Universidade Federal do Rio Grande do Sul
Instituto de Matemática

Measurement of the Latent Variable in Latent Trait Analysis of Binary Data

Maria T. Albanese

Cadernos de Matemática e Estatística
Série A, nº 15, ABR/91
Porto Alegre, abril de 1991

SUMMARY

1- *Introduction*

The aim of this paper is to present and discuss some old and new results about the measurement of the latent variable when fitting a single latent variable logit-probit model to binary response data.

Consider the response function for the logit or logit-probit model given by

$$\text{logit}[\pi_i(z)] = \alpha_{i,0} + \alpha_{i,1}z, \qquad (1)$$

where $z = H^{-1}(y)$ is logistic or normally distributed and Y is uniformly distribtuted in (0,1).

Models of this type for binary response data were popularised by Bartholomew (1987). Properties of these models were extensively investigated by Albanese (1990).

We start by given the main results about scaling the latent variable in a logit model given by Bartholomew (1980, 1981, 1984). After this, we present some new theoretical results about the relation between the posterior density $h(z|x)$, its mean $E(Z|x)$ and the component score $c_1(x) = \sum \alpha_{i,1}x_i$. Some findings complement and others contradict Bartholomew's results, depending on the pattern of the $\hat{\alpha}_{i,1}$'s.

Finally we investigate the shape of the posterior density $h(z|x)$ when at least one of the $\hat{\alpha}_{i,1}$ is very large, and we suggest a cluster analysis in the latent-space based on $h(z|x)$. A shorter version of the main new results is given by Knott and Albanese (1991).

2- *Theoretical results for the relation between* posterior mean *$E(Z|x)$ and* component score $\sum \alpha_{i,1}x_i$

We shall suppose that n individuals respond 0 or 1 (no/yes, disagree/agree, for example) to each of p items designed to measure a single latent variable. The response of individual j on item i is written x_{ij}. Individual j has a value z for a latent variable Z, and we suppose that the values Z are drawn independently from a function p(z). The response function for individual j on item i is given by

$$P(X_{ji}=1|z) = \pi_i(z)$$

where

$$\ln\left[\frac{\pi_i(z)}{1-\pi_i(z)}\right] = \alpha_{i,0} + \alpha_{i,1}\, z$$

and $Z = H^{-1}(Y)$ is a logistic function in a logit model and a standard normal function in a logit-probit model; Y is uniform in (0,1).

It is assumed that the latent variable Z explains completely the association between responses for an individual, in the sense that the probability of the response pattern $x_j = (x_{j1}, x_{j2}, \ldots, x_{jp})$ for individual j with latent variable value z is

$$g(x_j \mid z) = \prod_{i=1}^{p} (\pi_i(z))^{x_{ij}} (1 - \pi_i(z))^{1-x_{ij}}$$

Another assumption is that the response function $\pi_i(z)$ is monotonic nondecreasing in the latent variable ($\alpha_{i,1}>0$). This means that increasing any z, the probability of a positive response does not decrease.

The difficulty parameter $\alpha_{i,0}$ and the discrimination parameter $\alpha_{i,1}$, i=1,,...,p are estimated by marginal maximum likelihood using a modified E-M procedure (Bock and Aitkin,1982) available as Fortran programs FACONE (Shea (1985)) and TWOMISS (Albanese and Knott (1991)).

According to Bartholomew (1980,1981) the scaling of the latent variable Y should be done via the posterior density of y given the response pattern x. Thus, for example, he suggests the mean $E(Y|x)$ (or $E(Z|x)$), which may not be particularly appropriate when the posterior density $h(y|x)$ is highly skewed.

Bartholomew (1984) shows that $E(\Phi(y)|c_1(x))$ is a nondecreasing function of the component score $c_1(x) = \sum \alpha_{i,1} x_i$ for every nondecreasing function of $\Phi(y)$. In particular, $E(Y|c_1(x))$ or $E(Z|c_1(x))$ is an increasing function of $c_1(x)$. This means that the component score induces a stochastic ordering of the posterior distributions. Thus, for example, the rank of individuals given by the component score $c_1(x)$ is the same as given by the posterior means $E(Y|x)$ and $E(Z|x)$. Therefore, if we are only interested in the ranking of the individuals on the latent scale, we can use any one of these three measures, from which the component score is the easiest to be obtained.

Bartholomew pointed out that for the logit model $E(Y|x)$ is an approximately linear function of the component score $c_1(x) = \sum \alpha_{i,1}x_i$, which can be justified by a Taylor expansion if all $\alpha_{i,1}$'s are small. At the same time, when all $\alpha_{i,1}$'s are equal to 1 and π_i's are equal to 0.5 then the exact value of $E(Y|x)$ is $(1 + c_1(x)) | (2 + A)$, where $A = \sum\alpha_{i,1}$. He also found out from empirical work that the relationship between $E(Y|x)$ and $c_1(x)$ is approximately linear well outside the range of the validity of this later result. We show that this is often false when at least one of the $\hat{\alpha}_{i,1}$ is large (say $\geq 3/\sigma$, where σ is the standard deviation of the latent distribution).

For the logit model, Barthclomew (1984) shows that when π_i and $\alpha_{i,1}$ are fixed, the posterior density $h(y|x)$ depends on x only through the component score $c_1(x)$. And therefore, under this conditions $c_1(x)$ is a Bayesian sufficient statistic of y. This property is not shared, for example, by the probit model used by Bock and Liberman (1970). We shall show that $h(y|x)$ is a function of x only through $\sum \alpha_{i,1}x_i$ if no $\alpha_{i,1}$ is infinity.

Now we give three results, which summarise Albanese(1990, Chapter 7) findings and they are valid for both logit and logit-probit models.

Result 1

If no $\alpha_{i,1}$ is infinity and two response patterns have the same posterior mean $E(Z|x)$ then they have the same component score $\sum \alpha_{i,1}x_i$ and the same posterior density $h(z|x)$.

Let $$h(z|x) = \frac{g(x|z)\ h(z)}{f(x)} .$$

Then

$$g(x|z) = \prod_{i=1}^{p} [\ \pi_i(z)\ ]^{x_i}\ [\ 1-\pi_i(z)\ ]^{1-x_i}$$

$$= \prod_{i=1}^{p} \left[\frac{\pi_i(z)}{1-\pi_i(z)} \right]^{x_i} \left[1 - \pi_i(z) \right]$$

$$= \prod_{i=1}^{p} \left\{ \exp \left[\ln \left[\frac{\pi_i(z)}{1-\pi_i(z)} \right] \right] \right\}^{x_i} \left[1 - \pi_i(z) \right]$$

$$= \prod_{i=1}^{p} \left[\exp \left(\alpha_{i,0} + \alpha_{i,1} z \right) x_i \right] \left[1 - \pi_i(z) \right] \tag{2}$$

$$= \frac{\exp(c_1(\mathbf{x})z) \exp(c_0(\mathbf{x})) f(o,z)}{f(\mathbf{x})} \tag{3}$$

where $c_0(\mathbf{x}) = \sum_{i=1}^{p} \alpha_{i,0} x_i$ and $c_1(\mathbf{x}) = \sum_{i=1}^{p} \alpha_{i,1} x_i$.

And thus

$$f(\mathbf{x}) = f(0) \exp(c_0(\mathbf{x})) M_{z|0}(c_1(\mathbf{x})) \tag{4}$$

where $M_{z|0}(c_1(\mathbf{x}))$ is the moment generating function of the latent variable Z given a zero response on all items $c_1(\mathbf{x})$.

Substituting (4) in (3), we obtain that the posterior density of z given the response pattern **x** is

$$h(z|\mathbf{x}) = \frac{\exp(c_1(\mathbf{x})z)\ h(z|0)}{M_{z|0}(c_1(\mathbf{x}))} \quad \text{for every response pattern } \mathbf{x}. \tag{5}$$

From (5) and for every response pattern **x**, the moment generating function of the posterior distribution of Z given **x** is

$$M_{z|x}(t) = \frac{M_{z|0}(c_1(x)+t)}{M_{z|0}(c_1(x))} \tag{6}$$

Therefore from (6), the posterior density $h(z|x)$ is a function of x only through the component score $c_1(x)$, if no $\alpha_{i,1}$ is infinity. This result was first given by Bartholomew (1984), when assuming π_i and $\alpha_{i,1}$ are fixed for the logit model.

Furthermore from (6)

$$E(Z|x) = \frac{M'_{z|0}(c_1(x))}{M_{z|0}(c_1(x))} = \frac{\partial}{\partial t}\left[\log M_{z|0}(c_1(x))\right]$$

And therefore

$$E(Z|x) = \frac{\partial}{\partial t} K_{z|0}(t) \Big|_{t=c_1(x)} \tag{7}$$

and

$$Var(Z|x) = \frac{\partial^2}{\partial t^2} K_{z|0}(c_1(x)+t) \Big|_{t=0} \tag{8}$$

where $K_{z|0}$ is the cumulant generating function of $Z|0$ and

$$c_1(x) = \sum_{i=1}^{p} \alpha_{i,1}\, x_i .$$

But

$$\frac{\partial^2}{\partial t^2} K_{z|0}(t) = \frac{E(e^{zt})\,E(z^2 e^{zt}) - E(z\,e^{zt})\,E(z\,e^{zt})}{E(e^{zt})^2} > 0 \tag{9}$$

Since $E(e^{zt})^2 > 0$ and from the Cauchy inequality

$$E\left[ze^{\frac{1}{2}zt} \quad e^{\frac{1}{2}zt} \right]^2 < E\left[z^2 e^{zt} \right] E\left[e^{zt} \right],$$

since Z is a random variable.

It follows from (6) that $E(Z|x)$ is increasing in $\sum_{i=1}^{p} \alpha_{i,1} x_i$.

Therefore if for two response patterns x_1 and x_2

$$E(Z|x_1) = E(Z|x_2) \overset{(8)}{\Longrightarrow} K'_{Z|0}(c_1(x_1)) = K'_{Z|0}(c_1(x_2))$$

$\overset{(7)}{\Longrightarrow} c_1(x_1) = c_1(x_2)$ since $E(Z|x)$ is increasing in $c_1(x)$.

Finally, result 1 follows from (6) and (7).

Result 2

If the posterior density $h(z|x_s)$ is normal, then its mean $E(Z|x_s)$ is linear in the component score $c_1(x_s)$.

If the mean $E(Z|x_s)$ is linear in the component score $c_1(x_s)$, then the posterior density $h(z|x_s)$ will be close to the normal distribution.

Proof:

If $h(z|x_s)$ is normal then the mean $E(Z|x_s)$ is linear in the component score $c_1(x_s)$ from (7).

If, on the other hand, the posterior mean is linear in the component score, then for some fixed a_0 and a_1,

$E(Z|x_s) = a_0 + a_1 c_1(x_s)$, so from (7) for all response patterns x_s

$$a_0 + a_1 c_1(x_s) = \frac{\partial}{\partial t} K_{z|0}(t) \Big|_{t=c_1(x_s)}$$

For typical choices of the parameters $\alpha_{i,}$, this will mean that there are distinct values t_s, $s=1,\ldots,2^p$ for which

$$a_0+a_1 t_s = \frac{\partial}{\partial t} K_{z|0}(t) \Big|_{t=t_s}.$$

This does not quite amount to the property of linearity in t which would imply a normal distribution for the posterior $h(z|x_s)$, but it comes as close as is possible to that with $K_{z|0}(t)$ determined only at a finite number of values. Fixing $K'_{z|0}(t_s)$ leads to $K_{z|0}(t_s)$ having the value appropriate for a normal distribution, for all s. If p is large, the posterior distribution is therefore constrained to be close to the normal distribution.

Bartholomew(1984) shows that if Z has a standard logistic distribution, then in some circumstances the relation between the posterior mean and the component score is linear. He conjectures that an approximate linear relation is often valid for such prior distribution of Z. The results here show that one may think of normality of posterior distributions instead of linearity.

An application of this result can be seen when fitting a logit-probit model to the Law School Admission Test, section 6, (LSAT VI), as shown below.

Law School Admission Test, Section VI

LSAT VI consists of 5 items taken by 1000 individuals designed to measure a single latent variable.

Table 1 - Frequency distribution and scores obtained by fitting the logit-probit model to the Law School Admission Test Section VI data.

Response pattern	Frequency observed	Frequency expected	Total score	Posterior mean
00000	3	2.3	0	-1.90
00001	6	5.9	1	-1.48
00010	2	2.6	1	-1.46
01000	1	1.8	1	-1.43
10000	10	9.5	1	-1.37
00100	1	0.7	1	-1.32
00011	11	8.9	2	-1.03
01001	8	6.4	2	-1.01
10001	29	34.6	2	-0.94
10010	14	15.6	2	-0.92
00101	1	2.6	2	-0.90
11000	16	11.3	2	-0.90
00110	3	1.2	2	-0.88
10100	3	4.7	2	-0.79
01011	16	13.6	3	-0.55
10011	81	76.6	3	-0.48
11001	56	56.1	3	-0.46
00111	4	6.0	3	-0.44
11010	21	25.7	3	-0.44
01101	3	4.4	3	-0.42
01110	2	2.0	3	-0.40
10101	28	25.0	3	-0.35
10110	15	11.5	3	-0.33
11100	11	8.4	3	-0.30
11011	173	173.3	4	0.01
01111	15	13.9	4	0.05
10111	80	83.5	4	0.12
11101	61	62.5	4	0.15
11110	28	29.1	4	0.17
11111	298	296.7	5	0.65

The fit of the logit-probit model provided a goodness of fit measure χ^2 equal to 15.30 on 13 degrees of freedom, which indicates a very satisfactory fit. We may conclude that the items are measuring a single latent variable. The parameter estimates $\hat{\alpha}_{i,1}$, $i=1,\ldots,5$, are equal to 0.83, 0.72, 0.89, 0.69 and 0.66, respectively.

This data is also well fitted by the Rasch model (Rasch(1960)), using RASCHMIS program (Albanese and Knott (1991)). The Rasch model is a special case of a the logit-probit model when all $\alpha_{i,1}$ are equal.

Figure 1 below shows clearly that the posterior mean E(Z|x) is a linear function of the component score $c_1(x) = \sum \alpha_{i,1} x_i$.

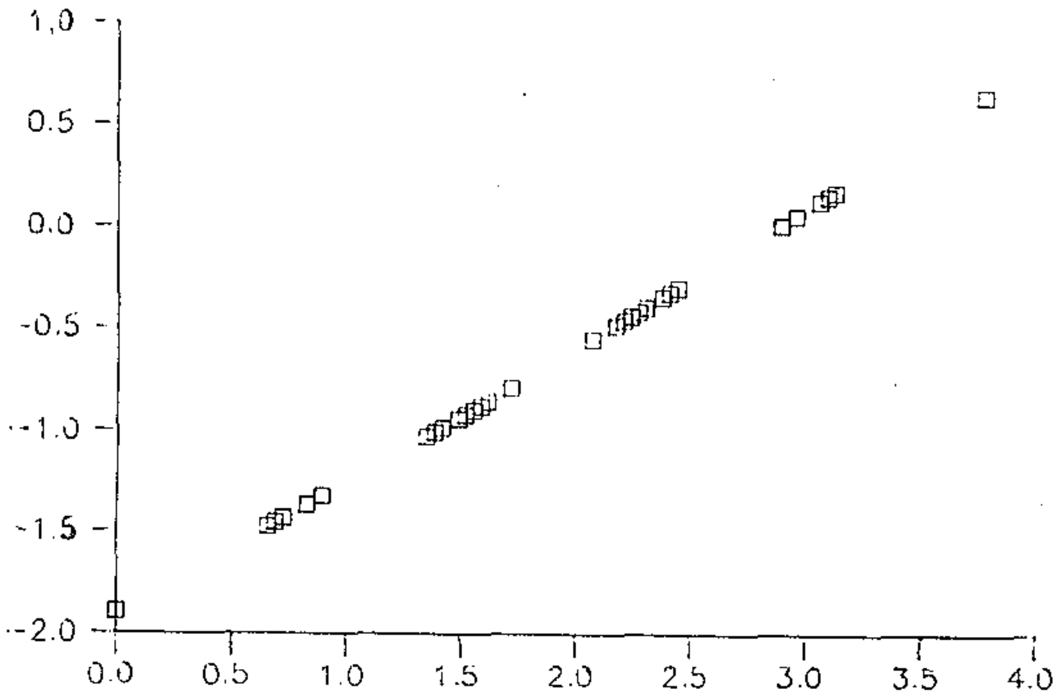

Figure 1- Relation between E(Z|x) and $\sum \alpha_{i,1} x_i$ when fitting a logit-probit model to the LSAT VI.

From Figure 1 we can see that the response patterns are distributed into 6 groups along the line $-1.92 + 0.67\, c_1(x)$. Table 2 shows that they correspond to the 6 different values assumed by $\sum x_i$. As the number of positive responses increases by one unit, both posterior means, E(Z|x) and E(Y|x), and the component score $c_1(x)$ jump to higher values.

Table 2- Estimates of $E(Y|x)$, $E(Z|x)$ and the component score $\sum \alpha_{i,1}x_i$ when fitting a logit-probit model to the LSAT VI.

group	$E(Y\|x)$	$E(Z\|x)$	$\sum \alpha_{i,1}x_i$	$\sum x_i$
1	0.007	-1.90	0.00	0
2	0.12 to 0.15	-1.47 to -1.32	0.66 to 0.89	1
3	0.21 to 0.27	-1.03 to -0.79	1.34 to 1.72	2
4	0.33 to 0.41	-0.55 to -0.30	2.07 to 2.44	3
5	0.50 to 0.55	0.01 to 0.17	2.89 to 3.13	4
6	0.69	0.64	3.79	5

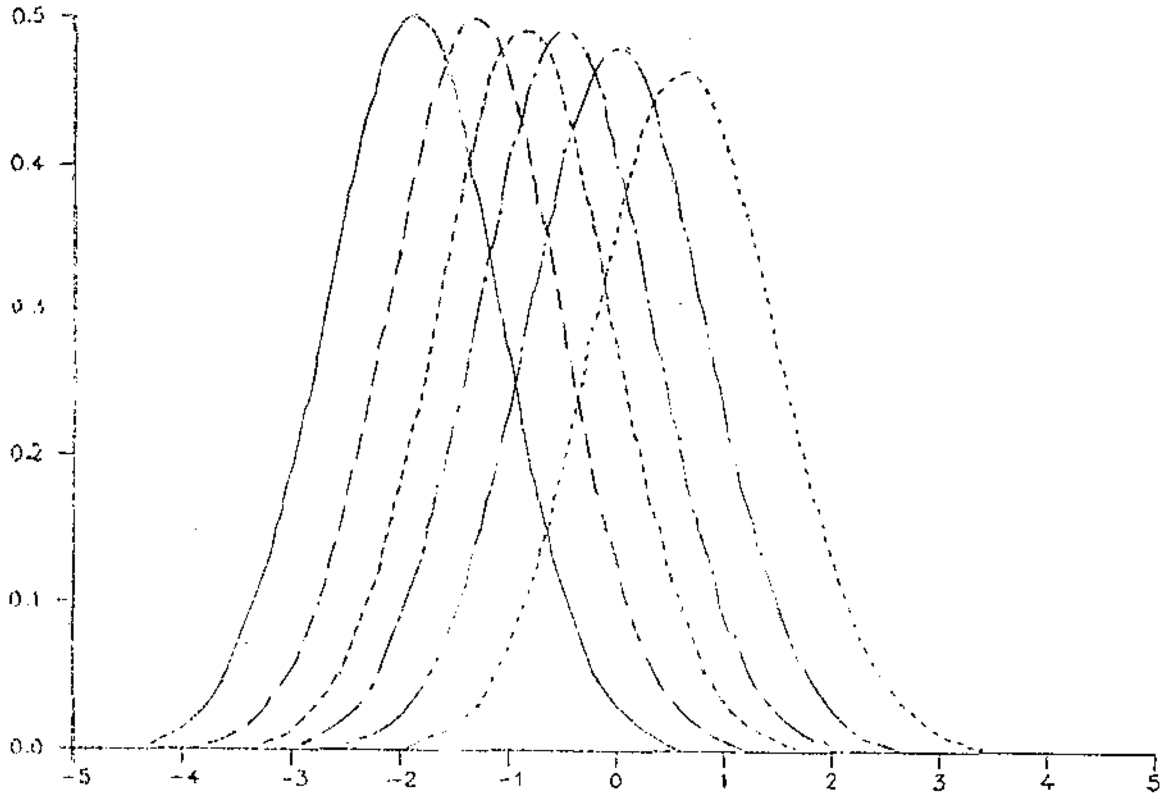

Figure 2- Posterior densities $h(z|x)$ when fitting a logit-probit model to the LSAT VI, for the response patterns '00000', '01000', '00101', '01101', '10111' and '11111'.

As $E(Z|x)$ is a linear function of $c_1(x)$ then from result 2 and for every response pattern x_s, $s=1,\ldots,32$, the posterior density $h(z|x_s)$ is approximately normal. Besides as the discrimination parameter estimates $\hat{\alpha}_{i,1}$ are nearly the same for all items, the posterior distributions have approximately the same variances (Figure 2).

Result 3

For the logit-probit (or logit) model the posterior density $h(z|\mathbf{x})$ is not a function of $\mathbf{x}$ through $\sum_{i=1}^{p} \alpha_{i,1} x_i$ if at least one of the $\alpha_{i,1}$'s is equal to infinity.

Proof:

Assume that $\alpha_{1,1}$ is equal to infinity so that

$$\pi_1(z) = \begin{cases} 0 & \text{if } z \leqslant z_0 \\ 1 & \text{if } z > z_0 \end{cases}$$

Then

$$g(\mathbf{x}|z) = \prod_{i=1}^{p} [\ \pi_i(z)\ ]^{x_i}\ [\ 1 - \pi_i(z)\ ]^{1-x_i}$$

$$= \prod_{i=1}^{p} [\ \pi_i(z)\ ]^{x_i} [\ 1 - \pi_i(z)\ ]^{1-x_i} \quad * \begin{cases} 0 & \text{if } \begin{cases} z \leqslant z_0 \text{ and } x_1 = 1 \\ z > z_0 \text{ and } x_1 = 0 \end{cases} \\ 1 & \text{if } \begin{cases} z \leqslant z_0 \text{ and } x_1 = 0 \\ z > z_0 \text{ and } x_1 = 1 \end{cases} \end{cases}$$

From (2), $g(\mathbf{x}|z)$ can also be written as

$$= \exp\left[\sum_{i=2}^{p} \alpha_{i,0} x_i + z \sum_{i=2}^{p} \alpha_{i,1} x_i \right] \prod_{i=2}^{p} [1 - \pi_i(z)] * \begin{cases} 0 & \text{if } \begin{cases} z \leqslant z_0 \text{ and } x_1 = 1 \\ z > z_0 \text{ and } x_1 = 0 \end{cases} \\ 1 & \text{if } \begin{cases} z \leqslant z_0 \text{ and } x_1 = 0 \\ z > z_0 \text{ and } x_1 = 1 \end{cases} \end{cases} \qquad (10)$$

Substituting $g(x|z)$ given by (10) in $h(z|x) = \dfrac{g(x|z)\, h(z)}{f(x)}$,

it follows that $h(z|x)$ is <u>not</u> a function of x through $\sum_{i=1}^{p} \alpha_{i,1} x_i$

if at least one of the $\alpha_{i,1}$'s is equal to infinity.

3- Applications showing the *relation between $E(Z|x)$ and $\sum \alpha_{i,1}x_i$, when at least one of the $\hat{\alpha}_{i,1}$'s is large*

One of the consequences of result 3 is that the relation between the posterior mean $E(Z|x)$ and the component score $c_1(x)$ may not be linear, if at least one of the $\hat{\alpha}_{i,1}$'s is large (say $\geq 3|\sigma$, where σ is the standard deviation of the latent distribution). This situation is illustrated using two tests with 18 to 40 items, and different number of large $\hat{\alpha}_{i,1}$.

3.1- Test 12

This test was applied by the National Foundation for Educational Research in order to measure the reading ability of English, Welsh and Irish children aged 11 in 1983. Test 12 consists of 18 items and it was answerd by 502 children. This data set is reasonable fitted by a logit-probit model with parameter estimates given in Table 3 below.

Table 3- Parameter estimates and asymptotic standard deviations from fitting a logit/probit model to Test 12.

Item i	$\hat{\alpha}_{i,1}$	SD($\hat{\alpha}_{i,1}$)	$\hat{\alpha}_{i,0}$	SD($\hat{\alpha}_{i,0}$)	$\hat{\pi}_i$
1	1.17	0.18	2.34	0.19	0.91
2	1.62	0.19	1.11	0.14	0.75
3	1.26	0.15	0.05	0.11	0.51
4	1.61	0.18	0.95	0.14	0.72
5	2.08	0.23	0.45	0.15	0.61
6	1.34	0.17	-0.88	0.13	0.29
7	1.49	0.17	0.36	0.12	0.59
8	2.20	0.26	2.18	0.23	0.90
9	1.49	0.17	0.84	0.13	0.70
10	0.87	0.13	-0.13	0.10	0.47
11	0.62	0.12	-0.35	0.10	0.41
12	2.02	0.22	0.62	0.15	0.65
13	1.24	0.18	-1.58	0.16	0.17
14	1.65	0.23	-1.76	0.18	0.15
15	4.50	0.83	3.72	0.51	0.98
16	4.37	0.70	2.51	0.40	0.92
17	1.75	0.20	1.55	0.17	0.82
18	1.58	0.18	0.66	0.13	0.66

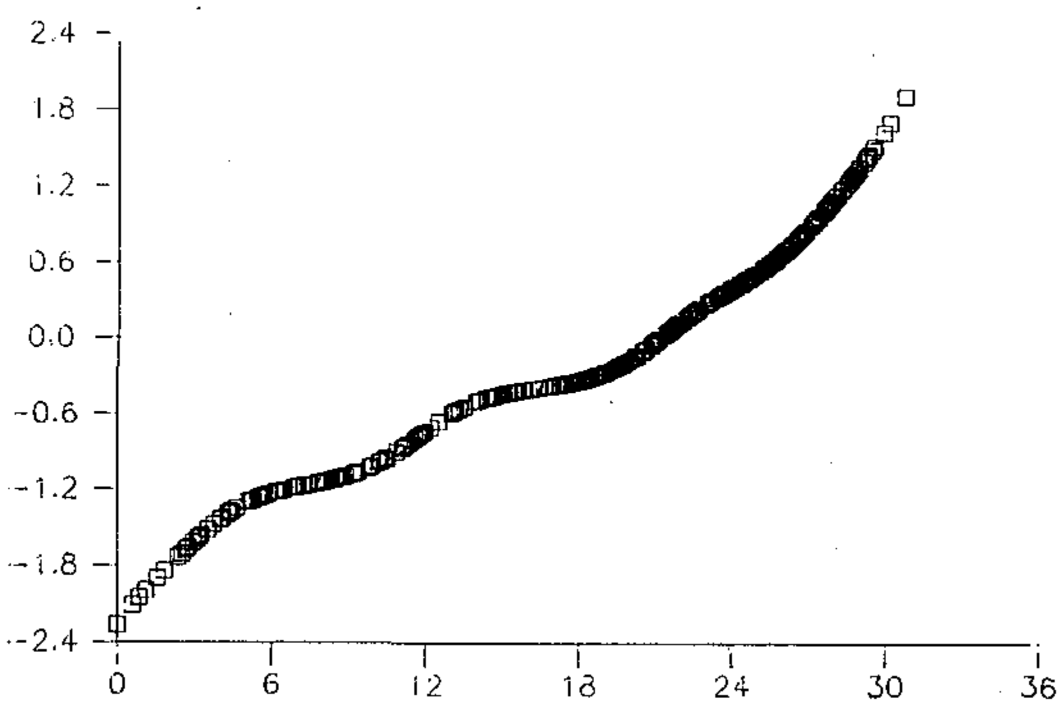

Figure 3- Relation between E(Z|x) and $\sum \alpha_{i,1} x_i$ when fitting a logit-probit model to Test 12.

Table 4- Estimates of E(Y|x), E(Z|x) and the component score $\sum \alpha_{i,1} x_i$ when fitting a logit-probit model to Test 12.

E(Y|x)	E(Z|x)	$\sum \alpha_{i,1} x_i$	$\sum x_i$	n (%)
[0.02;0.10)	[-2.26;-1.35)	[0.00; 4.65)	0 to 4	51 (10)
[0.10;0.20)	[-1.35;-0.90)	[4.65;10.91)	3 to 7	55 (11)
[0.20;0.30)	[-0.90;-0.55)	[10.91;13.36)	5 to 10	25 (5)
[0.30;0.40)	[-0.55;-0.26)	[13.36;19.13)	6 to 12	71 (14)
[0.40;0.50)	[-0.26; 0.00)	[19.13;21.16)	9 to 12	45 (9)
[0.50;0.60)	[0.00; 0.27)	[21.16;22.97)	10 to 14	34 (7)
[0.60;0.70)	[0.27; 0.57)	[22.97;25.30)	12 to 15	77 (15)
[0.70;0.80)	[0.57; 0.94)	[25.30;27.22)	14 to 16	65 (13)
[0.80;0.90)	[0.94; 1.48)	[27.22;30.91)	15 to 17	67 (13)
[0.90;0.95]	[1.48; 1.91]	[30.91;32.88]	17 to 18	12 (2)

Figure 3 shows that the relation between $E(Z|x)$ and $c_1(x)$ is linear only for a partition of $E(Z|x)$ in 5 specific sections. Each one of the first 4 sections corresponds approximately to the first 4 intervals for $E(Z|x)$ given in Table 4.

In the first interval ($-2.26 \leqslant E(Z|x) < -1.35$) we observe that 98% of the individuals have answered '0' to both items 15 and 16.

In the second and fourth intervals, there is a greater change in component score $c_1(x)$ than in posterior mean $E(Z|x)$, which is shown by two slightly flat sections. The highest proportion of answers to items 15 and 16, in the second interval 52.71% to '00', while in the fourth interval is 70.4% to '11'.

In the third interval, all individuals answered '1' to at least one of the items 15 and 16, and the higher proportion of patterns is 44% to '11'.

Considering all the intervals together, for which $E(Z|x) \geqslant -0.26$ or $c_1(x) \geqslant 19.13$ the relation between these two measures is linear and 98.3% of the individuals have answered '11' to items 15 and 16.

From these results we can conclude that the non-linearity between the posterior mean $E(Z|x)$ and the component score $c_1(x)$ over all values assumed by them is due to the whole response pattern, instead of only due to the items with large $\hat{\alpha}_{i,1}$ ($i=15,16$).

Consider that the actual values of $\hat{\alpha}_{15,1}$ and $\hat{\alpha}_{16,1}$ are infinity, and therefore, $g(x|z)$ can be written as (9). Now Figure 4, instead of Figure 3, shows the relation between $E(Z|x)$ and the component score, which is also not linear.

Figure 4 shows roughly three curves, which one corresponding to an specific pattern for '$x_{15}x_{16}$', the answers to the items with $\alpha_{i,1}$ equal to infinity. From the top to the bottom, the first curve is given by the 359 response patterns with '$x_{15}x_{16}$' = '11', the second one by the 39 and 2 response patterns with '10' and '01', and finally, the last one for the 106 patterns with '00'.

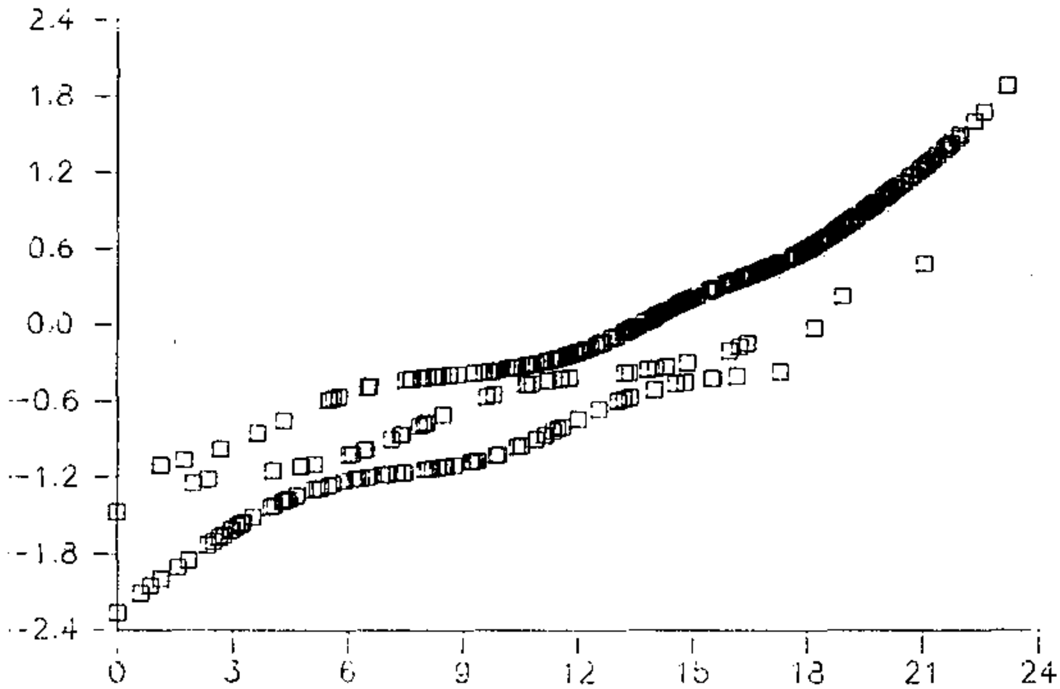

Figure 4- Relation between E(Z|x) and $\sum \alpha_{i,1}x_i$, assuming $\alpha_{15,1}$ and $\alpha_{16,1}$ equal to infinity, when fitting a logit-probit model to Test 12.

Comparing Figures 3 and 4 we can conclude that for a specific answer to the items with $\alpha_{i,1}$ equal to infinity, the relation between E(Z|x) and the component score $c_1(x)$ is closer to linearity than when taking $c_1(x)$ over all items and $\alpha_{i,1}$'s not equal to infinity.

3.2- Test 13

Test 13 was also applied by the National Foundation of Education Research in order to measure the reading ability of pupils of aged 11 in 1983. The sample size was 498 and the test length 40 items. The distribution of the discrimination parameter estimates $\hat{\alpha}_{i,1}$, i=1,...,40, when fitting a logit-probit model may be given by

$\hat{\alpha}_{i,1}$	count
[0.31; 1.00)	10
[1.00; 2.00)	20
[2.00; 3.00)	4
> 3.00	6

Therefore the fitting of a logit-probit model to Test 13 (length 40) provides six parameter estimates $\hat{\alpha}_{i,1}$ bigger than 3.0.

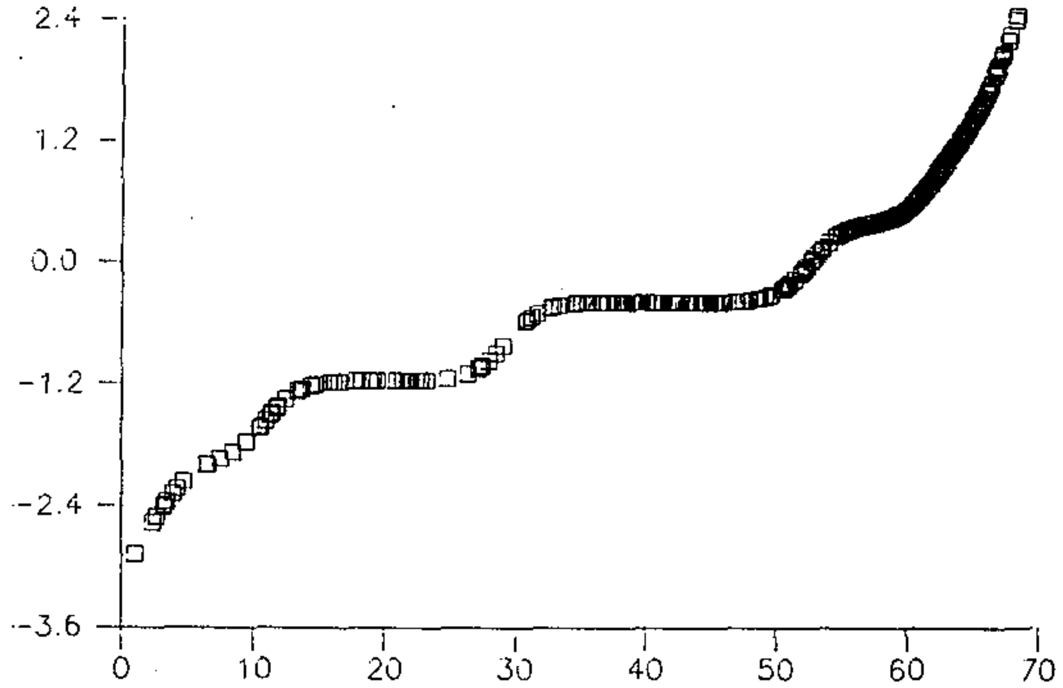

Figure 5- Relation between E(Z|x) and $\sum \alpha_{i,1} x_i$ when fitting a logit-probit model to Test 13.

Figure 5 shows that the relation between E(Z|x) and the component score $c_1(x)$ is not linear. As in Figure 3, the dark parts of the curve represent great concentration of individuals with different response patterns in a small range of E(Z|x) and $c_1(x)$ values.

Table 5 below was constructed in such way that it reflects the different aspects of the relationship between E(Z|x) and $\sum \alpha_{i,1} x_i$ displayed in Figure 5. Thus, for example, the second and fourth intervals represent the two flat parts of the curve, in which E(Z|x) remains approximately constant, while $c_1(x)$ increases significantly. In the second interval, for 21 individuals with different response patterns, E(Z|x) ranges only from -1.18 to -1.09, while $c_1(x)$ increases significantly from 15.72 to 26.42. In the fourth interval, for a large number of individuals (75) with different response patterns, E(Z|x) remains almost constant (-0.40 to -0.38), while $c_1(x)$ increases from 34.61 to 46.86.

Table 5- Estimates of E(Y|x), E(Z|x) and the component score $\sum \alpha_{i,1} x_i$ when fitting a logit-probit model to Test 13.

E(Y|x)	E(Z|x)	$\sum \alpha_{i,1} x_i$	$\sum x_i$	n (%)
[0.005;0.120)	[-2.87;-1.18)	[1.02;15.72)	1 to 13	30 (6)
[0.120;0.143)	[-1.18;-1.09)	[15.72;26.42)	9 to 18	21 (4)
[0.143;0.346)	[-1.09;-0.40)	[26.42;34.61)	13 to 21	21 (4)
[0.346;0.353)	[-0.40;-0.38)	[34.61;46.86)	17 to 30	75 (15)
[0.353;0.400)	[-0.38;-0.26)	[46.86;50.49)	21 to 30	15 (3)
[0.400;0.510)	[-0.26; 0.02)	[50.49;52.71)	22 to 27	20 (4)
[0.510;0.601)	[0.02; 0.26)	[52.71;54.62)	24 to 27	10 (2)
[0.601;0.701)	[0.26; 0.56)	[54.62;60.26)	25 to 32	118 (24)
[0.701;0.801)	[0.56; 0.92)	[60.26;62.56)	30 to 36	85 (17)
[0.801;0.900)	[0.92; 1.42)	[62.56;65.10)	31 to 37	63 (13)
[0.900;0.982]	[1.42; 2.43]	[65.10;68.53]	34 to 39	40 (8)

The curve also changes its slope significantly, but is less flat than in the second and fourth intervals, when E(Z|x) ranges from 0.26 to 0.56 and $c_1(x)$ from 54.62 to 60.26. In this interval, there is a great concentration of individuals (24% against the expected 10%), all of them with different response patterns.

The investigation of reasons why flat parts occur led us to look at the relation between the distribution of the number of positive responses given to the 6 items with large $\hat{\alpha}_{i,1}$ (≥3.0) and the slope of the curve. A selection of the results is displayed in Table 6.

Table 6- Frequency distribution of the number of positive responses given to the six items with $\hat{\alpha}_{i,1} \geq 3.0$ for some intervals of E(Z|x).

E(Z|x)	0	1	2	3	4	5	6	total
[-2.87;-1.18)	25	5						30
[-1.18;-1.09)	9	8	3	1				21
[-1.09;-0.40)	5	4	2	9	0	1		21
[-0.40;-0.38)	0	3	4	49	4	9	6	75
[0.26; 0.56)	0	0	0	2	21	44	51	118

Table 6 shows that there is a great combination of possible results for the 6 items with large $\hat{\alpha}_{i,1}$, even in the flat parts of the curve. Thus, for example, where 75 response patterns have approximately the same E(Z|x), -0.40 to -0.38, the only possible result that does not happen is all 6 items answered '0'. This means that the response patterns are not concentrated on a specific configuration for the items with large $\hat{\alpha}_{i,1}$.

Moreover, in the flat parts of the curve we found response patterns with the same response to the 6 items with large $\hat{\alpha}_{i,1}$, have component score values significantly different. For example, two response patterns, in which all these 6 items were answered '0', were associated to either a component score equal to 15.71 or 24.77 for nearly the same expected value (-1.18 and -1.14).

This implies, that at least for these response patterns, the greater relative difference between the component scores than between the posterior means E(Z|x) is not due to the items with large $\hat{\alpha}_{i,1}$.

For an expected number of individuals equal to 10% and E(Z|x) between 0.26 and 0.56, it was observed 23.7% of the sample, of whom 90.5% have answered '1' to five or to the six items with large $\hat{\alpha}_{i,1}$. From this point E(Z|x) and $\sum \alpha_{i,1} x_i$ increases faster and most of the individuals have answered '1' to the 6 items with large $\hat{\alpha}_{i,1}$.

These results combined with those from Table 6 indicate that when at least one of the $\hat{\alpha}_{i,1}$'s is very large, for some response patterns it may occur that the posterior mean practically does not change while the component score increases significantly, even when the response to the items with large $\hat{\alpha}_{i,1}$ is fixed.

4- *Distribution of the individuals on the latent scale according to the posterior density $h(z|x)$*

Very often, in practice, we are not only interested in the ranking of the individuals, which is obtained either from the component scores or from the posterior means $E(Z|\mathbf{x})$ or $E(Y|\mathbf{x})$. Thus, for example, in Educational Testing, we may be interested in comparing the lower with the higher ability group of individuals. The criterion for the distribution (allocation) of the respondents in groups is usually based on an arbitrary percentage, for example 20%.

If we know the distribution of the individuals along the latent scale, then we can use this information to partition the sample in groups. One way to do this is to use the information given by the posterior density $h(z|\mathbf{x})$ or even the mean $E(Z|\mathbf{x})$.

If we intend to use the mean $E(Z|\mathbf{x})$ as the measure of comparison between the position of the individuals on the latent scale then we must have information about the shape of $h(z|\mathbf{x})$, at least in terms of skweness and spread.

Let us consider two individuals with different response patterns $\mathbf{x}_1$ and $\mathbf{x}_2$ and the posterior densities $h(z|\mathbf{x}_1)$ and $h(z|\mathbf{x}_2)$, which are not skew and have nearly the same dispersion. If $h(z|\mathbf{x}_1)$ and $h(z|\mathbf{x}_2)$ have roughly the same mean then $\mathbf{x}_1$ and $\mathbf{x}_2$ lead to the same beliefs about the value of Z.

In these situations the mean $E(Z|\mathbf{x})$ is a reliable measure to compare individuals according to their position on the latent scale.

The main goal of this section is to present the results from the investigation of the shape of the $h(z|\mathbf{x})$ we have found so far in practice. This will be done using two real data sets for tests with 18 and 40 items, for which the fittings of a logit-probit model yield two and six large $\hat{\alpha}_{i,1}$'s (bigger than 3.0).

4.1- **Test 12**

Test 12 has 18 items and was answered by 502, which have provided 417 different response patterns. Therefore for each one of these 417 response patterns there is one posterior density h(z|x).

In order to investigate the shape of the posterior densities h(z|x) and how they are distributed along Z, we have we have selected a representative sample of observed h(z|x)'s, from which we have chosen to display here the following three sets (Figures 6 to 8).

Figure 6- Posterior densities h(z|x) for the first ten different response patterns of test 12, for which $-2.26 \leqslant E(z|x) \leqslant -1.67$.

Figure 6 displays the posterior distributions h(z|x) for the first 10 different response patterns with the smallest E(Z|x) (or E(Y|x)). For these sets of h(z|x), the mean E(Z|x) assumes values from -2.26 to -1.67 while E(Y|x) ranges from 0.02 to 0.06. For these response patterns most of the items were answered '0', including items 15 and 16 for which $\hat{\alpha}_{i,1}$ are large.

The continuous line represents h(z|x) when an individual has answered '0' to all items and corresponds to the lowest observed ability. It also presents the biggest dispersion and is skewed to the left. As E(Z|x) increases, h(z|x) becomes less skew, less spread and similar posterior means represent individuals with nearly the same h(z|x).

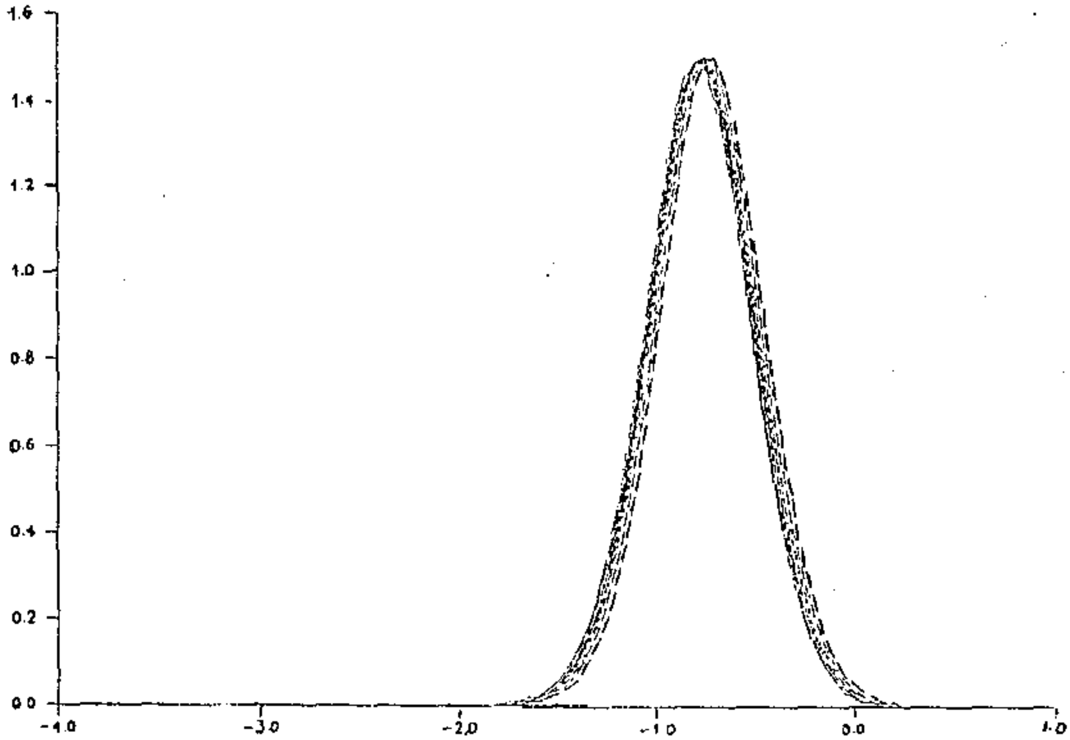

Figure 7- Posterior densities h(z|x) for some response patterns of Test 12, for which -0.81 $\leqslant$ E(Z|x) $\leqslant$ -0.66.

Figure 7 displays the h(z|x) for ten different response patterns, for which E(Z|x) assumes values from -0.81 to -0.66 (or E(Y|x) ranges from 0.22 to 0.27). The normal probability plots have shown that h(z|x)'s are approximately normal distributions with the same dispersion. This implies that the difference between h(z|x) is only in terms of location and these individuals lead to approximately the same beliefs about the value of Z. This was also found to be true for response patterns with similar posterior means, but which are not located in the ten higher observed positions on the latent scale.

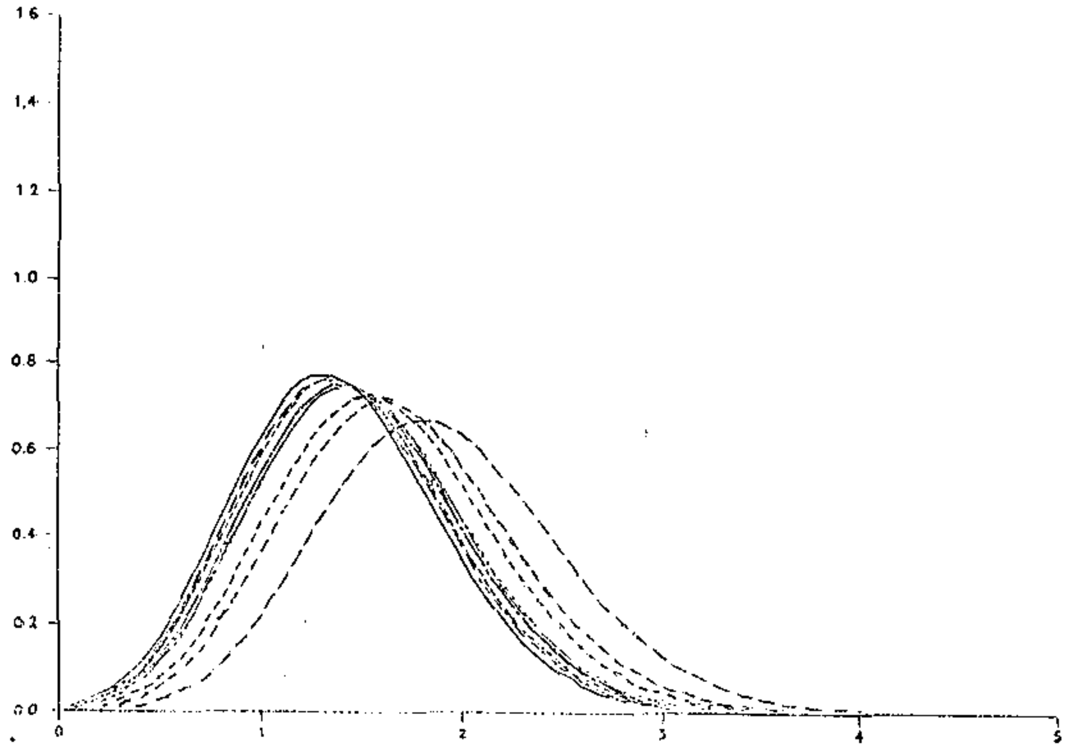

Figure 8- Posterior densities h(z|x) for the last ten different response patterns of Test 12, for which $1.39 \leqslant E(Z|x) \leqslant 1.90$.

Figure 8 displays the ten observed response patterns, which provide the ten different largest posterior means ($1.39 \leqslant E(Z|x) \leqslant 1.90$ and $0.89 \leqslant E(Y|x) \leqslant 0.95$). For these response patterns most of the items were answered '1', including as expected the items with large $\hat{\alpha}_{i,1}$. Now the posterior densities h(z|x) are slightly skew to the right and the dispersion is increasing as E(Z|x) increases.

From Figures 6 to 8 and many others not represented here, we can conclude that the means E(Z|x) (or E(Y|x)) represent very well the position of the individuals on the latent scale, since the posterior distribution is approximately normal. There are some restrictions on the extremes, where h(z|x) is slightly skew to the left or to the right depending on the responses to the items with large $\hat{\alpha}_{i,1}$.

These results lead us to conclude that we do not need to determine all the $h(z|x)$'s to have a clear idea about the distribution of $h(z|x)$ along Z. Instead, we can select a representative sample of $h(z|x)$, selecting **x** so that the whole set of values assumed by the $E(Y|x)$ (or $E(Z|x)$) is covered. Using this criterion we shall determine the $h(z|x)$ for the two response patterns, which provide the observed lowest and highest position on the latent scale and for those which corresponds $E(Y|x)$ equal to 0.10, 0.20,....,0.90.

Figure 9 displays a representative collection of the 417 observed posterior densities $h(z|x)$. Based on this figure we detect groups of response patterns (or individuals) who have nearly the same posterior densities $h(z|x)$, differing only on the location parameter.

Thus, for example, the position on the latent scale of individual with $E(Z|x)$ from -1.35 to 0.00 can be measured more precisely than those with $E(Z|x)$ ranging from -2.26 to -1.35 or from 0.57 to 1.91.

Therefore if we desire to make groups of individuals according to their distribution on the latent scale, we can combine the information obtained from Figure 9, which gives the shape of $h(z|x)$ and its location along Z, with the results from Table 4, which provides the observed frequency distribution of those $h(z|x)$.

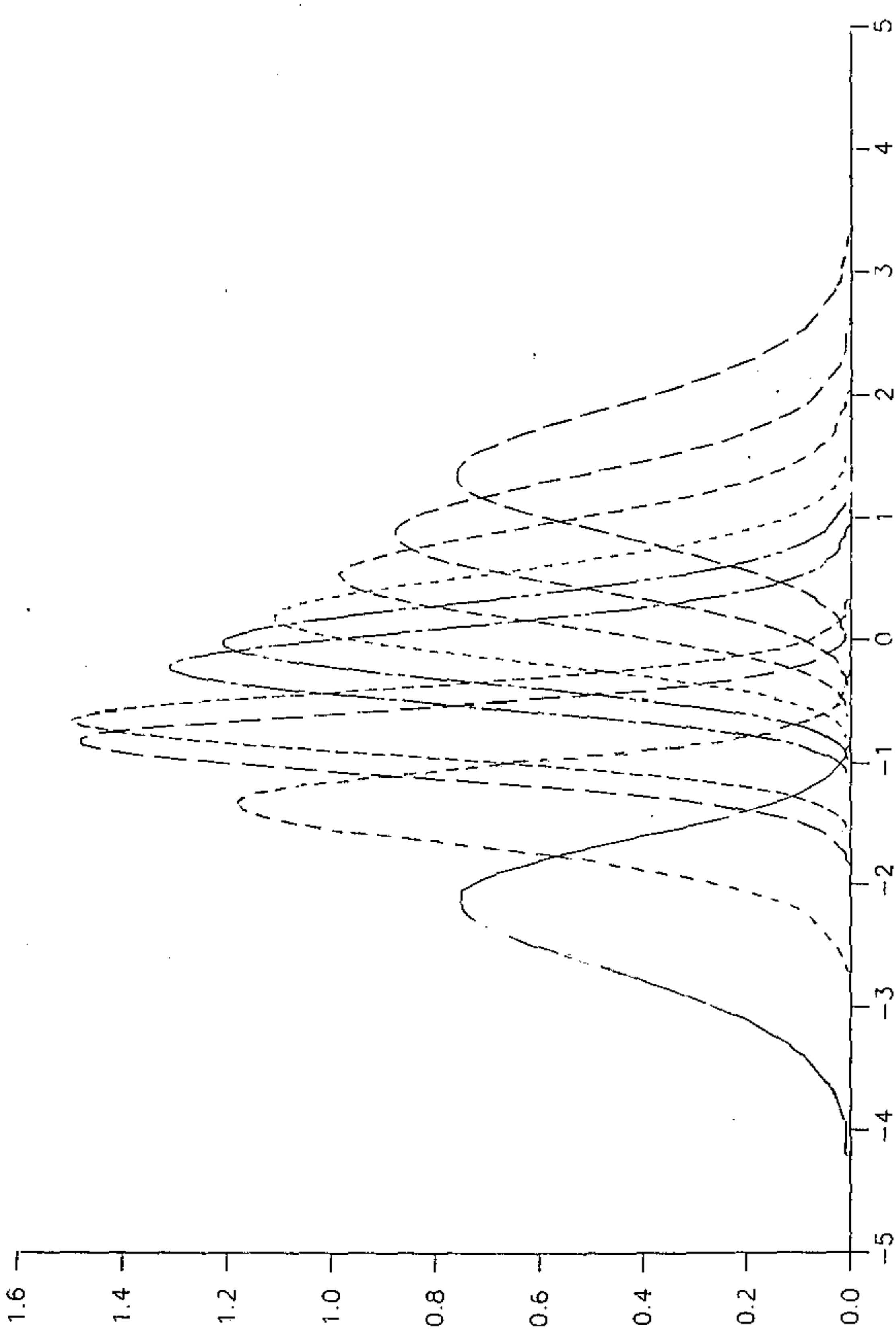

Figure 9- Representative collection of posterior densities h(z|x) for the observed response patterns of Test 12.

4.2- **Test 13**

As described in the previous section, Test 13 has 40 items and the fitting by a logit-probit model yielded 6 items with $\hat{\alpha}_{i,1} \geq 3.0$. The 498 individuals who answered the test provided 488 different response patterns to which one corresponds one posterior density $h(z|x)$.

As for Test 12, we have determined the posterior densities $h(z|x)$ for a significant number of observed response patterns, so that the $E(Z|x)$'s are distributed along the whole latent scale Z. More precisely, for each interval of $E(Z|x)$ in Table 5 we have selected at least 10 different response patterns and we have determined their $h(z|x)$'s.

The results agree with those from Test 12 in terms of

(1) the equivalence between similar $E(Z|x)$'s (or $E(Y|x)$) and nearly equal $h(z|x)$'s. Similar $E(Z|x)$'s (or $E(Y|x)$'s) come from nearly equal $h(z|x)$'s, specially for those individuals who are not located in the extrems left and right of the latent scale;

(2) representativity of the whole set of $h(z|x)$ through few $h(z|x)$, taking into account the two response patterns which are in the lowest and highest position of the latent scale and at least 9 response patterns, for which their $E(Y|x)$ are distributed along (0,1).

Figure 10 displays a representative collection of observed posterior densities $h(z|x)$ for Test 13, for which $E(Y|x)$ are equal to 0.05, 0.10, 0.20, 0.35, 0.40, 0.50, 0.60, 0.70, 0.80, 0.90 and 0.98 or $E(Z|x)$ are equal to -2.97, -1.35, -0.90, -0.38, -0.25, 0.00, 0.25, 0.56, 0.92, 1.42 and 2.43.

Looking at Figure 10 we can see how $h(z|x)$ changes in terms of shape and dispersion along Z. It also provides a measure of precision for comparing individuals located in different points of along the latent scale Z.

In Figure 10, although most of the consecutive means $E(Y|x)$ are equidistant, the posterior densities $h(z|x)$ corresponding to $E(Y|x)$ equal to 0.60 is closer(more similar) to that with mean 0.50 than 0.70.

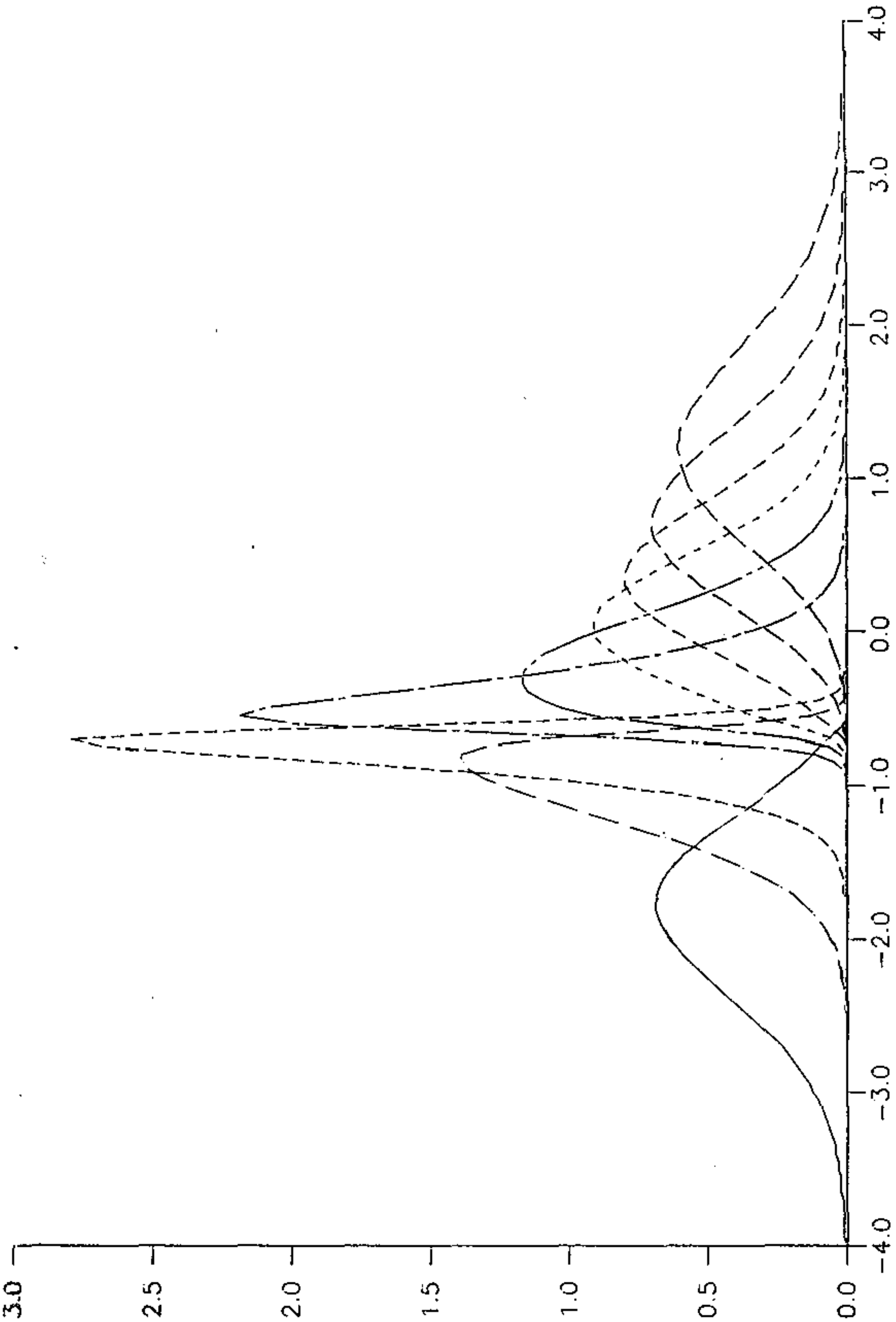

Figure 10. Representative collection of posterior densities h(z|x) for the observed response patterns of Test 13.

Furthermore, combining the information given in Table 5 and Figure 10 we will be able to make groups using the information given by the $h(z|x)$'s. Thus, for example, the 21 (or 75) individuals who have $E(Y|x)$ between -1.18 and -1.09 (or -0.40 and -0.38) should belong to the same group, since they have nearly the same $h(z|x)$ and therefore they lead to the same beliefs about the latent variable.

We have also look at diagrams like Figures 9 and 10 for many tests with smaller number of items (10 or less), which yielded one or two large $\hat{\alpha}_{i,1}$ (≥ 3.0) when fitted by a logit-probit model. The posterior densities tend to be skew to the left or to the right depending on the responses to the items with large $\hat{\alpha}_{i,1}$ and the variances of the distributions differ. When the $\hat{\alpha}_{i,1}$ are close to each other the posterior distributions are approximately normal distributed, which confirm result 2.

5- *Conclusions*

The investigations carried out in this paper for the logit and logit-probit models lead to the following conclusions:

(1) If no $\alpha_{i,1}$ is infinity and two response patterns have the same mean $E(Z|x)$ then they have the same component score $\sum \alpha_{i,1}x_i$ and the same posterior density $h(z|x)$.

(2) If the posterior density $h(z|x)$ is normal, then its mean $E(Z|x)$ is linear in the component score $c_1(x)$. If the mean $E(Z|x)$ is linear in $c_1(x)$, then the posterior density $h(z|x)$ will be close to the normal distribution.

(3) The posterior density $h(z|x)$ is <u>not</u> a function of x through the component score $c_1(x)$ if at least one of the $\alpha_{i,1}$'s is equal to infinity.

(4) The relation between the posterior mean, E(Y|x) or E(Z|x), and the component score is unlikely to be linear when at least one of the $\hat{\alpha}_{i,1}$ is large (say $\geq 3|\sigma$, where σ is the standard deviation of the latent distribution). This may be due to the fact that the component scores are strongly dependent on the values of $\alpha_{i,1}$ while E(Y|x) (or E(Z|x)) depends on π_i, which is nearly the same for all $\hat{\alpha}_{i,1} \geq 3|\sigma$, independently of $\hat{\alpha}_{i,0}$.

(5) The greater the test length, the greater the possible number of different response patterns and configuration of $\hat{\alpha}_{i,1}$'s can occur and the less likely the linearity between the posterior mean and the component score seems to be.

(6) Significant differences between component scores do not always reflect different positions on the latent scale, according to the E(Y|x) or E(Z|x). They are shown through flat sections or jumps in the curve obtained when plotting the component scores against the means E(Y|x) (or E(Z|x)).

(7) The occurrence of flat sections seems to depend on the number of items with large $\hat{\alpha}_{i,1}$ and test length. At the same time, we expect that the effect of 2 large $\hat{\alpha}_{i,1}$ in a test with 40 items is smaller than in a test with 20 items. Usually, they do not present a specific pattern for the items with large $\hat{\alpha}_{i,1}$.

(8) Consider a test for which the sample size not small compared with the number of items, for example Test 12 and 13. It seems that even though when fitting a logit-probit model some items have large $\hat{\alpha}_{i,1}$, similar E(Z|x)'s (or E(Y|x)'s) come from nearly equal h(z|x)'s, which are approximately normal distributions, specially for those individuals who are not located at the extreme left and right of the latent scale. For a smaller number of items, h(z|x) tends to be skew to the left or to the right depending on the responses to the items with large $\hat{\alpha}_{i,1}$, and the variances are different.

Therefore the general pattern that emerges is that as the number of items increases, the posterior distributions look more normal and less skew, though with different variances. This is even true if there are several $\alpha_{i,1}$'s estimated as large, and the relation between the posterior means and the component scores is far from linear. This implies that in general the posterior mean is a reliable measure of the latent variable, and it has better behaviour than the component scores.

(9) We do not need to determine all the $h(z|\mathbf{x})$'s to have a clear idea about the distribution of $h(z|\mathbf{x})$ along the latent scale Z. Instead, we can select a representative sample of $h(z|\mathbf{x})$, selecting the response pattern $\mathbf{x}$ so that the whole set of values assumed by $E(Y|\mathbf{x})$ (or $E(Z|\mathbf{x})$) is covered.

(10) If we desire to make groups of individuals according to their distribution on the latent scale, we can combine the information obtained from the shape of $h(z|\mathbf{x})$'s for all $\mathbf{x}$ (Figure 10, for example) with the observed frequency distribution of these $h(z|\mathbf{x})$ (Table 5, for example).

6 - References

Albanese,M.T.(1990) Latent variable models for binary response. Ph.D. thesis. University of London.

Albanese,M.T. and Knott,M.(1991) TWOMISS: a computer program for fitting a two-factor logit-probit model to binary response data when observations may be missing.

Albanese,M.T. and Knott,M.(1991) RASCHMISS: a computer program for fitting the Rasch model to binary response data when observations may be missing.

Bartholomew,D.J.(1980) Factor analysis for categorical data. J. Roy. Statist. Soc., B, 42, 293-321.

Bartholomew,D.J.(1981) Posterior analysis of the factor model. Br. J. Math. and Statist. Psychol., 34, 93-99.

Bartholomew,D.J.(1984) Scaling binary data using a factor model. J. Roy. Statist. Soc., B, 46, 120-123.

Knott,M. and Albanese, M.T.(1990) Posterior distribution of factor values ans scaling for binary data. Submitted for publication.

Rasch,G.(1960) Probabilistic modelsfor some intelligence and attainment tests. Copenhagen: Paedagogiske Institute.

Shea,B.L.(1984) FACONE: A computer program for fitting the logit latent model by maximum likelihood. Department of Statistics, London Scholl of Econ

Publicações do Instituto de Matemática da UFRGS
Cadernos de Matemática e Estatística

Série A: Trabalho de Pesquisa

1. Marcos Sebastiani - Transformation des Singularités - MAR/89.
2. Jaime Bruck Ripoll - On a Theoremof R. Langevin About Curvature and Complex Singularities - MAR/89.
3. Eduardo Cisneros, Miguel Ferrero e Maria Inés Gonzales - Prime Ideals of Skew Polynomial Rings and Skew Laurent Polynomial Rings - ABR/89.
4. Oclide José Dotto - ϵ - Dilations - JUN/89.
5. Jaime Bruck Ripoll - A Characterization of Helicoids - JUN/89.
6. Mark Thompson, V. B. Moscatelli - Asymptotic Distribution of Liusternik-Schnirelman Eigenvalues for Elliptic Nonlinear Operators -JUL/89.
7. Mark Thompson - The Formula of Weyl for Regions with a Self- Similar Fractal Boundary - JUL/89.
8. Jaime Bruck Ripoll - A Note on Compact Surfaces with Non Zero Constant Mean Curvature - OUT/89.
9. Jaime Bruck Ripoll - Compact ϵ - Convex Hypersurfaces - NOV/89.
10. Jandyra Maria G. Fachel - Coeficientes de Correlação Tipo -Contingência - JAN/90.
11. Jandyra Maria G. Fachel - The Probability of Ocurrence of Heywood Cases - JAN/90.
12. Jandyra Maria G. Fachel - Heywood Cases in Unrestricted Factor Analysis - JAN/90.
13. Julio Cesar R. Claeyssen, Tereza Tsukazan - Dynamical Solutions of Linear Matrix Differential Equations - JUL/90.
14. Maria T. Albanese - Behaviour of the Likelihood Function in Latent Trait Analysis of Binary Data - ABR/91.
15. Maria T. Albanese - Measurement of the Latente Variable in Latent Trait Analysis of Binary Data - ABR/91.
16. Maria T. Albanese - Adequacy of the Asymptotic Variance-Covariance Matrix using Bootstrap and Jackknife Techniques in Latent Trait Analysis of Binary Data - ABR/91.
17. Maria T. Albanese - Latent Variable Models for Binary Response - ABR/91.

Universidade Federal do Rio Grande Sul
Reitor: Professor Tuiskon Dick

Instituto de Matemática
Diretor: Professor Aron Taitelbaum
Núcleo de Atividades Extra Curriculares
Coordenador: Professora Jandyra G. Fachel
Secretária: Rosaura Monteiro Pinheiro

Os Cadernos de Matemática e Estatística publicam as seguintes séries:
Série A: Trabalho de Pesquisa
Série B: Trabalho de Apoio Didático
Série C: Colóquio de Matemática SBM/UFRGS
Série D: Trabalho de Graduação
Série E: Dissertações de Mestrado
Série F: Trabalho de Divulgação
Série G: Textos para Discussão

Toda correspondência com solicitação de números publicados e demais informações deverá ser enviada para:

NAEC - Núcleo de Atividades Extra Curriculares
Instituto de Matemática - UFRGS
Av. Bento Gonçalves, 9500
91.500 - Agronomia - POA/RS
Telefone: 36.11.59 ou 36.17.85 Ramal: 252