The Korean Communications
in Statistics Vol. 5, No. 3, 1998
pp. 847-853

On the Comparison of Two Non-hierarchical Log-linear Models[1]

Min Gweon Oh[2], Chong Sun Hong[3] and Donguk Kim[4]

Abstract

Suppose we want to compare following non-hierarchical log-linear models, $H_0: f(x; \theta \in \Theta_\alpha)$ vs. $H_1: g(x; \theta \in \Theta_\beta)$ for $\Theta_\alpha, \Theta_\beta \subset \Theta$ such that $\Theta_\alpha \not\subset \Theta_\beta$. The goodness of fit test using the likelihood ratio test statistic for comparing these models could not be acceptable. By using the polyhedrons plots of Choi and Hong (1995), we propose a method to decide a better model between two non-hierarchical log-linear models $f(x; \theta \in \Theta_\alpha)$ and $g(x; \theta \in \Theta_\beta)$.

1. Introduction

For a random variable X and parameter space Θ, suppose we test $H_0: f(x; \theta \in \Theta_\alpha)$ vs. $H_1: g(x; \theta \in \Theta_\beta)$, where $\Theta_\alpha, \Theta_\beta \subset \Theta$ is such that $\Theta_\alpha \not\subset \Theta_\beta$. For testing this hypotheses, the generalized likelihood ratio (GLR) test could not be performed. Cox (1961, 1962) suggested the following test statistic for separate families, which is a modification of Maximum-likelihood (ML) ratio test by Neyman-Pearson,

$$T_f = \{L_f(\hat{\theta}) - L_g(\hat{\hat{\theta}})\} - E_{\hat{\theta}}\{(L_f(\hat{\theta}) - L_g(\hat{\hat{\theta}}))\},$$

where $\hat{\theta}$ and $\hat{\hat{\theta}}$ denote ML estimator for θ, $L_f(\hat{\theta})$ and $L_g(\hat{\hat{\theta}})$ denote the maximum value of the log-likelihood functions to the null and alternative hypotheses, respectively. A procedure suggested by Cox (1961, 1962) involves essentially four steps.

(step 1) Obtain the ML estimator $\hat{\theta}$ and $\hat{\hat{\theta}}$.

(step 2) Compute the log-likelihood ratio $L_f(\hat{\theta}) - L_g(\hat{\hat{\theta}})$.

(step 3) Find $E_{\hat{\theta}}\{L_f(\hat{\theta}) - L_g(\hat{\hat{\theta}})\}$.

(step 4) Derive the asymptotic variance of T_f.

---

1) This paper was supported by 63 Research Fund, Sungkyunkwan University, 1996.
2) Lecturer, Department of Statistics, Sungkyunkwan University, Seoul, 110-745, KOREA.
3) Professor, Department of Statistics, Sungkyunkwan University, Seoul, 110-745, KOREA.
4) Assistant Professor, Department of Statistics, Sungkyunkwan University, Seoul, 110-745, KOREA.

In this work we are interested in making inferences for categorical data. Lindsey (1974a, 1974b) developed a method of comparison and plausibility of fit of discrete probability distributions using the likelihood function. For testing two non-hierarchical log-linear models in the hypotheses

$$H_0: f(x;\theta\in\Theta_\alpha) \text{ vs. } H_1: g(x;\theta\in\Theta_\beta), \tag{1.1}$$

where $\Theta_\alpha \not\subset \Theta_\beta$, the test statistic T_f could not be calculated in (step 3), so not only the GLR test but also the modified GLR test are no longer acceptable methods (See Cox (1962 p.411)). But we can perform the GLR tests for the null hypothesis model and the alternative hypothesis model, separately. Since each model can be accepted or rejected, we consider four possible cases which are summarized in <Table 1>.

<Table 1> The results of the separate GLR test for non-hierarchical log-linear models

Case	Null hypothesis	Alternative hypothesis	Decision
①	reject	reject	Both rejected
②	accept	reject	Null accepted
③	reject	accept	Alternative accepted
④	accept	accept	Further analysis

For case ① in <Table 1>, both models are not fitted to the corresponding data. So it is meaningless to decide that which one is better between two models. Hence we do not need to consider the testing of the hypotheses (1.1). Either one model is well-fitted and other is not in case ② and ③, so a clear decision might be derived. But both models in case ④ are all well-fitted. Nonetheless it is worth while to choose the better one between two models, but it is not easy to decide which one is better-fitted. In this paper we are interested in case ④ and we will propose a method that can select a better model between two non-hierarchical well-fitted log-linear models by using "the polyhedrons plots" (see Choi and Hong (1995), and Hong (1995) for more detail).

2. Goodness-of-fit tests for non-hierarchical log-linear models

We illustrate the goodness-of-fit test for the non-hierarchical log-linear models using the well-known data in <Table 2>. This data was studied earlier by Ries and Smith (1963), and Hong(1995) among others.

<Table 2> Detergent preference data

water softness	brand preferences	used			
		yes		no	
		temperature			
		high	low	high	low
soft	X	19	57	29	63
	M	29	49	27	53
medium	X	23	47	33	66
	M	47	55	23	50
hard	X	24	37	42	68
	M	43	52	30	42

This is a cross classification of the degree of softness of the water they used (soft, medium, hard) (var. 1), brand preferences (X or M) (var. 4), whether they had used brand M previously (yes, no) (var. 2), and the temperature of the laundry water used (high, low) (var. 3). Suppose we are testing following hypothesis

$$H_0 : [3][12][24] \text{ vs. } H_1 : [13][24][34] . \tag{1.2}$$

Since these models are not nested, we can not use the GLR test directly. Hence we consider two separate GLR tests for the null and alternative hypothesis. The GLR statistic for testing H_0 has value $G^2 = 21.27$ based on d.f.=15. This statistic yields a p-value of 0.1284, hence the model [3][12][24] can be acceptable. Furthermore for the alternative, we have $G^2 = 11.89$ based on d.f.=14, and it's p-value is 0.6154. The model [13][24][34] fits the data very well. With these results, we knew that both models [3][12][24] and [13][24][34] fit the data. In order to find which one is a better model between two, and whether the difference of two models is significant, let us take two sets of appropriate hierarchical log-linear models including each model, respectively. For example, here is one set whose elements are six hierarchical log-linear models in <Table 3>, where the null model [3][12][24] locates in the middle of the set. Then the GLR tests for the hierarchical set have performed. <Table 3> shows the results of goodness-of-fit tests of hierarchical models including [3][12][24].

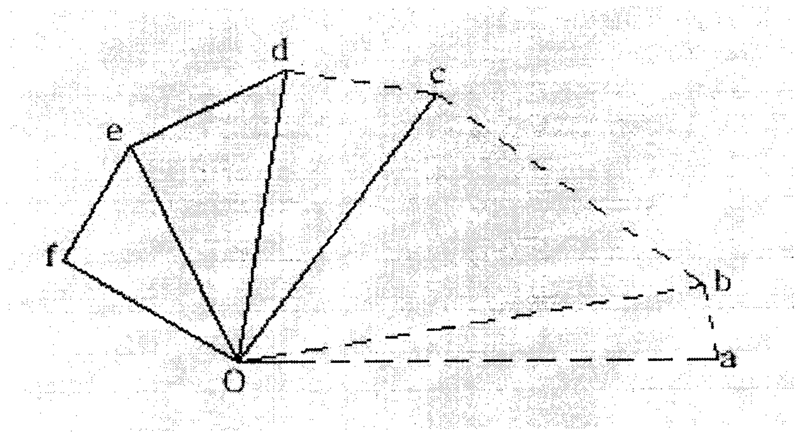
<Figure 1> contains the polyhedrons plot which represents the results in <Table 3>. In <Figure 1>, the length of $\overline{Oa}$, $\overline{Ob}$, $\overline{Oc}$, $\overline{Od}$, $\overline{Oe}$, and $\overline{Of}$ denote the square-rooted values of GLR test statistics and the length $\overline{ab}$, for example, denotes that of the difference of G^2 for model (a) and model (b). Note that the angles of abO, bcO, cdO, ..., efO are all 90° and

$\cos\theta = \frac{G(b)}{G(a)} = 1 - R^2$, where R^2 is the coefficient of determination for model (b). The dotted line (red line in the system) presents that the corresponding test statistic indicates significant, and the black solid line denotes not significant (see Choi and Hong (1995), and Hong (1995) for more detail).

<Table 3> The results of the GLR tests for an hierarchical set including the model of null hypothesis [3][12][24].

Model	d.f.	G^2	Difference of Models	d.f.	G^2
(a) : [1][2][3][4]	18	42.92*			
(b) : [3][4][12]	16	41.85*	(a) and (b)	2	1.07*
(c) : [3][12][24]	15	21.27	(b) and (c)	1	20.28*
(d) : [12][24][34]	14	16.91	(c) and (d)	1	4.36*
(e) : [34][124]	10	11.29	(d) and (e)	4	5.62
(f) : [124][134]	8	7.81	(e) and (f)	2	3.48

* indicates significant at 5% level.



<Figure 1> Polyhedrons plot for <Table 3>

From <Figure 1> the goodness-of-fit test statistics of model (c) : [3][12][24] and (d) : [12][24][34] are not significant at 5% significant level. So both are said to be well-fitted models. But the difference between two models is also statistically significant. Hence we can say that the model (c) : [3][12][24] is not the best model among the six models and this model is not enough to explain the data perfectly.

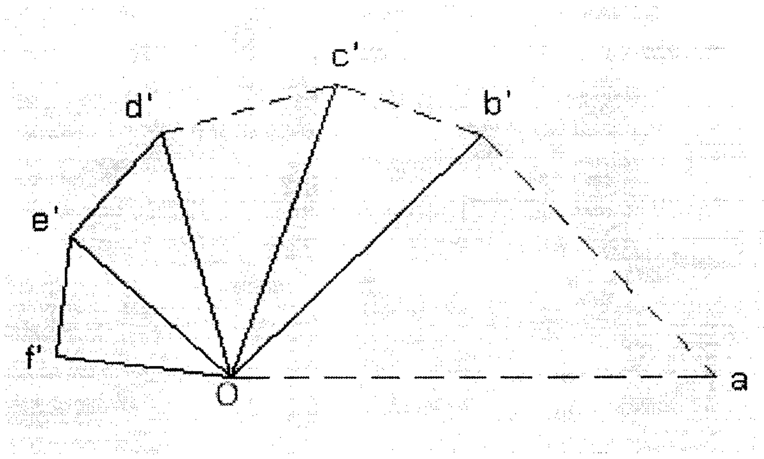
Now let us consider another hierarchical log-linear model set where the model of alternative hypothesis [13][24][34] locates in the middle of the set. <Table 4> summarizes the results of the GLR tests for the hierarchical log-linear model set including the model [13][24][34].

<Table 4> The results of the GLR tests for an hierarchical set including the model of alternative hypothesis [13][24][34].

Model	d.f.	G^2	Difference of Models	d.f.	G^2
(a) : [1][2][3][4]	18	42.92*			
(b′) : [1][3][24]	17	22.35	(a) and (b′)	1	20.57*
(c′) : [1][24][34]	16	17.99	(b′) and (c′)	1	4.36*
(d′) : [13][24][34]	14	11.89	(c′) and (d′)	2	6.1*
(e′) : [13][234]	12	8.41	(d′) and (e′)	2	3.48
(f′) : [123][234]	8	5.66	(e′) and (f′)	4	2.75

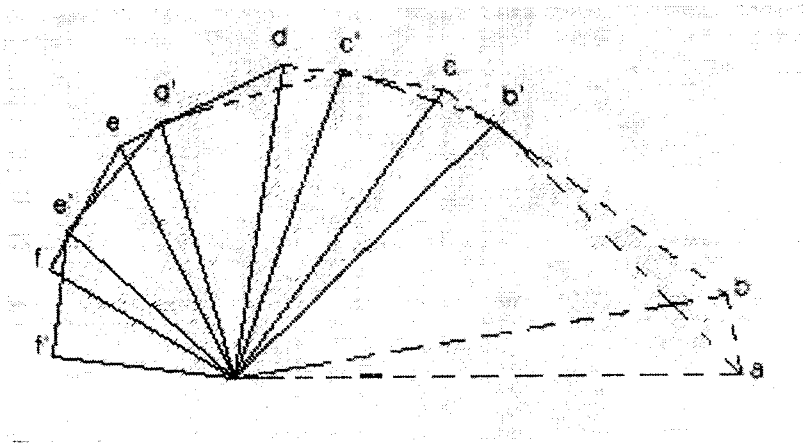
* indicates significant at 5% level.

The polyhedrons plot in <Figure 2> represents the contents of <Table 4>. With <Figure 2> and <Table 4>, the model (d′) : [13][24][34] fits the data well. Also the model (c′) and (e′) are well-fitted. Nonetheless, the difference between (c′) and (d′) is significant, and the difference between (d′) and (e′) is not. Hence we can conclude that the model (d′) : [13][24][34] is the best model among the model set in <Table 4>.



<Figure 2> Polyhedrons plot for <Table 4>

Now consider the testing hypotheses of (1.2), where these two models are not hierarchical. Even though both models are well-fitted, our best choice would be to accept the alternative model (d′) : [13][24][34] because there is no more improvement over the model, while we can have better model than the null hypothesis model.



<Figure 3> The overlapped polyhedrons plot of <Figures 2, and 3>

<Figure 3> shows the overlapped polyhedrons plots in <Figures 2, and 3>. If two polyhedrons plots which represent two sets of hierarchical log-linear models are overlapped like in <Figure 3>, then we can not only recognize each statistical results of both non-hierarchical models, but also the difference of two models. Particularly, it is obvious when the null model (c) locates in the right side of the plots in <Figure 3>, and the alternative model (d') locates in the left side. Hence in order to select the better model among some non-hierarchical log-linear models, the polyhedrons plots and their overlapped plots can help us to give reasonable solution.

Let us consider another kind of hypotheses test for non-hierarchical log-linear models. Suppose first one needs to compare model (c) : [3][12][24] in <Table 2> and model (c') : [1][24][34] in <Table 3>, which are not only non-hierarchical but also well-fitted models. From <Figure 3>, both model (c) and (c') are not the best model in each set of hierarchical models. Therefore one can not say that there exists a big difference between model (c) and (c').

Now we take another hypotheses models, for example, model (f) : [124][134] in <Table 2> and model (e') : [13][234] in <Table 3>, whose situations are similar as that of model (c) and (c'). But we can say that both model (f) and (e') are well-enough-fitted models. With the analogous arguments, model (f) and (e') are said to be indistinguishable. <Figure 3> shows us that the corresponding lines for model (c) and (c') locate in the right side, and the corresponding lines for model (f) and (e') are in the left side. Therefore the proposed method in this paper would give us some useful informations to compare two non-hierarchical log-linear models in two separate sets.

3. Conclusion

For testing $H_0: f(x; \theta \in \Theta_\alpha)$ vs. $H_1: g(x; \theta \in \Theta_\beta)$ where $\Theta_\alpha \not\subset \Theta_\beta$ among the non-hierarchical log-linear models, the ordinary GLR goodness-of-fit test could not be performed. Furthermore, if both models in the null and alternative hypotheses are separately acceptable, it is not easy to decide a better one between two well-fitted models. In this case, we have considered two sets of appropriate hierarchical models which contains the null and alternative hypothesis, respectively. Then the log-linear models in the two sets could be expressed by the polyhedrons plots of Choi and Hong (1995), and their overlapped plots. In this paper we propose a method that not only each statistical results of both non-hierarchical models, but also the difference of two models can be recognized.

If two sets of hierarchy is not well selected, then the comparison of two non-hierarchical log-linear models might not be possible. Hence it is important to select a proper hierarchy including null (or alternative) log-linear model. In practice, we need to compare all possible sets of hierarchy. But our software, that is developed to meet this purpose, will perform the process discussed in this paper and help us to analyze and compare two non-hierarchical log-linear models. The software used in this paper can be accessed by using http://stat.skku.ac.kr/~cshong/Categorical.html.

References

[1] Cox, D. R. (1961). *Tests of separate families of hypotheses*. Proc. 4th Berkeley Symp., Vol. 1, 105-123.

[2] Cox, D. R. (1962). Further results on tests of separate families of hypotheses. *J. Royal Statist.* Soc. B. Vol. 24, 406-423.

[3] Choi, H. J. and Hong, C. S. (1995). Graphical descriptions for hierarchical log-linear models, *The Korean Communications in Statistics*, Vol. 2, No. 2, 310-319.

[4] Lindsey, J, K. (1974a). Comparison of probability distributions. *J. Royal Statist. Soc.* B, Vol. 24, 38-47.

[5] Lindsey, J, K. (1974b). Construction and comparison of statistical models. *J. Royal Statist. Soc.,* Vol. 24, 418-425.

[6] Ries, P. N. and Smith, H. [1963], The use of chi-square for preference testing in multidimensional problems. Chemical Engineering Progress Vol. 59, 39-43.

[7] Hong, C. S. (1995). Log-linear Models, Freedom academy press.