Open Journal of Statistics, 2021, 11, 1062-1071
https://www.scirp.org/journal/ojs
ISSN Online: 2161-7198
ISSN Print: 2161-718X

Measure of Departure from Point-Symmetry for the Analysis of Collapsed Square Contingency Tables

Kiyotaka Iki, Kouji Yamamoto, Sadao Tomizawa

Department of Economics, Nihon University, Tokyo, Japan
Email: iki.kiyotaka@nihon-u.ac.jp

How to cite this paper: Iki, K., Yamamoto, K. and Tomizawa, S. (2021) Measure of Departure from Point-Symmetry for the Analysis of Collapsed Square Contingency Tables. *Open Journal of Statistics*, **11**, 1062-1071.
https://doi.org/10.4236/ojs.2021.116063

Received: November 29, 2021
Accepted: December 25, 2021
Published: December 28, 2021

Abstract

For square contingency tables with ordered categories, there may be some cases that one wants to analyze them by considering collapsed 3 × 3 tables with some adjacent categories combined in the original table. This paper considers the point-symmetry model (Wall and Lienert, 1976) for collapsed tables and proposes a measure to represent the degree of departure from point-symmetry for collapsed tables. Also it gives approximate confidence interval for the proposed measure.

Keywords

Collapsed Table, Diversity Index, Measure, Point-Symmetry, Power-Divergence

1. Introduction

Consider an $r \times r$ square contingency table with the same row and column classifications. Let p_{ij} denote the probability that an observation will fall in the *i*th row and *j*th column of the table ($i = 1, \cdots, r; j = 1, \cdots, r$). The point-symmetry (PS) model is defined by

$$p_{ij} = p_{i^* j^*} \quad \left(i = 1, \cdots, r; j = 1, \cdots, r\right),$$

where the symbol $*$ denotes $i^* = r + 1 - i$; see Wall and Lienert [1]. This indicates that the probability of an observation falling in (i, j) th cell is equal to the probability of the observation falling in point symmetric $\left(i^*, j^*\right)$ th cell with respect to the center cell (when *r* is odd) or center point (when *r* is even). Now, we consider the $\left[(r-1)/2\right]$ ways of collapsing the $r \times r$ original table with ordered categories into a 3 × 3 table by choosing cut points after *h*th and $h'(= r - h)$ th rows and after hth and h'th columns for $h=1,\cdots,\left[(r-1)/2\right]$, where

$$\left[\frac{r-1}{2}\right]=\begin{cases}\dfrac{r-1}{2} & (r \text{ is odd}),\\ \dfrac{r-2}{2} & (r \text{ is even}).\end{cases}$$

We refer to each collapsed 3 × 3 table as the $T_{hh'}$ $\left(h=1,\cdots,\left[(r-1)/2\right]\right)$ table. In the collapsed $T_{hh'}$ table, let $G_{kl}^{(h,h')}$ denote the corresponding cumulative probability for row value $k\,(k=1,2,3)$ and column value $l\,(l=1,2,3)$; *i.e.*,

$$G_{11}^{(h,h')}=\sum_{i=1}^{h}\sum_{j=1}^{h}p_{ij},\quad G_{12}^{(h,h')}=\sum_{i=1}^{h}\sum_{j=h+1}^{h'}p_{ij},\quad G_{13}^{(h,h')}=\sum_{i=1}^{h}\sum_{j=h'+1}^{r}p_{ij},$$

$$G_{21}^{(h,h')}=\sum_{i=h+1}^{h'}\sum_{j=1}^{h}p_{ij},\quad G_{22}^{(h,h')}=\sum_{i=h+1}^{h'}\sum_{j=h+1}^{h'}p_{ij},\quad G_{23}^{(h,h')}=\sum_{i=h+1}^{h'}\sum_{j=h'+1}^{r}p_{ij},$$

$$G_{31}^{(h,h')}=\sum_{i=h'+1}^{r}\sum_{j=1}^{h}p_{ij},\quad G_{32}^{(h,h')}=\sum_{i=h'+1}^{r}\sum_{j=h+1}^{h'}p_{ij},\quad G_{33}^{(h,h')}=\sum_{i=h'+1}^{r}\sum_{j=h'+1}^{r}p_{ij}.$$

Then, Yamamoto *et al.* [2] considered the collapsed point-symmetry (CoPS) model as

$$G_{ij}^{(h,h')}=G_{i^\dagger j^\dagger}^{(h,h')}\quad\left(i=1,2,3;\,j=1,2,3;\,(i,j)\neq(2,2)\right),$$

for all $h=1,\cdots,\left[(r-1)/2\right]$, where the symbol † denotes $i^\dagger=4-i$. Note that the PS model implies the CoPS model, but the PS model is not equivalent to the CoPS model.

When the CoPS model does not hold, we are interested in measuring the degree of departure from CoPS. For square contingency tables with ordered categories, Tomizawa *et al.* [3] proposed a measure to represent the degree of departure from PS.

By the way, consider the data in **Table 1** taken from Hashimoto [4]. These data describe the cross-classification of father's and son's occupational status categories in Japan which were examined in 1975 and 1995. For the data in **Table 1(a)** & (**Table 1(b)**) having five categories, there may be a case that we want to combine the occupational status into the simpler three categories, namely, "high", "middle" and "low". For example, the collapsed 3 × 3 table T_{14} has "high" category which is "(1) Capitalist" category in the original 5 × 5 table, "middle" category which is obtained by combing "(2) New middle", "(3) Working" and "(4) Self-employed" categories in the original table, and "low" category which is "(5) Farming" category in it. Similarly, we can consider the collapsed 3 × 3 table T_{23}, which has "high" category which is obtained by combing "(1) Capitalist" and "(2) New middle" categories in the original 5 × 5 table, "middle" category which is "(3) Working" category in the original table, and "low" category which is obtained by combing "(4) Self-employd" and "(5) Farming" categories in it. **Table 2** and **Table 3** give the collapsed 3 × 3 tables T_{14}, T_{23} (for observations) for the data in **Table 1(a)** and (**Table 1(b)**, respectively. Now, we are interested in seeing what degree the departure from PS is for each of tables T_{14} and T_{23}. So, the present paper proposes a measure which represents the degree of departure from

Table 1. Occupational status for Japanese father-son pairs; from Hashimoto [4]. (a) examined in 1975; (b) examined in 1995.

(a)

Father's status	Son's status					Total
	(1)	(2)	(3)	(4)	(5)	
(1)	29	43	25	31	4	132
(2)	23	159	89	38	14	323
(3)	11	69	184	34	10	308
(4)	42	147	148	184	17	538
(5)	42	176	377	114	298	1007
Total	147	594	823	401	343	2308

(b)

Father's status	Son's status					Total
	(1)	(2)	(3)	(4)	(5)	
(1)	68	48	36	23	1	176
(2)	33	191	102	33	3	362
(3)	25	147	229	34	2	437
(4)	48	119	146	129	5	447
(5)	40	126	192	82	88	528
Total	214	631	705	301	99	1950

Note: Status (1) is Capitalist, (2) New middle, (3) Working, (4) Self-employed and (5) Farming.

Table 2. Collapsed tables T_{14} and T_{23} for the data in **Table 1(a)**. (a) T_{14} table; (b) T_{23} table.

(a)

Father's status	Son's status			Total
	High	Middle	Low	
High	29	99	4	132
Middle	76	1052	41	1169
Low	42	667	298	1007
Total	147	1818	343	2308

(b)

Father's status	Son's status			Total
	High	Middle	Low	
High	254	114	87	455
Middle	80	184	44	308
Low	407	525	613	1545
Total	741	823	744	2308

Table 3. Collapsed Tables T_{14} and T_{23} for the data in **Table 1(b)**. (a) T_{14} table; (b) T_{23} table

(a)

Father's status	Son's status			Total
	High	Middle	Low	
High	68	107	1	176
Middle	106	1130	10	1246
Low	40	400	88	528
Total	214	1637	99	1950

(b)

Father's status	Son's status			Total
	High	Middle	Low	
High	340	138	60	538
Middle	172	229	36	437
Low	333	338	304	975
Total	845	705	400	1950

CoPS by using collapsed 3 × 3 tables. For related research, see Iki *et al.* [5] and Balcha [6].

The new measures are introduced in Section 2. Section 3 presents an approximate variance and a confidence interval for the proposed measure. Section 4 gives examples. Finally, Section 5 concludes the paper.

2. Measure of Departure from Point-Symmetry for Collapsed Tables

Assume that $\left\{G_{ij}^{(h,h')} + G_{i^\dagger j^\dagger}^{(h,h')} \neq 0\right\}$. Let

$$D = \left\{(i,j) \mid i = 1,2,3; j = 1,2,3; (i,j) \neq (2,2)\right\},$$

and

$$\delta_{hh'} = 1 - G_{22}^{(h,h')}, \quad G_{ij}^{*(h,h')} = \frac{G_{ij}^{(h,h')}}{\delta_{hh'}} \quad \left((i,j) \in D\right),$$

$$Q_{ij}^{*(h,h')} = \frac{G_{ij}^{*(h,h')} + G_{i^\dagger j^\dagger}^{*(h,h')}}{2} \quad \left((i,j) \in D\right).$$

Consider a measure to represent the degree of departure from CoPS, defined by

$$\Psi(\lambda) = \frac{1}{\left\lceil \frac{r-1}{2} \right\rceil} \sum_{h=1}^{\left\lceil \frac{r-1}{2} \right\rceil} \Psi_{hh'}(\lambda) \quad (\lambda > -1),$$

where

$$\Psi_{hh'}(\lambda) = \frac{\lambda(\lambda+1)}{2^\lambda - 1} I_{hh'}(\lambda),$$

$$I_{hh'}(\lambda)=\frac{1}{\lambda(\lambda+1)}\sum_{(i,j)\in D}G_{ij}^{*(h,h')}\left\{\left(\frac{G_{ij}^{*(h,h')}}{Q_{ij}^{*(h,h')}}\right)^{\lambda}-1\right\},$$

and the value at $\lambda=0$ is taken to be continuous limit as $\lambda\to 0$. Namely

$$\Psi(0)=\frac{1}{\left[\frac{r-1}{2}\right]}\sum_{h=1}^{\left[\frac{r-1}{2}\right]}\Psi_{hh'}(0),$$

where

$$\Psi_{hh'}(0)=\frac{1}{\log 2}I_{hh'}(0),$$

$$I_{hh'}(0)=\sum_{(i,j)\in D}G_{ij}^{*(h,h')}\log\left(\frac{G_{ij}^{*(h,h')}}{Q_{ij}^{*(h,h')}}\right).$$

The submeasure $\Psi_{hh'}(\lambda)$ represents the degree of departure from PS for the collapsed $T_{hh'}$ table. We note that $I_{hh'}(\lambda)$ is the power-divergence between two probabilities $\left\{G_{ij}^{*(h,h')}\right\}$ and $\left\{Q_{ij}^{*(h,h')}\right\}$, and especially $I_{hh'}(0)$ is the Kullback-Leibler information between them. (For more details of the power-divergence $I_{hh'}(\lambda)$, see Cressie and Read [7]; Read and Cressie [8]).

Let

$$G_{ij}^{c(h,h')}=\frac{G_{ij}^{(h,h')}}{G_{ij}^{(h,h')}+G_{i^\dagger j^\dagger}^{(h,h')}}\quad\left((i,j)\in D\right).$$

Also let $E=\{(1,1),(1,2),(1,3),(2,1)\}$. Then the submeasure $\Psi_{hh'}(\lambda)$ is expressed as

$$\Psi_{hh'}(\lambda)=\frac{\lambda(\lambda+1)}{2^{\lambda}-1}\sum_{(i,j)\in E}\left(G_{ij}^{*(h,h')}+G_{i^\dagger j^\dagger}^{*(h,h')}\right)I_{ij}^{(h,h')}(\lambda)\quad(\lambda>-1),$$

where

$$I_{ij}^{(h,h')}(\lambda)=\frac{1}{\lambda(\lambda+1)}\left[G_{ij}^{c(h,h')}\left\{\left(\frac{G_{ij}^{c(h,h')}}{1/2}\right)^{\lambda}-1\right\}+G_{i^\dagger j^\dagger}^{c(h,h')}\left\{\left(\frac{G_{i^\dagger j^\dagger}^{c(h,h')}}{1/2}\right)^{\lambda}-1\right\}\right],$$

and the value at $\lambda=0$ is taken to be continuous limit as $\lambda\to 0$. Namely

$$\Psi_{hh'}(0)=\frac{1}{\log 2}\sum_{(i,j)\in E}\left(G_{ij}^{*(h,h')}+G_{i^\dagger j^\dagger}^{*(h,h')}\right)I_{ij}^{(h,h')}(0),$$

$$I_{ij}^{(h,h')}(0)=G_{ij}^{c(h,h')}\log\left(\frac{G_{ij}^{c(h,h')}}{1/2}\right)+G_{i^\dagger j^\dagger}^{c(h,h')}\log\left(\frac{G_{i^\dagger j^\dagger}^{c(h,h')}}{1/2}\right).$$

Moreover, the submeasure $\Psi_{hh'}(\lambda)$ is also expressed as

$$\Psi_{hh'}(\lambda)=1-\frac{\lambda 2^{\lambda}}{2^{\lambda}-1}\sum_{(i,j)\in E}\left(G_{ij}^{*(h,h')}+G_{i^\dagger j^\dagger}^{*(h,h')}\right)H_{ij}^{(h,h')}(\lambda),$$

where

$$H_{ij}^{(h,h')}(\lambda) = \frac{1}{\lambda}\left[1 - \left(G_{ij}^{c(h,h')}\right)^{\lambda+1} - \left(G_{i^\dagger j^\dagger}^{c(h,h')}\right)^{\lambda+1}\right],$$

and the value at $\lambda = 0$ is taken to be continuous limit as $\lambda \to 0$. Namely

$$\Psi_{hh'}(0) = 1 - \frac{1}{\log 2}\sum_{(i,j)\in E}\left(G_{ij}^{*(h,h')} + G_{i^\dagger j^\dagger}^{*(h,h')}\right)H_{ij}^{(h,h')}(0),$$

$$H_{ij}^{(h,h')}(0) = -G_{ij}^{c(h,h')}\log G_{ij}^{c(h,h')} - G_{i^\dagger j^\dagger}^{c(h,h')}\log G_{i^\dagger j^\dagger}^{c(h,h')}.$$

Note that $H_{ij}^{(h,h')}(\lambda)$ is Patil and Taillie's [9] diversity index of degree λ for $\left\{G_{ij}^{c(h,h')}\right\}$ and $\left\{G_{i^\dagger j^\dagger}^{c(h,h')}\right\}$, which includes the Shannon entropy (when $\lambda = 0$) in a special case.

We note that for all $h = 1,\cdots,\left[\frac{r-1}{2}\right]$ and $\lambda > -1$, (i) $0 \le H_{ij}^{(h,h')}(\lambda) \le \left(2^\lambda - 1\right)/\lambda 2^\lambda$, (ii) $H_{ij}^{(h,h')}(\lambda) = 0$ if and only if $G_{ij}^{c(h,h')} = 1$ (then $G_{i^\dagger j^\dagger}^{c(h,h')} = 0$) or $G_{i^\dagger j^\dagger}^{c(h,h')} = 1$ (then $G_{ij}^{c(h,h')} = 0$), and (iii) $H_{ij}^{(h,h')}(\lambda) = \left(2^\lambda - 1\right)/\lambda 2^\lambda$ if and if only if $G_{ij}^{c(h,h')} = G_{i^\dagger j^\dagger}^{c(h,h')} = 1/2$, that is, $G_{ij}^{(h,h')} = G_{i^\dagger j^\dagger}^{(h,h')}$.

We see that the measure $\Psi(\lambda)$ lies between 0 and 1. Also the submeasures $\Psi_{hh'}(\lambda)$ lie between 0 and 1 for $h = 1,\cdots,\left[\frac{r-1}{2}\right]$. For each $\lambda(>-1)$, there is the structure of CoPS if and only if $\Psi(\lambda) = 0$; and the degree of departure from CoPS is the largest, in the sense that $G_{ij}^{c(h,h')} = 1$ (then $G_{i^\dagger j^\dagger}^{c(h,h')} = 0$) or $G_{i^\dagger j^\dagger}^{c(h,h')} = 1$ (then $G_{ij}^{c(h,h')} = 0$) for $(i,j) \in E$ and $h = 1,\cdots,\left[\frac{r-1}{2}\right]$ if and only if $\Psi(\lambda) = 1$.

3. Approximate Confidence Interval for Measure

Let n_{ij} denote the observed frequency in *i*th row and *j*th column of the table $(i = 1,\cdots,r; j = 1,\cdots,r)$. The sample version of $\Psi(\lambda)$, that is, $\hat{\Psi}(\lambda)$, is given by $\Psi(\lambda)$ with $\{p_{ij}\}$ replaced by $\{\hat{p}_{ij}\}$, where $\hat{p}_{ij} = n_{ij}/n$ and $n = \sum\sum n_{ij}$. We assume that $\{n_{ij}\}$ result from full multinomial sampling. We consider an approximate standard error for $\hat{\Psi}(\lambda)$ and a large-sample confidence interval for $\Psi(\lambda)$. The term $\sqrt{n}\left(\hat{\Psi}(\lambda) - \Psi(\lambda)\right)$ has asymptotically (as $n \to \infty$) a normal distribution with mean zero and variance $\sigma^2[\Psi(\lambda)]$ by using the delta method. See **Appendix** for the details of $\sigma^2[\Psi(\lambda)]$.

Let $\hat{\sigma}^2[\Psi(\lambda)]$ denote $\sigma^2[\Psi(\lambda)]$ with $\{p_{ij}\}$ replaced by $\{\hat{p}_{ij}\}$. Then $\hat{\sigma}[\Psi(\lambda)]/\sqrt{n}$ is an estimated approximate standard error for $\hat{\Psi}(\lambda)$, and $\hat{\Psi}(\lambda) \pm z_{p/2}\hat{\sigma}[\Psi(\lambda)]/\sqrt{n}$ is an approximate $100(1-p)$ percent confidence interval for $\Psi(\lambda)$, where $z_{p/2}$ is the percentage point from the standard normal distribution corresponding to a two-tail probability equal to *p*.

4. Examples

Consider the data in Table 1(a) and Table 1(b) again. From Table 4(a) and Table 4(b), since the confidence intervals for $\Psi(\lambda)$ applied to the data in each of Table 1(a) and Table 1(b) do not include zero for all λ, these would indicate that there is not a structure of CoPS in each table. When the degrees of departure from CoPS in Table 1(a) and Table 1(b) are compared using the confidence interval for $\Psi(\lambda)$, it is greater for Table 1(a) than for Table 1(b).

We further analyze the data in Table 1(a) and Table 1(b) using submeasures

Table 4. Estimate of measure $\Psi(\lambda)$, approximate standard error for $\hat{\Psi}(\lambda)$ and approximate 95% confidence interval for $\Psi(\lambda)$, applied to Table 1(a) and Table 1(b).

(a)

Value of λ	Estimated measure	Standard error	Confidence interval
−0.5	0.222	0.014	(0.195, 0.248)
0	0.338	0.019	(0.301, 0.374)
1.0	0.421	0.021	(0.381, 0.461)
1.5	0.427	0.021	(0.386, 0.467)

(b)

Value of λ	Estimated measure	Standard error	Confidence interval
−0.5	0.147	0.013	(0.122, 0.173)
0	0.225	0.018	(0.189, 0.260)
1.0	0.282	0.021	(0.241, 0.322)
1.5	0.286	0.021	(0.245, 0.327)

Table 5. Estimate of submeasures $\{\Psi_{hh'}(\lambda)\}$ applied to Table 1(a) and Table 1(b).

(a)

Submeasure	Value of λ	Estimated submeasure	Confidence interval
$\hat{\Psi}_{14}(\lambda)$	−0.5	0.299	(0.260, 0.338)
	0	0.446	(0.396, 0.496)
	1.0	0.545	(0.492, 0.598)
	1.5	0.551	(0.498, 0.604)
$\hat{\Psi}_{23}(\lambda)$	−0.5	0.145	(0.124, 0.166)
	0	0.229	(0.198, 0.261)
	1.0	0.297	(0.259, 0.335)
	1.5	0.302	(0.264, 0.340)

(b)

Submeasure	Value of λ	Estimated submeasure	Confidence interval
$\hat{\Psi}_{14}(\lambda)$	−0.5	0.190	(0.150, 0.229)
	0	0.284	(0.232, 0.336)
	1.0	0.352	(0.292, 0.411)
	1.5	0.357	(0.297, 0.416)
$\hat{\Psi}_{23}(\lambda)$	−0.5	0.105	(0.085, 0.126)
	0	0.165	(0.136, 0.195)
	1.0	0.212	(0.177, 0.247)
	1.5	0.215	(0.180, 0.251)

$\Psi_{hh'}(\lambda)$. We see from **Table 5(a)** that for **Table 1(a)**, the degree of departure from point-symmetry in the collapsed table T_{23} is smaller than that in T_{14}. Thus it is seen that (i) when we combine the categories (2), (3) and (4) in **Table 1(a)**, the degree of departure from point-symmetry for collapsed table T_{14} is large, and (ii) when we combine the categories (1) and (2), and combine (4) and (5) in **Table 1(a)**, that for the collapsed table T_{23} is less than the case of (i). Similarly, we see from **Table 5(b)** that for **Table 1(b)**, the degree of departure from point-symmetry in the collapsed table T_{23} is smaller than that in T_{14}. Thus it is seen that (i) when we combine the categories (2), (3) and (4) in **Table 1(b)**, the degree of departure from point-symmetry for collapsed table T_{14} is large, and (ii) when we combine the categories (1) and (2), and combine (4) and (5) in **Table 1(b)**, that for the collapsed table T_{23} is less than the case of (i).

5. Conclusions

When the CoPS model does not hold for the original 5 × 5 table, we are interested in (i) seeing what degree the departure from point-symmetry is for each of tables T_{14} and T_{23}, (ii) seeing for which table of T_{14} and T_{23} the degree of departure from point-symmetry is larger, and (iii) seeing what degree the departure from CoPS is for the original 5 × 5 table. For (i) and (ii), the proposed $\{\Psi_{hh'}(\lambda)\}$ are useful, and for (iii) the proposed measure $\Psi(\lambda)$ is useful.

Since the collapsed tables are obtained by combing adjacent categories, it is meaning to consider collapsed 3 × 3 tables only when an original square contingency table has ordered categories. Therefore, a measure for CoPS in square ordinal tables should depend on the order of listing the categories. We note that it does not matter whichever submeasures for the collapsed tables are invariant or not invariant, because each collapsed 3 × 3 table obtained from an original square table is unique.

In addition, the measure $\Psi(\lambda)$ is expressed by using same weights $1\Big/\left[\frac{r-1}{2}\right]$ for submeasures $\{\Psi_{hh'}(\lambda)\}$. It seems useful to analyze an original square contingency table using the measure $\Psi(\lambda)$ when we cannot decide which collapsed 3×3 table is important.

Acknowledgements

The authors would like to thank the referee for their helpful comments.

Conflicts of Interest

The authors declare no conflicts of interest regarding the publication of this paper.

References

[1] Wall, K.D. and Lienert, G.A. (1976) A Test for Point-Symmetry in J-Dimensional Contingency-Cubes. *Biometrical Journal*, **18**, 259-264.

[2] Yamamoto, K., Murakami, S. and Tomizawa, S. (2013) Point-Symmetry Models and Decomposition for Collapsed Square Contingency Tables. *Journal of Applied Statistics*, **40**, 1446-1452. https://doi.org/10.1080/02664763.2013.786028

[3] Tomizawa, S., Yamamoto, K. and Tahata, K. (2007) An Entropy Measure of Departure from Point-Symmetry for Two-Way Contingency Tables. *Symmetry*: *Culture and Science*, **18**, 279-297.

[4] Hashimoto, K. (1999) Class Structure in Modern Japan: Theory, Method and Quantitative Analysis. Toshindo Press, Tokyo. (In Japanese)

[5] Iki, K., Okada, M. and Tomizawa, S. (2018) An Extended Bivariate T-Distribution Type Symmetry Model for Square Contingency Tables. *Open Journal of Statistics*, **8**, 249-257. https://doi.org/10.4236/ojs.2018.82015

[6] Balcha, A. (2020) Curve Fitting and Least Square Analysis to Extrapolate for the Case of COVID-19 Status in Ethiopia. *Advances in Infectious Diseases*, **10**, 143-159. https://doi.org/10.4236/aid.2020.103015

[7] Cressie, N. and Read, T.R.C. (1984) Multinomial Goodness-of-Fit Tests. *Journal of the Royal Statistical Society*, *Series B*, **46**, 440-464. https://doi.org/10.1111/j.2517-6161.1984.tb01318.x

[8] Read, T.R.C. and Cressie, N. (1988) Goodness-of-Fit Statistics for Discrete Multivariate Data. Springer-Verlag, New York. https://doi.org/10.1007/978-1-4612-4578-0

[9] Patil, G.P. and Taille, C. (1982) Diversity as a Concept and Its Measurement. *Journal of the American Statistical Association*, **77**, 548-561. https://doi.org/10.1080/01621459.1982.10477845

Appendix

Using the delta method, $\sqrt{n}\left(\hat{\Psi}(\lambda)-\Psi(\lambda)\right)$ has asymptotically variance $\sigma^2\left[\Psi(\lambda)\right]$ as follows:

$$\sigma^2\left[\Psi(\lambda)\right]=\sum_{k=1}^{r}\sum_{l=1}^{r}p_{kl}\left(\frac{1}{\left[\frac{r-1}{2}\right]}\sum_{h=1}^{\left[\frac{r-1}{2}\right]}\Delta_{kl}^{(h,h')}(\lambda)\right)^2,$$

where

$$\Delta_{kl}^{(h,h')}(\lambda)=\frac{2^\lambda}{1-2^\lambda}\frac{1}{\delta_{hh'}}A_{kl}^{(h,h')}(\lambda)+\frac{1-\Psi_{hh'}(\lambda)}{\delta_{hh'}}B_{kl}^{(h,h')}(\lambda),$$

$$\begin{aligned}A_{kl}^{(h,h')}(\lambda)=\sum_{(i,j)\in E}\Bigg[&C_{kl(ij)}\left\{1-\left(G_{ij}^{c(h,h')}\right)^\lambda-\lambda G_{i^\dagger j^\dagger}^{c(h,h')}\left(\left(G_{ij}^{c(h,h')}\right)^\lambda-\left(G_{i^\dagger j^\dagger}^{c(h,h')}\right)^\lambda\right)\right\}\\&+D_{kl(ij)}\left\{1-\left(G_{i^\dagger j^\dagger}^{c(h,h')}\right)^\lambda-\lambda G_{ij}^{c(h,h')}\left(\left(G_{i^\dagger j^\dagger}^{c(h,h')}\right)^\lambda-\left(G_{ij}^{c(h,h')}\right)^\lambda\right)\right\}\Bigg],\end{aligned}$$

$$B_{kl}^{(h,h')}(\lambda)=1-I(h+1\le k\le h')\cdot I(h+1\le l\le h'),$$

$$C_{kl(ij)}=\begin{cases}I(k\le h)\cdot I(l\le h) & (i,j)=(1,1),\\ I(k\le h)\cdot I(h+1\le l\le h') & (i,j)=(1,2),\\ I(k\le h)\cdot I(h'+1\le l) & (i,j)=(1,3),\\ I(h+1\le k\le h')\cdot I(l\le h) & (i,j)=(2,1),\end{cases}$$

$$D_{kl(ij)}=\begin{cases}I(h'+1\le k)\cdot I(h'+1\le l) & (i,j)=(1,1),\\ I(h'+1\le k)\cdot I(h+1\le l\le h') & (i,j)=(1,2),\\ I(h'+1\le k)\cdot I(l\le h) & (i,j)=(1,3),\\ I(h+1\le k\le h')\cdot I(h'+1\le l) & (i,j)=(2,1),\end{cases}$$

and $I(\cdot)$ is the indicator function, $I(\cdot)=1$ if true, 0 if not.