



JURNAL STATISTIKA DAN SAINS DATA

Volume 3 Nomor 1, Oktober, 2025, 1-20

<https://journal.student.uny.ac.id/index.php/jssd>

<https://dx.doi.org/10.21831/jssd.v3i1.19536>

Estimasi parameter Distribusi Mixture Normal dengan algoritma MCMC untuk optimasi portofolio

Ayuning Maretania Qurbani*, Prodi Statistika FMIPA UNY

Rosita Kusumawati, Prodi Statistika FMIPA UNY

*e-mail: ayuningmaretania.2019@student.uny.ac.id

Abstrak. Penelitian ini membentuk portofolio optimal melalui diversifikasi saham menggunakan pendekatan *Mean–Variance* (MV). Namun, asumsi kenormalan return pada metode MV tidak sepenuhnya terpenuhi pada data empiris. Oleh karena itu, penelitian ini menerapkan model *Mixture Normal* sebagai alternatif dengan estimasi parameter menggunakan algoritma Markov Chain Monte Carlo (MCMC) Metropolis–Hastings. Data yang digunakan mencakup saham Indeks Bisnis-27 periode 1 Januari 2020 hingga 31 Desember 2022. Pemilihan saham dilakukan melalui klasterisasi *K-means* berdasarkan rasio keuangan. Empat saham terdekat dengan pusat klaster dipilih sebagai kandidat portofolio. Estimasi parameter mixture normal digunakan dalam perhitungan bobot MV, menghasilkan komposisi portofolio 23,12% Charoen Pokphand Indonesia, 21,27% Kalbe Farma, 25,93% Tower Bersama Infrastructure, dan 29,68% Sumber Alfaria Trijaya. Hasil *backtesting* menunjukkan bahwa portofolio ini memberikan kinerja yang baik berdasarkan empat indikator pengukuran kinerja.

Kata kunci: *Mixture Normal, MCMC, Optimasi Portofolio.*

Abstract. This study constructs an optimal stock portfolio through diversification using the *Mean–Variance* (MV) approach. However, the normality assumption of returns in the MV method is not fully satisfied in empirical data. Therefore, this research employs the *Mixture Normal* model as an alternative, with parameter estimation conducted using the Markov Chain Monte Carlo (MCMC) Metropolis–Hastings algorithm. The dataset consists of stocks listed in the Bisnis-27 Index from 1 January 2020 to 31 December 2022. Stock selection is performed through *K-means* clustering based on financial ratios, yielding four stocks closest to the cluster centers as portfolio candidates. The estimated mixture normal parameters are incorporated into the MV optimization, resulting in portfolio weights of 23.12% Charoen Pokphand Indonesia, 21.27% Kalbe Farma, 25.93% Tower Bersama Infrastructure, and 29.68% Sumber Alfaria Trijaya. Backtesting results indicate that this portfolio demonstrates strong performance across four performance measurement indicators.

Keywords: *Mixture Normal, MCMC, Portfolio Optimization.*

PENDAHULUAN

Investasi merupakan pengalokasikan sejumlah dana dalam bentuk uang ataupun barang seperti surat berharga dengan harapan menghasilkan keuntungan di masa depan. Salah satu bentuk dari investasi adalah saham. Keuntungan dalam investasi saham dapat diperoleh dari tingkat return yang berupa *capital gain* yaitu selisih nilai jual investasi yang lebih tinggi dibandingkan dengan nilai beli (Anoraga & Pakarti, 2003), namun investor juga akan berpeluang untuk menghadapi kerugian atau risiko berupa *capital loss* yaitu kondisi nilai jual investasi yang lebih rendah dibandingkan nilai beli. Risiko saham menurut Jogiyanto (2014) terbagi menjadi dua jenis, yaitu risiko sistematis dan risiko tidak sistematis. Brigham dan Daves (2004) menyatakan bahwa risiko sistematis merupakan risiko yang tidak dapat dihindari karena hal ini disebabkan oleh pengaruh pasar secara keseluruhan, sedangkan risiko tidak sistematis dapat dihindari melalui diversifikasi saham dengan membentuk portofolio. Portofolio merupakan suatu kombinasi atau gabungan dari dua atau lebih saham individu atau surat berharga pasar uang yang dimiliki investor (Halim, 2005). Soelehan (2014) menyatakan bahwa para investor cenderung mengharapkan tingkat keuntungan (*return*) tinggi atau tingkat risiko yang rendah, hal tersebut menjadi dasar pembentukan portofolio optimal.

Pembentukan portofolio optimal dapat dilakukan menggunakan metode *Mean Variance* (MV) yang ditemukan pertama kali oleh Markowitz (1952) dengan tujuan memaksimumkan nilai harapan return atau meminimumkan tingkat risiko suatu portofolio. Markowitz (1959) mengatakan bahwa dalam pembentukan model MV membutuhkan estimasi return dengan asumsi bahwa estimasi return tersebut berdistribusi normal. Asumsi distribusi normal tidak selalu terpenuhi dalam data estimasi return saham yang sebenarnya, sehingga dibutuhkan suatu solusi untuk mengatasi permasalahan tersebut salah satunya adalah dengan membentuk model mixture normal. Pembentukan model mixture normal dapat dilakukan menggunakan metode Bayesian melalui suatu algoritma simulasi untuk mengestimasi suatu parameter. Pada metode Bayesian dibutuhkan pembentukan distribusi posterior yang dihasilkan melalui penggabungan antara distribusi data amatan dengan distribusi prior yang akan disusun sebagai fungsi likelihood (Horst, 2009). Destiarina et al. (2019) mengatakan bahwa sangat rumit untuk mendapatkan distribusi posterior tersebut, sehingga dibutuhkan sebuah pendekatan numerik. Pada penelitian yang dilakukan oleh Okky Herashanti et al. (2022) yang bertujuan untuk mengestimasi parameter model mixture dalam melakukan pengelompokan negara berdasarkan kasus stunting, mereka menggunakan metode Bayesian melalui suatu pendekatan numerik algoritma MCMC dan menghasilkan model mixture dengan 5 komponen. Abidin (2011) juga melakukan penelitian estimasi parameter model mixture untuk pengeluaran rumah tangga di Kota Semarang menggunakan metode bayesian melalui algoritma MCMC dan menghasilkan pemodelan mixture yang terdiri dari dua komponen. Irawan (2011) juga melakukan penelitian dengan metode yang sama untuk membentuk pemodelan mixture dalam pemilihan portfolio, penelitian tersebut menghasilkan kesimpulan bahwa estimasi parameter menggunakan MCMC akan mempermudah penggunaan struktur model mixture normal untuk penghitungan portofolio.

MCMC adalah sebuah metode yang membangkitkan sejumlah peubah acak berdasarkan pada rantai markov. Stephens (2004) menyatakan bahwa metode MCMC membutuhkan pembentukan sebuah rantai Markov $\{\theta_i\}$ dengan distribusi posterior $p(\theta|x)$ dari parameter θ . Sehingga akan diperoleh sebuah barisan sampel acak yang berkorelasi, yaitu nilai ke- t dari

suatu barisan $\{\theta_t\}$ dan dilakukan sampling dari sebuah distribusi peluang yang bergantung pada nilai sebelumnya $\{\theta_{t-1}\}$ (Gamerman, 1997). MCMC banyak digunakan sebagai metode dalam pembentukan model harga saham (Setiawani, 2015), seperti penelitian yang dilakukan oleh Johannes et al. (1999), Eraker et al. (2003), dan Landauskas & Valakevicius (2011) menggunakan metode tersebut dalam penentuan model harga saham. Sifat data saham terdiri dari urutan kejadian yang saling terikat, hal tersebut sesuai dengan sifat metode MCMC. Sehingga pada penelitian kali ini akan sesuai dalam pengimplementasian metode MCMC dengan menggunakan ilustrasi numerik dari data saham di Indonesia.

Pada penelitian ini bersesuaian dengan pembahasan sebelumnya akan dilakukan estimasi distribusi mixture normal yang menerapkan metode simulasi MCMC dengan algoritma Metropolis Hasting (MH) untuk pembentukan portofolio optimal dengan suatu ilustrasi numerik. Ilustrasi numerik yang digunakan adalah sebuah portofolio dari beberapa data saham yang termasuk pada Index Bisnis-27, pemilihan saham tersebut akan ditentukan melalui hasil klasterisasi *K-Means*. Pada penelitian ini *K-means* dilakukan berdasarkan rasio keuangan yaitu *current ratio*, *return on equity*, *return on assets*, *net profit margin*, *gross profit margin*, *debt to equity* dan *debt to assets*. Portofolio disusun melalui pemilihan satu saham pada setiap klaster yang memiliki jarak terdekat dengan pusat klaster itu sendiri. Portofolio optimal yang terbentuk pada penelitian ini akan diukur kinerjanya menggunakan metode backtesting. Pengukuran kinerja tersebut menggunakan beberapa ajenis pengukuran yaitu *sharpe ratio*, *max drawdown*, *annual return*, dan *annual volatility* (Palomar & Zhou, 2022).

METODE

Data diperoleh melalui pengambilan data melalui website Yahoo Finance dengan melakukan pemilihan data saham yang termasuk pada Bisnis-27 periode Februari-Juli 2023 dan data dipilih berdasarkan rentang periode 1 Januari 2020 s.d. 31 Desember 2022. Data yang digunakan adalah data pada laporan keuangan setiap saham berupa neraca (*Balance Sheet*) dan laporan laba rugi (*Income Statement*) untuk perhitungan rasio keuangan, dan data *adjusted close price* pada setiap saham untuk perhitungan return.

Seluruh analisis dilaksanakan dalam lingkungan R (R Core Team, 2023) dengan skrip yang disusun khusus untuk penelitian ini. Paket R utama yang digunakan meliputi: readxl/base read.csv untuk impor data, dplyr/tidyverse untuk manipulasi data, scale/stats::kmeans untuk klasterisasi, factoextra dan cluster untuk validasi klaster, quantmod/xts untuk pengambilan dan pengolahan harga saham, mixtools untuk inisialisasi campuran normal (EM), nortest untuk uji Anderson–Darling, serta paket grafis graphics/ggplot2 untuk visualisasi. Selain itu, sejumlah fungsi kustom (mis. statdef, sd.m, mix.d, MH.A, plot.post, mu.f, V.f, sig.f, bobot.MV) diprogram untuk kebutuhan statistik deskriptif, estimasi Metropolis–Hastings, dan optimasi Mean–Variance.

1. Laporan keuangan perusahaan diimpor dari file CSV/Excel dan dihitung rasio keuangan sesuai kategori Munawir (2004): likuiditas, profitabilitas, dan solvabilitas. Tujuh rasio yang dihitung adalah Current Ratio (CR), Return on Equity (ROE), Return on Assets (ROA), Net Profit Margin (NPM), Gross Profit Margin (GPM), Debt to Equity Ratio (DER), dan Debt to Assets (DA). Implementasi dilakukan dengan dplyr:::mutate() untuk komputasi rasio. Contoh ringkas:

```
dat <- read.csv("RasioKeuangan.csv")
dat_calc <- dat %>%
```

```
mutate(  
    CR = CurrentAssets / CurrentLiabilities,  
    ROE = NetIncome / Equity,  
    ROA = NetIncome / TotalAssets,  
    NPM = NetIncome / Revenue,  
    GPM = GrossProfit / Revenue,  
    DER = TotalDebt / Equity,  
    DA = TotalDebt / TotalAssets  
)
```

2. Melakukan *clustering* menggunakan metode K-Means dengan variabel komponen berupa 7 pengukuran rasio keuangan yang telah dihitung pada langkah 1. Pengelompokan data menggunakan metode K-Means dilakukan dengan membentuk nilai centeroid untuk setiap klaster (Everitt & Hothorn, 2011). Sebelum klasterisasi, variabel rasio dinormalisasi menggunakan scale() untuk menghilangkan perbedaan skala antar rasio. Jumlah klaster ($k = 4$) ditentukan melalui *elbow method* menggunakan factoextra::fviz_nbclust(..., method = "wss"). Klasterisasi dijalankan dengan fungsi kmeans() (paket stats) dan hasilnya disimpan pada objek hasil klaster (contoh: clust.4means dari kode). Validitas klaster dieksplorasi dengan ringkasan *silhouette* dan PCA (FactoMineR::PCA pada kode asli). Potongan kode inti yang dipakai:

```
ratio_mat <- dat_calc %>% select(CR, ROE, ROA, NPM, GPM, DER, DA)  
ratio_scaled <- scale(ratio_mat)  
set.seed(2023)  
clust.4means <- kmeans(ratio_scaled, centers = 4, nstart = 50)
```

3. Pembentukan portofolio yaitu dari setiap klaster dipilih satu saham yang jaraknya paling dekat ke pusat klaster (centroid). Jarak yang digunakan adalah jumlah absolut selisih komponen rasio (yang diimplementasikan di skrip sebagai selisih absolut per variable kemudian dijumlahkan — rowSums() pada sub-data klaster) sehingga saham dengan S.ED terkecil dipilih sebagai wakil klaster. Contoh alur di skrip:

```
# contoh untuk cluster 1  
DC1 <- dat.C[dat.C$clust.4means.cluster==1,]  
dc1 <- clust.4means$centers[1,]  
CR <- abs(DC1$Current.Ratio - dc1[1])  
...  
S.ED <- rowSums(DC1.ED)  
arrange(DC1.ED[, ...], S.ED)
```

4. Analisis statistika deskriptif dari setiap saham yang membentuk portofolio pada langkah 3 dengan perhitungan nilai return setiap saham. Perhitungan return dilakukan menggunakan rumus log return (Bannet dan Hugen, 2016)). Harga penutupan disesuaikan (Adjusted Close) untuk saham terpilih diunduh menggunakan quantmod::getSymbols() dan diolah menjadi *daily log returns* menggunakan PerformanceAnalytics::dailyReturn() atau fungsi diff(log(...)). Statistik deskriptif tiap series dihitung menggunakan fungsi kustom statdef()

serta fungsi sd.m() yang disertakan di lampiran skrip. Contoh penggunaan berdasarkan Bennett & Hugen (2016).

```
D1K <- getSymbols("CPIN.JK", src = "yahoo", from="2020-01-01",
to="2022-12-31", auto.assign = FALSE)
D1K.retD <- dailyReturn(D1K)
statdef(D1K.retD)
```

5. Pengujian distribusi normal dari data return setiap saham yang membentuk portofolio menggunakan tes Anderson-Darling membandingkan seberapa cocok kumpulan data return dengan distribusi normal. Tes dilakukan dengan menggunakan nortest::ad.test() sesuai langkah skrip. Keputusan uji (p-value) digunakan sebagai dasar pertimbangan model: jika normalitas ditolak, model Mixture Normal dipertimbangkan dan dieksplorasi lebih lanjut. nortest::ad.test(D1K.retD)
6. Melakukan estimasi parameter distribusi *mixture* normal menggunakan metode MCMC dengan algoritma metropolis hastings, model distribusi mixture normal normal ditunjukkan pada persamaan (1).

$$f_i(R|\boldsymbol{w}_i, \boldsymbol{\mu}_i, \sigma_i^2) = w_{i1}g_1(x|\mu_{i1}, \sigma_{i1}^2) + w_{i2}g_2(x|\mu_{i2}, \sigma_{i2}^2) \quad (1)$$

dengan

$$\begin{aligned} f_i(R|\boldsymbol{w}_i, \boldsymbol{\mu}_i, \sigma_i^2) &= \text{fungsi probabilitas mixture normal dari data saham } i \\ g_k(R|\theta_k) &= \text{fungsi probabilitas normal ke-}k \end{aligned}$$

berikut langkah dalam penerapan algoritma metropolis hastings (Gelman et al., 2013).

- a. Pilih titik awal di ruang parameter dan hasilkan urutan nilai parameter kandidat dari nilai parameter proposal yang ditentukan melalui pembentukan histogram data return setiap perusahaan yang dibentuk menggunakan dailyReturn(), lalu histogram return digunakan oleh mix.d() untuk memperoleh parameter proposal $\theta_0 = (w_1, \mu_1, \sigma_1, \mu_2, \sigma_2)$.
- ```
prop <- matrix(c(par.D1K$theta, par.D2K$theta, par.D3K$theta,
par.D4K$theta),
ncol=4, byrow=TRUE)
```
- b. Tentukan fungsi likelihood untuk distribusi mixture normal, yang merupakan probabilitas data amatan yang diberi nilai parameter.

$$L_{mix}(R|\theta) = \prod_{i_1=1}^{m_k} w_1 g_1(R_{i_1}|\theta_1) + \prod_{i_2=1}^{m_k} w_2 g_2(R_{i_2}|\theta_2) \quad (2)$$

Fungsi ini diimplementasikan di dalam fungsi custom MH.A() pada kode R untuk menghitung likelihood setiap kandidat parameter.

```
MH.A <- function(x, param, N) {
 sim.size = N
 theta.star = matrix(NA, nrow = sim.size, ncol=5)
 colnames(theta.star) = c('p', 'mu1', 'sigma1', 'mu2',
 'sigma2')
 theta.star[1,] = param
 for(i in 2:sim.size){
 if(i %% 100==0) cat('i = ', i, '\n')
 .
 .
 }
}
```

```

 theta.star = data.frame(theta.star)
 result = apply(theta.star,2,mean)
 return(list(MH = result,dat.sim= theta.star))

 }

```

- c. Tentukan distribusi prior untuk nilai parameter kandidat  $p(\theta' | R)$ . Pada penelitian ini digunakan asumsi yang dinyatakan oleh Richardson dan Green (1997).

$$\begin{aligned}
 w &\sim Dirichlet(1,1) \\
 \mu &\sim Normal(\mu_0, \sigma_0) \\
 \sigma^2 &\sim Inv-Gamma(0.1,0.01)
 \end{aligned} \tag{3}$$

```

logPrior = function(param){
 p = param[1]
 mu1 = param[2]
 sigma1 = param [3]
 mu2 = param[4]
 sigma2 = param[5]
 beta1 = dgamma(sigma1,shape=0.1, scale=100)
 beta2 = dgamma(sigma1,shape=0.1, scale=100)
 lp = ddirichlet(p,1) + dnorm(mu1, mean=mu1, sd=sigma1) +
 dnorm(mu2, mean=mu2, sd=sigma2) + dinvgamma(sigma1,
 shape=0.1, rate=beta1, log=TRUE) + dinvgamma(sigma2,
 shape=0.1, rate=beta2, log=TRUE)
 return(lp)
}

}

```

- d. Tentukan distribusi proposal untuk algoritma Metropolis-Hastings. Distribusi proposal digunakan untuk menghasilkan nilai parameter kandidat  $\theta'$  yang akan diuji. Pada penelitian ini menggunakan distribusi uniform dengan batas atas  $a$  dan batas bawah  $b$ .

$$\theta' \sim Uniform(a, b)$$

- e. Evaluasi rasio penerimaan untuk nilai parameter kandidat, dengan melakukan perhitungan probabilitas pada persamaan (4)

$$\alpha = \min\{1, r\} \tag{4}$$

dengan perhitungan rasio  $r$  ditunjukkan pada persamaan (5).

$$r = \frac{p(\theta' | R)}{p(\theta^{t-1} | R)} \tag{5}$$

dengan  $p(\theta' | R) \propto L_{mix}(R | \theta') p(\theta')$  dan  $p(\theta^{t-1} | R) \propto L_{mix}(R | \theta^{t-1}) p(\theta^{t-1})$ . Persamaan (16) berdasarkan langkah c dan pembentukan distribusi posterior  $p(\theta' | R)$  dan  $p(\theta^{t-1} | R)$  akan mendefinisikan distribusi prior untuk  $p(\theta')$  sebagai  $p(\theta^{t-1}, \theta')$  dan  $p(\theta^{t-1})$  sebagai  $p(\theta', \theta^{t-1})$  dengan  $p(\theta^{t-1}, \theta') = p(\theta', \theta^{t-1})$  (Gelman et al., 2013).

- f. Terima atau tolak nilai target berdasarkan rasio penerimaan, dengan melakukan perhitungan berdasarkan persamaan (6) , dengan  $u \sim Uniform(0,1)$

$$\theta^t = \begin{cases} \theta^t = \theta', & \text{jika } u \leq \alpha \\ \theta^t = \theta^{t-1}, & \text{selainnya} \end{cases} \quad (6)$$

```
prob = min(r,1)
if(runif(1) < prob){
 theta.star[i,'p'] = p.proposal
} else {
 theta.star[i,'p'] = theta.star[i-1,'p']
}
```

- g. Jika nilai kandidat diterima, gunakan sebagai nilai target saat ini untuk iterasi berikutnya. Jika ditolak, lanjutkan dengan nilai sebelumnya.
- h. Ulangi langkah d-g untuk sejumlah iterasi tertentu untuk menghasilkan urutan nilai parameter.
- i. Setelah rantai dihasilkan, hitung ringkasan statistik dari distribusi posterior, seperti rata-rata dan varian dari setiap komponen
- j. Plot distribusi posterior untuk memvisualisasikan ketidakpastian nilai parameter. Pada penelitian ini digunakan taraf kepercayaan sebesar 95% atau taraf signifikansi sebesar 5%. Perhitungan interval kepercayaan berdasarkan persamaan (7),

$$\hat{\theta}_i \pm 1.96 (sd(\theta_i)) \quad (7)$$

dengan nilai 1.96 merupakan nilai probabilitas titik persentil 97.5 berdasarkan distribusi normal standar.

- k. Gunakan distribusi posterior untuk membuat prediksi tentang data masa depan atau untuk memperkirakan kuantitas yang tidak diketahui terkait dengan distribusi mixture normal.
7. Pembentukan portofolio optimal melalui metode *Mean Variance* (MV) dari hasil pembentukan model distribusi mixture normal sebelumnya. Portfolio optimal dengan metode MV yaitu (Elton et al., 2014)

$$\text{meminimumkan } \sigma(w_i) = \sum_{i=1}^n w_i^2 \sigma_i^2 + \sum_{i=1}^n \sum_{j=1, i \neq j}^n w_i w_j \sigma_{ij}, \quad (8)$$

dengan kendala  $\sum_{i=1}^n \mu_i w_i = \mu_p$  dan  $\sum_{i=1}^n w_i = 1$

Melalui peminimuman  $\sigma$  akan memperoleh cara perhitungan untuk setiap bobot saham pada portofolio optimal yang membutuhkan vektor mean dan matriks varian-kovarian ditunjukkan pada persamaan (9).

$$Aw = b_0 \quad (9)$$

dengan perhitungan rata-rata bersesuaian dengan persamaan (10) dan variansi bersesuaian dengan persamaan (11), serta perhitungan kovariansi bersesuaian dengan persamaan (12).

$$\mu_i = \sum_{k=1}^m w_k \mu_k \quad (10)$$

$$\sigma_i^2 = \sum_{k=1}^m w_k (\sigma_k^2 + \mu_k^2) - \left( \sum_{k=1}^m w_k \mu_k \right)^2 \quad (11)$$

$$\sigma_{ij} = E(R_1 - \mu_1)(R_2 - \mu_2), i \neq j \quad (12)$$

dengan,

|               |                                                 |
|---------------|-------------------------------------------------|
| $w_i$         | = bobot asset ke i                              |
| $\sigma_{ij}$ | = kovarian asset ke i dan asset ke j            |
| $n$           | = jumlah asset                                  |
| $R_i$         | = return saham perusahaan $i$                   |
| $w_k$         | = bobot komponen ke $k$                         |
| $\mu_i$       | = rata-rata mixture return saham perusahaan $i$ |
| $\mu_k$       | = hasil estimasi rata-rata komponen ke $k$      |
| $\sigma_i^2$  | = variansi mixture return saham perusahaan $i$  |
| $\sigma_k^2$  | = hasil estimasi variansi komponen ke $k$       |
| $m$           | = jumlah komponen                               |

- Melakukan proses *backtesting* dalam 4 jenis pengukuran untuk mengukur kinerja portofolio, yaitu *sharpe ratio*, *max drawdown*, *annual return*, dan *annual volatility* (Palomar & Zhou, 2022).

## HASIL DAN PEMBAHASAN

### Hasil

Hasil pada penelitian ini bersesuaian dengan setiap langkah yang telah disebutkan pada penejelasan sebelumnya. Langkah pertama yaitu perhitungan rasio keuangan dengan 7 jenis pengukuran, kedua klasterisasi data saham Bisnis-27, ketiga pembentukan portofolio, keempat analisis statistika deskriptif, kelima pengujian distribusi normal, keenam estimasi parameter distribusi *mixture* normal, ketujuh pembentukan portofolio optimal, dan terakhir melakukan proses *backtesting*.

#### 1. Rasio keuangan

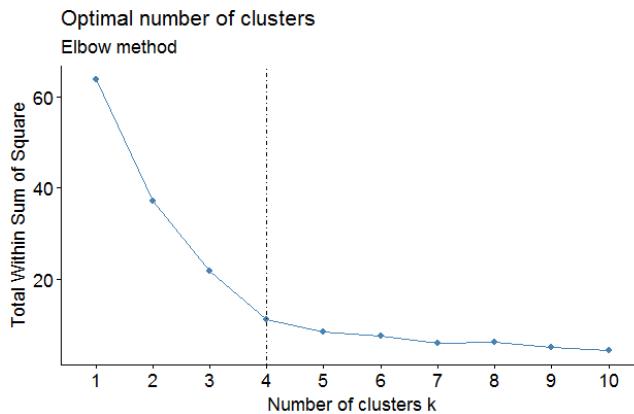
Tujuh perhitungan rasio keuangan yang dilakukan pada seluruh 27 data saham Bisnis-27 akan ditunjukkan pada Tabel 1.

**Tabel 1. Tabel 1.** Rasio Keuangan Saham Bisnis-27

| No | Kode | CR    | RoE   | RoA   | NPM   | GPM   | DE    | DA    |
|----|------|-------|-------|-------|-------|-------|-------|-------|
| 1  | ADRO | 2,173 | 0,382 | 0,231 | 0,308 | 0,574 | 0,242 | 0,147 |
| 2  | AKRA | 0,518 | 0,183 | 0,088 | 0,051 | 0,089 | 0,229 | 0,111 |
| 3  | AMRT | 0,901 | 0,226 | 0,084 | 0,028 | 0,208 | 0,266 | 0,099 |
| :  | :    | :     | :     | :     | :     | :     | :     | :     |
| 25 | TOWR | 0,338 | 0,282 | 0,052 | 0,319 | 0,729 | 3,895 | 0,714 |
| 26 | UNTR | 1,878 | 0,205 | 0,130 | 0,162 | 0,278 | 0,034 | 0,021 |
| 27 | UNVR | 0,608 | 1,342 | 0,293 | 0,130 | 0,463 | 0,321 | 0,070 |

#### 2. Klasterisasi data

Jumlah klaster untuk penggunaan metode *clustering K-Means* dapat ditentukan melalui metode elbow. Metode elbow mencari jumlah klaster terbaik berdasarkan nilai *Sum of Square Error* (SSE) yang mengalami penurunan signifikan atau berbentuk siku (Kosasih et al., 2016).



**Gambar 1.** Plot Nilai Metode Elbow dengan 10 Klaster

Pada Gambar 1 dapat diketahui bahwa jumlah klaster optimal yang dapat digunakan untuk penerapan metode *K-means* pada data rasio keuangan saham Bisnis-27 adalah 4 klaster.

**Tabel 2. Tabel 2.** Hasil K-Means clustering

| No | Kode      |           |           |           |
|----|-----------|-----------|-----------|-----------|
|    | Klaster 1 | Klaster 2 | Klaster 3 | Klaster 4 |
| 1  | ADRO      | INCO      | EXCL      | AKRA      |
| 2  | ANTM      | KLBF      | TBIG      | AMRT      |
| 3  | ASII      | MIKA      | TOWR      | BBCA      |
| 4  | CPIN      |           |           | BBNI      |
| 5  | CTRA      |           |           | BBRI      |
| 6  | ESSA      |           |           | BMRI      |
| 7  | INDF      |           |           | TLKM      |
| 8  | INKP      |           |           | UNVR      |
| 9  | MDKA      |           |           |           |
| 10 | PGAS      |           |           |           |
| 11 | PTBA      |           |           |           |
| 12 | SMGR      |           |           |           |
| 13 | UNTR      |           |           |           |

Tabel 2 menunjukkan anggota setiap klaster yang terbentuk. Pengelompokan tersebut dilakukan dengan bantuan program R pada fungsi “kmeans” yang juga menghasilkan kesimpulan bahwa total varians dalam kumpulan data yang dijelaskan oleh empat klaster tersebut memiliki nilai sebesar 83.1%.

**Tabel 3. Tabel 3.** Nilai Centeroid Setiap Klaster Terhadap Setiap Komponen Rasio Keuangan

| Klaster | Current Ratio | Return of Equity | Return on Asset | Net Profit Margin | Gross Profit Margin | Debt to Equity | Debt to Asset |
|---------|---------------|------------------|-----------------|-------------------|---------------------|----------------|---------------|
| 1       | 1,919         | 0,171            | 0,103           | 0,151             | 0,327               | 0,422          | 0,230         |
| 2       | 4,416         | 0,139            | 0,119           | 0,187             | 0,400               | 0,018          | 0,015         |
| 3       | 0,381         | 0,160            | 0,035           | 0,205             | 0,699               | 2,771          | 0,635         |
| 4       | 0,728         | 0,317            | 0,080           | 0,216             | 0,346               | 0,316          | 0,087         |

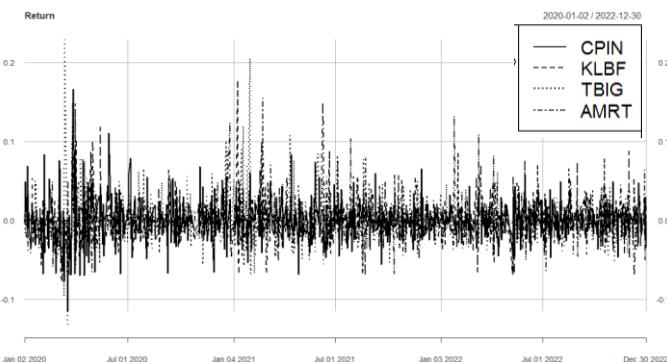
Pada Tabel 3 menunjukkan nilai pusat pada masing-masing klaster terhadap setiap pengukuran pada rasio keuangan yang digunakan. Klaster 1 merupakan klaster yang memiliki karakteristik nilai terkecil pada *net profit margin* dan *gross profit margin* dibandingkan klaster yang lain. Klaster 2 merupakan klaster yang memiliki karakteristik nilai tertinggi pada *current ratio* dan RoA, dan nilai terkecil pada *debt to equity* dan *debt to asset* dibandingkan klaster yang lain. Klaster 3 merupakan klaster yang memiliki karakteristik nilai terkecil pada current ratio dan RoA, dan nilai tertinggi pada *debt to equity* dan *debt to asset* dibandingkan klaster yang lain. Klaster 4 merupakan klaster yang memiliki karakteristik nilai tertinggi pada RoE dan *net profit of margin* dibandingkan klaster yang lain.

### 3. Pembentukan portofolio

Pembentukan portofolio yaitu dengan melakukan pemilihan satu saham pada setiap klaster terbentuk yang memiliki jarak terdekat dengan pusat klaster itu sendiri. Pada klaster 1 didapat saham dari perusahaan Charoen Pokphand Indonesia Tbk dengan kode CPIN yang memiliki jarak sebesar 0,479 dari titik pusat klaster 1. Pada klaster 2 didapat saham dari perusahaan Kalbe Farma Tbk dengan kode KLBF yang memiliki jarak sebesar 0,807 dari titik pusat klaster 2. Pada klaster 3 didapat saham dari perusahaan Tower Bersama Infrastructure Tbk dengan kode TBIG yang memiliki jarak sebesar 0,235 dari titik pusat klaster 3. Pada klaster 4 didapat saham dari perusahaan Sumber Alfaria Trijaya Tbk dengan kode AMRT yang memiliki jarak sebesar 0,650 dari titik pusat klaster 4.

### 4. Analisis statistika deskriptif

Pada Gambar 2, grafik menunjukkan nilai tertinggi dari return TBIG pernah mencapai di atas 0,2, nilai terendah dari return TBIG mencapai dibawah -0,1 dan juga ditunjukkan oleh CPIN yang memiliki nilai return terendah di bawah -0,1 .



**Gambar 2.** Return 4 Saham Penyusun Portofolio

Statistik deskriptif pada Gambar 2 ditunjukkan lebih jelas pada Tabel 4. Rata-rata return dari empat perusahaan memiliki nilai cukup kecil dengan tiga sampai empat angka di belakang koma dengan perusahaan CPIN menjadi perusahaan yang memiliki nilai terkecil sebesar 0,013%, sedangkan perusahaan AMRT memiliki rata-rata return tertinggi yaitu sebesar 0,192%.

**Tabel 4. Tabel 4.** Statistik Deskriptif Data Return

| Saham | Mean    | Variansi | Minimum  | Maksimum |
|-------|---------|----------|----------|----------|
| CPIN  | 0,00013 | 0,00069  | -0,11559 | 0,16627  |
| KLBF  | 0,00061 | 0,00054  | -0,07692 | 0,17725  |
| TBIG  | 0,00127 | 0,00086  | -0,13182 | 0,22905  |
| AMRT  | 0,00192 | 0,00083  | -0,06977 | 0,15517  |

Tabel 4 menunjukkan bahwa besar risiko dari masing-masing perusahaan memiliki nilai yang juga kecil bersesuaian dengan rata-rata return dari masing-masing saham. Perusahaan TBIG memiliki tingkat risiko paling tinggi dibandingkan tiga perusahaan yang lain dengan nilai variansi sebesar 0,086%. Perusahaan KLBF memiliki nilai risiko yang paling kecil dibandingkan tiga perusahaan yang lain dengan nilai variansi sebesar 0,086%.

##### 5. Pengujian asumsi distribusi normal

Pengujian statistik asumsi ditribusi normal menggunakan tes Anderson Darling dengan bantuan program R pada fungsi “ad.test”. Pengujian dilakukan menggunakan data return pada setiap perusahaan, dan hasil pengujian tersebut akan ditunjukkan pada Tabel 5.

**Tabel 5. Tabel 5.** Pengujian Distribusi Normal

| Perusahaan | Nilai Pengujian | P-value   |
|------------|-----------------|-----------|
| CPIN       | 739             | < 2,2e-16 |
| KLBF       | 721             | < 2,2e-16 |
| TBIG       | 776             | < 2,2e-16 |
| AMRT       | 699             | < 2,2e-16 |

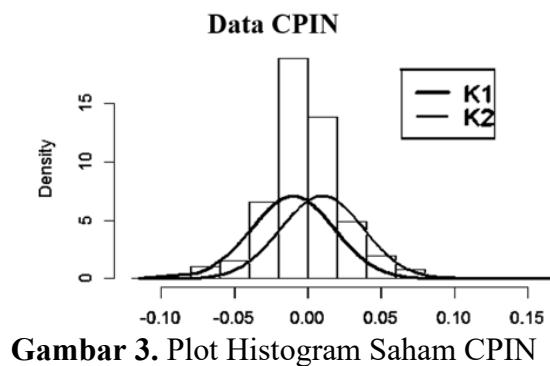
Tabel 5 menunjukkan bahwa data return dari ke empat perusahaan penyusun portofolio tersebut menghasilkan kesimpulan bahwa data return tidak memenuhi asumsi distribusi normal, karena nilai  $p$  – value kurang dari 0 yang memberikan keputusan bahwa asumsi data berdistribusi normal ditolak melalui Tes *Anderson-Darling*.

#### 6. Estimasi parameter distribusi mixture normal

Algoritma pada metode MCMC di penelitian kali ini akan menggunakan algoritma Metropolis Hasting, dengan ketentuan setiap distribusi mixture memiliki 2 komponen.

##### a. Parameter Proposal

Nilai parameter proposal bobot diasumsikan sebesar 0,5 akan ditunjukkan melalui gambar 3 dengan garis tebal untuk komponen 1 dan garis tipis untuk komponen 2.



**Gambar 3.** Plot Histogram Saham CPIN

Parameter  $\mu$  pada setiap komponen akan didasarkan terhadap nilai tengah pada 2 bar tertinggi di setiap histogram return saham yang terbentuk pada Gambar 3. Parameter proposal untuk bobot dari setiap komponen diasumsikan sebesar 0.5. Parameter proposal untuk sigma ditentukan melalui perhitungan berdasarkan nilai  $\mu$  yang sudah diperoleh dengan perhitungan melalui rumus pada persamaan (13).

$$\sigma_{0k} = \sqrt{\frac{\sum_{i=1}^n (R_i - \mu_{0k})^2}{N}} \quad (13)$$

Sehingga memperoleh nilai parameter proposal pada setiap saham yang ditunjukkan pada tabel 6.

**Tabel 6.** Tabel 6. Parameter Proposal

| Perusahaan | $w_1$ | $w_2$ | $\mu_1$ | $\sigma_1$ | $\mu_2$ | $\sigma_2$ |
|------------|-------|-------|---------|------------|---------|------------|
| CPIN       | 0,5   | 0,5   | -0,010  | 0,02821    | 0,010   | 0,02812    |
| KLBF       | 0,5   | 0,5   | -0,010  | 0,02548    | 0,010   | 0,02500    |
| TBIG       | 0,5   | 0,5   | -0,025  | 0,03937    | 0,025   | 0,03772    |
| AMRT       | 0,5   | 0,5   | -0,010  | 0,03119    | 0,010   | 0,02994    |

Tabel 6 menunjukkan bahwa nilai parameter proposal untuk  $\sigma$  (standart deviasi) di setiap saham memiliki kondisi yang sama yaitu nilai  $\sigma_1$  lebih tinggi dibandingkan

dengan nilai  $\sigma_2$ . Titik awal ruang parameter ditentukan berdasarkan nilai parameter proposal yang ditampilkan pada Tabel 6.

b. Fungsi likelihood distribusi mixture normal

Persamaan (2) merupakan fungsi likelihood mixture normal yang akan digunakan dalam estimasi parameter distribusi mixture normal dengan  $g_k(R_{i_k}|\theta_k)$  merupakan fungsi probabilitas normal komponen ke  $k$ . Sehingga dapat dibentuk persamaan  $L_{mix}(R|\theta^{t-1})$  yang akan bersesuaian dengan data sebelumnya dan persamaan  $L_{mix}(R|\theta')$  yang bersesuaian dengan data parameter kandidat.

c. Distribusi prior parameter

Distribusi prior  $p(\theta)$  bersesuaian dengan persamaan (3), nilai parameter yang ditentukan pada persamaan (3) bertujuan untuk membentuk hasil estimasi yang akan berpusat di sekitar distribusi prior.

d. Distribusi proposal untuk nilai kandidat

$$\theta' \sim U(\theta^{t-1} - 0,005, \theta^{t-1} + 0,005) \quad (14)$$

Distribusi proposal untuk algoritma MH dalam menentukan nilai kandidat merupakan distribusi uniform dengan batas bawah  $\theta^{t-1} - 0,005$  dan batas atas  $\theta^{t-1} + 0,005$  yang dapat ditulis sebagai persamaan (14), hal ini dilakukan melalui bantuan program R dengan fungsi “runif”, hasil tersebut akan ditunjukkan pada tabel 7.

**Tabel 7.** Tabel 7. Parameter  $\theta'$  MH iterasi 1

| Parameter        | $\theta_{min}^{(t)}$ | $\theta_{max}^{(t)}$ | $\theta'$ |
|------------------|----------------------|----------------------|-----------|
| $w_1^{(1)}$      | 0,49500              | 0,50500              | 0,4997    |
| $\mu_1^{(1)}$    | -0,01500             | -0,00500             | -0,01033  |
| $\mu_2^{(1)}$    | 0,00500              | 0,01500              | 0,00967   |
| $\sigma_1^{(1)}$ | 0,02321              | 0,03321              | 0,02788   |
| $\sigma_2^{(1)}$ | 0,02312              | 0,03312              | 0,02778   |

e. Evaluasi ratio

Perhitungan probabilitas untuk evaluasi rasio penerimaan nilai kandidat bersesuaian dengan persamaan (5). Hasil perhitungan  $r$  untuk setiap parameter pada iterasi 1 ditunjukkan pada Tabel 8

**Tabel 8.** Tabel 8. Distribusi Posterior  $\theta$

| Parameter        | $p(\theta' R)$ | $p(\theta^{t-1} R)$ | $r$   |
|------------------|----------------|---------------------|-------|
| $w_1^{(1)}$      | 1533           | 1533                | 0,999 |
| $\mu_1^{(1)}$    | 1532           | 1533                | 0,648 |
| $\mu_2^{(1)}$    | 1532           | 1532                | 1,416 |
| $\sigma_1^{(1)}$ | 1532           | 1532                | 0,527 |
| $\sigma_2^{(1)}$ | 1532           | 1532                | 0,918 |

f. Hasil nilai target

Keputusan tolak atau terima nilai target berdasarkan rasio penerimaan yang telah dituliskan pada persamaan (4). Bangkitkan  $u \sim U(0,1)$ , melalui bantuan program R dalam fungsi "runif" dan dihasilkan  $u = 0.4666139$ . Hasil keputusan pada iterasi 1-3 dalam penerimaan nilai parameter kandidat ditunjukkan pada tabel 9.

**Tabel 9.** Tabel 9. Hasil Sampel Parameter

| Parameter  | $\alpha_1$ | $\theta^1$ | $\alpha_2$ | $\theta^2$ | $\alpha_3$ | $\theta^3$ |
|------------|------------|------------|------------|------------|------------|------------|
| $w_1$      | 0,999      | 0,49666    | 1          | 0,49930    | 1          | 0,4990     |
| $\mu_1$    | 0,648      | -0,01033   | 0,6268     | -0,01067   | 0,605      | -0,01100   |
| $\mu_2$    | 1          | 0,00967    | 1          | 0,00933    | 1          | 0,00899    |
| $\sigma_1$ | 0,527      | 0,02788    | 0,464      | 0,02788    | 0,439      | 0,02788    |
| $\sigma_2$ | 0,918      | 0,02778    | 0,874      | 0,02745    | 0,829      | 0,02712    |

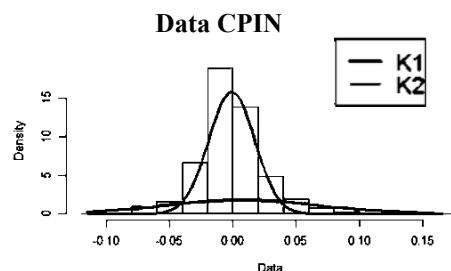
- g. Hasil  $\theta^{(1)}$  digunakan untuk iterasi kedua  $\theta^{(2)}$ , dan hasil  $\theta^{(2)}$  digunakan untuk iterasi ketiga  $\theta^{(3)}$ , hal terus berlanjut untuk iterasi berikutnya.
- h. Ulangi langkah d-g untuk 1000 iterasi dan untuk seluruh perusahaan penyusun portofolio.
- i. Hasil estimasi parameter

Hasil estimasi parameter mixture normal dapat diperoleh melalui perhitungan nilai rata-rata dari setiap hasil simulasi parameter pada 1000 iterasi.

**Tabel 10.** Tabel 10. Hasil Estimasi Parameter Distribusi Mixture Normal

| Perusahaan | $w_1$   | $w_2$   | $\mu_1$  | $\sigma_1$ | $\mu_2$  | $\sigma_2$ |
|------------|---------|---------|----------|------------|----------|------------|
| CPIN       | 0,28088 | 0,71910 | 0,00721  | 0,06145    | -0,00096 | 0,01819    |
| KLBF       | 0,21815 | 0,78185 | 0,02241  | 0,08541    | -0,00043 | 0,01633    |
| TBIG       | 0,72915 | 0,27085 | -0,00195 | 0,02492    | 0,01422  | 0,05830    |
| AMRT       | 0,73596 | 0,26404 | -0,00195 | 0,02410    | 0,01938  | 0,05493    |

Tabel 10 menunjukkan bahwa hasil estimasi melalui algoritma MH untuk seluruh parameter distribusi *mixture* normal cukup jauh berbeda dengan nilai parameter proposal pada Tabel 6. Hasil estimasi pada Tabel 10 juga menunjukkan bahwa komponen yang memiliki nilai  $\mu$  negatif memiliki nilai  $w$  yang lebih besar dan nilai  $\sigma$  yang lebih kecil dibandingkan dengan nilai  $\mu$  positif.

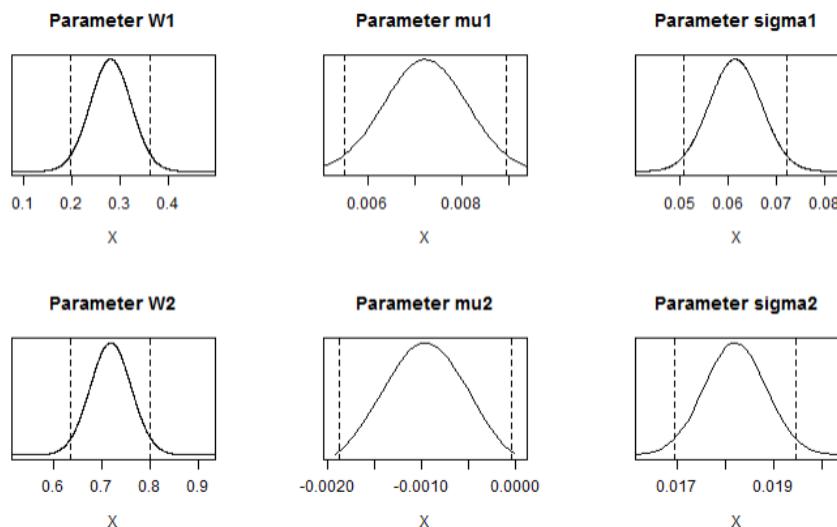


Gambar 4. Plot MH untuk Setiap Saham Portofolio

Hasil estimasi parameter distribusi mixture normal dalam bentuk grafik untuk data CPIN ditunjukkan pada Gambar 4. Bersejalan dengan hasil estimasi parameter bobot pada Tabel 10, pada perusahaan CPIN garis tipis untuk komponen kedua lebih tinggi dibandingkan dengan garis tebal untuk komponen pertama. Tabel 10 juga menunjukkan bahwa untuk perusahaan CPIN memiliki nilai bobot komponen 1 sebesar 28,09% dan komponen 2 sebesar 71,91%, untuk perusahaan KLBF memiliki nilai bobot komponen 1 sebesar 21,82% dan komponen 2 sebesar 78,19%, untuk perusahaan TBIG memiliki nilai bobot komponen 1 sebesar 72,92% dan komponen 2 sebesar 27,09%, untuk perusahaan AMRT memiliki nilai bobot komponen 1 sebesar 73,60% dan komponen 2 sebesar 26,40%.

j. Plot distribusi posterior

Plot distribusi posterior pada Gambar 5 merupakan hasil perhitungan pada persamaan (7). Gambar 5 menunjukkan selang kepercayaan 95% berdasarkan data dari distribusi posterior yang telah diperoleh untuk saham CPIN pada setiap parameter distribusi mixture normal.



Gambar 5. Selang Kepercayaan 95% Parameter Data CPIN.

Nilai interval pada Gambar 5 akan ditunjukkan lebih jelas pada Tabel 11.

**Tabel 11. Tabel 11.** Interval Konfidensi Estimasi Paramter Mixture Normal

| Saham | Parameter  | Nilai ( $\hat{\theta}$ ) | Ketidakpastian (95%) |
|-------|------------|--------------------------|----------------------|
| CPIN  | $w_1$      | 0,28088                  | 0,08285              |
|       | $w_2$      | 0,71910                  | 0,08285              |
|       | $\mu_1$    | 0,00721                  | 0,00172              |
|       | $\sigma_1$ | 0,06145                  | 0,01056              |
|       | $\mu_2$    | -0,00096                 | 0,00092              |
|       | $\sigma_2$ | 0,01819                  | 0,00125              |
| KLBF  | $w_1$      | 0,21815                  | 0,09931              |
|       | $w_2$      | 0,78185                  | 0,09931              |
|       | $\mu_1$    | 0,02241                  | 0,00186              |
|       | $\sigma_1$ | 0,08541                  | 0,00890              |
|       | $\mu_2$    | -0,00043                 | 0,00024              |
|       | $\sigma_2$ | 0,01633                  | 0,00137              |
| TBIG  | $w_1$      | 0,72915                  | 0,08761              |
|       | $w_2$      | 0,27085                  | 0,08761              |
|       | $\mu_1$    | -0,00195                 | 0,00049              |
|       | $\sigma_1$ | 0,02492                  | 0,00092              |
|       | $\mu_2$    | 0,01422                  | 0,00613              |
|       | $\sigma_2$ | 0,05830                  | 0,01156              |
| AMRT  | $w_1$      | 0,73596                  | 0,08350              |
|       | $w_2$      | 0,26404                  | 0,08350              |
|       | $\mu_1$    | -0,00195                 | 0,00092              |
|       | $\sigma_1$ | 0,02410                  | 0,00084              |
|       | $\mu_2$    | 0,01938                  | 0,00241              |
|       | $\sigma_2$ | 0,05493                  | 0,00806              |

Tabel 11 menunjukkan nilai ketidakpastian dari hasil estimasi pada setiap parameter distibusi normal yang diperoleh melalui penerapan algoritma MH. Tabel 11 memberikan kesimpulan 95% yakin bahwa nilai hasil estimasi parameter berada pada interval  $\hat{\theta}_i - 1,96(sd(\theta_i))$  dan  $\hat{\theta}_i + 1,96(sd(\theta_i))$ .

- k. Distribusi posterior digunakan untuk membuat prediksi tentang data masa depan atau untuk memperkirakan kuantitas yang tidak diketahui terkait dengan distribusi *mixture* normal.
- 7. Pembentukan portofolio optimal

Pembentukan portofolio optimal melalui metode MV dengan membentuk matriks  $A$  pada persamaan (9) untuk model *mixture* normal dibutuhkan perhitungan rata-rata dan variansi setiap saham, serta perhitungan kovarian menggunakan yang bersesuaian dengan persamaan (10) sampai (12). Sehingga memperoleh perhitungan yang ditunjukkan pada Tabel 12.

**Tabel 12.** Tabel 12. Statistika Deskriptif Berdasarkan Hasil Estimasi MH

| Saham | <i>mean</i> | <i>variance</i> | <i>StDev</i> |
|-------|-------------|-----------------|--------------|
| CPIN  | 0,00134     | 0,00131         | 0,03623      |
| KLBF  | 0,00456     | 0,00189         | 0,04346      |
| TBIG  | 0,00243     | 0,00143         | 0,03775      |
| AMRT  | 0,00369     | 0,00131         | 0,03623      |

Tabel 12 digunakan dalam perhitungan nilai bobot pada persamaan (9), hasil perhitungan tersebut ditunjukkan pada Tabel 13.

**Tabel 13.** Tabel 13. Bobot Portofolio Optimal

| Saham                            | MH     |
|----------------------------------|--------|
| Charoen Pokphand Indonesia Tbk   | 23,12% |
| Kalbe Farma Tbk                  | 21,27% |
| Tower Bersama Infrastructure Tbk | 25,93% |
| Sumber Alfaria Trijaya Tbk       | 29,68% |

Pada tabel 13 menunjukkan bahwa nilai bobot portofolio optimal yang terbentuk melalui estimasi parameter distribusi *mixture* normal menggunakan algoritma MH menghasilkan bobot yang hampir sama pada setiap saham, dengan bobot terbesar dimiliki oleh perusahaan AMRT dan bobot terkecil dimiliki oleh perusahaan KLBF.

#### 8. Backtesting

Perhitungan backtesting dapat dilakukan melalui penggunaan fungsi “portfolioBacktest” pada program R. Penggunaan perhitungan program R pada penelitian ini dijadikan sebagai alat perbandingan kinerja portofolio berdasarkan nilai bobot portofolio yang seimbang dengan nilai bobot portofolio optimal pada Tabel 14.

**Tabel 14.** Tabel 14. Hasil Backtesting melalui Program R

| Pengukuran        | Portfolio S | Portfolio MH |
|-------------------|-------------|--------------|
| Sharpe Ratio      | 0,072       | 0,075        |
| Max drawdown      | 0,077       | 0,086        |
| Annual return     | 0,248       | 0,269        |
| Annual volatility | 0,021       | 0,022        |

Tabel 14 menunjukkan bahwa nilai pengukuran pada portofolio MH cenderung lebih tinggi dibandingkan dengan nilai pengukuran dengan besar bobot yang seimbang. Hal tersebut menghasilkan kesimpulan yang sama dengan perhitungan manual yaitu portofolio optimal yang disusun oleh bobot hasil dari estimasi parameter *mixture* normal lebih baik dibandingkan dengan bobot seimbang.

## Pembahasan

Metode klasterisasi yang dilakukan didasarkan pada hasil perhitungan rasio keuangan pada setiap saham Bisnis-27 membentuk 4 klaster karakteristik yang berbeda pada setiap klasternya. Pembentukan portofolio yaitu dengan melakukan pemilihan satu saham pada setiap klaster terbentuk dengan jarak terdekat dengan pusat klaster itu sendiri. Pada klaster 1 didapatkan saham dari perusahaan Charoen Pokphand Indonesia Tbk dengan kode CPIN. Pada klaster 2 didapatkan saham dari perusahaan Kalbe Farma Tbk dengan kode KLBF. Pada klaster 3 didapatkan saham dari perusahaan Tower Bersama Infrastructure Tbk dengan kode. Pada klaster 4 didapatkan saham dari perusahaan Sumber Alfaria Trijaya Tbk dengan kode AMRT. Hasil klasterisasi berdasarkan rasio keuangan pada saham Bisnis-27 menghasilkan empat klaster dengan karakteristik berbeda, sejalan dengan temuan MacQueen (1967) bahwa pendekatan klaster mampu mengungkap struktur heterogen dalam data multivariat. Pemilihan saham representatif pada setiap klaster dilakukan berdasarkan kedekatan terhadap pusat klaster, sebagaimana juga dilakukan oleh Iriawan (2011) dalam pemodelan portofolio berbasis mixture.

Hasil dari penerapan metode MCMC pada pengestimasian parameter menggunakan algoritma MH diperoleh bahwa untuk seluruh parameter distribusi *mixture* normal cukup jauh berbeda dengan nilai parameter proposal. Nilai estimasi parameter  $w_1$  untuk saham CPIN dan KLBF mengalami penurunan sedangkan untuk saham TBIG dan AMRT mengalami peningkatan. Nilai estimasi parameter  $\mu_1$  dan  $\sigma_1$  untuk setiap saham mengalami peningkatan kecuali saham AMRT yang mengalami penurunan, begitu sebaliknya untuk parameter  $\mu_2$  dan  $\sigma_2$ . Hasil estimasi pada Tabel 10 juga menunjukkan bahwa komponen yang memiliki nilai  $\mu$  negatif memiliki nilai  $w$  yang lebih besar dan nilai  $\sigma$  yang lebih kecil dibandingkan dengan nilai  $\mu$  positif.

Pembentukan portofolio optimal dilakukan melalui perhitungan bobot menggunakan metode *mean-variance* berdasarkan parameter distribusi *mixture* normal menghasilkan kesimpulan bahwa portofolio optimal tersusun dari 23,12% saham Charoen Pokphand Indonesia Tbk, 21,27% saham Kalbe Farma Tbk., 25,93% saham Tower Bersama Infrastructure Tbk, dan 29,68% perusahaan Sumber Alfaria Trijaya Tbk. Hasil ini mendukung teori portofolio Markowitz (1952, 1959) dan diperkuat oleh Elton et al. (2014) bahwa diversifikasi berbasis estimasi distribusi risiko-return dapat menghasilkan portofolio yang lebih efisien dibanding bobot seimbang. Evaluasi menggunakan backtesting pada beberapa ukuran performa menunjukkan bahwa portofolio optimal tersebut memiliki kinerja yang lebih baik daripada portofolio berbobot seimbang. Temuan ini sejalan dengan Soelehan et al. (2014), Ngamar (2022), dan Landauskas & Valakevicius (2011) yang menyatakan bahwa penggunaan estimasi probabilistik seperti MCMC dapat menghasilkan portofolio dengan performa lebih stabil dan lebih baik dalam jangka panjang. Kinerja portofolio optimal tersebut diukur menggunakan metode *backtesting* dengan 4 macam pengukuran, metode ini dilakukan untuk membandingkan kinerja portofolio antara bobot seimbang dan bobot optimal yang telah diperoleh sebelumnya. Hasil dari pengukuran tersebut memberikan kesimpulan bahwa portofolio optimal yang terbentuk dari hasil estimasi parameter *mixture* normal melalui algoritma MH memiliki kinerja portofolio yang baik. Hal tersebut berartikan bahwa algoritma MH dalam metode MCMC yang digunakan sebagai alat estimasi distribusi mixture normal dapat menghasilkan bobot portofolio optimal untuk kinerja portofolio yang lebih baik dibandingkan dengan penggunaan bobot seimbang.

## SIMPULAN

Hasil penelitian yang telah dilakukan terhadap data Bisnis-27 dalam hasil estimasi MCMC menunjukkan bahwa untuk saham CPIN terdiri dari komponen 1 dengan bobot sebesar 28,09% dan komponen 2 sebesar 71,91%, untuk perusahaan KLBF terdiri dari komponen 1 dengan bobot sebesar 21,82% dan komponen 2 sebesar 78,19%, untuk saham TBIG terdiri dari komponen 1 dengan bobot sebesar 72,92% dan komponen 2 sebesar 27,09%, untuk saham AMRT terdiri dari komponen 1 dengan bobot sebesar 73,60% dan komponen 2 sebesar 26,40%. Penelitian ini memberikan kesimpulan bahwa portofolio optimal yang terbentuk dari hasil estimasi parameter *mixture* normal melalui algoritma MH memiliki kinerja portofolio yang baik, dengan portofolio optimal yang terdiri dari 23,12% saham Charoen Pokphand Indonesia Tbk, 21,27% saham Kalbe Farma Tbk., 25,93% saham Tower Bersama Infrastructure Tbk, dan 29,68% perusahaan Sumber Alfaria Trijaya Tbk. Penelitian selanjutnya dapat menggunakan algoritma MCMC yang lain seperti *Gibbs Sampler* dan simulasi numerik pada penelitian selanjutnya dapat menggunakan metode klasterisasi yang berbeda seperti K-Medoids dan menggunakan data dengan index yang berbeda seperti saham IDX-30, IDX-80, dan Kompas100. Hal tersebut dapat dilakukan sehingga dapat diperoleh perbandingan hasil penelitian.

## DAFTAR PUSTAKA

- Anoraga, P., & Pakarti, P. (2003). *Pengantar pasar modal* (Edisi revisi). Rineka Cipta.
- Bennett, M. J., & Hugen, D. L. (2016). *Financial analytics with R: Building a laptop laboratory for data science*. Cambridge University Press.
- Brigham, E. F., & Daves, P. R. (2004). *Intermediate financial management* (8th ed.). McGraw-Hill.
- Destiarina, U., Hadijati, M., KomalaSari, D., & Fitriyani, N. (2019). Estimasi parameter distribusi mixture eksponensial dan Weibull dengan metode Bayesian Markov Chain Monte Carlo. *Eigen Mathematics Journal*, 1(1), 34–48. <https://doi.org/10.29303/emj.v1i1.30>
- Elton, E. J., Gruber, M. J., Brown, S. J., & Goetzmann, W. N. (2014). *Modern portfolio theory and investment analysis* (9th ed.). John Wiley & Sons.
- Eraker, B., Johannes, M., & Polson, N. (2003). The impact of jumps in volatility and returns. *The Journal of Finance*, 57(3), 1269–1300.
- Everitt, B., & Hothorn, T. (2011). *An introduction to applied multivariate analysis with R*. Springer.
- Gamerman, D. (1997). *Markov Chain Monte Carlo: Stochastic simulation for Bayesian inference*. Chapman & Hall.
- Gelman, A., Carlin, J. B., Stern, H. S., Dunson, D. B., Vehtari, A., & Rubin, D. B. (2013). *Bayesian data analysis* (2nd ed.). CRC Press.
- Halim, A. (2005). *Analisa investasi* (Edisi ke-2). Salemba Empat.
- Horst, R. (2009). *The Weibull distribution: A handbook*. Justus-Liebig University Giessen.
- Iriawan, N. (2011). Pemodelan mixture of mixture dalam pemilihan portofolio. *Prosiding Seminar Nasional Statistika Universitas Diponegoro*.
- Jogiyanto, H. (2014). *Teori portofolio dan analisis investasi* (Edisi ke-8). BPFE.

- Johannes, M., Kumar, R., & Polson, N. (1999). *State dependent jump models: How do US equity indices jump?* Working Paper, University of Chicago.
- Landauskas, M., & Valakevicius, E. (2011). Modelling of stock prices by the Markov Chain Monte Carlo method. *Intellectual Economics*, 5(1), 244–256.
- MacQueen, J. (1967). Some methods for classification and analysis of multivariate observations. In *Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability* (Vol. 1, pp. 281–297). University of California Press.
- Markowitz, H. (1952). Portfolio selection. *The Journal of Finance*, 7(1), 77–91.
- Markowitz, H. (1959). *Portfolio selection*. Yale University Press.
- Munawir, S. (2004). *Analisis laporan keuangan* (Edisi ke-4). Liberty.
- Ngamar, A. I. (2022). *Pembentukan portofolio dengan menggunakan metode mean variance dan CM-estimasi (constrained M-estimates): Studi kasus saham IDX30* (Undergraduate thesis). Universitas Islam Negeri Walisongo Semarang.
- Okky, H. A., Susanto, I., & Slamet, I. (2022). Pengelompokan negara berdasarkan kasus stunting dengan model finite mixture normal menggunakan pendekatan Bayesian. *UNEJ E-Proceeding*, 190–200.  
<https://jurnal.unej.ac.id/index.php/prosiding/article/view/33508>
- Palomar, D. P., & Zhou, R. (2022). *portfolioBacktest* (Version 0.4.1) [R package].  
<https://github.com/dppalomar/portfolioBacktest>
- Richardson, S., & Green, P. J. (1997). On Bayesian analysis with an unknown number of components. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 59(4), 731–792.
- Setiawani, P. A., Dharmawan, K., & Sumarjaya, I. W. (2015). Implementasi metode Markov Chain Monte Carlo dalam penentuan harga kontrak berjangka komoditas. *E-Jurnal Matematika*, 4(3), 122–126. <https://doi.org/10.24843/Mtk.2015.v04.i03.p099>
- Soelehan, A., Yudha, A., & Chandra, B. D. (2014). Analisis tingkat risiko dan imbal hasil saham terhadap pembentukan portofolio optimal. *Jurnal Ilmiah Manajemen Kesatuan*, 2(3), 191–199. <https://doi.org/10.37641/jimkes.v2i3.802>
- Stephens, M. (2004). Bayesian analysis of mixture models with an unknown number of components: An alternative to reversible jump methods. *Annals of Statistics*, 32(1), 40–74.
- Walpole, R. E., & Myers, R. H. (1995). *Ilmu peluang dan statistika untuk insinyur dan ilmuwan* (Edisi ke-4). Penerbit ITB.