

**Reports from the Research Laboratories
of the
Department of Psychiatry
University of Minnesota**

**“Detecting Latent Clinical Taxa, IV:
An Empirical Study of the Maximum Covariance Method
and the
Normal Minimum Chi-Square Method
Using Three MMPI Keys to Identify the Sexes”
by
ROBERT R. GOLDEN and PAUL E. MEEHL**

**MSDM
P95
QR311r
73-2**

THE LIBRARY



MSDM
P95
qR311r
73-2

University Archives

Reports from the Research Laboratories
of the
Department of Psychiatry
University of Minnesota

"Detecting Latent Clinical Taxa, IV:
An Empirical Study of the Maximum Covariance Method
and the
Normal Minimum Chi-Square Method
Using Three MMPI Keys to Identify the Sexes"¹

by

Robert R. Golden and Paul E. Meehl

Report Number PR-73-2

December 1973

¹This research was supported in part by grants from the Psychiatry Research Fund, the National Institute of Mental Health, Grant Number MH 24224, and the Schizophrenia Research Program, Scottish Rite, Northern Masonic Jurisdiction, U. S. A.

TABLE OF CONTENTS

I.	Introduction.....	1
II.	Development of Keys.....	2
III.	Maximum Covariance Method.....	5
IV.	Maximum Covariance Method Results.....	6
V.	Prediction of Biological Sex Using Maximum Covariance Parameter Estimates.....	8
VI.	Direct Estimation of Base-Rate P.....	11
VII.	Two Different Assumptions of Zero Within Taxa Covariance.....	12
VIII.	Estimation of the Hitmax Cut by Using the Output Variable Mean Above Less than Below a Sliding Cut on the Input Variable.....	15
IX.	Estimation of the Hitmax Cut by Using the Maximum of the Sum of the Latent Hit-Rates.....	17
X.	Estimation of the Hitmax Cut by a Relation Between the Manifest Frequencies Above and Below Hitmax Cut and the Latent Hit-Rates.....	17
XI.	Estimation of Parameter Estimate Error in the Maximum Covariance Method.....	18
XII.	Iterative Generalization of the Maximum Covariance Method.....	20
XIII.	The Normal Method.....	22
XIV.	Prediction of Biological Sex Using the Normal Minimal χ^2 Parameter Estimates.....	28
XV.	Evaluation of the Prediction of Biological Sex.....	29
XVI.	Estimation of the Hit-Miss Table.....	31
XVII.	Conclusions.....	33

1. Introduction

Previous contributions to this research report series (Meehl, 1965 and Meehl, 1968) provided the basis of several different interrelated methods for the estimation of a large variety of parameters concerning a dichotomous latent taxonomy. However, there are two distinctly different major assumptions strongly relied upon in this work: (a) the covariance between two indicators within the taxon and within its complementary extra-taxon class is zero (maximum covariance method) and (b) each single indicator distribution within the taxon and the extra-taxon class is normal (normal method). Many other results developed by Meehl can be used in conjunction with either of these two different general methods. It will be indicated below that the present trial illustrates how these two methods, although intuitively quite different in rationale, can be used as checks on each other when used with indicators such as MMPI keys.

The term "method" as used here can be regarded as equivalent to "model" or "theory". The term "assumption" is used in the usual sense in that it can be stated exactly in mathematical formulation and while it is never strictly true for any natural phenomena it can not be directly tested by the present data. It is only necessary that an assumption be approximately true, the closeness of the approximation required being determined by whatever the substantive problem requires in terms of accuracy of parameter estimation and major hypothesis testing such as "Is there one taxon plus the extra-taxon class?".

This report gives the results of an empirical trial of each of the two major methods and certain other results developed by Meehl. The raw data consisted of 1105 patients' MMPI item responses. Three twenty-item MMPI keys were developed so as to discriminate fairly well between the sexes and were used as three candidate taxonomic indicators, the purported taxonomy being that of biological sex. As was mentioned in a previous report of a male-female empirical trial of the normal method (Meehl et al., 1969) biological sex is one of the few true taxonomies that can provide good psychometric data for an empirical test of a taxometric method. This is because the determination of biological sex is completely objective and many personality measures have distributions which are bimodal or nearly so for mixed-sex samples.

II. Development of Keys

While the normal method requires only a single indicator such as scale 5 of the MMPI (which was, in fact, used in the report referred to above), the maximum covariance method requires three or more indicators. A first attempt at key development started with factor analyses (varimax solution) of the sixty items in scale 5 for the male sample, the female sample and the mixed sample. The three factor analyses each produced similar sets of four different factors, each accounting for 10-15% of the common variance and each being easily identified and matched with a factor from each of the other two analyses by similar patterns of the highest ten item loadings. The finding of

similar factor structure within the sexes is of substantive significance; however, the factor scales could not be used for the purpose of this study since it was found that while they possessed ample internal consistency "reliability" so as to be easily identified (especially the first two factors which were clearly "masculine interests" and "feminine interests") none discriminated adequately between the sexes. Key lengths of ten and twenty items both produced mean separations of about one standard deviation on each of the four scales. A Monte Carlo study (Golden and Meehl, 1973) shows that under otherwise optimal conditions, such a separation is too small for accurate taxonomic detection by the maximum covariance method. Other studies have indicated that about a two standard deviation separation is required for the normal method (Hasselblad, 1966; Hald, 1952). In short, the item selection by factor analysis produced keys that were too homogeneous in that homogeneity was obtained at the expense of discriminative power. This result can be explained in some mathematical detail as follows. Let the following notation be used:

Δ_i : the item plus-rate proportion for the females less that
for the males,

σ_{if} : the item variance for the females,

σ_{im} : the item variance for males, and

σ_{ij} : the within sex covariance (for simplicity, assumed here
to be the same for each sex)

where $i = 1, 2, 3, \dots, n$ (the number of items). If it is assumed (for simplicity here) that

$$\gamma = \sum_{i=1}^n \sigma_{if} = \sum_{i=1}^n \sigma_{im}$$

TABLE 1

Descriptive statistics for the male and female sample distributions on the three keys

	male		female			
frequency	430		675			
proportion (base-rate)	.389		.611			
<u>key</u>	mean	SD	mean	SD		
1	9.57	2.55	14.10	2.53		
2	7.31	2.42	11.68	2.47		
3	8.60	2.57	12.84	2.53		
	correlations			correlations		
key/	1	2	3	1	2	3
1				1		
2	.27			2	.33	
3	.53	.33		3	.53	.35
	covariances			covariances		
key/	1	2	3	1	2	3
1				1		
2	1.67			2	2.03	
3	3.45	2.08		3	3.37	2.18

for a set of n items comprising a key (with unit weights), then the difference in the male and female key means in terms of the common within sex standard deviation can be shown to be

$$\frac{\sum_{i=1}^n \Delta_i}{(\gamma + \sum_{\substack{i,j \\ (i \neq j)}} 2\sigma_{ij})^{\frac{1}{2}}}$$

Thus if keys are desired to discriminate maximally it is required that an item i be selected such that Δ_i is large and the σ_{ij} are small. That is, the best item for a key should not only be maximally discriminative itself, but also it should have minimal covariances with other items in the key. Factor analysis produces keys that have high interitem covariation within the sexes.

Adequately discriminative keys were produced by selecting from the total pool of 550 items the set of 60 most discriminative in terms of the plus-rate difference between the sexes, ordering the 60 items from most to least discriminative and assigning every third one to each of three keys so as to make each key approximately equally discriminative. As a further precaution, items with high covariances within a key were re-assigned to another key if this resulted in lower covariance contribution by the item to the latter key. In short, a non-optimal crude "guess and hope" method which considered both the Δ_i 's and the σ_{ij} 's was used to develop the three twenty item keys. See Table 1 for the basic descriptive statistics of the resulting keys which were used as the real data in the main portion of this study. It should be noted that the mean separations are slightly less than the desired two sigma differ-

ence and, therefore, provide for a moderately tough empirical trial of the two methods.

III. Maximum Covariance Method

The method is given in Section 1, pp. 2-7 of PR-68-4 as a revision of the original method given in Section 3, pp. 10-12 of PR-65-2. An outline of the method is given below.

A. Let w, x, and y be three indicators such that w is the input indicator and x and y are the output indicators. The latent taxa distributions on the input indicator are estimated by use of manifest relationships between the two output variables.

"Output" and "input" refer to statistical manipulations and have no psychological-causal connotation.

B. The covariance between x and y for cases lying within any interval of w is given by

$$\text{cov}_w(x,y) = p_w \text{cov}_{fw}(x,y) + q_w \text{cov}_{mw}(x,y) + p_w q_w \Delta \bar{x}_w \Delta \bar{y}_w \quad [1]$$

where

p_w is the proportion of individuals in w interval that are females,

q_w is the corresponding male proportion ($p_w + q_w = 1$),

$\text{cov}_{fw}(x,y)$ is the manifest conditional covariance between x and y for the females in interval w,

$\text{cov}_{mw}(x,y)$ is the corresponding male covariance,

$\Delta \bar{x}_w$ is the mean on x for the females in interval w less that for the males, and

$\Delta \bar{y}_w$ is the corresponding mean difference on y.

C. Under the assumptions

A₁: $\Delta\bar{x}_w \Delta\bar{y}_w = K$ (a constant) for all w, and

A₂: $\text{cov}_{fw}(x,y) = \text{cov}_{mw}(x,y)$ for all w,

It follows that $\max\{\text{cov}_w(x,y)\}$ occurs in the hitmax interval (where $p_w = q_w = \frac{1}{2}$ and the frequency distributions intersect) and is equal to the latent quantity $\pm\Delta\bar{x}_w \Delta\bar{y}_w = \pm K$.

D. Under a further extension of assumption A₂

A₃: $\text{cov}_{fw}(x,y) = \text{cov}_{mw}(x,y) = 0$ for all w

It follows that

$$p_w^2 - p_w + \frac{\text{cov}_w(x,y)}{\max\{\text{cov}_w(x,y)\}} = 0 \quad [2]$$

a quadratic with p_w and q_w as the two roots. In other words, the latent frequency distributions on w for each taxon are now estimated. From these, the latent taxa means, standard deviations, base-rates and any other distribution parameters are estimated.

E. With three indicators, the roles of input and output can be interchanged to produce three different arrangements as shown below.

<u>input indicator</u>	<u>output indicators</u>
key 1	key 2, key 3
key 2	key 1, key 3
key 3	key 1, key 2

IV. Maximum Covariance Method Results

The observed and estimated taxa frequency distributions and

the corresponding descriptive statistics for each of the three arrangements are given in Tables 2, 3, and 4. A perusal of these tables shows that there is considerable agreement between the actual and estimated frequencies. Usual tests of significance for comparing the actual and estimated frequency distributions are not strictly appropriate since population values are not known. However, if the model parameter estimates are considered as approximations of the true population values, then it is possible to check as to whether the observed values differ more than that for which sampling error would allow. Tests of goodness of fit between the expected (estimated) and observed distributions by the usual χ^2 test produced the following results.

χ^2 tests of goodness of fit of the within taxa
frequency distributions when estimated by the
maximum covariance method

	males			females		
	χ^2	d.f.	P	χ^2	d.f.	P
key 1	83.7	16	<.01	34.7	16	<.01
key 2	95.5	16	<.01	66.5	16	<.01
key 3	94.2	14	<.01	143.8	15	<.01

Although non-significant results are a desirable outcome they are not necessary for acceptance of the model. It is clear simply by "armchair inspection" that the parameter estimates are accurate enough for any applied or theoretical work in the area of personality measurement. (Physicists and chemists often proceed thus; why shouldn't psychologists?) It might be noted that here we have an example of one of the many

TABLE 2

The observed and estimated (by the maximum covariance method)
taxa frequency distributions for key 1

score	male		female	
	observed	estimated	observed	estimated
3	3	3	0	0
4	11	10	0	1
5	8	7	0	1
6	32	34	3	1
7	41	39	5	7
8	50	46	9	13
9	58	66	20	12
10	60	71	30	19
11	65	76	27	16
12	53	58	62	57
13	29	8	93	114
14	10	15	110	105
15	7	18	103	92
16	2	13	107	96
17	1	3	61	59
18	0	2	28	26
19	0	2	14	12
20	0	0	3	3
base-rate	.389	.426	.611	.574
mean	9.57	9.94	14.10	14.12
SD	2.55	2.84	2.56	2.50

TABLE 3

The observed and estimated (by the maximum covariance method)
taxa frequency distributions for key 2

score	male		female	
	observed	estimated	observed	estimated
1	1	1	0	0
2	7	7	0	0
3	13	9	0	4
4	33	26	3	10
5	48	36	3	15
6	63	44	13	32
7	67	61	17	23
8	62	59	27	30
9	59	59	59	59
10	36	29	87	94
11	22	52	110	80
12	11	26	97	82
13	6	32	91	65
14	1	18	83	66
15	0	5	51	46
16	1	10	26	17
17	0	1	5	4
18	0	0	2	2
19	0	0	1	1
base-rate	.389	.430	.611	.570
mean	7.31	8.66	11.68	10.97
SD	2.42	3.17	2.47	2.94

TABLE 4

The observed and estimated (by the maximum covariance method)
taxa frequency distributions for key 3

score	male		female	
	observed	estimated	observed	estimated
2	4	4	0	0
3	3	3	0	0
4	12	11	0	1
5	26	24	2	4
6	54	53	3	4
7	53	60	8	1
8	57	64	16	9
9	71	92	27	6
10	46	58	69	57
11	43	13	83	113
12	31	22	91	100
13	17	14	93	96
14	9	23	109	95
15	4	4	71	71
16	0	3	56	53
17	0	0	27	27
18	0	0	15	15
19	0	0	5	5
base-rate	.389	.405	.611	.595
mean	8.60	8.61	12.84	12.95
SD	2.57	2.62	2.53	2.40

differences between hypothesis testing and parameter estimation. A model need not fit the data perfectly as an H_0 test would require given unit power. It only is required that the important parameters be estimated with an adequate degree of accuracy.

In the present trial, 90%+ accuracy would seem to be adequate in the estimation of the various means, variances and base-rates. (% accuracy = $\frac{\text{observed value} - \text{estimated value}}{\text{observed value}} \times 100$.)

Inspection of Tables 2, 3, and 4 show the observed hit-max intervals for keys 1, 2, and 3 to be 12, 9, and 10. The manifest covariance functions between the pairs of output variables are given in Table 5 and it is seen that the corresponding maximum covariances occur in intervals 12, 9, and 10; thus in perfect agreement.

It should be noted that the sampling irregularities of the output covariance functions do indicate that some sort of curve smoothing will be desirable in other empirical investigations.

V. Prediction of Biological Sex Using Maximum Covariance Parameter Estimates

Strengthening the assumptions A_2 and A_3 further to

A_4 : The indicators are independent in the strongest sense within taxa

will allow for classification of individuals. That is, for any three intervals w , x , and y on the three keys 1, 2, and 3, the density (the proportion of the individuals in the taxon with the scores w , x , and y) $\phi(w,x,y)$ is equal to $\phi_1(w)\phi_2(x)\phi_3(y)$ where, for example, $\phi_1(w)$ is the taxon density for score w on key 1.

TABLE 5

The mixed group manifest covariance function between output indicators for each of the three input indicators

score	Input indicator		
	key 1	key 2	key 3
1	--	0.00	--
2	--	-0.53	1.50
3	-0.22	4.27	0.00
4	1.48	3.91	1.69
5	2.56	3.94	1.98
6	0.78	4.54	1.18
7	2.08	3.71	0.28
8	2.66	4.18	1.76
9	1.95	4.62	0.94
10	2.51	3.32	3.98
11	2.12	4.39	1.40
12	3.68	3.37	2.31
13	0.83	4.02	1.75
14	1.59	2.99	2.43
15	1.97	1.43	0.80
16	1.45	4.14	0.56
17	0.64	0.68	-0.06
18	0.61	-0.50	-0.96
19	0.99	0.00	-0.88
20	-1.33	--	--
total	6.55	7.98	6.58

Then the probability that an individual is a female given a vector of key scores (w,x,y) is

$$\begin{aligned} \text{Pr}(\text{female}|w,x,y) &= \frac{P\phi_f}{P\phi_f + Q\phi_m} \\ &= \frac{P\phi_{fw1}\phi_{fx2}\phi_{fy3}}{P\phi_{fw1}\phi_{fx2}\phi_{fy3} + Q\phi_{mw1}\phi_{mx2}\phi_{my3}} \end{aligned}$$

where

P is the base-rate for females,

Q (= 1 - P) is the base-rate for males,

$\phi_f = \phi(w,x,y)$ for the females,

$\phi_m = \phi(w,x,y)$ for the males, and

ϕ_{fw1} = the female density function value for interval w indicator 1, for example.

Then if the total misclassification rate is to be minimized it can be shown that the required classification rule is

"Classify as female if $\text{Pr}(\text{female}|w,x,y) > .5$, and classify as male otherwise".

The base-rate was estimated for each of the three keys giving close but, of course, somewhat different results. For use in the classification formula the simple arithmetic average of the three estimates was used. The estimated taxon density functions were determined directly from the corresponding estimated frequency functions. The predicted sex can then be compared with the actual sex in the form of a hits-misses table as given below.

	<u>Biological Sex</u>		
	male	female	total
<u>predicted sex</u>			
male	361	86	447
female	69	589	658
total	430	675	1105

The proportion of "female" predictions which were correct was .90, similarly the "male" hit-rate was .81 and the overall hit-rate was .86.

In PR-65-2, Section 10, pp. 34-37 it was proposed that the above classification procedure be applied after the indicators are each dichotomized at the hitmax cuts. Such a classification procedure does not make use of all of the information available for density estimation and results in a less detailed estimation of the taxon density function. In the present example, the "interval proportions" method allowed for 20^3 point densities to be estimated (of course, less than 1105 were actually ever required, or, in fact, calculated) whereas the "hitmax cut" method provides only $2^3 = 8$ point densities. The former method has the advantage of providing estimates of the density function at each point whereas the latter provides fewer but more stable density "pattern" estimates. An optimal procedure would appear to be one that combines certain contiguous indicator intervals so as to combine the advantages of both methods.

The hits-misses table resulting from using the hitmax cut classification method is given below.

<u>predicted sex</u>	<u>Biological Sex</u>		
	male	female	total
male	333	68	401
female	97	607	704
total	430	675	1105

The valid female rate, valid male rate and overall hit-rate were .86, .83, and .85 respectively. The proportion of the predictions which were "female" was .64. Comparison of the three hit-rates of the two methods and the general patterns of the hits-misses tables indicates that for the present empirical trial the interval proportions method has only a very slight edge in accuracy.

VI. Direct Estimation of Base-Rate P

In PR-68-4, pp. 24-26, it is suggested that the base-rate P can also be estimated (for consistency testing) by solving the covariance mixture equation when written for the total compound sample (not for a w interval of the input variable as was done in equation [1]):

$$\text{cov}(x,y) = P\text{cov}_f(x,y) + Q\text{cov}_m(x,y) + PQ\Delta\bar{x}\Delta\bar{y} \quad [3]$$

where

P is the base-rate of the females,

Q is the base-rate of the males,

cov(x,y) is the covariance between x and y for the total group,

cov_f(x,y) is the manifest covariance between x and y for the females,

$cov_m(x,y)$ is the corresponding male covariance,

$\Delta\bar{x}$ is the mean on x for the females less that for the males,

and

$\Delta\bar{y}$ is the corresponding mean difference on y.

If we assume that

$$A_5: cov_f(x,y) = cov_m(x,y) = 0$$

then [3] becomes

$$cov(x,y) = PQ\Delta\bar{x}\Delta\bar{y} \text{ or}$$

$$P = \frac{1}{2} \pm \sqrt{\frac{1}{4} - \frac{cov(x,y)}{\Delta\bar{x}\Delta\bar{y}}} \quad [4]$$

Since $cov(x,y)$ is directly observable and under A_1 $\Delta\bar{x}\Delta\bar{y}$ can be estimated from the hitmax interval data, for example, [4] provides another method to estimate P. For example, using key 2 and key 3 as output variables x and y, we find

$$\frac{cov(x,y)}{\Delta\bar{x}\Delta\bar{y}} = \frac{5.37}{(10.97-8.66)(12.95-8.61)} = .535.$$

Since this quantity is greater than $\frac{1}{4}$, equation [4] gives imaginary values for P. Thus, if sampling error can not explain this result then the conjunction A_1 and A_5 is false. From the actual male-female sample statistics given in the Appendix it can be seen that while A_1 is approximately true, A_5 is blatantly incorrect for the present data. Pending further study it appears that A_5 is too strong to make this a useful consistency test.

VII. Two Different Assumptions of Zero Within Taxa Covariance

It was shown above that the within interval (of the input variable) covariances can be assumed to be zero (assumption A_2) in the context of the maximum covariance method without giving

rise to serious errors in the parameter estimation (the within taxa indicator frequency distributions, the taxa base-rates, and the taxon membership of each individual). The present data (see Appendix) then illustrate that assumptions A_2 and A_5 are not equivalent; further, examination of special cases such as the one of Figure 1 shows geometrically that the assumptions are not equivalent. Further clarification of the relationship between the two assumptions can be obtained by writing the covariance mixture equation for two input variable intervals w_i and $w_{i'}$, for just one of the taxa; say, the females

$$\begin{aligned} \text{cov}_{f_{w_i+w_{i'}}}(x,y) &= P_{w_i} \text{cov}_{f_{w_i}}(x,y) + P_{w_{i'}} \text{cov}_{f_{w_{i'}}}(x,y) \\ &\quad + P_{w_i w_{i'}} \Delta \bar{x}_{w_{i'}} \Delta \bar{y}_{w_{i'}} \end{aligned}$$

where

$\text{cov}_{f_{w_i+w_{i'}}}(x,y)$ is the covariance between x and y for the females in w_i and $w_{i'}$,

P_{w_i} is the proportion of the females in w_i and $w_{i'}$ which are in w_i ,

$P_{w_{i'}}$ is the corresponding proportion for $w_{i'}$, ($= 1 - P_{w_i}$),

$\text{cov}_{f_{w_i}}(x,y)$ is the covariance between x and y for the females in w_i ,

$\text{cov}_{f_{w_{i'}}}(x,y)$ is the corresponding covariance for $w_{i'}$,

$\Delta \bar{x}_{w_{i'}}$ is the mean on x of the females in w_i less that for the females in $w_{i'}$, and

$\Delta \bar{y}_{w_{i'}}$ is the corresponding mean difference on y .

Under A_2 , $\text{cov}_{f_{w_i}}(x,y) = \text{cov}_{f_{w_{i'}}}(x,y) = 0$ and it follows that

$\text{cov}_{f_{w_i+w_{i'}}}(x,y) = P_{w_i} P_{w_{i'}} \Delta \bar{x}_{f_{w_{i'}}} \Delta \bar{y}_{f_{w_{i'}}}$. Thus it is seen that

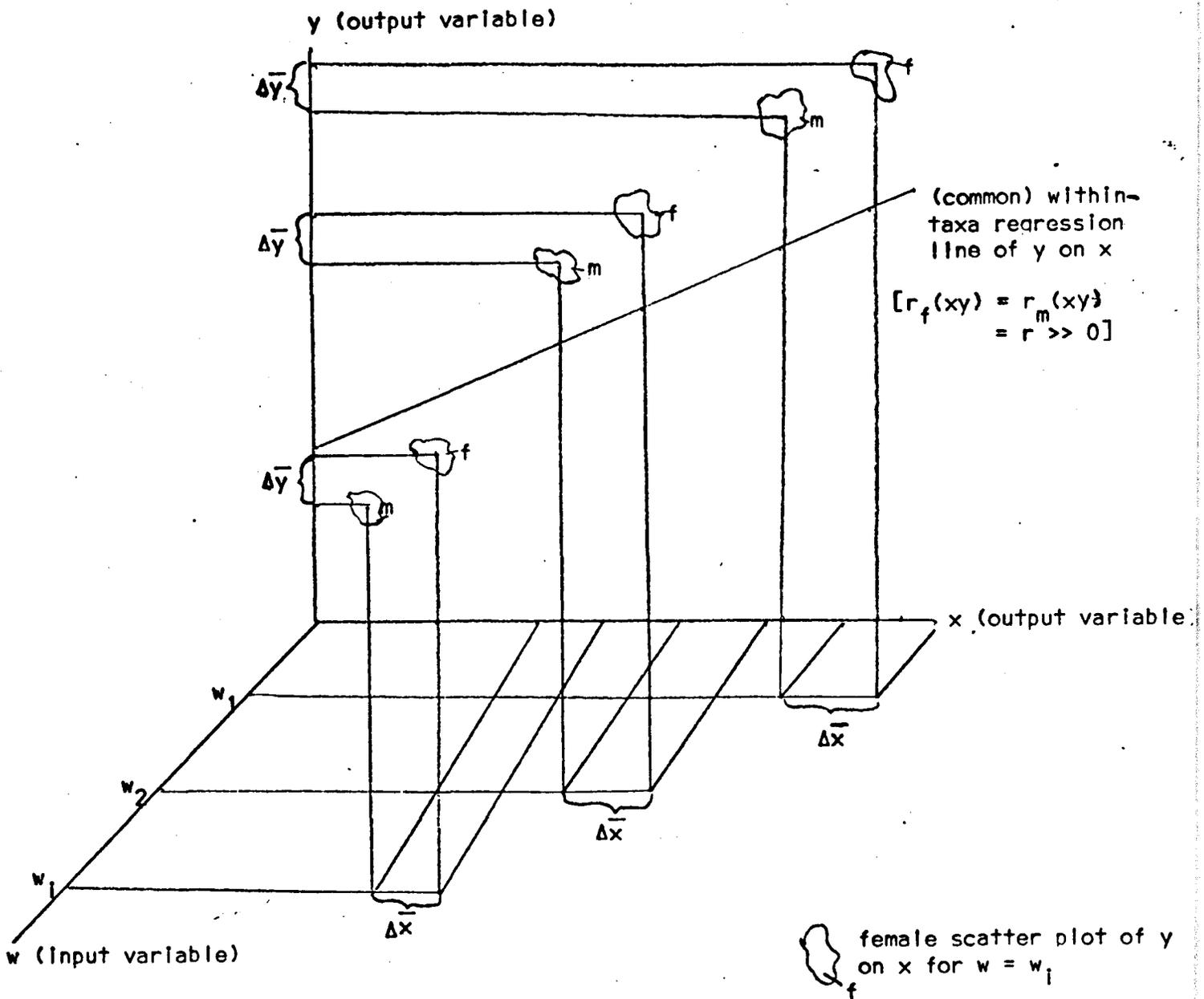


Figure 1. Illustration of case where within-taxa within-input variable covariances are zero and the within-taxa covariances are greater than zero

$\text{cov}_{f_{w_1+w_1'}}(x,y)$ is zero if and only if $\Delta\bar{x}_{fw_1} = 0$ and/or $\Delta\bar{y}_{fw_1} = 0$.

The latter condition can be written as $\bar{x}_{fw_1} = \bar{x}_{fw_1}$, and/or $\bar{y}_{fw_1} =$

\bar{y}_{fw_1} . It can be shown by the method of mathematical induction

then that for n input variable intervals $w_1, w_2, w_3, \dots, w_n$,

$\text{cov}_{f_{\sum w_i}}(x,y) = \text{cov}_f(x,y) = 0$ if $\bar{x}_{fw_1} = \bar{x}_f$ and/or $\bar{y}_{fw_1} = \bar{y}_f$ for all

i. Thus the further condition required for A_5 to follow from A_2

can be given in words as "the within taxa regression curve of x on

w or y on w is a line of zero slope". That this is not the case for

the present data can be seen simply by inspection (see Appendix).

All of the within taxa regression lines have large positive slopes.

For example, the means on key 2 for the females in intervals 6-20

of key 1 go steadily up from 7.7 to 12.0. With such steeply slop-

ing regression lines it is not difficult to see that the method

of direct estimation of the base-rate would lack adequate robust-

ness with respect to assumption A_5 . For the simplified maximum

covariance method to provide adequately accurate parameter estima-

tion there must be sufficient robustness with respect to A_2 .

Perusal of the within taxa covariances as functions of input

variable values (see Appendix) shows that they are somewhat posi-

tive, rather irregular (evidently due to sampling error) and with-

out any other discernible properties. As a typical example, the

covariances between key 2 and key 3 for the females in intervals

6-20 of key 1 are -1.7, 3.2, .1, 3.2, 1.4, .1, 1.4, .0, 1.1, 1.5,

1.5, .7, .6, 1.0, and -1.3. It is not clear to the eye that these

values are sufficiently close to zero as required by A_2 ; however,

In view of the final parameter estimates of the maximum covariance

method it is concluded that the method has adequate robustness

with respect to A_2 to tolerate such deviations.

VIII. Estimation of the Hitmax Cut by Using the Output Variable Mean Above Less Than Below a Sliding Cut on the Input Variable

In PR-68-4, Section 2, part b (pp. 9-24) it is suggested that the mean on the output variable y for those individuals in the mixed distribution with a score on the input variable greater than x , \bar{y}_{ax} , less than for below x , \bar{y}_{bx} , has a maximum approximately at the hitmax cut. As the results in Table 6 show, the method did not correctly locate the hitmax cuts with the possible exception of example A. The other input/output combination examples not given in Table 6 are similar to examples B and C. One source of error was clearly that of sampling and a fair test of the method should use some curve smoothing procedure. The sampling variance

$\bar{y}_{dx} = \bar{y}_{ax} - \bar{y}_{bx}$ is given by

$$\text{var}(\bar{y}_{dx}) = \frac{\text{var}(\bar{y}_{ax})}{N_{ax}} + \frac{\text{var}(\bar{y}_{bx})}{N_{bx}} .$$

As the x cut goes further into either tail of the mixed distribution, N_{ax} or N_{bx} becomes smaller and $\text{var}(\bar{y}_{dx})$ becomes larger. Examination of example A shows the method worked fairly well if one considers only the relative maximum and if small differences in contiguous values such as .02 is enough for stability of the sign of the derivative with respect to x . For the latter condition to be even roughly met it is necessary, it would seem, that $\text{var}(\bar{y}_{dx})$ be less than .2. With this requirement and estimates of $\text{var}(\bar{y}_a)$ and $\text{var}(\bar{y}_b)$ the minimum allowable size of N_a or N_b can be determined. Let N_a be the smaller of the two so that the right

TABLE 6

Examples of the mean above less the mean below on output variable y as a function of the input variable interval value x

<u>input variable (x)</u> output variable (y)	<u>key 1</u> key 2	<u>key 2</u> key 1	<u>key 3</u> key 1
Input variable interval value		$\bar{y}_{ax} - \bar{y}_{bx}$	
0	---*	---	---
1	---	---	---
2	---	---	---
3	---	---	---
4	---	---	---
5	---	4.207	---
6	---	3.968	---
7	---	3.856	4.613
8	---	3.801	4.636
9	3.470	3.581	4.390
10	3.440	3.341	3.970
11	3.384	3.340	3.877
12	3.404	3.143	3.715
13	3.392	2.947	3.521
14	3.445	2.889	3.409
15	3.227	---	3.211
16	3.063	---	---
17	2.805	---	---
18	---	---	---
19	---	---	---
20	---	---	---

*Values for intervals with mixed distribution tail frequencies less than 100 are omitted.

hand tail is nearly all females. The within sex variances are about $(2.5)^2 = 6.25$, so $\text{var}(\bar{y}_a) = 6.25$. The mixed group variances are about 3.3, so $\text{var}(\bar{y}_b) = 10.98$. If $N_a + N_b = 1000$, then it follows that N_a should be greater than 100. Also, it is seen that $\min\{\text{SD}(\bar{y}_{dx})\}$ is greater than $\sqrt{\frac{(2.5)^2}{500} + \frac{(2.5)^2}{500}} = .16$.

Thus it is difficult to trust the appearances of a relative maximum near hitmax in example A to the extent of having confidence that replication would produce a similar result.

The main source of error is clearly that of not approximately meeting the assumption that the within-taxa within-interval output means are constant which is the key assumption in the quasi-derivation of the result (see equations 12 and 13, p. 12, PR-68-4). Monte Carlo study has shown that the method does work quite well when the within-taxa within-interval output means are constant (these results will be given in a forthcoming research report) and the lack of a precisely developed mathematical demonstration that $\max(\bar{y}_{dx})$ approximates the hitmax is not something of concern in view of the failure of the method with the present data.

The present results do suggest an interesting possibility to be investigated by Monte Carlo study in that the absence of any local maximum and the presence of monotonicity of the \bar{y}_{dx} function might be a good indicator of failure to meet assumption A_5 . Thus if the two above methods of estimation of the base-rates (section IIIb and V) give discrepant results, as is the case for the present data, the \bar{y}_d function could be used to check the validity of assumption A_5 .

IX. Estimation of the Hitmax Cut by Using the Maximum of the Sum of the Latent Hit-Rates

In PR-68-4, pp. 14-24, it is argued that the sum of the latent hit-rates has a maximum at hitmax. First we can inquire if this is true for the male and female sample distributions on each key. The result is that the quantity has vaguely defined relative maxima at 11, 8, and 9 on keys 1, 2, and 3 respectively; these values correspond well with the true hitmax cuts of 12, 9, and 10. Second, we can inquire if this is true for the maximum covariance method estimated taxa distributions on each key. The result is that the quantity has nicely defined relative maxima at 11, 10, and 11 on keys 1, 2, and 3 respectively which also correspond well with the true values.

X. Estimation of the Hitmax Cut by a Relation Between the Manifest Frequencies Above and Below Hitmax Cut and the Latent Hit-Rates

In PR-68-4, section 2, part d, pp. 26-31, it is shown that

$\frac{U_{ax}}{U_{bx}} - \frac{N_{ax}}{N_{bx}}$ (where U_{ax} is the proportion of the individuals above

x that belong to the upper group less one-half, N_{ax} is the total number of individuals with scores greater than x , and U_{bx} and N_{bx} are defined similarly) has a minimum and is approximately

zero at hitmax. Using the male and female sample distributions, the hitmax estimates are 11, 9, and 10 which compare well with the true values of 12, 9, and 10 and the same results are obtained using the maximum covariance method estimated frequency

distributions. The graph of $\left| \frac{U_{ax}}{U_{bx}} - \frac{N_{ax}}{N_{bx}} \right|$ in each case has a very

nically defined minimum value between 0 and 1 in each case. It should be noted that these minimum points are considerably more discernible than the maximum of the sum of the hit-rates and, therefore, this method might prove to be of more general value than the latter one.

XI. Estimation of Parameter Estimate Error in the Maximum Covariance Method

In section III above, [1] is solved for p_w , the proportion in the interval w which are females, for each value of w by making certain simplifying assumptions. From [1] it is clear that the error in the p_w estimate is due to errors in $C_{fw} = \text{cov}_{fw}(x,y)$, $C_{mw} = \text{cov}_{mw}(x,y)$, and $k_w = \Delta \bar{x}_w \Delta \bar{y}_w$. Ignoring sampling error, $\text{cov}_{fw}(x,y)$ and $\text{cov}_{mw}(x,y)$ are in error because of assumption A_3 , and k_w is in error because of assumptions A_1 and A_2 . Denoting the error in the dependent p_w variable by Δp_w and the errors in the independent variables by ΔC_{fw} , ΔC_{mw} , and Δk_w , the exact differential of p_w is given by

$$dp_w = \frac{\partial p_w}{\partial k_w} dk_w + \frac{\partial p_w}{\partial C_{fw}} dC_{fw} + \frac{\partial p_w}{\partial C_{mw}} dC_{mw}$$

from which we obtain the (approximate relationship between the errors

$$\Delta p_w \doteq \frac{\partial p_w}{\partial C_{fw}} \Delta C_{fw} + \frac{\partial p_w}{\partial C_{mw}} \Delta C_{mw} + \frac{\partial p_w}{\partial k_w} \Delta k_w \quad [5]$$

By implicit differentiation of [1], we obtain

$$\frac{\partial p_w}{\partial C_{fw}} = \frac{p_w}{d_w},$$

$$\frac{\partial p_w}{\partial C_{mw}} = \frac{1-p_w}{d_w}, \text{ and}$$

$$\frac{\partial p_w}{\partial k_w} = \frac{p_w q_w}{d_w}, \text{ where}$$

$$d_w = (p_w - q_w)k_w + C_{fw} - C_{mw}$$

and substitution into [5] gives

$$\Delta p_w = [p_w \Delta C_{fw} + (1 - p_w) \Delta C_{mw} + p_w q_w \Delta k_w] / d_w \quad [6]$$

Using the convention that the estimate of parameter θ , denoted

by $\hat{\theta}$, is given by $\hat{\theta} = \theta + \Delta\theta$ we have $\Delta C_{fw} = -C_{fw}$ and $\Delta C_{mw} = -C_{mw}$ because of assumption A_3 and $\Delta k_w = 4\max\{C_w\} - \frac{C_w}{p_w q_w} + \frac{C_{fw}}{q_w} +$

$\frac{C_{mw}}{p_w}$ because of assumptions A_1 and A_2 . From [6] we then have

$$\Delta p_w = \frac{4p_w q_w \max\{C_w\} - C_w}{(p_w q_w)k_w + C_{mw} - C_{fw}} \quad [7]$$

The resulting error in the overall base-rate ΔP is $\sum_w \Delta p_w \phi_w$ where

ϕ_w is the mixed group density for interval w . Substitution of

the method produced estimates for the parameters in [7] should

yield sufficiently small values for Δp_w and Δp . The maximum

covariance method produces estimates of all parameters except

C_{fw} and C_{mw} ; hence, it will be assumed that $C_{fw} - C_{mw} = 0$.

The resulting values of $\Delta p_w \phi_w$ for each interval of each key of

the present data are never greater than .002 and are usually

less than .001, and the values for Δp for keys 1, 2, and 3 are

.009, .009 and .027 respectively, which are all sufficiently

small. Also, it can be shown that $|C_{fw} - C_{mw}|$ can be as large

as 1 or 2, which is considerably larger than the true differ-

ences, without markedly changing the $\Delta p_w \phi_w$ and the Δp .

The result given in [7] can be considered as another con-

sistency test. Under both of the conditions: (a) $C_{mw} - C_{fw} = 0$,

and (b) $|C_{fw} - C_{mw}|$ is of maximum size as estimated by some method, the resulting Δp_{fw} and Δp should be sufficiently small. However, it is presumably possible that the maximum covariance method could produce parameter estimates which are grossly in error yet do not yield large Δp_{fw} and Δp values. It is only suggested that if the assumptions are adequately satisfied then the Δp_{fw} and Δp should be sufficiently small; if in a particular application they are not, then it is suggested that the model is not adequate in that instance.

XII. Iterative Generalization of the Maximum Covariance Method

In PR-68-4, pp. 54-58 it is suggested that after the procedure of section III above the following procedure could be used to relax the within taxa covariance assumptions A_2 and A_3 .

For each indicator:

1. Estimate $k = \Delta \bar{x} \Delta \bar{y}$ from the estimates of the latent means of the output variables.
2. Under the assumptions

$$A_5: \text{cov}_{fw}(x,y) = C_f \text{ and } \text{cov}_{mw}(x,y) = C_m \text{ for all } w, \text{ and}$$

$$A_6: \bar{x}_{fw} = \bar{x}_f, \bar{y}_{fw} = \bar{y}_f, \bar{x}_{mw} = \bar{x}_m, \bar{y}_{mw} = \bar{y}_m \text{ for all } w$$

it can be shown by the method of mathematical induction that if we consider the two mixed samples above and below some cut, say the median, on the input variable then

$$\text{cov}_{aw}(x,y) = P_{fa} C_f + P_{ma} C_m + P_{fa} P_{ma} k, \text{ and}$$

$$\text{cov}_{bw}(x,y) = P_{fb} C_f + P_{mb} C_m + P_{fb} P_{mb} k,$$

where

$\text{cov}_{aw}(x,y)$ is the manifest covariance of the mixed group

above the cut,

$cov_{bw}(x,y)$ is the corresponding covariance of the mixed group below the cut,

P_{fa} is the proportion of those above the cut which are females, and

P_{ma} , P_{fb} , P_{mb} are defined similarly.

All quantities in [8] can be estimated directly from the results of section III except C_f and C_m and we have then two independent equations which can be solved simultaneously for these two unknowns.

It might be suspected right off that this method would not work well for the present data since it has already been shown that A_6 is grossly incorrect. However, there occurs a more serious error in the estimation of k . The k estimated from the mean estimates is considerably too small; that is, the taxa means are estimated to be too close together. When k is estimated by $4max\{cov\}$ it is also too small but considerably more accurate.

(See table below.)

<u>input variable</u>	<u>true sample value of k^*</u>	<u>$4max\{cov\}$ hitmax estimate of k</u>	<u>taxa means estimate of k</u>
key 1	15.70	14.75	6.05
key 2	19.61	17.60	7.19
key 3	19.50	15.95	13.74

*by using the male and female sample values

The hitmax estimate of k is, of course, larger than the value obtained when the hitmax interval male and female means are used, because the within sex-within hitmax interval covariances of equation [1] are positive. In any event, solution of equation [1]

for P_{fw} works well when hitmax estimate is used for k but not well at all when the estimate from the mean estimates is used. The covariances obtained by solution of [8] using the revised estimates of k were usually negative or less accurate than the value zero; however those obtained by using the hitmax estimates of k were fairly accurate as shown below.

<u>Input variable</u>	<u>C_f</u>	<u>C_m</u>
key 1	1.71	1.89
key 2	1.08	0.77
key 3	2.08	1.30

These values compare fairly well with the within-sex true sample values of the interval covariances given in the Appendix; they are at least closer than zero is. When these are used along with the hitmax estimate of k in equation [1] the results of the first iteration are completely off the mark. Apparently the reason is that assumptions A_1 , A_2 , and A_3 are all quite incorrect for the present data but counterbalance one another when used simultaneously; relaxation of one, rather than making estimates better, makes them worse.

The proposed iterative scheme above is being thoroughly revised and other iterative methods are being tested; the results will be given in a forthcoming research report. Suffice it to say here that the proposed method did not work because of violation of assumption A_6 .

XIII. The Normal Method

In PR-68-4, pp. 47-54 it is suggested that another approach

results from assuming that the within taxa indicators distributions are normally distributed and considering each indicator singly. The latent parameters for each indicator are then

\bar{x}_f, \bar{x}_m : the female and male means,

s_f, s_m : the female and male standard deviations, and

P: the base-rate of the females.

The method proposed is to try arbitrary values for the parameters \bar{x}_f, s_f, P (a method is given for calculation of "consistent" values of \bar{x}_m and s_m from these) and calculate the goodness of fit χ^2 for the resulting mixed distribution when compared with the observed one. The set of parameter values yielding the smallest χ^2 value are chosen as the estimates.

This method is actually a trial and error version of the minimal χ^2 estimator method where the χ^2 function is written in terms of the latent parameters and the minimum χ^2 parameter estimates are found by solving the following set of simultaneous equations:

$$\frac{\partial \chi^2}{\partial P} = 0 \quad \frac{\partial \chi^2}{\partial \bar{x}_f} = 0 \quad \frac{\partial \chi^2}{\partial \bar{x}_m} = 0 \quad \frac{\partial \chi^2}{\partial s_f} = 0 \quad \frac{\partial \chi^2}{\partial s_m} = 0.$$

(See Cramér, 1946) Such minimal χ^2 estimates are asymptotic approximations of the maximum-likelihood estimates which have been found by Hasselblad (1966). Experimentation has shown the resulting Hasselblad calculation scheme is considerably less time consuming than the proposed trial and error χ^2 method, that being the major shortcoming of the latter. Since empirical trials of the Hasselblad method and of a multivariate generalization of the Hasselblad method have been very encouraging and will be given

In a forthcoming research report, the present analysis of an empirical test of the trial and error method is less detailed than that given in another empirical trial of the χ^2 method (Meehl et al, 1969). It is of interest to note that another method of solution of the normal model was given by Pearson (1894) where the method of moments was used. This resulted in a complicated set of nonic equations -- the solution of which was very subject to sampling error.

When using indicators that are keys made of MMPI items it is not unreasonable to assume normality within taxa in view of the generalized version of the central limit theorem given by Liapounoff (Von Mises; 1964, p. 302). It will be shown that:

The sum of n independent Bernoulli random variables tends toward a normal distribution as $n \rightarrow \infty$.

Proof (a modification of that outlined by Von Mises, p. 304)

Let population parameters for the Bernoulli variables β_i be p_i where $0 < p_i < 1$ and $i = 1, 2, 3, \dots, n$. Then according to the Theorem of Liapounoff, the distribution of $S = \sum \beta_i$ tends toward the normal distribution as $n \rightarrow \infty$, if $\frac{\sum_{i=1}^n E[|\beta_i|^k]}{S_n^k} \rightarrow 0$ as $n \rightarrow \infty$, for some $k > 2$, where $E[|\beta_i|^k]$ is the absolute moment about mean of order k for the i th variable and $S_n^k = (\sum p_i q_i)^{k/2}$. Since $E(\beta_i) = p_i$ and choosing $k = 3$, we have

$$E[|\beta_i|^3] = p_i(q_i)^3 + q_i(p_i)^3 = p_i q_i (p_i^2 + q_i^2) < p_i q_i.$$

Thus,

$$\lim_{n \rightarrow \infty} E \sum_{i=1}^n [|\beta_i|^3] / S_n^3 = \lim_{n \rightarrow \infty} \frac{\sum_{i=1}^n p_i q_i (p_i^2 + q_i^2)}{(\sum_{i=1}^n p_i q_i)^{3/2}}$$

$$\lim_{n \rightarrow \infty} \frac{\sum_{i=1}^n p_i q_i}{(\sum p_i q_i)^{3/2}} = \lim_{n \rightarrow \infty} \frac{1}{(\sum p_i q_i)^{1/2}}$$

The Liapounoff condition is satisfied if this limit is zero, in other words if $\sum p_i q_i$ is divergent. Since $p_i > 0$ suppose there exists some $\epsilon > 0$ such that $p_i > \epsilon > 0$. If N is an arbitrarily large number, then the sum of $\sum_{i=1}^n p_i q_i$ for any n of the items is

greater than $n\epsilon(1 - \epsilon)$ which is greater than N if $n > \frac{N}{\epsilon(1 - \epsilon)}$.

Thus $\sum_{i=1}^n p_i q_i$ is divergent as $n \rightarrow \infty$. In the case where $p_i = p$ for all i then the sum is a binomial random variable and it is well known that it is approximated well by a normal distribution if $n > 30$. The above result removes the restriction that the p_i are equal and pending Monte Carlo study, it will be assumed that the sum is approximated adequately by a normal distribution for any plausible values of the vector $\{p_i\}$.

The trial and error minimal χ^2 method was applied to each of three keys of the present data. The trial values for the frequency of the female group was allowed to go from 400 to 800 in increments of 50, the mean for the female group was allowed to go from 8.0 to 18.0 in increments of 1.0 and the standard deviation of the female group was allowed to go from 1.50 to 3.50 in increments of .2. Thus $9 \times 11 \times 11 = 1089$ different sets of parameter values were tried and the corresponding χ^2 values calculated for each key. The sets of values giving the minimal χ^2 values are given in the table below.

minimal χ^2 estimates

key	Males			Females			χ^2
	\bar{x}_m	s_m	1 - P	\bar{x}_f	s_f	P	
1	10.83	3.12	.355	15.00	1.70	.645	39.9
2	7.19	2.37	.410	12.00	2.10	.590	48.4
3	8.88	2.65	.433	13.00	2.50	.567	33.2

Comparison of these values with true sample values in Table 1 shows that the mean estimates are accurate to nearly within one interval which is the coarseness of the trial values, the base-rate estimates are also each within $50/1105 = .045$, the coarseness of the trial values, and the standard deviation estimates, while appearing to be the least accurate, are probably not too far off for most personality research. The three chi-square values are each significant at the .01 level; however, this does not mean the idealized model is to be rejected as this decision is correctly made only in terms of parameter estimation accuracy. For most personality research the parameter estimates are accurate enough and, accordingly, the normal model is acceptable as a description of the real taxonomy. The most alarming shortcoming of the method was the presence of several other minima of the χ^2 function, with χ^2 values very close to the minima of the minima given in the table above, and yet sometimes resulting from completely erroneous parameter estimates. Consideration of the various minima of the χ^2 values for each of the three keys can be done by using the common estimates of the base-rates. The method was modified by fixing the base-rates at the average of the three originally produced values and then repeating the procedure

of allowing \bar{x}_f and s_f to vary but the results were considerably worse than the original estimates.

The minimum χ^2 method as described by Cramér implies that there are not constraints to be imposed on any of the parameters by values assigned to others; thus, the method was tried where \bar{x}_m was allowed to vary over the same values as \bar{x}_f and s_m over the same values as s_f . The number of χ^2 values calculated for each key was 132,569 thereby resulting in a sizable calculation expense but the results were encouraging as shown in the table below.

Minimum χ^2 estimates when all parameters are allowed to vary independently

p	Males			Females			χ^2
	\bar{x}_m	s_m	1 - P	\bar{x}_f	s_f	P	
.27	8.83	2.30	.323	14.00	2.30	.677	16.75
.53	7.08	2.30	.417	12.00	2.30	.503	14.72
.63	8.61	2.50	.417	13.00	2.50	.583	11.61

The χ^2 values have associated probabilities of .27, .53, and .63 and can, therefore, be considered as non-significant. Comparison of the parameter estimates with true sample values in Table 1 show that they are substantially better than the estimates of the original method and, indeed, are nearly as good as the trial value coarseness allowed. It is notable also that there were no other locally minimum χ^2 values close to those values in the table above and resulting from totally erroneous parameter values. For key 1, there was one χ^2 value very close to the minima of 16.75 but it resulted from a more accurate set of parameter values.

TABLE 7

The observed and estimated (by the normal minimal χ^2 method)
taxa frequency distributions for key 1

score	male		female	
	observed	estimated	observed	estimated
3	3	4	0	0
4	11	7	0	0
5	8	16	0	0
6	32	29	3	0
7	41	45	5	1
8	50	57	9	5
9	58	61	20	13
10	60	54	30	29
11	65	39	27	56
12	53	24	62	89
13	29	12	93	118
14	10	5	110	129
15	7	2	103	118
16	2	1	107	89
17	1	0	61	56
18	0	0	28	29
19	0	0	14	13
20	0	0	3	5
base-rate	.389	.322	.611	.679
mean	9.57	8.83	14.10	13.99
SD	2.55	2.33	2.56	2.29

TABLE 8

The observed and estimated (by the normal minimal χ^2 method)
taxa frequency distributions for key 2

score	male		female	
	observed	estimated	observed	estimated
1	1	3	0	0
2	7	7	0	0
3	13	17	0	0
4	33	32	3	0
5	48	52	3	1
6	63	70	13	4
7	67	78	17	11
8	62	72	27	25
9	59	56	59	48
10	36	35	87	77
11	22	19	110	102
12	11	8	97	112
13	6	3	91	102
14	1	1	83	77
15	0	0	51	48
16	1	0	26	25
17	0	0	5	11
18	0	0	2	4
19	0	0	1	1
base-rate	.389	.410	.611	.586
mean	7.31	7.09	11.68	11.90
SD	2.42	2.30	2.47	2.41

TABLE 9

The observed and estimated (by the normal minimal χ^2 method)
taxa frequency distributions for key 3

score	male		female	
	observed	estimated	observed	estimated
2	4	4	0	0
3	3	7	0	0
4	12	16	0	0
5	26	29	2	1
6	54	44	3	2
7	53	59	8	6
8	57	66	16	15
9	71	64	27	31
10	46	53	69	53
11	43	37	83	79
12	31	22	91	100
13	17	12	93	100
14	9	5	109	100
15	4	2	71	79
16	0	1	56	53
17	0	0	27	31
18	0	0	15	15
19	0	0	5	6
base-rate	.389	.381	.611	.614
mean	8.60	8.24	12.84	12.97
SD	2.57	2.48	2.53	2.48

The procedure of holding the base-rate fixed at the average of the three estimates and then repeating the trial and error process was then applied; the resulting frequency distributions These tables show the estimated hitmax cuts (11, 10, 10) which agree well with the true samples values (12, 9, 10).

The results of this empirical trial then indicate that all parameters should be allowed to vary independently. After the above calculations were performed it was realized that the taxa frequencies need not total to 1105 as this apparently is an undesirable constraint on one of the taxa parameters also. The constraints would necessarily hold for a population of values but not for a sample no matter whether it is of apparently large size.

XIV. Prediction of Biological Sex Using the Normal Minimal χ^2 Parameter Estimates

As in section V the strong independence within taxa assumption allows one to apply Bayes' Rule to predict the biological sex.

Using the "interval proportions" method yields the following hits-misses table.

	Biological Sex		
	male	female	total
predicted sex			
male	360	64	424
female	70	611	681
total	430	675	1105

The proportion of "female" predictions which were correct was .91, similarly the "male" prediction hit-rate was .84 and the overall

hit-rate was .88. The proportion of the predictions which were "female" was .616 which very closely agrees with the observed base-rate of .611. These results compare with those of the maximum covariance method; both methods performed equally well.

The hitmax cut method of prediction yields a very similar hits-misses table as was the case for the maximum covariance method.

Presumably the hit-rate in the prediction of taxon membership could be improved if the strong independence within taxa assumption could be weakened. Rulon et al. (1967) provide a prediction method which assumes that the multi-indicator distributions within taxa are multivariate normal and requires estimates of the taxon indicator mean vectors and the within taxa covariance matrices. Possibly the iterative maximum covariance method could be used to provide the latter.

XV. Evaluation of the Prediction of Biological Sex

In order to evaluate the level of accuracy of the prediction of biological sex by the maximum covariance and normal methods, various methods of prediction of biological sex as a dependent or criterion variable with the indicator variables as the independent variables were used. The first method used was that of Fisher's multiple linear discriminant function analysis. As expected one highly significant linear function was found and this yielded, via the hitmax cut method, a hit-rate of .910. The second method used consisted of the following steps: (1) Each of the indicator scores was dichotomized at the true sample hitmax

cut yielding a 3-tuple vector having eight values:

1. +++
2. ++-
3. +-+
4. +-- +: score greater than hitmax
5. -++ -: score less than or equal to hitmax
6. -+-
7. --+
8. ---

(2) For each pattern, the number of individuals of each sex was tallied. (3) The prediction rule then was to determine the pattern for the individual and predict the sex that was most frequent for that pattern. This method yielded a hit-rate of .856. The third method used was a slight generalization of the second; the three hitmax cuts were determined simultaneously. The three hitmax cuts were varied over trial values to determine which set of values yielded the highest hit-rate of sex by predicting the most frequent sex just as in the second method. The set of values 12, 9, and 9 produced a maximum hit-rate of .873. It is of interest to note that a large variety of other cut score values (12, 9, 7), (12, 10, 7), (12, 9, 8), (13, 9, 8), (12, 10, 8), (11, 10, 9), (12, 10, 9), and (12, 10, 13) produced hit-rates over .865. A fourth method consisted of using the same prediction scheme referred to in sections III and XIV as the "interval proportions" method except that true sample within sex distributions were used rather than the method estimated ones. The resulting hit-rate was .910. The fifth method was similar to the fourth except the "hitmax cut" method was used. This

method produced a surprisingly low hit-rate value of .801. None of the above methods except the Fisher discriminate function, make direct use of the within taxa covariances between indicators. Hence, the last method used was the "centour method" described by Rulon et al. (1967) where the within taxa distributions are assumed to be trivariate normal. Thus, the within taxa tri-variate normal density function required the calculation of the true within sex sample values of the covariance matrices. This method then yielded a hit-rate of .885.

From these results it then is evident that the taxonomic methods predicted the actual sex just as well as the well-known criterion variable methods did. The classification method used with the taxonomic parameter methods might have been thought to be rather inadequate since it relied on the strong statistical independence assumption A_5 , but it is clear that for the present trial this was not the case.

XVI. Estimation of the Hit-Miss Table

After a taxonomic method has been applied so that it is felt that the marginal latent distributions have been accurately estimated by the use of various consistency tests, then it will normally be desirable to know how accurately individuals can be classified with a given set of indicators. In other words, an estimate of the hit-miss table is needed. A simple method for doing this goes as follows: Consider an indicator score vector (x,y,z) and let $\phi_1(x,y,z)$ be the density of taxon 1 at (x,y,z) ; similarly, let $\phi_2(x,y,z)$ be that for taxon 2. Then it follows, for example, the proportion of the mixed group which are members

of taxon 1 and are correctly classified as such is obtained as follows. The proportion of the individuals at (x,y,z) which are members of taxon 1 is

$$\theta(x,y,z) = \frac{P\phi_1(x,y,z)}{P\phi_1(x,y,z) + Q\phi_2(x,y,z)}$$

If we consider the subspace of scores such that $\theta(x,y,z) > .5$, call it S, then the desired proportion is

$$\sum_S \theta(x,y,z) \frac{n(x,y,z)}{N}$$

If $\theta(x,y,z)$ and $n(x,y,z)$ are approximated by sample values, then this value can be calculated by determining $\theta_i(x,y,z)$ for each individual ($i = 1, 2, 3, \dots, N$) and calculating $\sum_{\theta_i > .5} \theta_i$. Similarly, the proportion of the total group that are correctly classified as members of taxon 2 is $\sum_{(1-\theta_i) > .5} (1-\theta_i)$, the proportion

of taxon 1 and misclassified as taxon 2 is $\sum_{\theta_i \leq .5} \theta_i$, and the

proportion of taxon 2 and misclassified as taxon 1 is $\sum_{(1-\theta_i) \leq .5} (1-\theta_i)$.

Using assumption A₅ and the interval proportions classification method the estimated hit-miss tables for each method are given below.

Estimated hit-miss table for the maximum covariance method

	Actual Sex			proportion
	male	female	total	
predicted sex				
male	409	38	447	.405
female	45	613	658	.595
total	454	651	1105	
proportion	.411	.589	overall hit-rate:	.925

Estimated hit-miss table for the modified normal minimal χ^2 method

	Actual Sex			
	male	female	total	proportion
predicted sex				
male	393	31	424	.304
female	31	650	681	.616
total	424	681	1105	
proportion	.384	.616		
overall hit-rate:	.955			

Thus the estimates of the various hit-rates tend to be consistently a few points too high but of close magnitude.

XVII: Conclusions

(1) Both the maximum covariance and the (modified) normal minimal χ^2 methods augmented by a simple classification method certainly worked well enough to justify further study and elaboration; they worked as well as could possibly be hoped for in that criterion variable approaches were not more accurate and the overlap of the two taxa distributions was somewhat less than 2 sigma units previously thought to be a lower bound for accurate taxometric detection.

(2) The data illustrate that the assumption that the interval latent means be constant is too strong. The maximum covariance method only requires the assumption that $k_i = \Delta\bar{x}\Delta\bar{y} = k$ for all i but several of the consistency tests require that interval latent means be constant. There are two general reasons why serious consideration must be given to removal of the constant interval mean assumption. First, the maximum covariance

and the simple classification methods work very well even though the assumption is grossly violated. Second, there is no reason to regard the present data to be unusual in this respect especially with reference to other MMPI key indicator taxonomies. Further, it would seem to be quite plausible that corresponding to any latent taxonomy there are latent factors that maximally account for the indicator covariance (such as produced by factor analysis) and that these would be approximately isomorphic to those within each of the taxa; that is, the correlational structure for the mixed group differs only from that within the taxa in terms of magnitude. As the factor analyses mentioned in section II illustrated, those factors that maximally account for the item covariation of a male plus female group are the same as those for the male group and those for the female group. Discriminative indicators of a taxonomy might reasonably be expected to correlate not only due to the taxonomy, but also, possibly to a somewhat lesser extent, because of non-taxonomic sources. An example of the latter in the present data may be that of a response set where some individuals may develop a set to respond to these items, mostly which concern interests in certain activities such as planting flowers and hunting where the sex direction is easily discernable "as a man would" or "as a woman would" irregardless of the actual sex of the individual.

Another argument toward the same result is that indicators that discriminate between two taxa can also discriminate between individuals in either of the taxa; that is, things that mainly tend to be of a kind still can tend to be of a degree. The point

Is not that all taxonomies are of this nature but that some evidently are and in further taxonomic work with the MMPI this is the kind of situation to expect in view of the results of the present study.

(3) The extraordinary close fit of distributions by the normal model is evidence in favor of the accuracy and the robustness of an assumption previously viewed with some skepticism (Meehl, PR-65-2). The result of Von Mises given in section XIV shows the analytical reasonableness of the assumption, although further Monte Carlo investigation of robustness with respect to interitem correlation and of short key length is required. Further, especially for MMPI key studies, the normal and the maximum covariance methods need not be regarded as alternative methods but as consistency checks of each other. Some sort of union of the two methods might prove to be most powerful in taxonomic detection.

(4) While the assumption of constant within taxa-within interval covariance holds well, the within taxa covariances are of higher value because of the steady increasing, with respect to interval value, of the within taxa interval means. One approach would be to use the covariance mixture equation of the total group given by [3], and assuming that the within taxa covariances are equal. This assumption, although generally not a clearly adequate approximation of the truth, might still be good enough for classification methods such as centroid analysis or discriminate function analysis developed by Anderson (1946). It is mainly for purposes of classification that the estimates of the covariances might be required, but they would

also be useful as first, consistency tests in that they should not be too large and second, as information for the substantive interpretation of the taxonomy.

(5) The improvement of the parameter estimation due to the modification of the χ^2 normal method to allow all parameters to vary thereby giving approximation of the maximum-likelihood estimates makes the maximum-likelihood solution by Hasselblad of higher value, since it reduces the calculation time by a factor of 10 or more, thereby making extended Monte Carlo robustness study feasible. Use of the Hasselblad method on the present and other data show it be a very encouraging one.

(6) Certain minor modifications of the maximum covariance method of calculations are clearly desirable. The manifest covariance curves should be smoothed by methods such as moving averages. A least-squares polynomial was used with the present data and while the overall results were not improved, this will not generally be the case. The interval frequencies in the distribution tails for the present study are clearly too small. Probably, intervals should be redefined such that the mixed distribution frequencies are about equal. This latter modification could also be used in the normal model.

(7) The overestimation of the hit-rates for the estimated hit-miss table apparently results from the strong independence with taxa assumption. With accurate estimates of the within taxa covariance matrices, methods such as those of Rulon or Anderson could be used to improve the hit-rate estimates.

APPENDIX

Descriptive statistics of the output variables
as a function of the input variable
for the mixed-sex, male, and female samples

TABLE 1

Descriptive statistics for the mixed-sex group when key 1 is the input variable
and key 2 and key 3 are the output variables

Input variable Interval	value	N	P	cov(key 2, key 3)	key 2 mean	key 3 mean	key 2 SD	key 3 SD
1	3	3	.00	-0.22	3.66	4.33	0.94	0.47
2	4	11	.01	1.47	5.72	5.09	2.21	2.10
3	5	8	.02	2.56	6.62	6.50	2.23	1.93
4	6	35	.05	0.77	6.02	6.88	2.04	1.99
5	7	46	.09	2.08	7.34	7.19	2.70	2.01
6	8	59	.14	2.66	7.71	8.00	3.09	2.38
7	9	78	.21	1.95	7.80	9.09	2.50	2.29
8	10	90	.29	2.50	8.42	9.37	2.84	2.47
9	11	92	.38	2.12	8.80	10.18	2.51	2.31
10	12	115	.48	3.67	9.52	10.90	3.01	2.65
11	13	122	.59	0.82	10.10	11.46	2.53	2.17
12	14	120	.70	1.58	11.40	12.71	2.45	2.28
13	15	110	.80	1.97	11.84	13.36	2.59	2.07
14	16	109	.90	1.44	12.33	13.63	2.32	2.07
15	17	62	.95	0.64	12.90	14.32	2.17	2.01
16	18	28	.98	0.61	12.78	15.35	1.76	2.27
17	19	14	.99	0.98	12.64	15.35	1.79	1.75
18	20	3	1.00	-1.33	12.00	14.66	2.16	2.05
TOTAL		1105		6.54	9.97	11.19	3.24	3.28

TABLE 2

Descriptive statistics for the male and the female samples when key 1 is the input variable and key 2 and key 3 are the output variables

Input variable					key 2	key 3	key 2	key 3
<u>Males</u>								
Interval	value	N	P	cov(key 2, key 3)	mean	mean	SD	SD
1	3	3	.00	-0.22	3.66	4.33	0.94	0.47
2	4	11	.03	1.47	5.72	5.09	2.21	2.10
3	5	8	.05	2.56	6.62	6.50	2.23	1.93
4	6	32	.12	0.81	5.87	6.78	2.04	1.94
5	7	41	.22	1.53	7.04	7.04	2.36	1.98
6	8	50	.33	2.18	7.12	7.76	2.75	2.41
7	9	58	.47	0.47	7.22	8.63	2.23	2.17
8	10	60	.61	1.27	7.48	8.75	2.67	2.26
9	11	65	.76	1.33	7.83	9.69	1.84	2.21
10	12	53	.88	1.12	7.43	9.54	2.21	2.25
11	13	29	.95	-0.08	8.20	10.06	2.21	2.42
12	14	10	.97	0.82	8.20	10.90	1.24	1.37
13	15	7	.99	0.38	7.57	11.57	2.55	0.90
14	16	2	.99	0.00	13.00	11.50	0.00	0.50
15	17	1	1.00	0.00	12.00	15.00	0.00	0.00
	TOTAL	430		2.07	7.30	8.59	2.42	2.57
<u>Females</u>								
4	6	3	.00	-1.66	7.66	8.00	1.24	2.16
5	7	5	.01	3.28	9.80	8.40	3.81	1.85
6	8	9	.02	0.11	11.00	9.33	2.78	1.70
7	9	20	.05	3.25	9.50	10.40	2.48	2.15
8	10	30	.09	1.44	10.30	10.63	2.14	2.40
9	11	27	.13	0.09	11.14	11.37	2.35	2.12
10	12	62	.23	1.36	11.30	12.06	2.38	2.40
11	13	93	.36	0.02	10.69	11.90	2.33	1.89
12	14	110	.53	1.08	11.69	12.88	2.32	2.27
13	15	103	.68	1.52	12.13	13.48	2.32	2.07
14	16	107	.84	1.50	12.32	13.67	2.33	2.07
15	17	61	.93	0.66	12.91	14.31	2.19	2.02
16	18	28	.97	0.61	12.78	15.35	1.76	2.27
17	19	14	.99	0.98	12.64	15.35	1.79	1.75
18	20	3	1.00	-1.33	12.00	14.66	2.16	2.05
	TOTAL	675		2.17	11.67	12.84	2.47	2.52

TABLE 3

Descriptive statistics for the mixed-sex group when key 2 is the input variable and key 1 and key 3 are the output variables

Input variable					key 1	key 3	key 1	key 3
Interval	value	N	P	cov(key 1, key 3)	mean	mean	SD	SD
1	1	1	.00	0.00	6.00	6.00	0.00	0.00
2	2	7	.00	-0.53	8.85	7.28	1.72	1.16
3	3	13	.01	4.26	8.53	7.07	3.02	2.33
4	4	36	.05	3.91	8.08	6.75	2.79	2.20
5	5	51	.09	3.94	8.84	8.11	2.60	2.73
6	6	76	.16	4.53	9.71	8.39	2.80	2.69
7	7	84	.24	3.70	10.25	9.60	2.39	2.57
8	8	89	.32	4.18	10.79	9.44	2.67	2.60
9	9	118	.43	4.62	11.89	10.55	3.00	2.71
10	10	123	.54	3.31	12.75	11.53	2.59	2.40
11	11	132	.66	4.38	13.00	12.12	2.78	2.72
12	12	108	.75	3.36	14.09	12.51	2.71	2.45
13	13	97	.84	4.02	14.51	13.81	2.66	2.45
14	14	84	.92	2.98	14.64	13.19	2.51	2.53
15	15	51	.96	1.42	15.39	14.15	1.91	2.50
16	16	27	.99	4.13	14.70	13.59	2.38	2.28
17	17	5	.99	0.68	15.20	14.60	0.74	1.74
18	18	2	.99	-0.50	15.50	17.00	0.50	1.00
19	19	1	1.00	0.00	17.00	13.00	0.00	0.00
TOTAL		1105		7.97	12.33	11.19	3.36	3.28

TABLE 4
 Descriptive statistics for the male and the female samples when key 2 is the input variable
 and key 1 and key 3 are the output variables

Input variable					key 1	key 3	key 1	key 3
<u>Males</u>								
Interval	value	N	P	cov(key 1, key 3)	mean	mean	SD	SD
1	1	1	.00	0.00	6.00	6.00	0.00	0.00
2	2	7	.01	-.053	8.85	7.28	1.72	1.16
3	3	13	.04	4.26	8.53	7.17	3.02	2.33
4	4	33	.12	3.52	7.78	6.60	2.68	2.10
5	5	48	.23	3.76	8.79	8.08	2.59	2.69
6	6	63	.38	3.28	9.27	7.88	2.47	2.37
7	7	67	.54	2.45	9.71	8.97	2.15	2.35
8	8	62	.68	2.85	10.08	8.64	2.40	2.40
9	9	59	.82	1.80	10.18	9.23	2.28	2.44
10	10	36	.90	1.57	10.55	9.97	2.02	2.16
11	11	22	.95	1.37	9.45	9.68	2.23	2.26
12	12	11	.98	6.23	10.90	10.54	3.05	3.05
13	13	6	.99	0.75	11.83	10.50	3.53	1.25
14	14	1	.99	0.00	8.00	4.00	0.00	0.00
16	16	1	1.00	0.00	10.00	8.00	0.00	0.00
	TOTAL	430		3.45	9.57	8.59	2.55	2.57
<u>Females</u>								
4	4	3	.00	2.55	11.33	8.33	1.70	2.62
5	5	3	.00	6.22	9.66	8.66	2.49	3.30
6	6	13	.02	4.28	11.84	10.84	3.30	2.76
7	7	17	.05	2.01	12.35	12.11	2.11	1.74
8	8	27	.09	2.86	12.44	11.29	2.52	2.03
9	9	59	.18	2.93	13.59	11.88	2.65	2.30
10	10	87	.31	2.02	13.66	12.18	2.23	2.18
11	11	110	.47	2.90	13.70	12.61	2.29	2.54
12	12	97	.61	2.25	14.45	12.74	2.41	2.27
13	13	91	.75	3.61	14.69	14.03	2.49	2.35
14	14	83	.87	2.27	14.72	13.30	2.41	2.34
15	15	51	.95	1.42	15.39	14.15	1.91	2.50
16	16	26	.98	3.24	14.88	13.80	2.24	2.03
17	17	5	.99	0.68	15.20	14.60	0.74	1.74
18	18	2	.99	-0.50	15.50	17.00	0.50	1.00
19	19	1	1.00	0.00	17.00	13.00	0.00	0.00
	TOTAL	675		3.37	14.10	12.84	2.52	2.52

TABLE 5

Descriptive statistics for the mixed-sex group when key 3 is the input variable
and key 1 and key 2 are the output variables

Input variable					key 1	key 2	key 1	key 2
Interval	value	N	P	cov(key 1, key 2)	mean	mean	SD	SD
1	2	4	.00	1.50	6.00	4.75	1.58	1.29
2	3	3	.00	0.00	4.00	6.33	0.00	1.24
3	4	12	.01	1.69	6.33	5.91	1.70	2.87
4	5	28	.04	1.98	7.25	6.21	1.90	1.93
5	6	57	.09	1.18	8.49	6.49	2.25	2.54
6	7	61	.14	0.27	9.45	6.96	2.18	2.37
7	8	73	.21	1.75	9.27	8.04	2.27	2.83
8	9	98	.30	0.94	10.62	8.06	2.30	2.47
9	10	115	.40	3.98	11.97	9.70	2.69	2.81
10	11	126	.52	1.40	12.19	10.17	2.66	2.71
11	12	122	.63	2.31	13.20	10.80	2.41	2.62
12	13	110	.73	1.74	14.00	11.25	2.19	2.65
13	14	118	.83	2.42	14.41	11.51	2.03	2.60
14	15	75	.90	0.79	15.05	12.04	2.17	2.18
15	16	56	.95	0.56	15.73	12.62	1.90	2.19
16	17	27	.98	-0.06	16.66	14.03	1.27	1.55
17	18	15	.99	-0.96	15.46	13.20	1.85	1.79
18	19	5	1.00	-0.88	16.80	13.60	2.48	0.80
TOTAL		1105		6.59	12.33	9.97	3.36	3.24

TABLE 6

Descriptive statistics for the male and female samples when key 3 is the input variable and key 1 and key 2 are the output variables

<u>Males</u>					key 1	key 2	key 1	key 2
Input variable				cov(key 1, key 2)	mean	mean	SD	SD
Interval	value	N	P					
1	2	4	.00	1.50	6.00	4.75	1.58	1.29
2	3	3	.01	0.00	4.00	6.33	0.00	1.24
3	4	12	.04	1.69	6.33	5.91	1.70	2.87
4	5	26	.10	2.34	7.23	6.15	1.92	1.91
5	6	54	.23	1.23	8.40	6.48	2.23	2.55
6	7	53	.35	-0.06	9.34	6.56	2.09	1.96
7	8	57	.48	1.11	8.91	7.49	2.22	2.67
8	9	71	.65	0.37	10.22	7.42	2.21	2.24
9	10	46	.75	0.15	10.23	7.37	2.03	2.20
10	11	43	.85	-0.01	10.79	7.95	2.42	2.20
11	12	31	.93	0.95	11.12	9.09	2.12	1.83
12	13	17	.97	-1.38	11.41	8.94	1.71	1.98
13	14	9	.99	0.65	11.44	7.77	0.95	1.61
14	15	4	1.00	3.00	14.00	9.00	1.73	3.08
TOTAL		430		1.67	9.57	7.30	2.55	2.42
<u>Females</u>								
4	5	2	.00	-3.00	7.50	7.00	1.50	2.00
5	6	3	.00	0.00	10.00	6.66	2.16	2.49
6	7	8	.01	0.09	10.25	9.62	2.58	3.03
7	8	16	.04	0.81	10.56	10.00	1.93	2.47
8	9	27	.08	0.02	11.66	9.74	2.21	2.25
9	10	69	.18	2.03	13.13	11.26	2.44	1.98
10	11	83	.30	-0.31	12.91	11.32	2.49	2.18
11	12	91	.44	1.15	13.91	11.38	2.07	2.59
12	13	93	.58	1.02	14.47	11.67	1.93	2.53
13	14	109	.74	1.58	14.66	11.82	1.89	2.42
14	15	71	.84	0.48	15.11	12.21	2.18	1.98
15	16	56	.93	0.56	15.73	12.62	1.90	2.19
16	17	27	.97	-0.06	16.66	14.03	1.27	1.55
17	18	15	.99	-0.96	15.46	13.20	1.85	1.79
18	19	5	1.00	-0.88	16.8	13.6	2.48	0.80
TOTAL		675		2.03	14.10	11.67	2.52	2.47

REFERENCES

- Anderson, T. W. An introduction to multivariate statistical analysis. New York: Wiley, 1958.
- Cramér, H. Mathematical methods of statistics. Princeton: University Press, 1946.
- Golden, R. R. and Meehl, P. E. "Detecting latent clinical taxa, V: A Monte Carlo study of the maximum covariance method and associated consistency tests". Reports from the Research Laboratories of the Department of Psychiatry, University of Minnesota, Report No. PR-73-3, 1973.
- Hald, A. Statistical theory with engineering applications. New York: Wiley and Sons, 1952.
- Hasselblad, V. "Estimation of parameters for a mixture of normal distributions", Technometrics, 1966, 8, 431-444.
- Meehl, P. E. "Detecting latent clinical taxa by fallible quantitative indicators lacking an accepted criterion". Reports from the Research Laboratories of the Department of Psychiatry, University of Minnesota, Report No. PR-65-2, 1965.
- Meehl, P. E. "Detecting latent clinical taxa, II: A simplified procedure, some additional hitmax cut locators, a single-indicator method, and miscellaneous theorems. Reports from the Research Laboratories of the Department of Psychiatry, University of Minnesota, Report No. PR-68-4, 1968.

- Meehl, P. E., Lykken, D. T., Burdick, M. R. and Schoener, G. R.
"Identifying latent clinical taxa, III: An empirical trial of the normal single-indicator method, using MMPI Scale 5 to identify the sexes". Reports from the Research Laboratories of the Department of Psychiatry, University of Minnesota, Report No. PR-69-4, 1969.
- Pearson, K. "Contributions to the mathematical theory of evolution", Philos. Trans. Roy. Soc. London, 1894, 185, 71-110.
- Rulon, P. J. (Ed.) Multivariate statistics for personnel classification. New York: Wiley, 1967.
- Von Mises, R. Probability, statistics and truth. Translated by J. Neyman, D. Sholl and E. Rabinowitsch. New York: McMillan Company, 1939.