

**CONFIGURAL FREQUENCY ANALYSIS UNTUK MELIHAT
KARAKTERISTIK CALON INVESTOR POTENSIAL
PT BURSA EFEK INDONESIA DI JAWA TIMUR**

RESA SEPTIANI PONTOH

*Jurusan Statistika FMIPA Universitas Padjadjaran
Jl. Bandung-Sumedang Km. 21 Jatinangor
Email : echachicha@yahoo.com*

Abstrak

Configural Frequency Analysis (CFA) adalah salah satu metode yang digunakan untuk menganalisis data kategori. Banyak metode yang digunakan dalam mempelajari pola asosiasi antara sekelompok variabel kategori seperti log-linear dan regresi logistik. Pada pengaplikasiannya, kedua metode tersebut mempunyai perbedaan dengan metode CFA. Dalam CFA, parameter bukanlah fokus dari pengujian yang dilakukan serta tidak melihat apakah suatu model sudah dapat menjelaskan data dengan baik seperti yang terdapat pada metode log-linear. 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. Dalam pengaplikasiannya, metode ini akan digunakan untuk mengetahui secara jelas karakteristik dari calon investor potensial PT Bursa Efek Indonesia di daerah Jawa Timur.

Kata Kunci: *Configural Frequency Analysis*, CFA, Data Kategori, Resa Septiani Pontoh

PENDAHULUAN

Salah satu fenomena yang sering kita amati adalah terjadinya suatu kejadian di luar apa yang kita harapkan. Banyak kejadian terjadi tidak sesuai dengan apa yang sudah diprediksikan, seperti misalnya cuaca yang tidak menentu, prediksi jumlah wisatawan yang berbeda jauh dengan jumlah yang sudah diperkirakan, investor pasar modal yang ternyata sering berubah-ubah jumlah dan aktivitasnya, dan sebagainya. Kondisi seperti ini perlu ditangani lebih lanjut apakah fenomena tersebut secara signifikan memang benar adanya serta perlu juga dipikirkan alternatif penyelesaian masalah yang dapat dilakukan. Pendekatan yang sering dilakukan adalah melihat pola asosiasi diantara faktor-faktor yang ada. Data kategori sering digunakan dalam penelitian-penelitian tersebut.

Salah satu fenomena yang menarik untuk dipelajari adalah dunia pasar modal. Seperti yang kita ketahui bahwa pasar modal di Indonesia belum terlalu banyak diminati. Dari data sub-rekening di Kustodian, diketahui bahwa jumlah investor Indonesia saat ini hanya sekitar 80.000 investor. Bandingkan dengan Singapura yang mempunyai penduduk sekitar 3 juta jiwa tetapi mempunyai investor sebanyak 1 juta-an (kurang lebih sebesar 33%). Demikian juga Malaysia, 16% dari populasi penduduknya yang berjumlah sekitar 18 juta jiwa merupakan investor yang cukup aktif di pasar modal. (Herwidayatmo). Peluang pertambahan jumlah investor di Indonesia sebenarnya masih terbuka lebar walaupun tantangan pada saat ini adalah terjadinya krisis global yang tentu saja membuat para pelaku pasar modal harus lebih giat dalam melakukan pengembangan jumlah investor. Di Indonesia, sangat sulit untuk memprediksi jumlah investor saham retail serta sulit untuk melihat karakteristik calon investor saham retail yang potensial.

Untuk keperluan ini, terdapat beberapa alternatif metode yang dapat dipergunakan diantaranya: Regresi Logistik, Log-Linear, dan CFA. Dari ketiga metode tersebut, CFA merupakan alternatif yang menarik untuk diaplikasikan karena metode ini dapat melihat apakah suatu kejadian lebih sering atau jarang terjadi dari apa yang telah diprediksikan.

METODE PENELITIAN

Dalam bagian ini akan dibahas tentang Configural Frequency Analysis (CFA) kaitannya dengan keinginan PT Bursa Efek Indonesia untuk melakukan mapping karakteristik calon investor saham retail. Untuk penganalisisan data, digunakan data survey yang telah dilakukan oleh PT Bursa Efek Indonesia pada wilayah Jawa Timur.

Data yang digunakan adalah data survey tiga wilayah yaitu Jawa Barat sebanyak 457 responden, Jawa Tengah sebanyak 566 responden dan Jawa Timur sebanyak 605 responden yang berbentuk kategori dikotomus. Pada penelitian untuk melihat karakteristik calon investor retail terhadap bentuk investasi saham yang dilakukan oleh PT Bursa Efek Jakarta pada tahun 2004, terdapat 4 faktor yang akan dilihat keterkaitannya yaitu:

1. Ketertarikan Responden terhadap bentuk investasi saham
Unit pengamatan dibagi menjadi dua kategori yaitu tertarik akan saham (1) dan tidak tertarik akan saham (2).
2. Usia Responden
Usia unit pengamatan diukur dari rentang usia di bawah di bawah 40 tahun (1) dan usia di atas 40 tahun (2).
3. Jenis Pekerjaan
Jenis Pekerjaan dibagi menjadi dua kategori yaitu pegawai negeri (1) dan non pegawai negeri (2).
4. Penghasilan
Penghasilan dibagi menjadi dua kategori yaitu di bawah Rp. 5.000.000,00 (1) dan di atas Rp. 5.000.000,00 (2).

Configural Frequency Analysis

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. *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.

Langkah-Langkah pengujian dengan menggunakan *CFA* (Von eye, 2002)

Dalam pengujian konfigurasi dengan menggunakan *CFA*, terdapat lima langkah yang perlu dilakukan, yaitu sebagai berikut:

1. Pemilihan *base model* untuk *CFA* dan pengestimasian frekuensi harapan dari suatu sel.
2. Pemilihan suatu konsep penyimpangan dari suatu model.
3. Pemilihan tes untuk melihat signifikansi.
4. Penjabaran hasil pengujian signifikansi dan pengidentifikasian apakah konfigurasi masuk ke dalam *type* atau *antitype*.
5. Penginterpretasian *type* dan *antitype*.

Pemilihan base model untuk *CFA* dan pengestimasian frekuensi harapan dari suatu sel.

Log-Linear sering digunakan sebagai *base model* dalam *CFA*, dimana model umum dari log-linear adalah sebagai berikut (Von eye, 2002):

$$\log E = \lambda_0 + \sum_{\text{main effects}} \lambda_i + \sum_{\text{first order interaction}} \lambda_{ij} + \sum_{\text{second order interaction}} \lambda_{ijk} + \dots,$$

dimana λ_0 adalah intercept dan i, j, k adalah indeks variabel.

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.

Dalam CFA, *base model* digunakan untuk merefleksikan asumsi teorikal dari sifat suatu parameter apakah semua variabel mempunyai status yang sama, atautkah terbagi menjadi prediktor dan kriteria. Berikut ini akan dijelaskan beberapa contoh model log-linear yang biasa digunakan.

Jika tidak ada variabel yang mempengaruhi model (*zero-order*), model log-linear secara umum (Von eye, 2002) adalah sebagai berikut:

$$\text{Log}E(Y) = \hat{\mu}$$

dimana $E(Y)$ = frekuensi diharapkan dalam setiap sel.

$\hat{\mu}$ = *Intercept* atau *constant* atau rata-rata umum

Jika semua variabel mempunyai status yang sama, dan hanya *Main effect* atau efek utama yang digunakan (*first-order*), model log-linear secara umum (Von eye, 2002) adalah sebagai berikut:

$$\text{Log}E(Y_{ij\dots}) = \hat{\mu} + \hat{\alpha}_i + \hat{\beta}_j + \dots$$

dimana $E(Y_{ij\dots})$ = frekuensi diharapkan dalam setiap sel.

$\hat{\mu}$ = *Intercept* atau *constant* atau rata-rata umum

$\hat{\alpha}_i$ = parameter pengaruh tingkat ke-i faktor α

$\hat{\beta}_j$ = parameter pengaruh tingkat ke-j faktor β

Base Model yang digunakan dalam penganalisisan data survey PT Bursa Efek Jakarta adalah dengan menggunakan Metode Log-Linear. Base model pada model Log-Linear yang digunakan dalam penganalisisan data survey yang telah dilakukan oleh PT Bursa Efek Indonesia, terbagi menjadi tiga prediktor (usia, jenis pekerjaan, dan penghasilan) dan satu kriteria (ketertarikan terhadap saham). Karena itu, pada penelitian ini akan melihat karakteristik calon investor saham retail PT Bursa Efek Indonesia dalam tiga variabel yaitu usia, jenis pekerjaan, dan penghasilan dilihat dari ada atau tidaknya ketertarikan responden terhadap saham.

Model Log-Linear untuk data survey tersebut adalah sebagai berikut:

$$\log E(Y_{ijkl}) = \hat{\mu} + \hat{\alpha}_i + \hat{\beta}_j + \hat{\gamma}_k + \hat{\delta}_l + \hat{\alpha}\hat{\beta}_{ij} + \hat{\alpha}\hat{\gamma}_{il} + \hat{\beta}\hat{\gamma}_{jl} + \hat{\alpha}\hat{\beta}\hat{\gamma}_{ijk} \dots\dots (1)$$

dimana $E(Y_{ijkl})$ = frekuensi yang diharapkan dalam setiap sel.

$\hat{\mu}$ = *Intercept* atau *constant* atau parameter rata-rata umum

$\hat{\alpha}_i$ = Parameter pengaruh tingkat ke-i (i = 1 (pegawai negeri) dan i = 2 (non pegawai negeri) faktor α (Jenis Pekerjaan)

$\hat{\beta}_l$ = Parameter pengaruh tingkat ke-j (j = 1 (di bawah 5 juta) dan j = 2 (di atas 5 juta) faktor β (Penghasilan)

$\hat{\gamma}_k$ = Parameter pengaruh tingkat ke-k (k = 1 (usia di bawah 40 tahun) dan k = 2 (Usia di atas 40 tahun) faktor γ (Usia Responden).

$\hat{\delta}_l$ = Parameter pengaruh tingkat ke-l (l = 1 (tertarik) dan l = 2 (tidak tertarik) faktor δ (Ketertarikan terhadap saham)

$\hat{\alpha}\hat{\beta}_{ij}$ = Parameter pengaruh interaksi tingkat ke-i dan ke-j faktor α (Jenis Pekerjaan) dan β (Penghasilan).

$\hat{\alpha}\hat{\gamma}_{il}$ = Parameter pengaruh interaksi tingkat ke-i dan ke-k faktor α (Jenis Pekerjaan) dan γ (usia).

$\hat{\beta}\hat{\gamma}_{jk}$ = Parameter pengaruh interaksi tingkat ke-j dan ke-k faktor β (Penghasilan) dan γ (usia).

$\hat{\alpha}\hat{\beta}\hat{\gamma}_{ijk}$ = Parameter pengaruh interaksi tingkat ke-i, ke-j dan ke-k faktor α (Jenis Pekerjaan), β (Penghasilan) dan γ (usia).

Untuk melihat kebenaran adanya interaksi antar prediktor, dapat dilakukan dengan menggunakan pengujian main effect antar prediktor itu sendiri. Jika terdapat type ataupun antitype, artinya memang benar terdapat interaksi antar prediktor. Model Log-Linear untuk melihat ada atau tidaknya interaksi antar prediktor adalah sebagai berikut:

$$\log E(Y_{ijk}) = \hat{\mu} + \hat{\alpha}_i + \hat{\beta}_j + \hat{\gamma}_k \quad \dots\dots (2)$$

dimana $E(Y_{ijk})$ = frekuensi diharapkan dalam setiap sel.

$\hat{\mu}$ = Intercept atau constant atau parameter rata-rata umum

$\hat{\alpha}_i$ = Parameter pengaruh tingkat ke-i (i = 1 (pegawai negeri) dan i = 2 (non pegawai negeri) faktor α (Jenis Pekerjaan)

$\hat{\beta}_j$ = Parameter pengaruh tingkat ke-j (j = 1 (di bawah 5 juta) dan j = 2 (di atas 5 juta) faktor β (Penghasilan)

$\hat{\gamma}_k$ = Parameter pengaruh tingkat ke-k (k = 1 (usia di bawah 40 tahun) dan k = 2 (Usia di atas 40 tahun) faktor γ (Usia Responden).

Penaksiran terhadap ekspektasi frekuensi konfigurasi ($E(Y_{ijkl})$) digunakan dengan menggunakan metode penaksiran Maksimum likelihood.

Fungsi dari Distribusi Multinomial dengan frekuensi sel Y_1, \dots, Y_N , dengan peluang tiap sel

adalah $\theta_1, \dots, \theta_N$, $n = \sum_{i=1}^N y_i$ dan nilai n telah ditentukan sebelumnya (fixed), adalah sebagai berikut:

$$f(y; \theta|n) = n! \prod_{i=1}^N \theta_i^{y_i} / y_i! \quad \dots\dots(3).$$

Untuk mengestimasi nilai $E(Y_i)$ dengan menggunakan metode maksimum likelihood adalah:

$$L = \log n! + \sum_i (y_i \log \theta_i - \log y_i!)$$

$$L = \text{constant} + \sum_i y_i \log \theta_i \quad \dots\dots(4)$$

dimana $n = \sum_{i=1}^N y_i$ dan $\sum \theta_i = 1$.

Untuk memaksimalkan model, penaksir maksimum likelihood dari parameter θ_i diperoleh dengan memaksimalkan fungsi likelihood dengan constraint nya yaitu $\sum \theta_i = 1$ dan $n = \sum_{i=1}^N y_i$, yang dapat dilakukan dengan menggunakan lagrange multiplier yang akan meminimalkan λ dan θ_i dari persamaan berikut:

$$t = \text{constant} \quad + \sum_i y_i \log \theta_i - \lambda (\sum \theta_i - 1) \quad \dots\dots(5)$$

solusinya adalah $\frac{\partial t}{\partial \lambda} = 0$ dan $\frac{\partial t}{\partial \theta_i} = 0$ untuk $i = 1, \dots, N$ diperoleh:
 $\lambda = n$

dengan mensubstitusi $\lambda = n$ ke dalam Persamaan 4.4 maka diperoleh:

$$t = \text{constant} \quad + \sum_i y_i \log \theta_i - n (\sum \theta_i - 1) \quad \dots\dots(6)$$

akan ditaksir nilai dari parameter θ_i dan diperoleh:

$$\frac{\partial t}{\partial \theta_i} = \frac{y_i}{\theta_i} - n = 0$$

$$\hat{\theta}_i = \frac{Y_i}{n} \quad n \hat{\theta}_i = Y_i$$

Jadi dapat disimpulkan bahwa maksimum likelihood estimator dari $E(Y_i)$ adalah

$$n \hat{\theta}_i = Y_i \quad (\text{Dobson, 1982}) \quad \dots\dots(7)$$

Sehingga $E(Y_{ijkl}) = n \hat{\theta}_{...l} \hat{\theta}_{ijk}$ dan $E(Y_{ijk}) = n \hat{\theta}_{i..} \hat{\theta}_{.j.} \hat{\theta}_{..k}$

Pemilihan suatu konsep penyimpangan dari suatu model.

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.

Samplng scheme pada penganalisisan data survey ini adalah dengan menggunakan multinomial, maka uji hipotesis yang digunakan adalah dengan menggunakan uji independensi yaitu dengan hipotesis sebagai berikut:

$$H_0 : E(Y_{ijkl}) = n \hat{\theta}_{...l} \hat{\theta}_{ijk}$$

$$H_1 : E(Y_{ijkl}) \neq n \hat{\theta}_{...l} \hat{\theta}_{ijk}$$

Diketahui bahwa, $\hat{\theta}_{ijk.} = \frac{y_{ijk.}}{n}$, dan $\hat{\theta}_{...l} = \frac{y_{...l}}{n}$, sehingga

$$e_{ijkl} = \frac{y_{ijk.} \cdot y_{...l}}{n} \quad \text{dimana } i = 1, 2; j = 1, 2; k = 1, 2; l = 1, 2. \quad \dots\dots (8)$$

Model log-linear di bawah H_0 adalah sebagai berikut:

$$\eta_{ijkl} = \log E(Y_{ijkl}) = \hat{\mu} + \hat{\alpha}_i + \hat{\beta}_j + \hat{\gamma}_k + \hat{\delta}_l + \hat{\alpha}\hat{\beta}_{ij} + \hat{\alpha}\hat{\gamma}_{il} + \hat{\beta}\hat{\gamma}_{jl} + \hat{\alpha}\hat{\beta}\hat{\gamma}_{ijk} \quad \dots\dots (9)$$

dengan asumsi bahwa

$$\sum_i \hat{\alpha}_i = \sum_j \hat{\beta}_j = \sum_k \hat{\gamma}_k = \sum_l \hat{\delta}_l = \sum_{i,j} \hat{\alpha}\hat{\beta}_{ij} = \sum_{i,l} \hat{\alpha}\hat{\gamma}_{il} = \sum_{j,l} \hat{\beta}\hat{\gamma}_{jl} = \sum_{i,j,k} \hat{\alpha}\hat{\beta}\hat{\gamma}_{ijk} = 0$$

log eijkl diestimasi oleh η_{ijkl}

$$\log \hat{\theta}_{...l} + \log \hat{\theta}_{ijk.} - \log n = \hat{\mu} + \hat{\alpha}_i + \hat{\beta}_j + \hat{\gamma}_k + \hat{\delta}_l + \hat{\alpha}\hat{\beta}_{ij} + \hat{\alpha}\hat{\gamma}_{il} + \hat{\beta}\hat{\gamma}_{jl} + \hat{\alpha}\hat{\beta}\hat{\gamma}_{ijk} \quad \dots\dots(10)$$

dengan mensubstitusikan $\sum_i \sum_j \sum_k \sum_l$ ke dalam Persamaan 4.10, akan diperoleh estimasi varians minimum yaitu sebagai berikut:

$$\hat{\mu} = \frac{1}{IJK} \sum_k \log y_{ijk.} + \frac{1}{L} \sum_l \log y_{...l} - \log n \quad \dots\dots(11)$$

$$\hat{\alpha}_i = \frac{1}{L} \sum_l \log y_{...l} + \frac{1}{JK} \sum_{j,k} \log y_{ijk.} - \log n - \hat{\mu} \quad \dots\dots (12)$$

Dengan cara yang sama di atas, diperoleh nilai $\hat{\beta}_j, \hat{\gamma}_k, \hat{\delta}_l$ sebagai berikut:

$$\hat{\beta}_j = \frac{1}{L} \sum_l \log y_{...l} + \frac{1}{IK} \sum_{i,k} \log y_{ijk.} - \log n - \hat{\mu} \quad \dots\dots (13)$$

$$\hat{\gamma}_k = \frac{1}{L} \sum_l \log y_{...l} + \frac{1}{IJ} \sum_{i,j} \log y_{ijk.} - \log n - \hat{\mu} \quad \dots\dots (14)$$

$$\hat{\delta}_l = \log y_{...l} + \frac{1}{IJK} \sum_{i,j,k} \log y_{ijk.} - \log n - \hat{\mu} \quad \dots\dots (15)$$

$$\hat{\alpha}\hat{\beta}_{ij} = \frac{1}{L} \log y_{...l} + \frac{1}{K} \sum_k \log y_{ijk.} - \log n - \hat{\mu} - \hat{\gamma} - \hat{\delta} \quad \dots\dots (16)$$

$$\hat{\alpha}\hat{\gamma}_{ik} = \frac{1}{L} \log y_{...l} + \frac{1}{J} \sum_j \log y_{ijk.} - \log n - \hat{\mu} - \hat{\beta} - \hat{\delta} \quad \dots\dots (17)$$

$$\hat{\beta}\hat{\gamma}_{jk} = \frac{1}{L} \log y_{...l} + \frac{1}{I} \sum_i \log y_{ijk.} - \log n - \hat{\mu} - \hat{\alpha} - \hat{\delta} \quad \dots\dots (18)$$

$$\hat{\alpha}\hat{\beta}\hat{\gamma}_{ijk} = \frac{1}{L} \log y_{...l} + \frac{1}{IJK} \sum_k \log y_{ijk.} - \log n - \hat{\mu} - \hat{\delta} \quad \dots\dots (19)$$

(Dobson: 1982).

Statistik Uji yang digunakan adalah chi kuadrat dengan rumusan sebagai berikut:

$$\chi^2 = \sum_i \sum_j \sum_k \sum_l \frac{(n_{ijkl} - e_{ijkl})^2}{e_{ijkl}} \quad \dots\dots (20) \text{ (Dobson: 1982)}$$

dengan degree of freedom adalah (IJK-1)(L-1)

Kriteria Uji:

Tolak H_0 jika χ^2 hitung $\geq \chi^2_\alpha$ dengan kata lain terdapat asosiasi antar variabel dan terima H_0 jika χ^2 hitung $< \chi^2_\alpha$ dengan kata lain model $\log E(Y_{ijkl}) = \hat{\mu} + \hat{\alpha}_i + \hat{\beta}_j + \hat{\gamma}_k + \hat{\delta}_l + \hat{\alpha}\hat{\beta}_{ij} + \hat{\alpha}\hat{\gamma}_{il} + \hat{\beta}\hat{\gamma}_{jl} + \hat{\alpha}\hat{\beta}\hat{\gamma}_{ijk}$ diterima.

Karena pada CFA difokuskan pada konfigurasi kategori antar variabel dan tidak difokuskan pada nilai dari parameternya dan kecocokan model, maka digunakan χ^2 untuk masing-masing konfigurasi dengan rumusan sebagai berikut:

$$\chi^2 = \frac{(n_t - e_t)^2}{e_t} \quad \dots\dots (21)$$

dengan χ^2 tabel adalah $\chi^2_{(IJK-1)(L-1)}$, t = konfigurasi ke-(1,2,...).

Dengan kriteria Uji:

Tolak H_0 jika χ^2 hitung $\geq \chi^2_\alpha$ dengan kata lain akan muncul type atau antitype dan terima H_0 jika χ^2 hitung $< \chi^2_\alpha$ dengan kata lain model $\log E(Y_{ijkl}) = \hat{\mu} + \hat{\alpha}_i + \hat{\beta}_j + \hat{\gamma}_k + \hat{\delta}_l + \hat{\alpha}\hat{\beta}_{ij} + \hat{\alpha}\hat{\gamma}_{il} + \hat{\beta}\hat{\gamma}_{jl} + \hat{\alpha}\hat{\beta}\hat{\gamma}_{ijk}$ diterima.

Untuk pengujian keindepedenan dari prediktor, menggunakan main effect dengan model sebagai berikut:

$$H_0: E(Y_{ijk}) = n\hat{\theta}_{i..}\hat{\theta}_{.j.}\hat{\theta}_{..k}$$

$$H_1: E(Y_{ijk}) \neq n\hat{\theta}_{i..}\hat{\theta}_{.j.}\hat{\theta}_{..k}$$

Diketahui bahwa, $\hat{\theta}_{i..} = \frac{y_{i..}}{n}$, $\hat{\theta}_{.j.} = \frac{y_{.j.}}{n}$ dan $\hat{\theta}_{..k} = \frac{y_{..k}}{n}$, sehingga

$$e_{ijkl} = \frac{y_{i..}y_{.j.}y_{..k}}{n^2} \quad \text{dimana } i = 1, 2; j = 1, 2; k = 1, 2. \quad \dots\dots (22)$$

$$\eta_{ijkl} = \log E(Y_{ijkl}) = \hat{\mu} + \hat{\alpha}_i + \hat{\beta}_j + \hat{\gamma}_k \quad \dots\dots (23)$$

dengan asumsi bahwa

$$\sum_i \hat{\alpha}_i = \sum_j \hat{\beta}_j = \sum_k \hat{\gamma}_k = 0$$

$\log e_{ijkl}$ diestimasi oleh η_{ijkl}

$$\log y_{i..} + \log y_{.j.} + \log y_{..k} - 2\log n = \hat{\mu} + \hat{\alpha}_i + \hat{\beta}_j + \hat{\gamma}_k \quad \dots\dots(24)$$

$$\sum_i \sum_j \sum_k \sum_l$$

dengan memasukkan ke persamaan di atas, akan diperoleh estimasi varians minimum yaitu sebagai berikut:

$$\hat{\mu} = \frac{1}{i} \sum_i \log y_{i..} + \frac{1}{j} \sum_j \log y_{.j.} + \frac{1}{k} \sum_k \log y_{..k} - 2\log n \quad \dots\dots(25)$$

Nilai $\hat{\alpha}$ dapat diperoleh dengan memasukkan $\sum_j \sum_k$ ke dalam Persamaan 24

$$\hat{\alpha}_i = \sum_i \log y_{i..} - \frac{1}{I} \sum_i \log y_{i..} \quad \dots\dots (26)$$

Dengan cara yang sama di atas, diperoleh nilai $\hat{\beta}_j, \hat{\gamma}_k$ sebagai berikut:

$$\hat{\beta}_j = \sum_j \log y_{.j.} - \frac{1}{J} \sum_j \log y_{.j.} \quad \dots\dots (27)$$

$$\hat{\gamma}_k = \sum_k \log y_{..k} - \frac{1}{K} \sum_k \log y_{..k} \quad \dots\dots (28)$$

(Annette J. Dobson).

Statistik Uji yang digunakan adalah chi kuadrat dengan rumusan sebagai berikut:

$$\chi^2 = \sum_i \sum_j \sum_k \frac{(n_{ijk} - e_{ijk})^2}{e_{ijk}} \quad \dots\dots (29)$$

(Dobson: 1982)

dengan degree of freedom adalah (I-1)(J-1)(K-1)

Kriteria Uji:

Tolak H_0 jika χ^2 hitung $\geq \chi^2_\alpha$ dengan kata lain terdapat asosiasi antar variabel dan terima H_0 jika χ^2 hitung $< \chi^2_\alpha$ dengan kata lain model $\log E(Y_{ijk}) = \hat{\mu} + \hat{\alpha}_i + \hat{\beta}_j + \hat{\gamma}_k$ diterima.

Karena pada CFA difokuskan pada konfigurasi kategori antar variabel dan tidak difokuskan pada nilai dari parameternya dan kecocokan model, maka digunakan χ^2 untuk masing-masing konfigurasi dengan rumusan sebagai berikut:

$$\chi^2 = \frac{(n_t - e_t)^2}{e_t} \dots\dots (4.30)$$

dengan χ^2 tabel adalah $\chi^2_{(I-1)(j-1)(k-1)}$, t = konfigurasi ke-(1,2,...).

Dengan kriteria Uji:

Tolak H_0 jika χ^2 hitung $\geq \chi^2_\alpha$ dengan kata lain akan muncul type atau antitype dan terima H_0

jika χ^2 hitung $< \chi^2_\alpha$ dengan kata lain model $\log E(Y_{ijk}) = \hat{\mu} + \hat{\alpha}_i + \hat{\beta}_j + \hat{\gamma}_k$ diterima.

$$z = \frac{N_t - \hat{E}_t}{\sqrt{\hat{E}_t}}$$

Dengan pendekatan normal, maka diperoleh rumusan

Metode Bonferroni untuk Melihat Signifikansi Konfigurasi

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:

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

Sebelumnya, perlu diketahui nilai α sebagai batasan dalam pengujian. Karena nilai α untuk tiap konfigurasi berbeda dengan α keseluruhan, maka digunakan metode Bonferroni dengan membagi $\alpha = 0.05$ dengan banyaknya konfigurasi yang terjadi.

$$\alpha^* = \frac{\alpha}{\sum t} = \frac{0.05}{16} = 0.003125$$

dan untuk melihat interaksi antar prediktor adalah:

$$\alpha^* = \frac{\alpha}{\sum t} = \frac{0.05}{8} = 0.00625$$

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:

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

kriteria uji:

Jika z-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-value $< \alpha^*$

maka H_0 ditolak atau akan muncul type dan antitype, dengan kata lain model tersebut tidak mewakili keberadaan dari konfigurasi tersebut.

HASIL PENELITIAN DAN PEMBAHASAN

Dalam penganalisisan data, penulis menggunakan software CFA 2002 (Von eye, 2001). Ada atau tidaknya interaksi antar prediktor dapat dilihat dari output di bawah ini:

Tabel 4.3.1 output CFA Jawa Timur first-order

Configuration	fo	fe	statistic	p	
111	23.	30.063	-1.288	.09884732	
112	21.	9.051	3.972	.00003570	Type
121	33.	48.334	-2.206	.01370627	
122	25.	14.552	2.739	.00308297	Type
211	159.	148.251	.883	.18867259	
212	29.	44.635	-2.340	.00963644	
221	250.	238.352	.754	.22528585	
222	65.	71.762	-.798	.21237086	

Output tersebut memperlihatkan bahwa terdapat adanya interaksi antar prediktor. Karena itu, pengujian untuk melihat ada tidaknya interaksi antar prediktor dan kriteria dapat dilakukan dengan output sebagai berikut:

Tabel 4.3.2 output CFA Jawa Timur first-order

Configuration	fo	fe	statistic	p	
11	22.	20.707	.284	.38818773	
12	32.	33.293	-.224	.41137273	
21	20.	16.489	.865	.19363825	
22	23.	26.511	-.682	.24766717	
31	1.	.767	.266	.39507230	
32	1.	1.233	-.210	.41688055	
41	1.	1.150	-.140	.44423683	
42	2.	1.850	.111	.45596739	
51	116.	128.463	-1.100	.13575739	
52	219.	206.537	.867	.19291751	
61	15.	19.557	-1.030	.15139771	
62	36.	31.443	.813	.20820072	
71	43.	28.377	2.745	.00302466	Type
72	31.	45.623	-2.165	.01519567	
81	14.	16.489	-.613	.26993411	
82	29.	26.511	.483	.31438535	

Dari output tabel 4.3.2 dapat dijelaskan bahwa type muncul pada konfigurasi 71. Dapat pula disimpulkan pula bahwa tidak terlihat adanya suatu pola tertentu dari masyarakat Jawa Tengah sebagai calon investor saham retail dilihat dari ketertarikan akan saham, namun dapat diamati bahwa ternyata masyarakat wilayah Jawa Timur yang selama ini tertarik dengan saham adalah masyarakat yang bekerja sebagai non pegawai negeri dengan penghasilan di atas Rp. 5 juta dan berusia di bawah 40 tahun.

KESIMPULAN

1. Keinginan PT Bursa Efek Indonesia untuk melihat karakteristik calon investor saham retail dapat diteliti dengan menggunakan CFA, namun berdasarkan hasil pengolahan data yang telah dilakukan, ternyata karakteristik calon investor saham di wilayah Jawa Timur yang selama ini tertarik dengan saham adalah masyarakat yang bekerja sebagai non pegawai negeri dengan penghasilan di atas Rp. 5 juta dan berusia di bawah 40 tahun
2. Terdapat perbedaan antara CFA dan Log Linear yaitu sebagai berikut:
 - Pada CFA, analisis difokuskan pada ketidakcocokan (discrepancies) antara base model dan data yang diperlihatkan dengan ada tidaknya type dan antitype berdasarkan konfigurasi yang terjadi sedangkan pada log-linear, analisis difokuskan pada kecocokan model yang memperhatikan ada atau tidaknya interaksi antar variabel.
 - Pada CFA, parameter dari base model tidak terlalu menjadi perhatian karena diasumsikan bahwa base model tidak menjelaskan data dengan baik sedangkan pada Log Linear, parameter baru dapat dijelaskan jika model sesuai dengan data.

SARAN

1. Kebenaran interpretasi atas type dan antitype sangat didasari oleh tepat atau tidaknya pemilihan base model dari CFA, karena itu penulis sangat menganjurkan agar pemilihan base model harus sesuai dengan tujuan penelitian.
2. Software CFA 2000 (Von eye 2001) hanya bisa menampung paling banyak 10 variabel, karena itu jika jumlah variabel terlalu besar disarankan menggunakan software lainnya seperti SPSS, SAS, SYSTAT, LEM

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.
- Agung, I Gusti Ngurah. 2002. *Analisis Hubungan Kausal Berdasarkan Data Kategorik*, PT RajaGrafindo Persada, Jakarta.
- Agresti, Alan. 1976. *An Introduction to Categorical Data Analysis*, Jhon Wiley and Sons, Inc., New York.
- Gujarati, D. 1978. *Basic Econometrics*, McGraw-Hill Companies, Inc., New York.
- Sudjana. 1996. *Metode Statistika*, Penerbit Tarsito, Bandung.
- The Jakarta Stock Exchange WebSite, <http://www.jsx.co.id>
- Supranto, J. 2004. *Analisis Multivariat*. Rineka Cipta, Jakarta.
- <http://www.google.com>