Economic Review (Otaru University of Commerce),
Vo.59, No.4, 41-48, March, 2009

Two-term Edgeworth expansions of the distributions of fit indexes under fixed alternatives in covariance structure models*

Haruhiko Ogasawara

Abstract. Asymptotic expansions of the distributions of thirteen fit indexes used in covariance structure analysis in practice are obtained. The fit indexes include the usual log likelihood ratio statistic for a posited model and the functions of this statistic and the corresponding statistic of the so-called baseline model of uncorrelated observed variables. The results are derived by the two-term Edgeworth expansion under fixed alternatives for possibly nonnormally distributed data. A numerical example using a misspecified factor analysis model is shown to see the behavior of the asymptotic results in finite samples.

Keywords: Fixed alternatives, nonnormal distributions, Edgeworth expansion, structural equation modeling, RMSEA

1 Purpose

In covariance structure analysis, various indexes are used to assess the goodness-of-fit of a posited model. Among these fit indexes, the classic log likelihood ratio chi-square statistic is a basic one with the assumption of multivariate normality for observed variables under the null hypothesis of a true model. Let F be the discrepancy function for Wishart maximum likelihood estimation, i.e. $F = F(\mathbf{S}, \mathbf{\Sigma}) = \ln|\mathbf{\Sigma}| - \ln|\mathbf{S}| + \mathrm{tr}(\mathbf{\Sigma}^{-1}\mathbf{S}) - p$, where $\mathbf{\Sigma} = \mathbf{\Sigma}(\boldsymbol{\theta})$ is the covariance matrix for p observed variables described by a $q \times 1$ parameter vector $\boldsymbol{\theta}$ and $\mathbf{S}$ is the $p \times p$ unbiased sample covariance matrix. Then, the vector of the Wishart maximum likelihood

* Partially supported by Grant-in-Aid for Scientific Research from the Japanese Ministry of Education, Culture, Sports, Science and Technology No.2500249.
Author’s address: Department of Information and Management Science, Otaru University of Commerce, 3-5-21, Midori, Otaru 047-8501 Japan. Email: hogasa@res.otaru-uc.ac.jp

estimators $\hat{\boldsymbol{\theta}}$ is obtained by minimizing F over the parameter space. Let $\hat{F}$ denote F when $\boldsymbol{\Sigma}$ is given by $\hat{\boldsymbol{\Sigma}} = \boldsymbol{\Sigma}(\hat{\boldsymbol{\theta}})$. Then, it is well known that under normality with a true model $n\hat{F}$ is asymptotically chi-square distributed with the degrees of freedom (*df*) being $\upsilon \equiv p(p+1)/2 - q$, where $n+1=N$ is the sample size. When the model is slightly misspecified or under a local alternative in normal samples, $n\hat{F}$ has asymptotically the noncentral chi-square distribution. Under normality this can be used for testing models. In the local alternative the population covariance matrix is assumed to be a function of *n*, which is a technical one to have the noncentral distribution and is not a realistic assumption.

In practice, it is known that when the sample size is large, $n\hat{F}$ tends to become large, which indicates that fixed alternative hypotheses are more realistic. In practice, data are usually nonnormally distributed. Ogasawara (2007a) derived the asymptotic distributions of various fit indexes including the chi-square statistic by the single-term Edgeworth expansion under nonnormality and fixed alternatives. In this paper the results are extended to the two-term Edgeworth expansion with simulations for confirmation with a finite sample size.

2 Fit indexes

As in Ogasawara (2007a), thirteen fit indexes including $\hat{F}$ are dealt with in this article. The members of the first group ([1] to [8] given below) are the fit indexes as functions of $\hat{F}$ while the members of the second group ([9] to [15]) use the so-called baseline model of uncorrelated observed variables as well as $\hat{F}$. They are defined as

[1] $\hat{F} = \ln|\hat{\boldsymbol{\Sigma}}| - \ln|\mathbf{S}| + \mathrm{tr}(\hat{\boldsymbol{\Sigma}}^{-1}\mathbf{S}) - p$; [2] $\hat{F}_{\mathrm{B}} = \ln|\mathrm{Diag}(\mathbf{S})| - \ln|\mathbf{S}|$;

[3] $\mathrm{GFI} = 1 - \dfrac{2\hat{F}}{2\hat{F}+p} = \dfrac{p}{2\hat{F}+p}$; [4] $\mathrm{AGFI} = 1 - \dfrac{p(p+1)}{2\upsilon}(1-\mathrm{GFI})$;

[5] $\mathrm{Abs.GFI} = \exp\left\{-\dfrac{1}{2}\left(\hat{F} - \dfrac{\upsilon}{n}\right)\right\}$; [6] $\mathrm{RMSEA} = \sqrt{\mathrm{Max}\left(\dfrac{\hat{F}}{\upsilon} - \dfrac{1}{n},\ 0\right)}$;

[7] $\hat{\Gamma}_1 = 1 - \dfrac{2\{\hat{F}-(\upsilon/n)\}}{2\{\hat{F}-(\upsilon/n)\}+p} = \dfrac{p}{2\{\hat{F}-(\upsilon/n)\}+p}$;

[8] $\hat{\Gamma}_2 = 1 - \frac{p(p+1)}{2\upsilon}(1-\hat{\Gamma}_1)$;

$$I = \frac{\hat{F}_{\mathrm{B}} - k_1\hat{F} - k_2}{\hat{F}_{\mathrm{B}} - k_0} = 1 - \frac{k_1\hat{F} + k_2 - k_0}{\hat{F}_{\mathrm{B}} - k_0}$$ with

[9] NFI, $k_0 = 0$, $k_1 = 1$, $k_2 = 0$; [10] $\mathrm{IFI}(\Delta_2)$, $k_0 = \upsilon / n$, $k_1 = 1$, $k_2 = 0$; [11] ρ_1, $k_0 = 0$, $k_1 = \upsilon_{\mathrm{B}} / \upsilon$, $k_2 = 0$; [12] ρ_2, $k_0 = \upsilon_{\mathrm{B}} / n$, $k_1 = \upsilon_{\mathrm{B}} / \upsilon$, $k_2 = 0$; [13] FI (RNI), $k_0 = \upsilon_{\mathrm{B}} / n$, $k_1 = 1$, $k_2 = (\upsilon_{\mathrm{B}} - \upsilon)/n$; where $\hat{F}$ is as before; $\hat{F}_{\mathrm{B}}$ is the value of $\hat{F}$ for the baseline model of uncorrelated observed variables with $df = \upsilon_{\mathrm{B}} \equiv p(p+1)/2 - p = p(p-1)/2$; $\mathrm{Diag}(\cdot)$ denotes a diagonal matrix whose diagonal elements are those of the argument matrix; and for details of the fit indexes see Ogasawara (2007a).

3 Asymptotic expansion of the fit indexes

Let $I = I(\hat{F}, \hat{F}_{\mathrm{B}})$ be the function of $\hat{F}$ and/or $\hat{F}_{\mathrm{B}}$. We assume that the following Taylor expansion with the differentiability of I with respect to sample variances and covariances up to the third order is available:

$$\begin{aligned} I = I_0 &+ \frac{\partial I}{\partial \mathbf{s}'}\Big|_{\mathbf{s}=\boldsymbol{\sigma}_{\mathrm{T}}} (\mathbf{s} - \boldsymbol{\sigma}_{\mathrm{T}}) + \frac{1}{2}\left(\frac{\partial}{\partial \mathbf{s}'}\right)^{<2>} I\Big|_{\mathbf{s}=\boldsymbol{\sigma}_{\mathrm{T}}} (\mathbf{s} - \boldsymbol{\sigma}_{\mathrm{T}})^{<2>} \\ &+ \frac{1}{6}\left(\frac{\partial}{\partial \mathbf{s}'}\right)^{<3>} I\Big|_{\mathbf{s}=\boldsymbol{\sigma}_{\mathrm{T}}} (\mathbf{s} - \boldsymbol{\sigma}_{\mathrm{T}})^{<3>} + O_p(n^{-2}), \end{aligned} \tag{1}$$

where $I_0 = I(F_0, F_{\mathrm{B}0})$, $F_0 = \ln|\boldsymbol{\Sigma}_0| - \ln|\boldsymbol{\Sigma}_{\mathrm{T}}| + \mathrm{tr}(\boldsymbol{\Sigma}_0^{-1}\boldsymbol{\Sigma}_{\mathrm{T}}) - p$, $\boldsymbol{\Sigma}_0 = \boldsymbol{\Sigma}(\boldsymbol{\theta}_0)$, $F_{\mathrm{B}0} = \ln|\mathrm{Diag}(\boldsymbol{\Sigma}_{\mathrm{T}})| - \ln|\boldsymbol{\Sigma}_{\mathrm{T}}|$, $\boldsymbol{\Sigma}_{\mathrm{T}} = \mathrm{E}(\mathbf{S})$, $\mathbf{s} = \mathrm{v}(\mathbf{S})$, $\boldsymbol{\sigma}_{\mathrm{T}} = \mathrm{v}(\boldsymbol{\Sigma}_{\mathrm{T}})$, $\mathrm{v}(\cdot)$ is the vectorizing operator taking the nonduplicated elements of a symmetric matrix, $\mathbf{s}^{<p>} = \mathbf{s} \otimes \mathbf{s} \otimes \cdots \otimes \mathbf{s}$ (*p* times) is the *p*-fold Kronecker product of a vector, and $\boldsymbol{\theta}_0$ is given by fitting the (misspecified) model to the true covariance matrix $\boldsymbol{\Sigma}_{\mathrm{T}}$.

Let $w = \sqrt{n}(I - I_0)$. From (1), under a fixed alternative assumption with possibly nonnormal data, when the cumulants exist they are given by

$$
\begin{aligned}
\kappa_1(w) &= \mathrm{E}(w) = n^{-1/2}\alpha_1 + O(n^{-3/2}), \\
\kappa_2(w) &= \mathrm{E}[\{w - \mathrm{E}(w)\}^2] = \alpha_2 + n^{-1}\Delta\alpha_2 + O(n^{-2}), \\
\kappa_3(w) &= \mathrm{E}[\{w - \mathrm{E}(w)\}^3] = n^{-1/2}\alpha_3 + O(n^{-3/2}), \\
\kappa_4(w) &= \mathrm{E}[\{w - \mathrm{E}(w)\}^4] - 3\{\kappa_2(w)\}^2 = n^{-1}\alpha_4 + O(n^{-2}).
\end{aligned}
\tag{2}
$$

Using the asymptotic cumulants of (2), it is known that the following distribution function by the two-term Edgeworth expansion is valid for nonnormal data under regularity conditions:

$$
\Pr\left(\frac{w}{\alpha_2^{1/2}} \le z\right) = \Phi(z) - n^{-1/2}\left\{\frac{\alpha_1}{\alpha_2^{1/2}} + \frac{\alpha_3}{6\alpha_2^{3/2}}(z^2-1)\right\}\phi(z)
$$

$$
-n^{-1}\left\{\frac{1}{2}(\Delta\alpha_2 + \alpha_1^2)\frac{z}{\alpha_2} + \left(\frac{\alpha_4}{24} + \frac{\alpha_1\alpha_3}{6}\right)\frac{z^3-3z}{\alpha_2^2} + \frac{\alpha_3^2(z^5-10z^3+15z)}{72\alpha_2^3}\right\}\phi(z)
$$

$$
+o(n^{-1}),
$$

where $\phi(z) = (1/\sqrt{2\pi})\exp(-z^2/2)$ and $\Phi(z) = \int_{-\infty}^{z}\phi(t)dt$.

The asymptotic cumulants are given by the moments of the observable variables up to the eighth order with the partial derivatives of the fit indexes with respect to the sample variances and covariances of the observable variables up to the third order (see Ogasawara, 2007a, b). The remaining asymptotic cumulant other than those available in Ogasawara (2007a) is α_4 , which is the most complicated one in (2). This can be given by using the formula of Ogasawara (2006):

$$
\alpha_4 = \sum_{a\ge b}\sum_{c\ge d}\sum_{e\ge f}\sum_{g\ge h}\left[\frac{\partial I_0}{\partial\sigma_{\mathrm{T}ab}}\frac{\partial I_0}{\partial\sigma_{\mathrm{T}cd}}\frac{\partial I_0}{\partial\sigma_{\mathrm{T}ef}}\frac{\partial I_0}{\partial\sigma_{\mathrm{T}gh}}\right.
$$

$$
\times\left(\kappa_{abcdefgh} + \sum^{24}\kappa_{ac}\kappa_{bdefgh} + \sum^{32}\kappa_{ace}\kappa_{bdfgh}\right.
$$

$$
+\sum^{8}\kappa_{aceg}\kappa_{bdfh} + \sum^{24}\kappa_{abeg}\kappa_{cdfh} + \sum^{96}\kappa_{ac}\kappa_{be}\kappa_{dfgh} + \sum^{48}\kappa_{ac}\kappa_{eg}\kappa_{bdfh}
$$

$$
\left.+\sum^{96}\kappa_{ac}\kappa_{beg}\kappa_{dfh} + \sum^{48}\kappa_{bc}\kappa_{de}\kappa_{fg}\kappa_{ha} - \sum^{6}\kappa_{abcd}M(ef,gh)\right)
$$

$$+\sum_{j\geq k} 2\frac{\partial^2 I_0}{\partial\sigma_{\mathrm{T}ab}\partial\sigma_{\mathrm{T}cd}}\frac{\partial I_0}{\partial\sigma_{\mathrm{T}ef}}\frac{\partial I_0}{\partial\sigma_{\mathrm{T}gh}}\frac{\partial I_0}{\partial\sigma_{\mathrm{T}jk}}\sum^{10} M(ab,cd)M(ef,gh,jk)$$

$$+\sum_{j\geq k}\sum_{l\geq m}\left(\frac{3}{2}\frac{\partial^2 I_0}{\partial\sigma_{\mathrm{T}ab}\partial\sigma_{\mathrm{T}cd}}\frac{\partial^2 I_0}{\partial\sigma_{\mathrm{T}ef}\partial\sigma_{\mathrm{T}gh}}+\frac{2}{3}\frac{\partial^3 I_0}{\partial\sigma_{\mathrm{T}ab}\partial\sigma_{\mathrm{T}cd}\partial\sigma_{\mathrm{T}ef}}\frac{\partial I_0}{\partial\sigma_{\mathrm{T}gh}}\right)$$

$$\left.\times\frac{\partial I_0}{\partial\sigma_{\mathrm{T}jk}}\frac{\partial I_0}{\partial\sigma_{\mathrm{T}lm}}\sum^{15} M(ab,cd)M(ef,gh)M(jk,lm)\right]$$

$$-(4\alpha_1\alpha_3+6\alpha_2\Delta\alpha_2+6\alpha_2\alpha_1^2),$$

where the partial derivatives denote those evaluated at the population values; $a\geq b$ is $p\geq a\geq b\geq 1$; *M*(*ab*, *cd*) and *M*(*ef*, *gh*, *jk*) are $n\,\mathrm{E}\{(s_{ab}-\sigma_{\mathrm{T}ab})(s_{ab}-\sigma_{\mathrm{T}cd})\}$ and $n^2\,\mathrm{E}\{(s_{ef}-\sigma_{\mathrm{T}ef})(s_{gh}-\sigma_{\mathrm{T}gh})(s_{jk}-\sigma_{\mathrm{T}jk})\}$ up to $O(1)$, respectively with $\mathbf{S}=\{s_{ab}\}$ and $\mathbf{\Sigma}_{\mathrm{T}}=\{\sigma_{\mathrm{T}ab}\}$; and $\overset{k}{\Sigma}$ is the sum of *k* similar terms.

4 Numerical illustration

A numerical example using a misspecified one-factor model as in Ogasawara (2007a) is presented, where the true covariance matrix is given by the two-factor model:

$$\mathbf{\Sigma}_{\mathrm{T}}=\mathbf{\Lambda}_{\mathrm{T}}\mathbf{\Lambda}_{\mathrm{T}}'+\mathbf{\Psi}_{\mathrm{T}},\ \mathbf{\Lambda}_{\mathrm{T}}=\begin{bmatrix}.6 & .6 & .6 & .6 & .6 & .6\\ .3 & .3 & .3 & -.3 & -.3 & -.3\end{bmatrix}',$$

$$\mathbf{\Psi}_{\mathrm{T}}=\mathrm{diag}(.55,\ .55,\ .55,\ .55,\ .55,\ .55,\ .55).$$

Normal and nonnormal observations with sample size *N*=300 were randomly generated by $\mathbf{x}=\mathbf{\Gamma}_{\mathrm{T}}\mathbf{f}$, where $\mathbf{\Gamma}_{\mathrm{T}}$ is a lower-triangular matrix such that $\mathbf{\Sigma}_{\mathrm{T}}=\mathbf{\Gamma}_{\mathrm{T}}\mathbf{\Gamma}_{\mathrm{T}}'$ and **f** is a random vector with independent elements. For nonnormal observations the elements of **f** have standardized chi-square distributions with *df*=3. The simulation was performed with 1,000,000 replications. Table 1 shows asymptotic and simulated cumulants except α_2. The simulated $\Delta\alpha_2$ is defined by $n^2(\mathrm{SD}^2-n^{-1}\alpha_2)$, where SD is the usual standard deviation for a fit index given by 1,000,000 estimates. Table 2 shows the root mean square errors of the approximate distribution functions of the standardized estimators ($n^{1/2}(I - I_0)/\alpha_2^{1/2}$) given by the single- and two-term Edgeworth expansions and Hall's (1992) variable transformation. The true distribution functions were given by the simulations. It is known that Hall's method is asymptotically equivalent to the single-term Edgeworth expansion. We see that on average the two-term Edgeworth expansion has reduced the errors.

References

Hall, P. (1992). On the removal of skewness by transformation. *Journal of the Royal Statistical Society, B, 54*, 221-228.

Ogasawara, H. (2006). Asymptotic expansion of the sample correlation coefficient under nonnormality. *Computational Statistics and Data Analysis, 50*, 891-910.

Ogasawara, H. (2007a). Higher-order approximations to the distributions of fit indexes under fixed alternatives in structural equation models. *Psychometrika*, *72*, 227-243.

Ogasawara, H. (2007b). Higher-order estimation error in structural equation modeling. *Economic Review (Otaru University of Commerce), 47*(4), 131-160. http://www.res. otaru-uc.ac.jp/~hogasa/

Table 1. Theoretical and simulated cumulants for the misspecified one-factor model (N=300)

	Normal		χ^2 (df=3)		Normal		χ^2 (df=3)	
Fit index	Th.	Sim.	Th.	Sim.	Th.	Sim.	Th.	Sim.
	$\Delta\alpha_2$ (higher-order added variance)				α_1 (bias)			
F	.17	1.97	-8.53	-2.62	7.74	7.87	8.44	8.59
F_B	31.18	36.72	82.84	86.18	15.00	15.18	18.71	18.76
GFI	-1.20	-1.00	-2.54	-1.92	-2.15	-2.17	-2.33	-2.36
AGFI	-6.54	-5.43	-13.85	-10.43	-5.02	-5.07	-5.43	-5.50
Abs.GFI	-2.43	-1.97	-5.48	-4.03	-3.41	-3.45	-3.70	-3.74
RMSEA	-5.58	-4.97	-9.29	-7.77	2.44	2.48	2.56	2.63
$\hat{\Gamma}_1$	-1.26	-1.05	-2.66	-2.00	-2.19	-2.21	-2.37	-2.40
$\hat{\Gamma}_2$	-6.85	-5.69	-14.47	-10.91	-5.11	-5.16	-5.54	-5.60
NFI	-12.81	-11.71	-27.32	-22.51	-4.35	-4.28	-4.37	-4.31
IFI	-13.57	-12.42	-29.12	-24.01	-4.60	-4.52	-4.65	-4.58
ρ_1	-35.59	-32.52	-75.90	-62.52	-7.25	-7.13	-7.29	-7.18
ρ_2	-38.90	-35.61	-83.62	-68.95	-7.77	7.64	-7.85	-7.73
FI (RNI)	-14.00	-12.82	-30.10	-24.82	-4.66	-4.58	-4.71	-4.64
	α_3 (skewness)				α_4 (kurtosis)			
F	7.89	7.11	13.9	12.3	73.5	51.0	204	142
F_B	54.0	55.9	130.0	134.3	377	491	2898	3083
GFI	-.14	-.11	-.26	-.20	.1	.0	.6	.2
AGFI	-1.79	-1.45	-3.27	-2.57	4.0	.9	16.9	7.2
Abs.GFI	-.58	-.48	-1.06	-.85	1.2	.5	4.4	2.1
RMSEA	-.10	-.08	-.08	-.05	.5	.2	.8	.3
$\hat{\Gamma}_1$	-.15	-.12	-.27	-.21	.1	.0	.6	.3
$\hat{\Gamma}_2$	-1.88	-1.53	-.3.45	-.2.70	4.2	.9	18.0	7.6
NFI	-.3.26	-.2.53	-.4.40	-.3.09	22.0	9.9	27.6	7.1
IFI	-.3.36	-2.60	-4.58	-3.21	22.9	10.2	29.4	7.3
ρ_1	-15.1	-11.7	-20.4	-14.3	170	76.4	213	55.0
ρ_2	-16.5	-12.8	-22.4	-15.7	194	88.3	247	64.3
FI (RNI)	-3.56	-2.76	-4.83	-3.39	25.1	11.4	32.1	8.3

Note. Th.=Theoretical or asymptotic values. Sim.=Simulated values.

Table 2. $10^5 \times$ root mean square errors of the approximate distribution functions of the standardized estimators for the misspecified one-factor model (*N*=300)

	N*	E1	E2	Hall	N*	E1	E2	Hall
	Normally distributed data				Chi-square distributed data with *df*=3			
F	7690	864	248	254	7377	746	456	368
F_B	5776	626	66	118	6114	701	80	150
GFI (AGFI)	7473	669	245	303	7138	547	450	469
Abs.GFI	7513	702	244	283	7183	581	451	443
RMSEA	6858	318	233	724	6467	331	409	897
$\hat{\Gamma}_1$ ($\hat{\Gamma}_2$)	7471	667	245	304	7136	545	450	470
NFI (ρ_1)	5561	399	332	499	5181	628	487	753
IFI	5818	399	338	508	5432	623	501	772
ρ_2 (FI, RNI)	5791	403	339	511	5405	629	501	774

Note. N*=Normal approximation, E1=Single-term Edgeworth expansion, E2=Two-term Edgeworth expansion, Hall=Hall's method by variable transformation.