)$ must be equal to 0 almost everywhere with respect to Lebesgue measure on the set $[p_k, p_{k+1}]$. This is a consequence of theorem B, page 104 in Halmos, [2], as $(x - p_k) f_{X'}(x) \geq 0$ for $x \in [p_k, p_{k+1}]$ and the Lebesgue measure of $[p_k, p_{k+1}]$ equals $p_{k+1} - p_k$ which is greater than zero as $p_k < p_{k+1}$ for $k = 1,2,\ldots,n$. But for $k = 1,2,\ldots,n$, $(x - p_k) f_{X'}(x) > 0$ for $x \in (p_k, p_{k+1})$ and the Lebesgue measure of this set is equal to $p_{k+1} - p_k$ which is greater than zero. Therefore $(x - p_k) f_{X'}(x) \neq 0$ almost everywhere and hence a contradiction has been reached. Therefore $\bar{p}_k \neq p_k$ for $k = 1,2,\ldots,n$. Hence the claim has been proved.

The Lemma therefore follows from "a", "b" and "c" above.

2. Sufficiency of solution given in theorem 2.

Picking the item difficulties and the estimators to satisfy

$$p_k = \frac{y_k + y_{k-1}}{2} \qquad \text{for } k = 1,2,\ldots,n$$

and $y_k = E(X' | p_k < X' \leq p_{k+1})$ for $k = 0,1,2,\ldots,n$ is sufficient to minimize $E(Y - X')^2$ subject to $0 = p_0 < p_1 < \ldots < p_n < p_{n+1} = 1$.

Proof: To show this it suffices to show that $E(Y - X')^2$ is minimized for each parameter, holding the other parameters constant at the minimizing relations given by theorem 2 where the parameters are $y_0, y_1, \ldots, y_n$, $p_1, p_2, \ldots, p_n$. This will be accomplished in two parts.

a. Choose any p_k, $k = 1,2,\ldots,n$. p_k appears in $E(Y - X')^2$ only in

$$g(p_k) = \int_{p_{k-1}}^{p_k} (y_{k-1} - x)^2 f_{X'}(x)\,dx + \int_{p_k}^{p_{k+1}} (y_k - x)^2 f_{X'}(x)\,dx.$$

It must therefore be shown that

$$g(p_k) \geq g(\frac{y_k + y_{k-1}}{2}) \quad \text{if}$$

(i) $p_k > \frac{y_k + y_{k-1}}{2}$ or if

(ii) $p_k < \frac{y_k + y_{k-1}}{2}$

(i) Assume $p_k > \frac{y_k + y_{k-1}}{2}$. Examine

$$\int_{\frac{y_k + y_{k-1}}{2}}^{p_k} (y_{k-1} - x)^2 f_{X'}(x)dx.$$

Since in the evaluation of this integral $\frac{y_k + y_{k-1}}{2} \leq x$

and since $y_{k-1} < y_k$ in the solution given by theorem 2 it follows that $y_{k-1} - x < y_k - x \leq x - y_{k-1}$. Therefore $(y_k - x)^2 \leq (x - y_{k-1})^2$ for $x \in [\frac{y_k + y_{k-1}}{2}, p_k]$. Hence

$$\int_{\frac{y_k + y_{k-1}}{2}}^{p_k} (y_{k-1} - x)^2 f_{X'}(x)dx \geq \int_{\frac{y_k + y_{k-1}}{2}}^{p_k} (y_k - x)^2 f_{X'}(x)dx.$$

Therefore

$$g(p_k) = \int_{p_{k-1}}^{p_k} (y_{k-1} - x)^2 f_{X'}(x)dx + \int_{p_k}^{p_{k+1}} (y_k - x)^2 f_{X'}(x)dx$$

$$= \int_{p_{k-1}}^{\frac{y_k + y_{k-1}}{2}} (y_{k-1} - x)^2 f_{X'}(x)dx + \int_{\frac{y_k + y_{k-1}}{2}}^{p_k} (y_{k-1} - x)^2 f_{X'}(x)dx + \int_{p_k}^{p_{k+1}} (y_k - x)^2 f_{X'}(x)dx$$

$$\geq \int_{p_{k-1}}^{\frac{y_k + y_{k-1}}{2}} (y_{k-1} - x)^2 f_{X'}(x)dx + \int_{\frac{y_k + y_{k-1}}{2}}^{p_k} (y_k - x)^2 f_{X'}(x)dx + \int_{p_k}^{p_{k+1}} (y_k - x)^2 f_{X'}(x)dx$$

$$= \int_{p_{k-1}}^{\frac{y_k + y_{k-1}}{2}} (y_{k-1} - x)^2 f_{X'}(x)dx + \int_{\frac{y_k + y_{k-1}}{2}}^{p_{k+1}} (y_k - x)^2 f_{X'}(x)dx$$

$$= g\left(\frac{y_k + y_{k-1}}{2}\right)$$

Hence $g(p_k) \geq g\left(\frac{y_k + y_{k-1}}{2}\right)$ if $p_k > \frac{y_k + y_{k-1}}{2}$

(ii) Assume $p_k < \frac{y_k + y_{k-1}}{2}$.

Examine

$$\int_{p_k}^{\frac{y_k + y_{k-1}}{2}} (y_k - x)^2 f_{X'}(x)dx$$

Since in the evaluation of this integral $x \leq \frac{y_k + y_{k-1}}{2}$ and since $Y_{k-1} < y_k$ in the solution given by theorem 2 it follows that

$$x - y_k < x - y_{k-1} \leq y_k - x$$

Therefore $(y_{k-1} - x)^2 \leq (y_k - x)^2$ for $x \in [p_k, \frac{y_k + y_{k-1}}{2}]$.

Hence

$$\int_{p_k}^{\frac{y_k + y_{k-1}}{2}} (y_k - x)^2 f_{X'}(x)dx \geqq \int_{p_k}^{\frac{y_k + y_{k-1}}{2}} (y_{k-1} - x)^2 f_{X'}(x)dx.$$

Therefore

$$g(p_k) = \int_{p_{k-1}}^{p_k} (y_{k-1} - x)^2 f_{X'}(x)dx + \int_{p_k}^{p_{k+1}} (y_k - x)^2 f_{X'}(x)dx$$

$$= \int_{p_{k-1}}^{p_k} (y_{k-1} - x)^2 f_{X'}(x)dx + \int_{p_k}^{\frac{y_k + y_{k-1}}{2}} (y_k - x)^2 f_{X'}(x)dx + \int_{\frac{y_k + y_{k-1}}{2}}^{p_{k+1}} (y_k - x)^2 f_{X'}(x)dx$$

$$\geqq \int_{p_{k-1}}^{p_k} (y_{k-1} - x)^2 f_{X'}(x)dx + \int_{p_k}^{\frac{y_k + y_{k-1}}{2}} (y_{k-1} - x)^2 f_{X'}(x)dx + \int_{\frac{y_k + y_{k-1}}{2}}^{p_{k+1}} (y_k - x)^2 f_{X'}(x)dx$$

$$= \int_{p_{k-1}}^{\frac{y_k + y_{k-1}}{2}} (y_{k-1} - x)^2 f_{X'}(x)dx + \int_{\frac{y_k + y_{k-1}}{2}}^{p_{k+1}} (y_k - x)^2 f_{X'}(x)dx = g\left(\frac{y_k + y_{k-1}}{2}\right)$$

Hence $g(p_k) \geqq g\left(\frac{y_k + y_{k-1}}{2}\right)$ if $p_k < \frac{y_k + y_{k-1}}{2}$

b. Choose any y_k, $k = 0,1,2,\ldots,n$. y_k contributes to $E(Y - X')^2$ only in

$$g(y_k) = \int_{p_k}^{p_{k+1}} (y_k - x)^2 f_{X'}(x)dx.$$

Let $c_k = E[X' | p_k < X' \leqq p_{k+1}] = \dfrac{\int_{p_k}^{p_{k+1}} x f_{X'}(x)dx}{\int_{p_k}^{p_{k+1}} f_{X'}(x)dx}$

It must be shown that $g(y_k) \geq g(c_k)$ if

(i) $y_k > c_k$ or if

(ii) $y_k < c_k$.

(i) Assume that $y_k > c_k$.

Then

$$y_k - c_k = y_k - \frac{\int_{p_k}^{p_{k+1}} x f_{X'}(x)dx}{\int_{p_k}^{p_{k+1}} f_{X'}(x)dx}$$

This implies that

$$(*) \quad (y_k - c_k) \int_{p_k}^{p_{k+1}} f_{X'}(x)dx = \int_{p_k}^{p_{k+1}} (y_k - x)\, f_{X'}(x)dx$$

Now $$g(c_k) = \int_{p_k}^{p_{k+1}} (c_k - x)^2\, f_{X'}(x)dx$$

$$= \int_{p_k}^{p_{k+1}} (c_k - y_k)^2\, f_{X'}(x)dx + 2\int_{p_k}^{p_{k+1}} (c_k - y_k)(y_k - x) f_{X'}(x)dx + \int_{p_k}^{p_{k+1}} (y_k - x)^2 f_{X'}(x)dx$$

$$= (y_k - c_k)^2 \int_{p_k}^{p_{k+1}} f_{X'}(x)dx - 2(y_k - c_k) \int_{p_k}^{p_{k+1}} (y_k - x) f_{X'}(x)dx + g(y_k)$$

$$= g(y_k) - (y_k - c_k)^2 \int_{p_k}^{p_{k+1}} f_{X'}(x)dx \quad \text{by } (*)$$

$$< g(y_k) \quad \text{since} \int_{p_k}^{p_{k+1}} f_{X'}(x)dx > 0$$

Therefore $g(y_k) > g(c_k)$ if $y_k > c_k$.

(ii) Assume that $y_k < c_k$.

Then $$c_k - y_k = \frac{\int_{p_k}^{p_{k+1}} x f_{X'}(x)dx}{\int_{p_k}^{p_{k+1}} f_{X'}(x)dx} - y_k$$

This implies that

$$(**) \quad (c_k - y_k) \int_{p_k}^{p_{k+1}} f_{X'}(x)dx = \int_{p_k}^{p_{k+1}} (x - y_k) f_{X'}(x)dx$$

Now $$g(c_k) = \int_{p_k}^{p_{k+1}} (c_k - x)^2 f_{X'}(x)dx$$

$$= \int_{p_k}^{p_{k+1}} (c_k - y_k)^2 f_{X'}(x)dx + 2\int_{p_k}^{p_{k+1}} (c_k - y_k)(y_k - x) f_{X'}(x)dx + \int_{p_k}^{p_{k+1}} (y_k - x)^2 f_{X'}(x)dx$$

$$= (c_k - y_k)^2 \int_{p_k}^{p_{k+1}} f_{X'}(x)dx - 2(c_k - y_k) \int_{p_k}^{p_{k+1}} (x - y_k) f_{X'}(x)dx + g(y_k)$$

$$= g(y_k) - (c_k - y_k)^2 \int_{p_k}^{p_{k+1}} f_{X'}(x)dx \quad \text{by } (**)$$

$$< g(y_k) \quad \text{since} \quad \int_{p_k}^{p_{k+1}} f_{X'}(x)dx > 0.$$

Therefore $g(y_k) > g(c_k)$ if $y_k < c_k$. Hence by "a" and "b" the solution given by theorem 2 is sufficient for minimizing $E(Y - X')^2$.

3. Sufficiency of solution given in theorem 3. Picking the item difficulties and the estimators to satisfy

$$p_k = \frac{y_k + y_{k-1}}{2} \quad \text{for } k = 1,2,\ldots,n$$

and

$$F_{X'}(y_k) = \frac{F_{X'}(p_k) + F_{X'}(p_{k+1})}{2} \quad \text{for } k = 0,1,\ldots,n \text{ is}$$

sufficient to minimize $E|Y-X'|$ subject to $0 = p_0 < p_1 < \ldots < p_n < p_{n+1} = 1$.

Proof: As in the preceding proof, to show the above it suffices to show that $E|Y-X'|$ is minimized for each parameter, holding the other parameters constant at the minimizing relations given by theorem 3, where the parameters are $y_0, y_1, \ldots, y_n$; $p_1, p_2, \ldots, p_n$. This will be accomplished in two parts.

a. Choose any p_k, $k = 1,2,\ldots,n$. p_k contributes to $E|Y-X'|$ only in

$$g(p_k) = \int_{y_{k-1}}^{p_k} (x - y_{k-1})\, f_{X'}(x)dx + \int_{p_k}^{y_k} (y_k - x)\, f_{X'}(x)dx$$

It must be shown that

$$g(p_k) \geqq g\left(\frac{y_k + y_{k-1}}{2}\right) \quad \text{if}$$

$$\text{(i)} \quad p_k > \frac{y_k + y_{k-1}}{2} \quad \text{or if}$$

$$\text{(ii)} \quad p_k < \frac{y_k + y_{k-1}}{2}$$

(i) Assume that $p_k > \dfrac{y_k + y_{k-1}}{2}$.

Examine

$$\int_{\frac{y_k + y_{k-1}}{2}}^{p_k} (x - y_{k-1})\, f_{X'}(x)dx$$

In the evaluation of this integral $x \geqq \dfrac{y_k + y_{k-1}}{2}$ and hence $x - y_{k-1} \geqq y_k - x$. Therefore

$$\int_{\frac{y_k + y_{k-1}}{2}}^{p_k} (x - y_{k-1})\, f_{X'}(x)dx \geqq \int_{\frac{y_k + y_{k-1}}{2}}^{p_k} (y_k - x)\, f_{X'}(x)dx.$$

Hence

$$g(p_k) = \int_{y_{k-1}}^{p_k} (x - y_{k-1})\, f_{X'}(x)dx + \int_{p_k}^{y_k} (y_k - x)\, f_{X'}(x)dx$$

$$= \int_{y_{k-1}}^{\frac{y_k + y_{k-1}}{2}} (x-y_{k-1})\, f_{X'}(x)dx + \int_{\frac{y_k+ y_{k-1}}{2}}^{p_k} (x-y_{k-1})\, f_{X'}(x)dx + \int_{p_k}^{y_k} (y_k-x)f_{X'}(x)dx$$

$$\geq \int_{y_{k-1}}^{\frac{y_k + y_{k-1}}{2}} (x - y_{k-1})\, f_{X'}(x)dx + \int_{\frac{y_k+ y_{k-1}}{2}}^{p_k} (y_k- x)f_{X'}(x)dx + \int_{p_k}^{y_k} (y_k- x)\, f_{X'}(x)dx$$

$$= \int_{y_{k-1}}^{\frac{y_k + y_{k-1}}{2}} (x - y_{k-1})\, f_{X'}(x)dx + \int_{\frac{y_k+ y_{k-1}}{2}}^{y_k} (y_k- x)\, f_{X'}(x)dx = g(\frac{y_k + y_{k-1}}{2})$$

Therefore $g(p_k) \geq g(\frac{y_k + y_{k-1}}{2})$ if $p_k > \frac{y_k + y_{k-1}}{2}$.

(ii) Assume that $p_k < \frac{y_k + y_{k-1}}{2}$.

Examine $\int_{p_k}^{\frac{y_k + y_{k-1}}{2}} (y_k - x)f_{X'}(x)dx.$

In the evaluation of this integral $x \leq \frac{y_k + y_{k-1}}{2}$ and hence $x - y_{k-1} \leq y_k - x.$

Therefore

$$\int_{p_k}^{\frac{y_k + y_{k-1}}{2}} (x - y_{k-1})\, f_{X'}(x)dx \leq \int_{p_k}^{\frac{y_k + y_{k-1}}{2}} (y_k - x)\, f_{X'}(x)dx.$$

Hence $g(p_k) = \int_{y_{k-1}}^{p_k} (x - y_{k-1})\, f_{X'}(x)dx + \int_{p_k}^{y_k} (y_k - x)\, f_{X'}(x)dx$

$$= \int_{y_{k-1}}^{p_k} (x - y_{k-1})\, f_{X'}(x)dx + \int_{p_k}^{\frac{y_k + y_{k-1}}{2}} (y_k - x)\, f_{X'}(x)dx + \int_{\frac{y_k + y_{k-1}}{2}}^{y_k} (y_k - x) f_{X'}(x)dx$$

$$\geq \int_{y_{k-1}}^{p_k} (x - y_{k-1})\, f_{X'}(x)dx + \int_{p_k}^{\frac{y_k + y_{k-1}}{2}} (x - y_{k-1})\, f_{X'}(x)dx + \int_{\frac{y_k + y_{k-1}}{2}}^{y_k} (y_k - x)\, f_{X'}(x)dx$$

$$= \int_{y_{k-1}}^{\frac{y_k + y_{k-1}}{2}} (x - y_{k-1}) f_{X'}(x)dx + \int_{\frac{y_k + y_{k-1}}{2}}^{y_k} (y_k - x)\, f_{X'}(x)dx = g\left(\frac{y_k + y_{k-1}}{2}\right)$$

Therefore $g(p_k) \geq g\left(\frac{y_k + y_{k-1}}{2}\right)$ if $p_k < \frac{y_k + y_{k-1}}{2}$.

b. Choose any y_k, $k = 0,1,\ldots,n$. y_k contributes to $E|Y-X'|$ only in

$$g(y_k) = \int_{p_k}^{y_k} (y_k - x)\, f_{X'}(x)dx + \int_{y_k}^{p_{k+1}} (x - y_k)\, f_{X'}(x)dx.$$

Let c_k be such that $F_{X'}(c_k) = \frac{F_{X'}(p_{k+1}) + F_{X'}(p_k)}{2}$

Then it must be shown that $g(y_k) \geq g(c_k)$ if

(i) $y_k > c_k$ or if

(ii) $y_k < c_k$.

Since $F_{X'}(c_k) = \dfrac{F_{X'}(p_k) + F_{X'}(p_{k+1})}{2}$ it follows that

$$(*) \qquad \int_{p_k}^{c_k} f_{X'}(x)dx = \int_{c_k}^{p_{k+1}} f_{X'}(x)dx.$$

(i) Assume that $y_k > c_k$.

$$\text{Then } g(y_k) = \int_{p_k}^{y_k} (y_k - x) f_{X'}(x)dx + \int_{y_k}^{p_{k+1}} (x - y_k) f_{X'}(x)dx$$

$$= \int_{p_k}^{c_k} (y_k-x)f_{X'}(x)dx + \int_{c_k}^{y_k} (y_k-x)f_{X'}(x)dx + \int_{c_k}^{p_{k+1}} (x-y_k)f_{X'}(x)dx - \int_{c_k}^{y_k}(x-y_k)f_{X'}(x)dx.$$

$$= \int_{p_k}^{c_k} (y_k-x) f_{X'}(x)dx + \int_{c_k}^{p_{k+1}} (x-y_k) f_{X'}(x)dx + 2\int_{c_k}^{y_k} (y_k- x) f_{X'}(x)dx$$

$$\geq \int_{p_k}^{c_k} (y_k- x) f_{X'}(x)dx + \int_{c_k}^{p_{k+1}} (x-y_k)f_{X'}(x)dx \quad \text{as} \quad \int_{c_k}^{y_k} (y_k- x) f_{X'}(x)dx \geq 0$$

for $(y_k - x) \geq 0$ when $x \in [c_k, y_k]$

$$\text{so} \quad g(y_k) \geq \int_{p_k}^{c_k} (y_k- x) f_{X'}(x)dx + \int_{c_k}^{p_{k+1}} (x-y_k) f_{X'}(x)dx$$

$$= \int_{p_k}^{c_k} (c_k - x)\, f_{X'}(x)dx + \int_{p_k}^{c_k} (y_k - c_k)\, f_{X'}(x)dx$$

$$+ \int_{c_k}^{p_{k+1}} (x - c_k)\, f_{X'}(x)dx + \int_{c_k}^{p_{k+1}} (c_k - y_k)\, f_{X'}(x)dx$$

$$= \int_{p_k}^{c_k} (c_k - x)\, f_{X'}(x)dx + \int_{c_k}^{p_{k+1}} (x - c_k) f_{X'}(x)dx + (y_k - c_k)\left[\int_{p_k}^{c_k} f_{X'}(x)dx - \int_{c_k}^{p_{k+1}} f_{X'}(x)dx\right]$$

$= g(c_k)$ by (*). Therefore $g(y_k) \geq g(c_k)$ if $y_k > c_k$.

(ii) Assume that $y_k < c_k$. Then

$$g(y_k) = \int_{p_k}^{y_k} (y_k - x)\, f_{X'}(x)dx + \int_{y_k}^{p_{k+1}} (x - y_k)\, f_{X'}(x)dx$$

$$= \int_{p_k}^{c_k} (y_k - x) f_{X'}(x)dx - \int_{y_k}^{c_k} (y_k - x) f_{X'}(x)dx + \int_{c_k}^{p_{k+1}} (x - y_k) f_{X'}(x)dx + \int_{y_k}^{c_k} (x - y_k) f_{X'}(x)dx$$

$$= \int_{p_k}^{c_k} (y_k - x)\, f_{X'}(x)dx + \int_{c_k}^{p_{k+1}} (x - y_k)\, f_{X'}(x)dx + 2\int_{y_k}^{c_k} (x - y_k)\, f_{X'}(x)dx$$

$$\geq \int_{p_k}^{c_k} (y_k - x)\, f_{X'}(x)dx + \int_{c_k}^{p_{k+1}} (x - y_k)\, f_{X'}(x)dx \quad \text{as} \quad \int_{y_k}^{c_k} (x - y_k)\, f_{X'}(x)dx \geq 0$$

for $(x - y_k) \geq 0$ when $x \in [y_k, c_k]$.

$$\text{So} \quad g(y_k) \geq \int_{p_k}^{c_k} (y_k - x)\, f_{X'}(x)\,dx + \int_{c_k}^{p_{k+1}} (x - y_k)\, f_{X'}(x)\,dx$$

$$= \int_{p_k}^{c_k} (c_k - x)\, f_{X'}(x)\,dx + \int_{p_k}^{c_k} (y_k - c_k)\, f_{X'}(x)\,dx$$

$$+ \int_{c_k}^{p_{k+1}} (x - c_k)\, f_{X'}(x)\,dx + \int_{c_k}^{p_{k+1}} (c_k - y_k)\, f_{X'}(x)\,dx$$

$$= \int_{p_k}^{c_k} (c_k - x)\, f_{X'}(x)\,dx + \int_{c_k}^{p_{k+1}} (x - c_k)\, f_{X'}(x)\,dx$$

$$+ (c_k - y_k)\left[\int_{c_k}^{p_{k+1}} f_{X'}(x)\,dx - \int_{p_k}^{c_k} f_{X'}(x)\,dx\right]$$

$$= g(c_k) \qquad \text{by } (*).$$

Therefore $g(y_k) \geq g(c_k)$ if $y_k < c_k$.

Hence by "a" and "b" the solution given by theorem 3 is sufficient for minimizing $E|Y - X'|$.

4. Calculations of solutions for a good test when X' has a Beta distribution with parameters (1/2, 1) and n=2.

a. Loss function $(Y - X')^2$.

The equations to be solved are

$$\text{(i)} \quad y_k = 1/3\left[\frac{p_{k+1}^{3/2} - p_k^{3/2}}{p_{k+1}^{1/2} - p_k^{1/2}}\right] = 1/3[p_{k+1} + (p_k p_{k+1})^{1/2} + p_k]$$

for k = 0,1,2.

and

$$\text{(ii)} \quad p_k = \frac{y_k + y_{k-1}}{2} \qquad \text{for } k = 1,2.$$

1. $y_0 = 1/3\ p_1$

Let k=0 in (i).

2. $y_1 = 5/3\ p_1$

From (ii) with k=1 it follows that

$y_1 = 2p_1 - y_0 = 5/3\ p_1$.

3. $p_2 = 2.44p_1$

From (i) with k=1 it follows that

$3y_1 = p_2 + (p_2p_1)^{1/2} + p_1$

or

$5p_1 = p_2 + (p_2p_1)^{1/2} + p_1$

or

$(4p_1 - p_2) = (p_2p_1)^{1/2}$

or

$p_2^2 - 8p_1p_2 + 16p_1^2 = p_1p_2$

or

$p_2^2 - 9p_1p_2 + 16p_1^2 = 0$

or

$$p_2 = \frac{9p_1 \pm \sqrt{81p_1^2 - 64p_1^2}}{2}$$

$$= \frac{p_1}{2}[9 \pm \sqrt{17}] = 2.44p_1$$

or $6.56p_1$ (extraneous).

4. $y_2 = 3.21p_1$

From (ii) with k=2 it follows that

$$y_2 = 2p_2 - y_1$$

$$= 2(2.44)p_1 - 1.67p_1 = 3.21p_1.$$

5. $p_1 = .247$

From (i) with k=2 it follows that

$$3y_2 = 1 + (p_2)^{1/2} + p_2$$

or

$$3(3.21)p_1 = 1 + \sqrt{2.44p_1} + 2.44p_1$$

or

$$7.19p_1 = 1 + 1.56\sqrt{p_1}$$

or

$$7.19p_1 - 1.56\sqrt{p_1} - 1 = 0.$$

Therefore $\sqrt{p_1} = \frac{1.56 \pm \sqrt{2.44 + 4(7.19)}}{14.38}$

$= .497$ or $-.28$ (extraneous)

Therefore $p_1 = (.497)^2 = .247.$

It therefore follows from the above that

$y_0 = 1/3\ p_1 = \frac{.247}{3} = .082.$

$p_1 = .247.$

$y_1 = 5/3\ p_1 = \frac{5(.247)}{3} = .412.$

$p_2 = 2.44p_1 = (2.44)(.247) = .603.$

$y_2 = 3.21p_1 = (3.21)(.247) = .793.$

b. Loss function $|Y - X'|$.

The equations to be solved are

(i) $$y_k^{1/2} = \frac{p_{k+1}^{1/2} + p_k^{1/2}}{2} \qquad k = 0,1,2.$$

and

(ii) $$p_k = \frac{y_k + y_{k-1}}{2} \qquad k = 1,2.$$

1. $y_0 = .25p_1$

Let k=0 in (i). Then

$$y_0^{1/2} = \frac{p_1^{1/2}}{2} \quad \text{or} \quad y_0 = \frac{p_1}{4}.$$

2. $y_1 = 1.75p_1$

Let k=1 in (ii). Then

$$y_1 = 2p_1 - y_0 = 1.75p_1.$$

3. $p_2 = 2.71p_1$

From (i) with k=1 it follows that

$$2y_1^{1/2} = p_2^{1/2} + p_1^{1/2}$$

$$2\sqrt{1.75}\, p_1^{1/2} = p_2^{1/2} + p_1^{1/2}$$

$$1.646\, p_1^{1/2} = p_2^{1/2}.$$

Therefore $p_2 = 2.71p_1$

4. $y_2 = 3.67p_1$

From (ii) with k=2 it follows that

$$y_2 = 2p_2 - y_1 = 2(2.71)p_1 - 1.75p_1$$
$$= 3.67p_1.$$

5. $p_1 = .209$

Let k=2 in (i). Then

$$2y_2^{1/2} = 1 + p_2^{1/2}$$

$$(2\sqrt{3.67} - \sqrt{2.71})p_1^{1/2} = 1$$

or $4.79p_1 = 1$.

Therefore $p_1 = \frac{1}{4.79} = .209$.

It therefore follows from the above that

$$y_0 = .25p_1 = .052.$$

$$p_1 = .209.$$

$$y_1 = 1.75p_1 = .366.$$

$$p_2 = 2.71p_1 = .566.$$

$$y_2 = 3.67p_1 = .767.$$

<u>Choice of Populations for an Individual</u>

Values from Tables of the Incomplete Beta function and Linear Interpolation for the example in which

X' has the Beta Distribution with parameters (6, 5)

and

X" has the Beta Distribution with parameters (4, 3).

a. Most powerful test with $\alpha = .05$.

	From Incomplete Beta Tables		Linear Interpolation
c_1	.21	.22	.2119
$F_{X'}(c_1)$	.0082	.0104	.0086
$F_{X'}(1-c_1)$	.9601	.9521	.9586
α	.0481	.0583	.0500

where $\alpha = 1 - F_{X'}(1-c_1) + F_{X'}(c_1)$.

Linear Interpolation. Let $0 \leq \Delta \leq .01$. Choose Δ such that

$$.05 = 1 - F_{X'}(1-.21-\Delta) + F_{X'}(.21+\Delta).$$

Now $\frac{1}{.01}(F_{X'}(.22) - F_{X'}(.21)) = .22$ and therefore with linearity assumption $F_{X'}(.21+\Delta) = F_{X'}(.21) + \Delta(.22)$.

Similarly $\frac{1}{.01}(F_{X'}(.78) - F_{X'}(.79)) = -.8$ and therefore with linearity assumption $F_{X'}(1-.21-\Delta) = F_{X'}(1-c_1) - \Delta(.8)$.

So choose Δ such that

$$.05 = 1 - F_{X'}(1-c_1) + \Delta(.8) + F_{X'}(.21) + \Delta(.22)$$

or

$$.05 = .0481 + \Delta(1.01)$$

so

$$\Delta = \frac{.0019}{1.01} = .0019.$$

Then

$$F_{X'}(.2119) = F_{X'}(.21) + \Delta(.22)$$
$$= .0082 + .0004 = .0086$$
$$F_{X'}(1-.2119) = F_{X'}(1-.21) - \Delta(.8)$$
$$= .9601 - .0015 = .9586$$

and therefore

$$\alpha = 1 - F_{X'}(1-.2119) + F_{X'}(.2119) = .05.$$

b. Minimax solution; most powerful test with $\alpha = \beta$.

	From Incomplete Beta Tables		Linear Interpolation
c_1	.37	.38	.37725
$F_{X'}(c_1)$	.1205	.1348	.1309
$F_{X'}(1-c_1)$	.7061	.6823	.6888
$F_{X''}(c_1)$	.1404	.1527	.1493
$F_{X''}(1-c_1)$	.6063	.5857	.5914
α	.4144	.4525	.4421
β	.4659	.4330	.4421

where $\alpha = 1 - F_{X'}(1-c_1) + F_{X'}(c_1)$

and

$$\beta = F_{X''}(1-c_1) - F_{X''}(c_1).$$

Linear Interpolation. Let $0 \leq \Delta \leq .01$. Choose Δ such that $\alpha = \beta$ i.e.

$$1 - F_{X'}(1-.37-\Delta) + F_{X'}(.37+\Delta) = F_{X''}(1-.37-\Delta) - F_{X''}(.37+\Delta).$$

$\frac{1}{.01}(F_{X'}(.38) - F_{X'}(.37)) = 1.43$ so with linearity assumption $F_{X'}(.37+\Delta) = F_{X'}(.37) + 1.43\Delta$.

$\frac{1}{.01}(F_{X'}(1-.38) - F_{X'}(1-.37)) = -2.38$ so with linearity assumption $F_{X'}(1-.37-\Delta) = F_{X'}(1-.37) - 2.38\Delta$.

$\frac{1}{.01}(F_{X''}(.38) - F_{X''}(.37)) = 1.23$ so with linearity assumption $F_{X''}(.37+\Delta) = F_{X''}(.37) + 1.23\Delta$.

$\frac{1}{.01}(F_{X''}(1-.38) - F_{X''}(1-.37)) = -2.06$ so with linearity assumption $F_{X''}(1-.37-\Delta) = F_{X''}(1-.37) - 2.06\Delta.$

So choose Δ such that

$$1 - F_{X'}(1-.37) + 2.38\Delta + F_{X'}(.37) + 1.43\Delta$$

$$= F_{X''}(1-.37) - 2.06\Delta - F_{X''}(.37) - 1.23\Delta$$

or

$$4.144 + 3.81\Delta = .4659 - 3.29\Delta$$

or

$$\Delta = \frac{.0515}{7.1} = .00725.$$

Then

$$F_{X'}(.37725) = F_{X'}(.37) + 1.43(.00725)$$

$$= .1205 + .0104 = .1309.$$

$$F_{X'}(1-.37725) = F_{X'}(1-.37) - 2.38(.00725)$$

$$= .7061 - .0173 = .6888.$$

$$F_{X''}(.37725) = F_{X''}(.37) + 1.23(.00725)$$

$$= .1404 + .0089 = .1493.$$

$$F_{X''}(1-.37725) = F_{X''}(1-.37) - 2.06(.00725)$$

$$= .6063 - .0149 = .5914.$$

$$\alpha = 1 - F_{X'}(1-.37725) + F_{X'}(.37725)$$

$$= 1 - .6888 + .1309 = .4421.$$

$$\beta = F_{X''}(1-.37725) - F_{X''}(.37725)$$

$$= .5914 - .1493 = .4421.$$

Hence $\alpha = \beta.$

REFERENCES

[1] Apostol, Tom M. (1960). Mathematical Analysis. Addison-Wesley Publishing Company, Inc., Reading, Mass.

[2] Halmos, Paul R. (1962). Measure Theory. D. Van Nostrand Company, Inc., Princeton, N. J.

[3] Kolmogorov, A. N. and Fomin, S. V. (1961). Elements of the Theory of Functions and Functional Analysis, Vol 2, Graylock Press, Albany, N. Y.

[4] Lehmann, E. L. (1959). Testing Statistical Hypothesis. John Wiley and Sons, Inc., New York.

[5] Lindgren, B. W. (1962). Statistical Theory. The Macmillan Company, New York.

[6] Pearson, Karl F.R.S. (1934). Tables of the Incomplete Beta Function. The Biometrika Office, University College, London.

[7] Rudin, Walter (1953). Principles of Mathematical Analysis. McGraw-Hill Book Company, Inc., New York.

[8] Sitgreaves, Rosedith. Chapters 1-3. [9]

[9] Solomon, Herbert (1961). Studies in Item Analysis and Prediction. Stanford University Press, Stanford, California.

[10] Wald, Abraham (1942). On the Choice of the Number of Class Intervals in the Application of the Chi-Square Test. The Annals of Mathematical Statistics. Vol. X111, No. 3.