Universidade Federal do Rio Grande do Sul
Instituto de Matemática

The Probability of Ocurrence of Heywood Cases

Jandyra M.G. Fachel

Cadernos de Matemática e Estatística

Série A, nº 11, JAN/90
Porto Alegre, janeiro de 1990

The probability of occurrence of Heywood cases

A simulation study with the purpose of assessing the main causes of Heywood cases was performed (see Fachel, 1990). The results show that Heywood cases occurr even when a factor model with large residual variances generated the data. In this paper we consider the case of one-factor models and we show how the probability of occurrence may be calculated. We also show that this probability for the one-factor model depends on sample size, parameter vectors and number of variables in the model. We shall consider the MINRES method of fitting for factor analysis, that is, we shall suppose that the factor loadings are estimated under the condition that the sum of squares of the off-diagonal residuals is minimized.

The MINRES method was introduced by Harman and Jones (1966) although, as they point out, the idea of getting a factor solution by minimizing off-residual correlations was first posed by Thurstone (1954). They do not consider the minimization of the total residual matrix (including diagonal terms), which would lead to the principal-factor solution. We do not consider the numerical procedures for obtaining the MINRES solution. It is on the priciple of the method that we shall base our method for obtaining the probabilties of occurrence of Heywood cases.

Consider the one-factor model given by

$$x_i = \lambda_i f + e_i \qquad i=1,2,\ldots,p \tag{1.1}$$

with $E(f)=E(e_i)=0$; $E(fe_i)=0$. Suppose $Var(f)=1$, $E(e_i e_j)=0$, $i\neq j$ and $Var(e_i)=\psi_i \geq 0$, $i=1,\ldots,p$. Suppose the x_i's are standardized so that $var(x_i) = 1$. Hence,

$$\lambda_i^2 + \psi_i = 1, \qquad i = 1,2,\ldots,p \tag{1.2}$$

Then, we have

$$\lambda_i^2 \leq 1, \qquad i = 1,2,\ldots,p \tag{1.3}$$

For this model

$$corr(x_i,x_j) = \rho_{ij} = \lambda_i\lambda_j \,. \tag{1.4}$$

Suppose, therefore, we fit the model by minimizing

$$S = \sum_{i=1}^{p} \sum_{j=i+1}^{p} (r_{ij} - \lambda_i\lambda_j)^2 \tag{1.5}$$

where r_{ij} is the observed correlation between variables x_i and x_j. (Actually, as x_i and x_j are satandardized variables, r_{ij} and ρ_{ij} may be considered as covariances).

From a purely mathematical point of view, the problem of finding a minimum for the nonlinear function S is well defined. S is a function of the p(p-1)/2 off-diagonal residual correlations, which are dependent upon the elements of the factor matrix Λ(px1). The minimum value for S occurs at the point where its partial derivatives with respect to the p elements of Λ are zero and its matrix of second derivatives is positive definite. Thus,

$$S'(\lambda_i) = \frac{1}{2} \frac{\partial S}{\partial \lambda_i} = \sum_{\substack{j=1 \\ j\neq i}}^{p} \lambda_i r_{ij} + \lambda_i \sum_{\substack{j=1 \\ j\neq i}}^{p} \lambda_j^2 \qquad i=1,2,\ldots,p \tag{1.6}$$

we also have,

$$\frac{1}{2} \frac{\partial^2 S}{\partial \lambda_i^2} = \sum_{\substack{j=1 \\ j\neq i}}^{p} \lambda_j \geq 0$$

$$\frac{1}{2} \frac{\partial^2 S}{\partial \lambda_i \partial \lambda_i} = -r_{ij} + 2\lambda_i\lambda_j$$

If the system of equations

$$\frac{\partial S}{\partial \lambda_i} = 0 \qquad i=1,2\ldots,p$$

has a solution with $h_i^2 \leq 1$, then it is clearly a minimum by (1.7). However, there may not be such a solution. For fixed $\lambda_1,\ldots\lambda_{i-1},\lambda_{i+1},\ldots\lambda_p$, $S'(\lambda_i)$ is a linear increasing function of λ_i. We are interested in the behaviour of the function at the point $\lambda_i = 1$ and consequently $\psi_i = 0$. If the minimum of the function occurs at $\lambda_i = 1$, an exact Heywood case occurs. Let us consider the following cases:

1) Suppose that the derivative of S with respect to λ_i at the point $\lambda_i = 1$ is negative, that is $S'(1) < 0$.
At $\lambda_i = 1$, from (1.6) we have

$$\left[\frac{\partial S}{\partial \lambda_i} \right]_{\lambda_i = 1} = - \sum_{\substack{j=1 \\ j \neq 1}}^{p} \lambda_j r_{ij} + \sum_{\substack{j=1 \\ j \neq 1}}^{p} \lambda_j^2$$

If $S'(1)<0$, this would imply that $S'(0)<0$ and that there will be no intermediate value of λ_i for which it is zero. The minimum of S will thus occur at $\lambda_i=1$, because S' is a linear increasing function of λ_i. A proof of this is given bellow:
Suppose $S'(1)<0$ then

$$- \sum_{\substack{j=1 \\ j \neq i}}^{p} \lambda_j r_{ij} + \sum_{\substack{j=1 \\ j \neq i}}^{p} \lambda_j^2 < 0$$

$$\sum_{j \neq i}^{p} \lambda_j r_{ij} > \sum_{j \neq i}^{p} \lambda_j^2 > 0$$

Therefore

$$S'(0) = \left[\frac{\partial S}{\partial \lambda_i} \right]_{\lambda_i = 0} = -\sum \lambda_j r_{ij} < 0$$

We recall that we are assuming $0 \leq \lambda_i \leq 1$. It is known that for the one-factor model, Λ reduces to a column vector of p elements that is unique apart from a possible change of sign of all its elements, which corresponds merely to changing the sign of the factor. According to Lawley and Maxwell (1971), such changes are merely trivial.

2) Suppose now that $S'(1)$ is positive. Therefore $S'(0)$ may be negative or positive and the minimum of S will not occur at the point $\lambda_i = 1$.

3) Finally, if $S'(1)=0$, the minimum of S will occur at $\lambda_i=1$ and a Heywood case occur.

We are interested in evaluating the probability of an occurrence of a Heywood case for variable x_i. Therefore we need consider only cases 1) and 3) above. That is

$$\Pr\left\{S'(1) \le 0\right\} = \Pr\left\{\sum_{\substack{j=1\\j\neq i}}^{p} \lambda_j r_{ij} \ge \sum_{\substack{j=1\\j\neq i}}^{p} \lambda_j^2\right\} \qquad i=1,2,\ldots,p \qquad (1.8)$$

For evaluating this probability we need to know the distribution of $\sum\limits_{j\neq i}^{p} \lambda_j r_{ij}$. An approximation may be obtained as follows.

Since the x's have zero means and unit variances

$$r_{ij} = \frac{1}{n} \qquad \sum_{h=1}^{n} x_{ih} x_{jh}$$

where h indexes the sample members. (The terms of this sum are independent for h=1,2,....,n; for sufficiently large n, this will be approximately normal by the central limit theorem).

Consider

$$z_i = \sum_{\substack{j=1\\j\neq i}}^{p} \lambda_j r_{ij} = \frac{1}{n} \sum_{h=1}^{n} x_{ih} \sum_{j\neq i} x_{jh} \lambda_j$$

Let us denote $\sum\limits_{j\neq i} x_{jh} \lambda_j$ by y_{ih} say. Therefore

$$z_i = \frac{1}{n} \sum_{h=1}^{n} x_{ih} y_{ih}.$$

We shall now find the moments of x_{ih} and y_{ih}. By definition we have

$$E(x_{ih}) = 0, \; \text{var}\,(x_{ih}) = 1$$

Thus

$$E(y_{ih}) = 0$$

$$\text{var}(y_{ih}) = \text{var}\,\left(\sum_{j\neq i} x_{jh} \lambda_j\right) = \sum_{\substack{j=1\\j\neq i}}^{p} \lambda_j^2 + \sum_{\substack{j=1\\j\neq i\\j\neq k}}^{p} \sum_{\substack{k=1\\k\neq i}}^{p} \lambda_j \lambda_k \,\text{cov}\,(x_{jh}, x_{kh}) \qquad (1.9)$$

Now

$$\mathrm{cov}(x_{ih}, x_{kh}) = \lambda_j \lambda_k \tag{1.10}$$

thus

$$\mathrm{var}(y_{ih}) = \sum_{\substack{j=1 \\ j \neq i}}^{p} \lambda_j^2 + \sum_{\substack{j=1 \\ j \neq i}}^{p} \sum_{\substack{k=1 \\ k \neq i \\ j \neq k}}^{p} \lambda_j^2 \lambda_k^2 \tag{1.11}$$

$$\mathrm{cov}(x_{ih}, y_{ih}) = \sum_{\substack{j=1 \\ j \neq i}}^{p} \lambda_j \, \mathrm{cov}(x_{ih}, x_{jh}) = \lambda_i \sum_{\substack{j=1 \\ j \neq i}}^{p} \lambda_j^2 = E(x_{ih}, y_{ih}) \tag{1.12}$$

Let us suppose that the observed variables x_i, $i=1,\dots,p$ have a sum of p-1 dependent normal variables, is also normal with mean and variance given by (1.9). Hence x_{ih} and y_{ih} have a normal bivariate distribution. From Kendall and Stuart (1977, Vol. I, p.85) it is known that if (x_{ih}, y_{ih}) are normal bivariate we have, using the fact that $\sigma^2_{x_{ih}} = 1$

$$E(x_{ih}^2 . y_{ih}^2) = (1 + 2\,\rho^2_{x_{ih}\, y_{ih}})\, \sigma^2_{x_{ih}}\, \sigma^2_{y_{ih}} = \sigma^2_{y_{ih}} + 2\left[E(x_{ih}, y_{ih})\right]^2$$

Thus

$$\mathrm{Var}(x_{ih} y_{ih}) = E(x_{ih}^2 y_{ih}^2) - \left[E(x_{ih} y_{ih})\right]^2 = \sigma^2_{y_{ih}} + 2\left[E(x_{ih} y_{ih})\right]^2 -$$

$$-\left[E(x_{ih} y_{ih})\right]^2 = \sigma^2_{y_{ih}} + \left[E(x_{ih} y_{ih})\right]^2 \tag{1.13}$$

Now

$$z_i = \frac{1}{n} \sum_{h=1}^{n} x_{ih} y_{ih}$$

where $z_{ih} = x_{ih}\ y_{ih}$ are independent for $h=1,\dots,n$. Therefore, using (1.12)

$$E(z_i) = \frac{n}{n} E(x_{ih}\ y_{ih}) = \lambda_i \sum_{\substack{j=1 \\ j \neq i}}^{p} \lambda_j^2 \tag{1.14}$$

and using (1.11) and (1.13) we have

$$\text{Var}\,(z_i) = \frac{1}{n}\left[\ \sum_{\substack{j=1\\ j\neq i}}^{p} \lambda_j^2 + \sum_{\substack{j=1\\ j\neq i}}^{p}\ \sum_{\substack{k=1\\ k\neq i\\ j\neq k}}^{p}\lambda_j^2\lambda_k^2 + \lambda_i^2\,(\sum_{\substack{j=1\\ j\neq i}}^{p}\lambda_j^2)^2\right] \qquad (1.15)$$

For n sufficently large

$$z_i = \frac{1}{n}\sum_{h=1}^{n} x_{ih}y_{ih} = \sum_{\substack{j=1\\ j\neq i}}^{p}\lambda_j r_{ij}$$

is approximately normal, by the central limit theorem, with mean and variance given by (1.14) and (1.15) respectively. Therefore (1.8) may be calculated approximately as

$$1 - \phi\left[\frac{\sum_{j\neq i}\lambda_j^2 - \lambda_i\sum_{j\neq i}\lambda_j^2}{\frac{1}{\sqrt{n}}\left[\sum_{j\neq i}\lambda_j^2 + \lambda_i^2\left[\sum_{j\neq i}\lambda_j^2\right] + \sum_{j\neq i}\sum_{\substack{k\neq i\\ j\neq k}}\lambda_j^2\,\lambda_k^2\right]^{1/2}}\right]$$

$i=1,2,\ldots,p$ (1.16)

where $\phi(z)$ is the normal distribution function.

The above expression gives the asymptotic probability of occurrence of a Heywood case for variable x_i in function of the sample size, magnitude of the factor loading parameters and the number p of variables in the model. As can be seen, if $n \Rightarrow \infty$, the expression (1.16) converges to zero. If $\lambda_i \Rightarrow 1$, the probability of Heywood case converges to 0.5 for any values of $\lambda_j\,(j\neq i)$.

We now present tables of the probability of occurrence of a Heywood case for several values of sample size and different values of p, considering

1) $\lambda_i = 0.98$ and $\lambda_j = 0.5$ $(j\neq i)$

2) $\lambda_i = 0.90$ and $\lambda_j = 0.5$ $(j\neq i)$ $\quad i,j=1,2,\ldots p$

3) $\lambda_i = 0.80$ and $\lambda_j = 0.5$ $(j\neq i)$

4) $\lambda_i = \lambda_j = 0.50$ $\quad i,j=1,2,\ldots p$

5) $\lambda_i = \lambda_j = 0.90$ $i, j=1,2,\ldots p$

Table 1.10 - Asymptotic probability of occurrence of Heywood cases for various sample sizes (n), different number of variables (p) and different magnitude of the factor loading parameters (cases 1 to 5)

1) $\lambda_i = 0.98$ $\lambda_j = 0.5$ $(j \neq i)$ $i, j = 1, \ldots p$

p \ n	50	100	200	400	500	1000	2000	5000
5	.4658	.4516	.4318	.4040	.3929	.3504	.2935	.1952
10	.4628	.4474	.4258	.3958	.3838	.3380	.2772	.1750
15	.4618	.4461	.4240	.3932	.3809	.3341	.2722	.1689
20	.4613	.4453	.4230	.3917	.3793	.3320	.2695	.1656
30	.4608	.4447	.4220	.3904	.3779	.3300	.2670	.1627
40	.4605	.4443	.4215	.3897	.3770	.3289	.2655	.1610
50	.4604	.4441	.4211	.3892	.3766	.3283	.2647	.1600
100	.4601	.4438	.4207	.3887	.3759	.3274	.2635	.1586

2) $\lambda_i = 0.90$ $\lambda_j = 0.5$ $(j \neq i)$ $i, j=1, \ldots p$

p \ n	50	100	200	400	500	1000	2000	5000
5	.3293	.2660	.1884	.1057	.0812	.0241	.0026	.0000
10	.3146	.2473	.1671	.0860	.0634	.0154	.0012	.0000
15	.3097	.2411	.1602	.0800	.0580	.0131	.0009	.0000
20	.3070	.2380	.1567	.0770	.0555	.0121	.0008	.0000
30	.3047	.2349	.1534	.0740	.0531	.0112	.0007	.0000
40	.3034	.2334	.1517	.0727	.0518	.0107	.0006	.0000
50	.3026	.2324	.1508	.0719	.0511	.0104	.0005	.0000
100	.3012	.2306	.1487	.0703	.0497	.0099	.0005	.0000

3) $\lambda_i = 0.80$ $\lambda_j = 0.5$ $(j \neq i)$ $i, j = 1, 2, \ldots p$

p \ n	50	100	200	400	500	1000	2000	5000
5	.1802	.0980	.0339	.0048	.0018	.0000	.0000	.0000
10	.1572	.0773	.0220	.0022	.0007	.0000	.0000	.0000
15	.1495	.0710	.0189	.0016	.0005	.0000	.0000	.0000
20	.1457	.0679	.0174	.0014	.0004	.0000	.0000	.0000
30	.1421	.0650	.0161	.0013	.0003	.0000	.0000	.0000
40	.1405	.0635	.0155	.0011	.0003	.0000	.0000	.0000
50	.1392	.0626	.0151	.0011	.0003	.0000	.0000	.0000
100	.1370	.0609	.0143	.0010	.0003	.0000	.0000	.0000

4) $\lambda_i = \lambda_j = 0.50$ $i, j = 1, 2, \ldots p$

p \ n	50	100	200	400	500	1000	2000	5000
5	.0062	.0002	.0000	.0000	.0000	.0000	.0000	.0000
10	.0025	.0000	.0000	.0000	.0000	.0000	.0000	.0000
15	.0017	.0000	.0000	.0000	.0000	.0000	.0000	.0000
20	.0015	.0000	.0000	.0000	.0000	.0000	.0000	.0000
30	.0012	.0000	.0000	.0000	.0000	.0000	.0000	.0000
40	.0011	.0000	.0000	.0000	.0000	.0000	.0000	.0000
50	.0010	.0000	.0000	.0000	.0000	.0000	.0000	.0000
100	.0009	.0000	.0000	.0000	.0000	.0000	.0000	.0000

5) $\lambda_i = \lambda_j = 0.90 \qquad i, j = 1, 2, \ldots, p$

p \ n	50	100	200	400	500	1000	2000	5000
5	.3025	.2322	.1507	.0717	.0510	.0103	.0005	.0000
10	.3009	.2303	.1483	.0700	.0494	.0009	.0005	.0000
15	.3004	.2297	.1477	.0694	.0490	.0097	.0005	.0000
20	.3002	.2294	.1474	.0692	.0488	.0096	.0005	.0000
30	.3000	.2292	.1471	.0690	.0486	.0095	.0005	.0000
40	.2999	.2290	.1457	.0689	.0485	.0095	.0005	.0000
50	.2998	.2290	.1457	.0688	.0484	.0094	.0005	.0000
100	.2997	.2288	.1457	.0687	.0484	.0094	.0005	.0000

Final comments

The tables of probability of Heywood case are obtained from considerations about the principle of the minres method, that is, minimizing the sum of squares of the off-diagonal residual correlations. We also believe that some of the constraints in the algorithm for the MLFA method, of the main statistical packages, as for example stopping the iteration for any particular variable which $\psi_i=0$, could lead to a different proportion of Heywood cases in some situations, because of a possible bias in the method.

For large sample sizes, the probability of Heywood cases is very small. We observe in Table 1.10 that the probability of Heywood cases decreases as 1) the sample size increases; 2) the number of variables in the model increases but a small variation is observed from p=5 to p=100 for all tables. Related to the magnitude of the factor loading parameters, we observe that when only one factor loading increases (the others being equal to 0.5), the probability of Heywood cases also increases. Finally, it is interesting to note, comparing cases 2) and 5) in Table 1.10, that when only one loading is equal to 0.90, the probabilities of Heywood cases are slightly higher than when all loadings are equal to 0.90.

The approximate probability of the occurrence of Heywood cases for different parameter vectors for the one-factor model may be easily evaluated using expression (1.16). In Table 1.10 we have included only some cases.

REFERENCES

Anderson, J.C. and Gerbing, D.W. (1984). The effect of sampling error on convergence, improper solutions, and goodness of fit indices for maximum likelihood confirmatory factor analysis. Psychometrika, 49, 155-173.

Boomsma, A. (1985). Nonconvergence, improper solutions and starting values in LISREL maximum likelihood estimation. Psychometrika, 50, 229-242.

Dixon, W.J. and Brown, M.B. (eds) (1983). BMDP Statistical Software (1983 Printing with additions). Berkeley: University of California Press.

Francis, I. (1973) . Factor analysis: its purpose, practice and packaged programs. Invited paper, American Statistical Association. NY, December 1973.

Francis I. (1974). Factor analysis: fact or fabrication. Mathematical Chronicle, 3, 9-44.

Gweke, J.F. and Singleton, K.J. (1980). Interpreting the likelihood ratio statistic in factor models when sample size is small. Journal of the American Statistical Association, 75, 133-137.

Harman, H.H. and Fukuda, Y. (1966). Resolution of the Heywood case in the MINRES solution. Psychometrika, 31, 563-571.

Harman H.H. and Jones, W.H. (1966). Factor analysis by minimizing residuals (MINRES). Psychometrika, 31, 351-368.

Heywood, H.B. (1931). On finite sequences of real numbers. Proceedings of the Royal Society of London, 134, 486-501.

Joreskog, K.G. (1967). Some contributions to maximum likelihood factor analysis. Psychometrika, 32, 443-482.

Martin, J.V. and McDonald, R.P. (1975). Bayesian estimation in unrestricted factor analysis: A treatment for Heywood cases. Psychometrika, 40, 505-517.

SAS Institute (1985). SAS User's guide: Statistics, Version 5 edition. Cary, NC: SAS Institute Inc.

Tumura, Y. and Fukutomi, K. (1970). On the improper solutions in factor analysis. TRU Mathematics, 6, 63-71.

Tumura, Y.; Fukutomi, K. and Asoo, Y. (1968). On the unique convergence of iterative procedures in factor analysis. TRU Mathematics, 4, 52-59.

Van Driel, O.P. (1978). On various causes of improper solutions in maximum likelihood factor analysis. Psychometrika, 43, 225-243.

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.

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