

PREDICTION-CFA PADA CFA REGIONAL

Resa Septiani Pontoh

Jurusan Statistika, FMIPA, Universitas Padjadjaran

Email: resa.pontoh@yahoo.com

Abstrak

Banyak peneliti menggunakan data kategori pada penelitiannya. Salah satu metode yang digunakan adalah Analisis Konfigurasi Frekuensi. Pada analisis tersebut nantinya akan dilihat suatu pola konfigurasi apakah nilainya akan lebih besar dari yang diharapkan (*type*) atau lebih sedikit dari yang diharapkan (*antitype*). Analisis Konfigurasi Frekuensi adalah metode *nonparametric* yang digunakan untuk menunjukkan nilai suatu tanda/pola yang berbeda dari frekwensi yang ada pada analisis *multivariate* (Lienert, 1969; Krauth and Linert, 1973, 1982; Lienert dan Oeveste, 1985). Suatu pola yang terjadi lebih sering dari yang telah diekspektasikan disebut dengan *type* dan yang terjadi lebih jarang dari yang diekspektasikan disebut dengan *antitype*. Pada analisis variabel kategori, *covariates* dapat dijelaskan bersamaan dengan variabel yang terstratifikasi (Graham, 1995) serta dalam contoh lain, bisa juga dikatakan bahwa *covariates* adalah juga variabel kontinu. Penggunaan *covariates* biasanya menghasilkan frekwensi estimasi dari suatu sel yang nilainya berdekatan dengan frekwensi yang diobservasi karena lebih banyak informasi yang digunakan dalam proses estimasi (Von Eye dan Niedermeier, 1999). Penelitian ini akan menggunakan metode Analisis Konfigurasi Frekuensi yang akan diaplikasikan pada *categorical covariates* dan dilakukan pada data yang terstratifikasi.

Kata kunci: *Prediction-CFA*, CFA Regional, Data Kategori

PENDAHULUAN

Configural Frequency Analysis adalah metode nonparametric yang digunakan untuk menunjukkan nilai suatu tanda/pola yang berbeda dari frekwensi yang ada pada analisis *multivariate* (Lienert, 1969; Krauth and Linert, 1973, 1982; Lienert dan Oeveste, 1985). Suatu pola yang terjadi lebih sering dari yang telah diekspektasikan disebut dengan *type* dan yang terjadi lebih jarang dari yang diekspektasikan disebut dengan *antitype*.

Pada CFA, terdapat dua model yang sering digunakan yaitu Model CFA Global dan Model CFA Regional. Model CFA Regional tentu saja berbeda dengan model CFA global. Hal fundamental yang terlihat adalah bahwa Model CFA Global menganggap semua variabel berstatus sama dalam arti pada model tersebut tidak dibedakan antara prediktor dan kriteria sedangkan Model CFA Regional menunjukkan perbedaan status beberapa variabel yang ada.

Pada pengaplikasiannya, Model CFA Regional ini sering sekali terjadi jika dibandingkan dengan Model CFA Global. Pengelompokan variabel sering terjadi bahkan pengelompokan tersebut biasanya dibedakan antara variabel prediktor dan variabel kriteria. Salah satu metode yang sering digunakan pada model CFA Regional adalah Prediction CFA. Metode ini akan menganalisis kelompok variabel yang ada apakah terdapat perbedaan secara signifikan antara frekuensi hasil pengamatan (*observed frequencies*) dengan frekuensi yang diharapkan (*expected frequencies*).

Peneliti sangat tertarik untuk membahas lebih lanjut mengenai metode *Prediction-CFA* pada model CFA Regional ini. Oleh karena itu, penelitian ini akan melihat bagaimana tahapan penggunaan metode Prediction-CFA dan bagaimana pengaplikasiannya pada data riil.

PEMBAHASAN

Dalam Configural Frequency Analysis, terdapat dua metode yang sering digunakan oleh para peneliti yaitu CFA Global dan CFA Regional. Kedua metode ini memiliki tahapan yang menarik untuk dikaji pada CFA Regional adalah Metode *Prediction-CFA* karna memiliki cara pengelompokan yang berbeda dari CFA Global.

Seperti yang telah dijelaskan sebelumnya bahwa *Configural frequency Analysis* adalah suatu metode yang digunakan untuk mengidentifikasi pola (konfigurasi) dari variabel kategori apakah terjadi ketidakcocokan (*discrepancies*) dengan apa yang telah diekspektasikan sebelumnya. *Prediction-CFA* juga melihat adanya *Discrepancies* tersebut. *Discrepancies* ini terjadi jika :

1. Suatu peristiwa lebih sering terjadi atau jumlah peristiwa yang terjadi lebih besar dari yang diharapkan atau diekspektasikan (*CFA type*), dan
2. Suatu peristiwa lebih jarang terjadi atau jumlah peristiwa yang terjadi lebih kecil dari yang diharapkan atau diekspektasikan (*CFA antitype*).

Jenis data yang digunakan dalam *CFA* adalah pasangan kategori. Hal ini didasarkan atas pengertian dari konfigurasi (Lienert, 1969) yaitu pasangan kategori yang menjelaskan suatu sel dari suatu tabel silang. Dengan kata lain, maka *Prediction-CFA* pun adalah merupakan pasangan kategori yang nantinya akan dijelaskan dalam bentuk tabel silang.

CFA maupun *Prediction-CFA* dapat dapat dipergunakan untuk menjawab beberapa pertanyaan di bawah ini:

1. Bagaimana cara membandingkan antara frekuensi hasil pengamatan (*observed frequencies*) dengan frekuensi yang diharapkan (*expected frequencies*)?

Jawaban dari pertanyaan di atas adalah bergantung pada pemilihan *base model* yang akan digunakan untuk melihat perbandingan nilai dari frekuensi-frekuensi tersebut dimana nantinya akan diperoleh interpretasi mengenai ada tidaknya perbedaan antara frekuensi pengamatan suatu sel dan frekuensi harapan sel tersebut. Frekuensi harapan suatu sel ditaksir berdasarkan *base model* yang ditetapkan untuk menggambarkan hubungan diantara variabel. Dari kedua hal di atas, akan terlihat adanya suatu perbedaan nilai antara frekuensi pengamatan suatu sel dan frekuensi harapan sel tersebut apakah frekuensi pengamatan suatu sel lebih besar ataukah lebih kecil atau bahkan sama dengan nilai dari frekuensi harapan sel tersebut. Perbedaan antara frekuensi pengamatan suatu sel dan frekuensi harapan sel dari kumpulan kelompok variabel yang ada tersebut yang nantinya akan diteliti lebih lanjut dengan menggunakan *Prediction-CFA*.

2. Apakah perbedaan yang terjadi antara frekuensi pengamatan suatu sel dan frekuensi harapan sel tersebut secara statistik bersifat signifikan?

Dalam beberapa penelitian yang telah dilakukan dengan menggunakan *CFA*, jarang sekali terjadi adanya kesamaan hasil antara frekuensi pengamatan suatu sel dan frekuensi harapan sel tersebut. Dalam *CFA* maupun *Prediction CFA*, perbedaan nilai antara frekuensi pengamatan suatu sel dan frekuensi harapan sel tersebut dijelaskan dengan munculnya suatu *type* dan *antitype*. Jika frekuensi pengamatan suatu sel lebih besar dari pada frekuensi harapan sel tersebut, maka akan muncul suatu *type* namun jika frekuensi pengamatan suatu sel lebih kecil dari pada frekuensi harapan sel tersebut, maka akan muncul suatu *antitype*. Munculnya *type* dan *antitype* akan melalui suatu proses pengujian statistik tertentu.

3. Apakah terdapat hubungan antara *predictor* (variabel bebas) dan *criterion* (variabel tak bebas)?

Dalam *CFA*, suatu model dapat menjelaskan apakah dalam model tersebut variabel-variabel terbagi menjadi prediktor dan kriteria ataukah semua variabel dianggap mempunyai status yang sama. Berkaitan dengan pertanyaan di atas, *CFA* dapat menjelaskan

ada atau tidaknya hubungan antara prediktor dan kriteria yang ditandai dengan munculnya suatu *type* dan *antitype*. Jika *type* dan *antitype* muncul artinya bahwa suatu prediktor dapat memprediksi terjadinya suatu kriteria tertentu. Metode Prediction-CFA akan digunakan pada bahasan ini.

Hal yang perlu diketahui adalah bahwa P-CFA tidak melihat apakah variabel-variabel yang ada berhubungan antara yang satu dengan yang lainnya, interaksi, ataupun prediksi satu dengan lainnya. P-CFA melihat pola tertentu pada suatu kategori di variabel prediktor akan memungkinkan memprediksi apakah munculnya pola variabel kriteria tertentu berada di bawah atau di atas angka harapan yang telah diekspektasikan. Pada metode CFA yang digunakan biasanya tidak dibedakan status dari variabel yang ada. Pada P-CFA, variabel yang ada akan dibedakan menjadi prediktor dan kriteria. Pada P-CFA, base model yang ada akan memenuhi kriteria sebagai berikut (Lienert dan Krauth;1973):

1. Untuk menghindari munculnya *type* dan *antitype* karena disebabkan adanya hubungan antar prediktor, maka base model yang ada akan disaturasikan pada prediktor
2. Untuk menghindari munculnya *type* dan *antitype* karena disebabkan adanya hubungan antar kriteria, maka base model yang ada akan disaturasikan pada kriteria
3. Prediktor dan kriteria saling independen

Perlu diketahui bahwa P-CFA mengasumsikan suatu *base model* tidak dapat menjelaskan data dengan baik. Karena itu, parameter bukanlah fokus dari pengujian P-CFA, tetapi yang difokuskan dalam P-CFA adalah penyimpangan yang terjadi pada model ditandai dengan munculnya *type* dan *antitype* artinya bahwa hasil akhir dari P-CFA bukanlah melihat apakah suatu model sudah dapat menjelaskan data dengan baik seperti yang dilakukan dengan menggunakan metode log-linear. Oleh karena itu, *the goodness of fit*, nilai χ^2 yang menjelaskan tentang kecocokan model dengan data, tidak menjadi perhatian dalam P-CFA.

Tahapan yang akan dilakukan pada penelitian ini adalah sebagai berikut:

1. Pemilihan Base Model pada P-CFA
2. Tes untuk melihat Signifikansi
3. Penjabaran hasil pengujian signifikansi dan pengidentifikasian apakah konfigurasi masuk ke dalam *type* atau *antitype*.
4. Penginterpretasian *type* dan *antitype*.

Pemilihan Base Model pada P-CFA

Pada tahapan ini, terdapat tiga hal yang perlu diperhatikan yaitu base model untuk prediktor, base model untuk kriteria dan base model yang memperlihatkan hubungan antara prediktor dan kriteria yang berhubungan dengan munculnya *type* dan *antitype*.

Base model pada P-CFA biasanya akan dilakukan saturasi pada prediktor. Hal ini dilakukan untuk menghindari kemungkinan hasil yang diperoleh dari munculnya *type* dan *antitype* adalah karena adanya hubungan antar prediktor. Berbeda dengan kriteria pada base model P-CFA adalah seandainya peneliti mengasumsikan adanya hubungan yang terjadi antar kriteria disebabkan oleh adanya prediktor, maka pada hubungan yang terjadi tersebut akan memungkinkan munculnya *type* dan *antitype*. Namun jika tidak diinginkan adanya hubungan antara kriteria, maka akan dilakukan saturasi pada kriteria. Aspek terakhir yang diperhatikan adalah hubungan antar prediktor dan kriteria. Biasanya pada base model diasumsikan tidak ada hubungan antar keduanya atau dengan kata lain prediktor dan kriteria saling independen.

Secara garis besar, base model pada P-CFA pada lima variabel (dua prediktor dan tiga kriteria adalah sebagai berikut:

$$\text{Log } E = \lambda_0 + \lambda_i^{P_1} + \lambda_j^P + \lambda_{ij}^{P_1 P_2} + \lambda_k^{C_1} + \lambda_l^{C_2} + \lambda_m^{C_3} + \lambda_{kl}^{C_1 C_2} + \lambda_{km}^{C_1 C_3} + \lambda_{lm}^{C_2 C_3} + \lambda_{klm}^{C_1 C_2 C_3}$$

Pada model tersebut, type dan antitype akan muncul jika konfigurasi pada kelompok prediktor terjadi berurutan dengan konfigurasi pada kelompok kriteria lebih sering atau jarang dari apa yang telah diekspetasikan.

Penelitian ini kemudian akan melanjutkan dan memperdalam bagaimana tahapan penggunaan metode *Prediction-CFA* pada model CFA Regional dan akan diaplikasikan pada data riil dalam bidang kesehatan untuk melihat bagaimana beragam jenis pola tidur seseorang berkaitan dengan konsentrasi kerjanya. Dalam pengerjaannya akan digunakan software *Configural Frequency Analysis – Version 2000* (von Eye, A. : 2001)

Perlu diketahui bahwa *CFA* mengasumsikan suatu *base model* tidak dapat menjelaskan data dengan baik. Karena itu, parameter bukanlah fokus dari pengujian *CFA*, tetapi yang difokuskan dalam *CFA* adalah penyimpangan yang terjadi pada model ditandai dengan munculnya *type* dan *antitype* artinya bahwa hasil akhir dari *CFA* bukanlah melihat apakah suatu model sudah dapat menjelaskan data dengan baik seperti yang dilakukan dengan menggunakan metode log-linear. Oleh karena itu, *the goodness of fit*, nilai χ^2 yang menjelaskan tentang kecocokan model dengan data, tidak menjadi perhatian dalam *CFA*. *Maksimum likelihood estimator* dari $E(Y_i)$ adalah

$$n\hat{\theta}_i = Y_i \text{ (Dobson, 1982)}$$

Hal lain yang menjadi perbedaan antara *CFA* dan Log-linear adalah bahwa *CFA* tidak hanya melihat adanya hubungan antara variabel-variabel tertentu yang biasanya menjadi *output* pada model Log-linear tetapi juga lebih memperhatikan pada konfigurasi-konfigurasi mana yang saling berkaitan satu sama lain.

Untuk melindungi signifikansi nominal α terhadap kesalahan pengujian dikarenakan nilai α untuk tiap konfigurasi berbeda dengan α keseluruhan, penyesuaian dapat dilakukan dengan memperhitungkan total jumlah tampilan pengujian atau banyaknya konfigurasi yang terjadi yaitu dengan rumusan sebagai berikut:

$$\alpha^* = \alpha / \text{banyaknya konfigurasi}$$

Jika statistika hitung lebih kecil dari α^* , akan terdapat *type* atau *antitype* pada konfigurasi tersebut.

Hipotesis:

$$H_0: E[N_t] = E_t$$

$$H_1: E[N_t] > E_t \text{ atau } E[N_t] < E_t$$

Statistika uji:

$$z = \frac{N_t - \hat{E}_t}{\sqrt{\hat{E}_t}} \text{ (Von eye, 2002)}$$

kriteria uji:

Jika $z\text{-value} \geq \alpha^*$ maka H_0 diterima atau tidak akan muncul *type* atau *antitype*, dapat dikatakan base model telah mewakili keberadaan dari konfigurasi tersebut, sedangkan jika $z\text{-value} < \alpha^*$ maka H_0 ditolak atau akan muncul *type* dan *antitype*, dengan kata lain model tersebut tidak mewakili keberadaan dari konfigurasi tersebut.

Data yang digunakan pada penelitian ini adalah hasil penelitian yang telah dilakukan oleh Mahoney pada tahun 2000 dengan melakukan metode *cluster-analitic* untuk membuat empat buah klaster dari kelompok usia remaja. Klaster yang terbentuk adalah kelompok remaja baik putra maupun putri dengan ciri sebagai berikut :

1. Klaster 1
Klaster satu adalah remaja yang memiliki kompetensi yang baik dalam bidang akademik, kurang menunjukkan kedewasaan fisik, lebih muda, lemah dalam pertarungan, cukup populer, dan memiliki tingkat SES yang tinggi
2. Klaster 2
Klaster satu adalah remaja yang memiliki kompetensi yang baik dalam bidang akademik, mempunyai perkembangan kedewasaan fisik secara normal , lebih muda, lemah dalam pertarungan, cukup populer, dan memiliki tingkat SES yang rendah
3. Klaster 3
Klaster satu adalah remaja yang memiliki kompetensi yang rendah dalam bidang akademik, kurang menunjukkan kedewasaan fisik, lebih muda, cukup memiliki keingan yang kuat untuk pertarungan, kurang populer, dan memiliki tingkat SES yang rendah
4. Klater 4
Lebih tua dari teman sebayanya, memiliki keinginan bertarung yang kuat, kompetensi yang rendah dalam bidang akademik, kurang populer, dan memiliki tingkat SES yang rendah

Pada studi ini, klaster 1 dan 2 sangat diperhatikan karena memiliki karakteristik yang hampir mirip. Dimana dalam penelitian ini terdapat variabel yang terbebtuk kadalah sebagai berikut:

1. Pattern (P) adalah merupakan keempat klaster yang terbentuk dengan pengkategorian sebagai berikut: 1. Klaster 1 dan 2; 2. Klaster 3; 3. Klaster 4.
2. Gender (G) merupakan jenis kelamin dari responden : 1. Putra; 2. Putri
3. Status dikeluarkan dari sekolah (S) : 1. Tidak; 2. Ya
4. Terkena kasus kriminalitas (A) : 1. Tidak; 2. Ya

Base model yang akan digunakan pada P-CFA ini adalah [P,G][S][A]. Pada *base model* ini, baik type dan antitype akan muncul dari hubungan antara prediktor dan kriteria [P,G], maupun hubungan asosiasi dari kedua variablel krieria lainnya yaitu [S] dan [A]. Hasil yang diperoleh adalah sebagai berikut:

Tabel 4.1 Hasil Analisis P-CFA Mahoney-2000 mengenai perilaku remaja (Von Eye, 2002)

Cell Index	Observed Frequencies	P-CFA model [P,G][S][A]	
		expected	p(z)
GPSA			
1111	155	127.79	0.008
1112	9	17.01	0.0261
1121	6	24.89	0.0001 A
1122	3	3.31	0.4318
1211	63	67.96	0.2738
1212	10	9.05	0.3756
1221	11	13.23	0.2696
1222	8	1.76	<a* T
1311	26	44.32	0.003
1312	8	5.9	0.1936
1321	13	8.63	0.0685
1322	13	1.15	<a* T

2111	188	152.91	0.0023
2112	7	20.35	0.0015 A
2121	12	29.78	0.0005 A
2122	0	3.96	0.0232
/cont.			
Cell Index	Observed Frequencies	P-CFA model	
		[P,G][S][A]	
GPSA		expected	p(z)
2211	76	70.91	0.2729
2212	8	9.44	0.3197
2221	6	13.81	0.0178
222	6	1.84	0.0010 T
2311	20	39.15	0.0011 A
2312	0	5.21	0.0112
2321	25	7.62	<a* T
2322	8	1.02	<a* T

Dari perhitungan yang telah dilakukan, diperoleh nilai likelihood ratio χ^2 adalah 212.74 (df =16; $p < 0,01$). Nilai Bonferroni-adjusted α^* adalah 0,0020833. Hasil yang diperoleh dari tabel 4.1 memperlihatkan bahwa Type muncul pada konfigurasi 1222, 1322, 222, 2321, dan 2322 sedangkan Antitype muncul pada konfigurasi 1121, 2112, 2121, dan 2311.

Antitype pada konfigurasi yang terbentuk memperlihatkan bahwa apa yang terjadi jauh lebih rendah dari apa yang telah diekspektasikan. Konfigurasi Antitype yang terbentuk adalah sebagai berikut:

1. Konfigurasi 1121 memperlihatkan bahwa remaja putra dari klaster 1 maupun klaster 2 yang dikeluarkan dari sekolah tidak melakukan tindak kejahatan.
2. Konfigurasi 2112 mengindikasikan bahwa jumlah remaja putri dari klaster 1 dan 2 tetap bersekolah walaupun melakukan tindak kejahatan
3. Konfigurasi 2121 menginformasikan bahwa remaja putri pada klaster 1 dan 2 yang dikeluarkan dari sekolah tanpa melakukan tindak kejahatan.
4. Konfigurasi 2311 menjelaskan remaja perempuan pada klaster keempat yang bersekolah tetapi tidak melakukan tindak kejahatan

Type pada konfigurasi yang terbentuk memperlihatkan bahwa apa yang terjadi melampaui dari apa yang telah diekspektasikan. Konfigurasi Type yang terbentuk adalah sebagai berikut:

1. Konfigurasi 1222 yang mengindikasikan bahwa remaja putra dari klaster tiga yang dikeluarkan dari sekolah juga memiliki kasus kriminal.
2. Konfigurasi 1322 menginformasikan bahwa remaja putra dari klaster empat yang dikeluarkan dari sekolah juga memiliki kasus kriminal.
3. Konfigurasi 2222 menginformasikan bahwa remaja putri dari klaster tiga yang dikeluarkan dari sekolah juga memiliki kasus kriminal.
4. Konfigurasi 2321 menginformasikan bahwa remaja putri dari klaster empat yang dikeluarkan dari sekolah, tidak memiliki kasus kriminal.

DAFTAR PUSTAKA

Von Eye, A. 2002. *Configural Frequency Analysis*, Lawrence Erlbaum Associates, Inc., London.

Von Eye, A. 2001. *Configural Frequency Analysis – Version 2000 A Program for Bit Windows Operating System*, *Methods of Psychological Research Online*, Vol. 6, No. 2, 129-139, internet <http://www.pabst-publishers.de/mpr/>

Dobson, Annette J. 1982. *Introduction to Statistical Modelling*, Chapman and Hall, New York.

Agresti, Alan. 1976. *An Introduction to Categorical Data Analysis*, Jhon Wiley and Sons, Inc., New York.

