Eva Tesaříková

Statistical analysis of variant II of probabilistic model of the school-achievement test with double-choice response

STATISTICAL ANALYSIS OF VARIANT II OF PROBABILISTIC MODEL OF THE SCHOOL-ACHIEVEMENT TEST WITH DOUBLE-CHOICE RESPONSE

Eva TESAŘÍKOVÁ

Abstract

The purpose of the present paper is to estimate the parameters ν and δ of variant II of the probabilistic model of school-achievement test with double choice response described in article [3], and to evaluate its properties. Numerical results of point estimation are summarized

1   Introduction

This paper investigates variant II of the probabilistic model of school-achievement tests with double-choice response [3] where the probabilistic structure of the test was defined including certain simplifying assumptions. The test consists of n independent questions of the same difficulty, $q > 3$ is the number of offered alternatives.

Setting M_0 denote the random variable expressing the number of questions of the test to which no correct response was given by the tested person, M_1 the random variable expressing the number of questions to which only one correct response was given, and M_2 the random variable expressing the number of questions to which both correct responses were given we can write

$$P(M_0 = m_0, M_1 = m_1, M_2 = m_2) =$$

$$= \frac{n!}{m_0!\, m_1!\, m_2!} \left(\nu \frac{(q-2)(q-3)}{q(q-1)} \right)^{m_0} \left(\delta \frac{q-2}{q-1} + \nu \frac{4(q-2)}{q(q-1)} \right)^{m_1} \quad (1)$$

$$\left(1 - \delta \frac{q-2}{q-1} - \nu \frac{q(q-1)-2}{q(q-1)} \right)^{m_2}$$

The distribution of the random vector $M = (M_0, M_1, M_2)$ is dependent on parameters ν and δ which represent the proportion of the tested topic with which the examined person is quite unfamiliar or only partially familiar, respectively.

The purpose of the present paper is to estimate the parameters ν, δ and in this way also the parameter $\tau = \delta/2 + \nu$ representing the whole proportion of tested topic with which the examinee is unfamiliar.

2 Point estimation of the parameters

The parameters ν and δ will be estimated at first by the moment method using the theoretical and empirical moments of the variables M_0, M_1, M_2 on the basis of a single application of the test. Comparing theoretical and empirical mean values and using the formulas (5),(8),(11) of [3] we get the following equations

$$E(M_0) = n\nu \frac{(q-2)(q-3)}{q(q-1)} = m_0$$

$$E(M_1) = n \left(\delta \frac{q-2}{q-1} + \nu \frac{4(q-2)}{q(q-1)} \right) = m_1$$

$$E(M_2) = n \left(1 - \delta \frac{q-2}{q-1} - \nu \frac{q(q-1)-2}{q(q-1)} \right) = m_2$$

We can use arbitrery two of these equations. The first equation provides the estimator of the parameter ν in the form

$$\nu = \frac{m_0}{n} \frac{q(q-1)}{(q-2)(q-3)} \quad (2)$$

After substitution in the second equation we obtain the following point estimator of the parameter δ

$$\hat{\delta} = \frac{m_1(q-1)}{n(q-2)} - \frac{m_0 4(q-1)}{n(q-2)(q-3)} \quad (3)$$

The same solution can be obtained also by the maximum likelihood method on the basis of a single application of the test ussing the marginal distributions of the variables M_0 and M_1 denoted as (4),(7) in [3].

This solution is convenient also to the system of likelihood equations

$$\frac{\partial \ln L(\nu, \delta, m)}{\partial \nu} = 0 \qquad \frac{\partial \ln L(\nu, \delta, m)}{\partial \delta} = 0 \qquad (4)$$

derived from the multinomial distribution (1) of the random vector

$$M = (M_0, M_1, M_2).$$

Expressing the logarithm of likelihood of the vector $m = (m_0, m_1, m_2)$.

$$\ln L(\nu, \delta, m) = \ln \frac{n!}{m_0! \, m_1! \, m_2!} + m_0 \ln \left(\nu \frac{(q-2)(q-3)}{q(q-1)} \right) +$$

$$+ m_1 \ln \left(\delta \frac{q-2}{q-1} + \nu \frac{4(q-2)}{q(q-1)} \right) + m_2 \ln \left(1 - \delta \frac{q-2}{q-1} - \nu \frac{q(q-1)-2}{q(q-1)} \right)$$

and partial derivatives

$$\frac{\partial \ln L(\nu, \delta, m)}{\partial \nu} = \frac{m_0}{\nu} + \frac{m_1 \frac{4(q-2)}{q(q-1)}}{\delta \frac{q-2}{q-1} + \nu \frac{4(q-2)}{q(q-1)}} - \frac{m_2 \frac{q(q-1)-2}{q(q-1)}}{1 - \delta \frac{q-2}{q-1} - \nu \frac{q(q-1)-2}{q(q-1)}}$$

$$\frac{\partial \ln L(\nu, \delta, m)}{\partial \delta} = \frac{m_1 \frac{(q-2)}{(q-1)}}{\delta \frac{q-2}{q-1} + \nu \frac{4(q-2)}{q(q-1)}} - \frac{m_2 \frac{(q-2)}{(q-1)}}{1 - \delta \frac{q-2}{q-1} - \nu \frac{q(q-1)-2}{q(q-1)}}$$

we get system (4) of likelihood equations in the form

$$\frac{m_0}{\nu} + \frac{4m_1}{q\delta + 4\nu} - \frac{m_2(q(q-1) - 2)}{q(q-1) - \delta q(q-2) - \nu(q(q-1) - 2)} = 0,$$

$$\frac{q m_1}{q\delta + 4\nu} - \frac{m_2(q-2)q}{q(q-1) - \delta q(q-2) - \nu(q(q-1) - 2)} = 0,$$

After multiplication of the first equation by $q(q-2)$ and the second equation by $q(q-1) - 2$ we can easily substract these equations to obtain the equivalent expressions

$$\frac{m_0 q(q-2)}{\nu} + \frac{4m_1 q(q-2)}{q\delta + 4\nu} = \frac{m_1 q(q(q-1) - 2)}{q\delta + 4\nu},$$

$$\frac{m_0(q-2)}{\nu} = \frac{m_1(q(q-1) - 2 - 4(q-2))}{q\delta + 4\nu},$$

$$\frac{m_0(q-2)}{\nu} = \frac{m_1(q^2 - 5q + 6)}{q\delta + 4\nu},$$

which lead to the equality

$$\delta = \frac{\nu m_1 (q-3)}{m_0 q} - \frac{4\nu}{q}$$

Substituting the parameter ν by the expression (2) we get the same expression of the maximum likelihood estimator of the parameter δ as the result (3) obtained by moment method.

However, the derived relations

$$\hat{\nu} = \frac{m_0}{n} \frac{q(q-1)}{(q-2)(q-3)} \tag{5}$$

$$\hat{\delta} = \frac{m_1(q-1)}{n(q-2)} - \frac{m_0 4(q-1)}{n(q-2)(q-3)} \tag{6}$$

are convenient solutions only on the conditions that

$$0 \le \hat{\nu} \le 1, \quad 0 \le \hat{\delta} \le 1, \quad 0 \le \hat{\nu} + \hat{\delta} \le 1 \tag{7}$$

If the expressions (5), (6) do not satisfy the conditions (7), they can not be used. In this case we have only to use the numerical solution giving a pair of values of the parameters ν and δ which satisfy the conditions (7) and at the same time give the maximum value of likelihood function (1).

The point estimator $\hat{\tau}$ of the parameter τ, representing the whole proportion of the tested topic with which the examinee is unfamilier, is then obtained by the relation
$$\hat{\tau} = \hat{\delta}/2 + \hat{\nu}.$$

3   Numerical results

The values of the point estimator of τ for different q, n and x are summarized in tables on pp. 169–171. They were calculated with precision 0.02 only.

The computer program (written in GWBASIC) used to produce these tables of results is given on pp. 171–174.

Point estimator of parameter τ in the school-achievement test with
double choice response : **VARIANT II** (on pp. 169–171)

m_0	m_1	$q = 4$	$q = 5$	$q = 6$	$q = 7$	$q = 8$	$q = 9$
0	0	0.00	0.00	0.00	0.00	0.00	0.00
	1	0.08	0.07	0.06	0.06	0.06	0.06
	2	0.15	0.13	0.13	0.12	0.12	0.11
	3	0.23	0.20	0.19	0.18	0.18	0.17
	4	0.30	0.27	0.25	0.24	0.23	0.23
	5	0.38	0.33	0.31	0.30	0.29	0.29
	6	0.45	0.40	0.38	0.36	0.35	0.34
	7	0.50	0.47	0.44	0.42	0.41	0.40
	8	0.50	0.50	0.50	0.48	0.47	0.46
	9	0.50	0.50	0.50	0.50	0.50	0.50
	10	0.50	0.50	0.50	0.50	0.50	0.50

m_0	m_1	$q = 4$	$q = 5$	$q = 6$	$q = 7$	$q = 8$	$q = 9$
1	0	0.12	0.12	0.10	0.10	0.10	0.10
	1	0.24	0.22	0.22	0.21	0.20	0.19
	2	0.36	0.33	0.29	0.27	0.26	0.25
	3	0.48	0.40	0.35	0.33	0.32	0.30
	4	0.60	0.47	0.42	0.39	0.37	0.36
	5	0.68	0.53	0.48	0.45	0.43	0.42
	6	0.75	0.60	0.54	0.51	0.49	0.48
	7	0.83	0.67	0.60	0.57	0.55	0.53
	8	0.98	0.70	0.63	0.61	0.59	0.57
	9	1.00	0.80	0.65	0.62	0.60	0.59

m_0	m_1	$q = 4$	$q = 5$	$q = 6$	$q = 7$	$q = 8$	$q = 9$
2	0	0.24	0.22	0.22	0.21	0.20	0.20
	1	0.36	0.34	0.33	0.32	0.31	0.30
	2	0.48	0.45	0.43	0.42	0.40	0.38
	3	0.60	0.56	0.52	0.48	0.46	0.44
	4	0.72	0.67	0.58	0.54	0.51	0.50
	5	0.84	0.73	0.65	0.60	0.57	0.55
	6	0.96	0.80	0.71	0.66	0.63	0.61
	7	1.00	0.87	0.80	0.73	0.69	0.67
	8	1.00	0.97	0.90	0.78	0.70	0.68

m_0	m_1	$q = 4$	$q = 5$	$q = 6$	$q = 7$	$q = 8$	$q = 9$
3	0	0.36	0.34	0.33	0.32	0.31	0.30
	1	0.48	0.44	0.43	0.42	0.41	0.40
	2	0.60	0.56	0.54	0.52	0.51	0.50
	3	0.72	0.66	0.64	0.63	0.60	0.57
	4	0.84	0.78	0.75	0.69	0.65	0.63
	5	0.96	0.88	0.81	0.75	0.71	0.69
	6	1.00	0.98	0.88	0.81	0.77	0.74
	7	1.00	1.00	0.95	0.85	0.82	0.80

m_0	m_1	$q = 4$	$q = 5$	$q = 6$	$q = 7$	$q = 8$	$q = 9$
4	0	0.48	0.45	0.43	0.42	0.41	0.40
	1	0.60	0.56	0.54	0.52	0.51	0.50
	2	0.72	0.66	0.64	0.63	0.62	0.61
	3	0.84	0.78	0.74	0.73	0.72	0.70
	4	0.96	0.88	0.86	0.84	0.79	0.76
	5	1.00	1.00	0.96	0.90	0.85	0.82
	6	1.00	1.00	1.00	0.97	0.90	0.86

m_0	m_1	$q = 4$	$q = 5$	$q = 6$	$q = 7$	$q = 8$	$q = 9$
5	0	0.60	0.56	0.54	0.52	0.51	0.50
	1	0.72	0.66	0.64	0.63	0.62	0.61
	2	0.84	0.78	0.75	0.74	0.73	0.72
	3	0.96	0.88	0.86	0.84	0.83	0.82
	4	1.00	1.00	0.96	0.94	0.92	0.90
	5	1.00	1.00	1.00	1.00	0.98	0.95

m_0	m_1	$q = 4$	$q = 5$	$q = 6$	$q = 7$	$q = 8$	$q = 9$
6	0	0.72	0.66	0.64	0.63	0.62	0.61
	1	0.84	0.78	0.75	0.74	0.73	0.72
	2	0.96	0.88	0.86	0.84	0.83	0.82
	3	1.00	1.00	0.97	0.95	0.94	0.93
	4	1.00	1.00	1.00	1.00	1.00	1.00

m_0	m_1	$q = 4$	$q = 5$	$q = 6$	$q = 7$	$q = 8$	$q = 9$
7	0	0.84	0.78	0.75	0.74	0.73	0.72
	1	0.96	0.88	0.86	0.84	0.83	0.82
	2	1.00	1.00	0.97	0.95	0.94	0.93
	3	1.00	1.00	1.00	1.00	1.00	1.00

m_0	m_1	$q = 4$	$q = 5$	$q = 6$	$q = 7$	$q = 8$	$q = 9$
8	0	0.98	0.90	0.86	0.84	0.83	0.82
	1	1.00	1.00	0.96	0.95	0.94	0.92
	2	1.00	1.00	1.00	1.00	1.00	1.00

m_0	m_1	$q = 4$	$q = 5$	$q = 6$	$q = 7$	$q = 8$	$q = 9$
9	0	1.00	1.00	0.98	0.96	0.94	0.92
	1	1.00	1.00	1.00	1.00	1.00	1.00

m_0	m_1	$q = 4$	$q = 5$	$q = 6$	$q = 7$	$q = 8$	$q = 9$
10	1	1.00	1.00	1.00	1.00	1.00	1.00

$n = 10$ number of questions of the test

m_0 number of questions to which no correct response was given

m_1 number of questions to which only one correct response was given

m_2 number of question to which both correct responses were given
$$m_2 = n - (m_0 + m_1)$$

q number of offered alternatives

τ proportion of the topic with which examinee is unfamiliar

4 Computer program

```
100 REM -COMPUTATION PROGRAM FOR ESTIMATION OF PARAME-
TER τ IN SCHOOL-ACHIEVEMENT TEST WITH DOUBLE-CHOICE RE-
SPONSE VARIANT II
101 REM Tesarikova Eva - 1990
103 CLS:LOCATE 4,15:PRINT"Calculation of parameter τ in school
achievement "
104 LOCATE 6,24:PRINT"test with double-choice response - varia
nt II"
105 LOCATE 10,10:PRINT"1 Calculation per partes "
106 LOCATE 12,10:PRINT"2 Automatic calculation"
107 LOCATE 14,10:PRINT"3 End"
108 LOCATE 20,10:PRINT"Order a number of questions and number
of method"
109 PRINT
110 INPUT "number of method = ";V
```

```
115 ON V GOTO 300,117,500
117 INPUT "number of questions = ";N
120 CLS:LOCATE 10,15 :PRINT "Point estimator is calculated"
132 LPRINT :LPRINT" Point estimator of parameter τ in
the school-achievement test "
133 LPRINT " with double choice response - variant II"
134 lPRINT:LPRINT" x ";
135 FOR Q=4 TO 9
137 P0(Q)=(Q*(Q-1))/((Q-2)*(Q-3))
138 P1(Q)=(Q-1)/(Q-2)
139 P3(Q)=(4*(Q-1)/((Q-2)*(Q-3)))
140 NEXT Q
141 FOR M=0 TO N
142 LPRINT" m0      m1      q=4      q=5      q=6      q=7
q=8      q=9"
145 LPRINT
151 LPRINT USING "#########";M0;
155 FOR M1=0 TO N-M0
160 IF M1=0 THEN 168
165 LPRINT USING "\          \";"";
168 LPRINT USING"#######";M1;
169 FOR Q=4 TO 9
170 NI=M0*P0(Q)/N
172 DEL=(M1*P1(Q)/N)-(M0*P3(Q)/N)
180 IF DEL+NI>1 OR NI<0 OR DEL<0 THEN GOSUB 600
185 TAU(Q)=DEL/2+NI
187 LPRINT USING "####### .##";TAU(Q); 
188 NEXT Q
194 NEXT M1
195 LPRINT:LPRINT:LPRINT
197 NEXT MO
199 LPRINT" N=";N;"..number of questions in the test"
203 LPRINT" m0.......number of questions to which no corre
ct response was given"
204 LPRINT" m1.......number of questions to which only one
correct response was given"
205 LPRINT" m2......number of questions to which both cor
rect responses were given...m2=n-(m0+m1)"
206 LPRINT" Q........number of offered alternatives"
209 LPRINT" τ........proportion of the topic with which
examinee is unfamiliar
210 LPRINT:LPRINT
217 NEXT Q
220 GOTO 103
```

```
300 CLS:LOCATE 8,25:PRINT"Calculation of the point estimator
of τ "
304 INPUT"Order a number of questions";N
307 INPUT"Order the number of questions to which no correct
answer was given m0 =";M0
320 CLS:LOCATE 10,15 :PRINT "Point estimator is calculated"
332 LPRINT :LPRINT" Point estimator of parameter τ in
the school-achievement test "
333 LPRINT " with double choice response - variant II"
334 LRINT:LPRINT
335 FOR Q=4 TO 9
337 P0(Q)=(Q*(Q-1))/((Q-2)*(Q-3))
338 P1(Q)=(Q-1)/(Q-2)
339 P3(Q)=(4*(Q-1)/((Q-2)*(Q-3))
340 NEXT Q
342 LPRINT" m0      m1      q=4      q=5      q=6      q=7
q=8      q=9"
345 LPRINT
351 LPRINT USING "#########";M0;
355 FOR M1=0 TO N-M0
360 IF M1=0 THEN 368
365 LPRINT USING "\          \",""";
368 LPRINT USING"#######";M1;
369 FOR Q=4 TO 9
370 NI=M0*P0(Q)/N
372 DEL=(M1*P1(Q)/N)-(M0*P3(Q)/N)
380 IF DEL+NI>1 OR NI<0 OR DEL<0 THEN GOSUB 600
385 TAU(Q)=DEL/2+NI
387 LPRINT USING "####### .11";TAU(Q);
388 NEXT Q
394 NEXT M1
395 LPRINT:LPRINT:LPRINT
398 GOTO 415
399 LPRINT" N=";N;"..number of questions in the test"
403 LPRINT" m0....... number of questions to which no corre
ct response was given"
404 LPRINT" m1.......number of questions to which only one
correct response was given"
405 LPRINT" m2.......number of questions to which both cor
rect responses were given...m2=n-(m0+m1)"
406 LPRINT" Q........number of offered alternatives"
409 LPRINT" τ........proportion of the topic with which
examinee is unfamiliar
410 LPRINT:LPRINT:GOTO 103
```

```
415 LOCATE 22,1:PRINT"Do you want to continue for next m0?
(y/n)"
417 INPUT A$ :IF A$ ="Y" OR A$ ="y" THEN 421
420 GOTO 430
421 M0=MO+1
422 GOTO 342
430 LPRINT " n=";N;... number of questions in the test"
435 LPRINT: GOTO 103
500 END
600 MAX=0
615 FOR NI=0 TO 1 STEP 0.01
620 FOR DEL=0 TO 1-NI STEP 0.01
625 X=N:GOSUB 1000
630 KN=S
633 IF MO=0 THEN K0=1:GOTO 643
635 X=M0:GOSUB 1000
640 K0=S
643 IF M1=0 THEN K1=1:GOTO 653
645 X=M1:GOSUB 1000
650 K1=S
653 IF N-M0-M1=0 THEN K2=1:GOTO 665
655 X=N-MO-M1:GOSUB 1000
660 K2=S
665 CF=KN/(K0*K1*K2)
670 C0=(NI*(Q-2)*(Q-3)/(Q*(Q-1))^M0
675 C1=(DEL*(Q-2)/(Q-1)+NI*4*(Q-2)/(Q*(Q-1)))^M1
680 C2=(1-DEL*(Q-2)/(Q-1)-NI*(Q*(Q-1)-2)/(Q*(Q-1)))^(N-M0-M1)
685 PF=CF*C0*C1*C2
690 IF PF>MAX THEN MAX=PF:NIM=NI:DELM=DEL
695 NEXT DEL
700 NEXT NI
705 NI=NIM:DEL=DELM
710 RETURN
1000 S=X
1005 FOR C=X-1 TO 1 STEP -1
1010 S=S*C
1015 NEXT C
1020 RETURN
```

References

[1] Tesaříková, E.: *Probabilistic model of school-achievement test with double-choice response: VARIANT I.* Acta UPO, Fac. Rer. Nat., 105(1992), 119–126

[2] Tesaříková, E.: *Statistical analysis of probabilistic model of school-achievement test with double choice response — VARIANT I.* Acta UPO, Fac. Rer. Nat., 105(1992), 127–146

[3] Tesaříková, E.: *Variant II of probabilistic model of school-achievement test with double chioce response.* Acta UPO., Fac. Rer. Nat., 110(1993), 159–164

[4] Tesaříková, E.: *Probabilistic models of school - achievement test with double choice response and their statistical analysis.* The Statistician, London (in press)

Author's address: Department of Algebra and Geometry
Faculty of Science
Palacký University
Tomkova 38, Hejčín
779 00 Olomouc
Czech Republic