

PERBANDINGAN KOEFISIEN VARIASI ANTARA 2 SAMPEL DENGAN METODE BOOTSTRAP

(Studi Kasus Pada Analisis Inflasi Bulanan Komoditas Beras, Cabe Merah Dan Bawang Putih Di Kota Semarang)

Adi Setiawan

**Program Studi Matematika, Fakultas Sains dan Matematika, Universitas Kristen Satya
Wacana Jl. Diponegoro 52-60 Salatiga 50711, adi_setia_03@yahoo.com**

Abstrak

Untuk membandingkan persebaran data antara 2 kelompok data yang mempunyai satuan yang berbeda tidak bisa dilakukan hanya menggunakan simpangan baku tetapi menggunakan perbandingan koefisien variasi antara 2 kelompok data tersebut. Dalam makalah ini dijelaskan tentang bagaimana melakukan perbandingan koefisien variasi antara 2 sampel dengan menggunakan metode bootstrap. Metode yang diusulkan dikerjakan pada data inflasi bulanan komoditas beras, cabe merah dan bawang putih. Studi simulasi digunakan untuk memberikan gambaran jenis sampel yang seperti apa yang mengakibatkan koefisien variasi antara kedua sampel berbeda secara signifikan.

Kata kunci : koefisien variasi, metode bootstrap, indeks harga konsumen, inflasi.

Abstract

For comparing the distribution of data between the two sets of data have different units can not be done using only the standard deviation but using the coefficient of variation comparisons between the 2 groups of data. In this paper explained about how to perform the comparison the coefficient of variation between two samples using the bootstrap method. The proposed method works in the monthly inflation data for commodities rice, red pepper and garlic. Simulation studies are used to illustrate what kind of samples that resulted in a coefficient of variation between the two samples differ significantly.

Key Words : coefficient of variation, the bootstrap method, the consumer price index, inflation.

1. Pendahuluan

Indeks harga konsumen IHK suatu komoditas merupakan perbandingan antara harga komoditas tersebut pada suatu periode waktu terhadap harganya pada periode waktu yang telah ditentukan (tahun dasar). Dengan mengetahui IHK komoditas-komoditas dari suatu kelompok barang atau jasa beserta dengan bobot konsumsi dari komoditas tersebut maka akan dapat ditentukan IHK pada suatu periode waktu yang akan dapat digunakan untuk menentukan besarnya inflasi. Salah satu statistik yang digunakan dalam mengukur fluktuasi inflasi adalah koefisien variasi dari inflasi. Koefisien variasi yang relatif kecil menunjukkan bahwa data berpecah di sekitar rata-rata dengan sebaran yang kecil. Karena koefisien variasi tidak berdimensi maka kita bisa membandingkan koefisien variasi antara 2 sampel yang mempunyai satuan yang berbeda.

Karena distribusi dari koefisien variasi sampel sebagai estimasi dari koefisien variasi populasi sulit ditentukan maka distribusi dari koefisien variasi dan selisih antara koefisien variasi 2 sampel ditentukan dengan metode bootstrap. Metode bootstrap juga digunakan dalam penerapan koefisien variasi pada paper Amiri dan Zwanzig. [1] [2] Dalam makalah ini akan dijelaskan tentang bagaimana melakukan perbandingan koefisien variasi antara 2 sampel dengan menggunakan metode bootstrap.

2. Dasar Teori

Koefisien variasi (*coefficient of variation*) atau koefisien dispersi adalah ukuran persebaran yang dinormalkan dari suatu distribusi probabilitas. Kadang-kadang nilai dari koefisien variasi dinyatakan dalam prosen. [3] Harga mutlak dari koefisien variasi kadang-kadang dikenal dengan nama simpangan baku relatif (*relative standard deviation – RSD*). Koefisien variasi didefinisikan sebagai rasio dari standard deviasi σ dengan mean μ yaitu

$$c_v = \frac{\sigma}{\mu}$$

dan estimasi dari koefisien variasi digunakan

$$\hat{c}_v = \frac{s}{\bar{x}}$$

dengan $\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i$ dan $s = \sqrt{s^2} = \sqrt{\frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n-1}}$.

Contoh 1

Tabel 1. IHK komoditas beras di kota Semarang periode bulan Januari 2002 sampai dengan Desember 2002.

Bulan	Jan	Feb	Mar	Apr	Mei	Juni	Juli	Agt	Sept	Okt	Nop	Des
IHK	109,0	111,7	101,7	99,2	100,9	99,5	95,2	92,6	93,8	94,8	99,8	101,9
Inf		2,7	-10,0	-2,5	1,7	-1,4	-4,3	-2,5	1,1	1,1	5,0	2,1

Misalkan dimiliki data IHK komoditas beras di Kota Semarang periode bulan Januari 2002 sampai dengan Desember 2002 pada Tabel 1. Berdasarkan IHK tersebut dapat ditentukan inflasi bulanan komoditas beras (dalam persen) periode bulan Februari 2002 sampai dengan Desember 2002 pada baris ketiga. Karena simpangan baku dan rata-rata dari inflasi komoditas beras untuk periode Februari 2002 sampai dengan Desember 2002 masing-masing adalah 4,13 dan -0,65 maka koefisien variasi dari inflasi bulanan komoditas beras adalah -6,37. Dalam hal ini, koefisien variasi dari inflasi bulanan komoditas beras bernilai negatif yaitu -6,37 sehingga hal itu berarti rata-ratanya bernilai negatif sehingga dalam periode tersebut, komoditas beras cenderung mengalami deflasi (inflasi yang bernilai negatif) untuk setiap bulannya.

Untuk memberikan perbandingan antara koefisien variasi sampel 1 dan koefisien variasi sampel 2 maka pada Tabel 2 diberikan IHK komoditas beras periode Juni 2004 sampai dengan Mei 2005 sehingga dapat diperoleh inflasi komoditas beras periode bulan Juli 2004 sampai dengan Desember 2004. Koefisien variasi untuk inflasi komoditas beras periode ini adalah 2,41 sehingga rata-ratanya bernilai positif. Hal itu berarti dalam periode tersebut, komoditas beras cenderung mengalami inflasi. Di samping itu, harga mutlak koefisien variasi ini lebih kecil dari harga mutlak koefisien variasi untuk inflasi komoditas beras periode bulan Februari 2002 sampai dengan Desember 2002. Hal itu berarti bahwa inflasi bulanan komoditas beras di kota Semarang pada tahun 2002 cenderung lebih bervariasi dibandingkan pada periode Juli 2004 sampai Mei 2005 sehingga dengan kata lain inflasi bulanan komoditas beras pada tahun 2002 cenderung lebih berfluktuasi. Pertanyaan yang muncul adalah apakah terdapat perbedaan yang signifikan antara koefisien variasi inflasi komoditas beras periode Februari 2002 sampai dengan Desember 2002 dengan koefisien variasi inflasi komoditas beras periode bulan Juni 2004 sampai dengan Mei 2005? Hal itu berarti ingin dilakukan pengujian hipotesis nul : $H_0 : c_{v;1} = c_{v;2}$ atau $H_0 : c_{v;1} - c_{v;2} = 0$ yaitu apakah koefisien variasi sampel pertama yang diambil dari populasi pertama sama dengan koefisien variasi sampel kedua yang diambil dari populasi kedua.

Tabel 2. IHK komoditas beras di Jawa Tengah periode bulan Juni 2004 sampai dengan Mei 2005.

Bulan	Jun	Juli	Agt	Sept	Okt	Nop	Des	Jan	Feb	Mar	Apr	Mei
IHK	97,9	97,9	97,1	97,0	98,1	100,8	101,5	106,77	109,49	111,66	108,70	108,05
Inf		0,0	-0,8	-0,1	1,1	2,7	0,7	5,24	2,72	2,17	-2,96	-0,65

3. Metode Penelitian

Data yang diperoleh adalah data IHK komoditas beras, cabe merah dan bawang putih di kota Semarang pada periode bulan Januari 2002 sampai dengan Desember 2007 berdasarkan data BPS. Berdasarkan data IHK komoditas beras, cabe merah dan bawang putih. Selanjutnya, ditentukan inflasi bulanan komoditas masing-masing komoditas dan koefisien variasi untuk periode bulan Februari 2002 sampai dengan Desember 2007. Berdasarkan koefisien variasi untuk masing-masing komoditas maka dapat ditentukan dan dilakukan perbandingan apakah koefisien variasi antara inflasi bulanan komoditas beras berbeda secara signifikan dengan koefisien variasi bulanan komoditas cabe merah dan juga untuk pasangan komoditas yang lain yaitu antara komoditas cabe merah dengan bawang putih dan antara komoditas beras dengan komoditas bawang putih.

Apabila diinginkan untuk melakukan pengujian hipotesis nul : $H_0 : c_{v;1} = c_{v;2}$ atau $H_0 : c_{v;1} - c_{v;2} = 0$ yaitu apakah koefisien variasi sampel pertama yang diambil dari populasi pertama sama dengan koefisien variasi sampel kedua yang diambil dari populasi kedua. Hipotesis nul ditolak jika nilai-p dari hipotesis nul lebih kecil dari tingkat signifikansi (*level of significance*) α yang dipilih. Karena statistik yang digunakan merupakan statistik cukup sulit untuk ditentukan distribusinya maka digunakan pendekatan bootstrap [4] yang dengan menggunakan pendekatan metode bootstrap, nilai-p ditentukan dengan langkah-langkah sebagai berikut :

1. Misalkan dimiliki sampel 1 ukuran n yaitu x_1, x_2, \dots, x_n dan sampel 2 ukuran m yaitu y_1, y_2, \dots, y_m . Berdasarkan pada sampel 1 dapat ditentukan koefisien variasi $c_{v;1}$ dan berdasarkan sampel 2 dapat ditentukan koefisien variasi $c_{v;2}$. Selanjutnya ditentukan statistik $T = |c_{v;1} - c_{v;2}|$.
2. Sampel 1 dan sampel 2 digabung menjadi satu sehingga membentuk sampel gabungan yaitu $x_1, x_2, \dots, x_n, y_1, y_2, \dots, y_m$. Berdasarkan sampel gabungan diambil dengan pengembalian sampel 1 baru ukuran n yaitu $x_1^*, x_2^*, \dots, x_n^*$ dan sampel 2 baru ukuran m yaitu $y_1^*, y_2^*, \dots, y_m^*$. Berdasarkan pada sampel 1 baru dapat ditentukan koefisien variasi $c_{v;1}^*$ dan berdasarkan sampel 2 baru dapat ditentukan koefisien variasi $c_{v;2}^*$. Selanjutnya ditentukan statistik $T^* = |c_{v;1}^* - c_{v;2}^*|$.
3. Prosedur 2 diulang sebanyak bilangan besar B kali sehingga diperoleh $T_1^*, T_2^*, \dots, T_B^*$.
4. Nilai-p dihitung dengan proporsi dari $T_1^*, T_2^*, \dots, T_B^*$ yang lebih besar dari T .

Karena hipotesis alternatifnya merupakan hipotesis 2 sisi maka hipotesis nul akan ditolak jika nilai-p pendekatan lebih kecil dari $\alpha/2$ yang dipilih.

Studi simulasi digunakan untuk mengecek apakah dapat diperoleh kesimpulan yang menolak hipotesis nol jika dibangkitkan 2 sampel dari 2 populasi yang mempunyai distribusi dengan koefisien variasi yang relatif jauh berbeda dengan yang lain. Demikian juga diperoleh kesimpulan yang menerima hipotesis nol jika dibangkitkan 2 sampel dari 2 populasi yang mempunyai distribusi dengan koefisien variasi yang relatif dekat (relatif sama) dengan yang lain. Studi simulasi S1 dilakukan dengan prosedur sebagai berikut :

1. Mengambil sampel 1 ukuran $n = 50$ yaitu x_1, x_2, \dots, x_{50} dari distribusi normal dengan mean $\mu = 1$ dan simpangan baku $\sigma_1 = 3$.
2. Mengambil sampel 2 ukuran $m = 50$ yaitu y_1, y_2, \dots, y_{50} dari distribusi normal dengan mean $\mu = 1$ dan simpangan baku $\sigma_2 = 3$.
3. Menentukan nilai-p bootstrap untuk menguji hipotesis nol $H_0 : c_{v;1} = c_{v;2}$ dengan prosedur di atas dengan menggunakan $B = 10.000$.

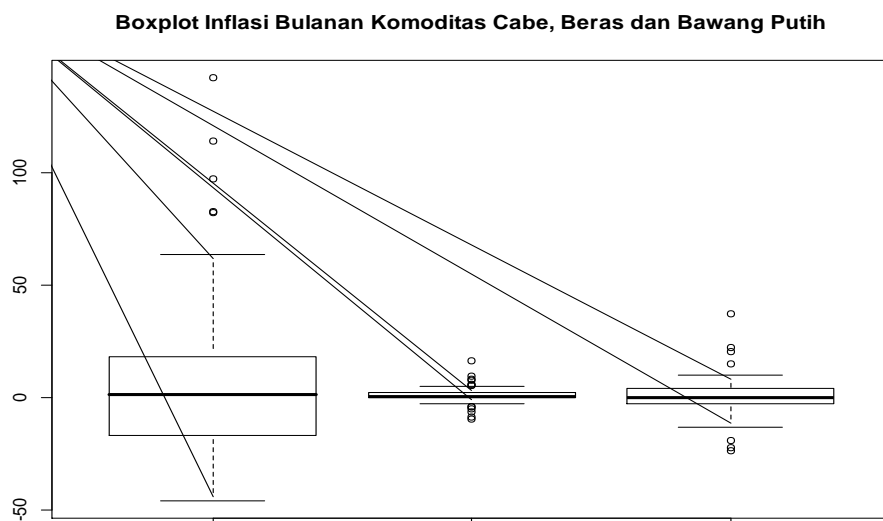
Mengulangi prosedur simulasi S1 dengan menggunakan ukuran sampel $n = m = 100, 500, 1000$ dan 5000 dengan melakukan perubahan $\sigma_2 = 1, 1.2, 1.4, 1.6, 1.8, 2.2, 2.4, 2.6, 2.8, 3.2, 3.4, 3.6, 3.8, 4.0, 4.2$.

4. Hasil Analisis dan Pembahasan

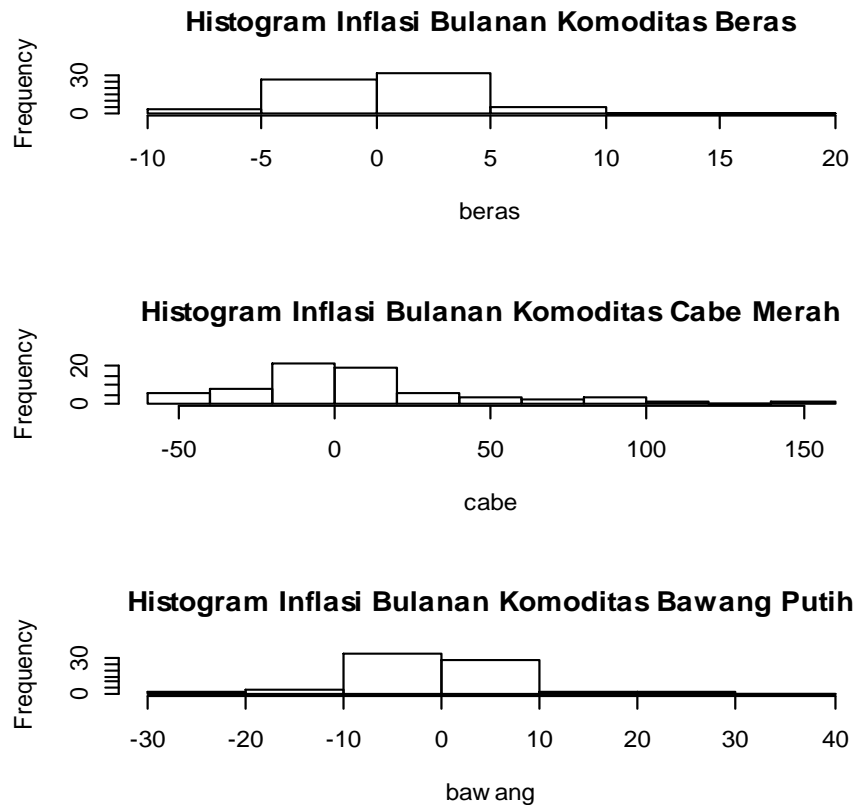
Tabel 1. Statistik deskriptif inflasi bulanan komoditas beras, cabe merah dan bawang putih periode Februari 2002 sampai dengan Desember 2007.

Statistik	Beras	Cabe Merah	Bawang Putih
Minimum	-9.46	-46	-23.9
Kuartil 1	-0.11	-17.08	-3.02
Median	0.49	1.19	0
Mean/Rata-rata	0.79	6.381	0.58
Kuartil 3	1.99	17.87	4.04
Maksimum	16.2	142.2	36.98
Simpangan baku	3.68	37.73	8.89
Skewness	0.72	1.36	0.64
Kurtosis	5.06	2.24	4.6

Tabel 1 menyatakan hasil statistik deskriptif dari ketiga komoditas yang menjadi perhatian yaitu beras, cabe merah dan bawang putih. Inflasi bulanan komoditas cabe merah mempunyai jangkauan yang lebih besar dibandingkan dengan yang lain. Karena untuk ketiga komoditas cukup banyak mempunyai *outlier* (Gambar 1) maka akan lebih baik menggunakan median untuk menyatakan kecenderungan memusat dari data inflasi bulanan ketiga komoditas tersebut. Hal itu berarti, inflasi bulanan komoditas beras, cabe merah dan bawang putih berturut-turut berkisar pada berkisar pada 0,49, 1.19 dan 0 persen. Kemencengan (*skewness*) dari inflasi bulanan komoditas beras, cabe merah dan bawang putih berturut-turut adalah 0.72, 1.36 dan 0.64 sehingga berarti kemencengan untuk inflasi bulanan ketiga komoditas bernilai positif. Hal itu berarti bahwa inflasi bulanan untuk masing-masing komoditas lebih banyak yang bernilai relatif kecil dibandingkan yang bernilai besar. Hal tersebut juga didukung oleh histogram inflasi bulanan untuk ketiga komoditas tersebut yang diberikan pada Gambar 2.



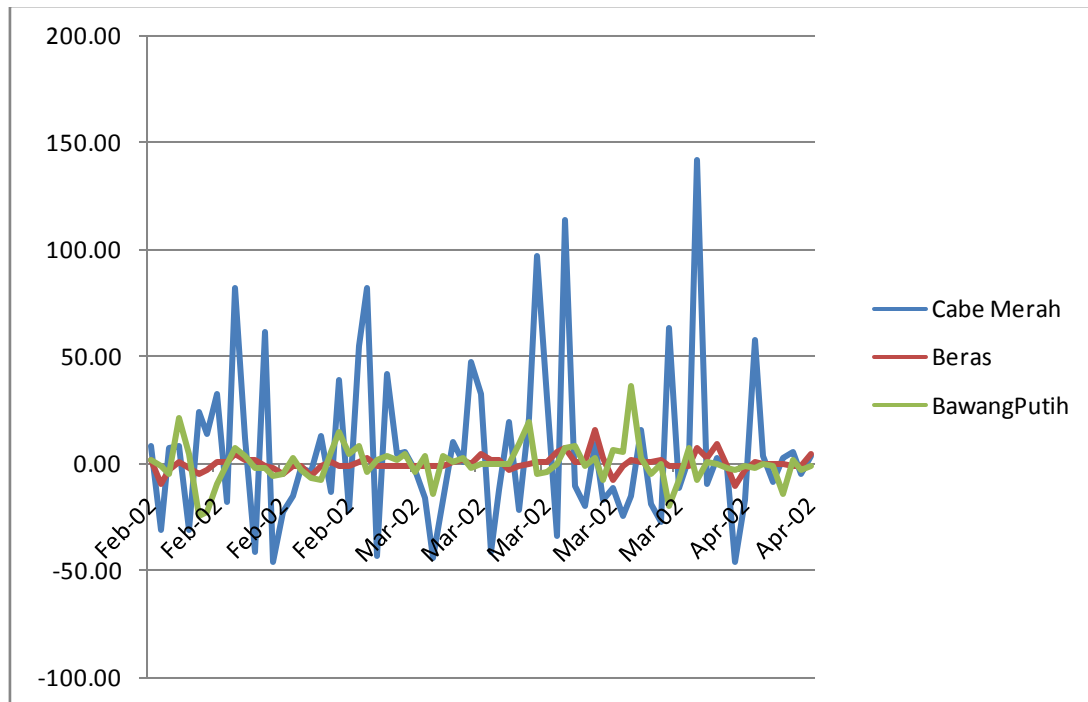
Gambar 1. Boxplot inflasi bulanan komoditas beras, cabe merah dan bawang putih periode Februari 2002 sampai dengan Desember 2007.



Gambar 2. Histogram inflasi bulanan komoditas beras, cabe merah dan bawang putih periode Februari 2002 sampai dengan Desember 2007.

Gambar 3 memperlihatkan perbandingan inflasi bulanan komoditas beras dan inflasi komoditas cabe merah dan perbandingan inflasi komoditas cabe merah dan inflasi komoditas bawang putih kota Semarang periode Februari 2002 sampai Desember 2007. Berdasarkan besaran inflasi pada Gambar 3, nampak seperti terlihat bahwa inflasi bulanan komoditas cabe merah lebih berfluktuasi dibandingkan dengan inflasi bulanan komoditas beras dan bawang putih sedangkan inflasi bulanan komoditas bawang putih lebih berfluktuasi dibandingkan inflasi bulanan komoditas beras. Akan tetapi inflasi bulanan komoditas beras periode bulan Februari 2002 sampai dengan Desember 2007, mempunyai koefisien variasi 4,63 sedangkan untuk periode yang sama, komoditas cabe merah dan bawang putih berturut-turut mempunyai koefisien variasi 5,91 dan 15,22. Karena koefisien variasi inflasi komoditas cabe merah lebih besar dari pada koefisien variasi inflasi komoditas beras maka hal itu berarti bahwa inflasi bulanan komoditas cabe merah lebih berfluktuasi dibandingkan dengan inflasi bulanan komoditas beras sedangkan inflasi komoditas bulanan bawang putih lebih berfluktuasi dibandingkan dengan inflasi bulanan komoditas cabe merah dan beras.

Dalam hal ini, akan dilakukan pengujian apakah koefisien variasi dari inflasi bulanan komoditas cabe merah berbeda secara signifikan dengan koefisien variasi dari inflasi bulanan komoditas bawang putih maupun koefisien variasi dari inflasi bulanan komoditas beras. Dengan menggunakan metode bootstrap diperoleh bahwa nilai-p pendekatan untuk hipotesis nol pertama dan hipotesis nol kedua berturut sebesarnya 0,4773 dan 0,8439. Hal itu berarti koefisien variasi antara inflasi bulanan komoditas cabe merah dan koefisien variasi antara inflasi bulanan komoditas bawang putih tidak berbeda secara signifikan dan kesimpulan yang sama juga berlaku untuk koefisien variasi antara inflasi bulanan komoditas cabe merah dan koefisien variasi antara inflasi bulanan komoditas beras.



Gambar 3. Perbandingan inflasi bulanan komoditas cabe merah, beras dan bawang putih kota Semarang periode Februari 2002 sampai Desember 2007.

Hasil studi simulasi dengan menggunakan prosedur studi simulasi yang telah dijelaskan di atas dinyatakan pada Tabel 3. Dengan membuat perubahan pada simpangan baku σ_2 dari distribusi normal yang digunakan untuk membangkitkan sampel kedua maka koefisien variasi pada sampel kedua diharapkan akan cenderung hanya tergantung pada σ_2 yang dipilih bervariasi dari $\sigma_2 = 1$ sampai dengan $\sigma_2 = 5$. Semakin dekat σ_2 dengan nilai $\sigma_1 = 3$ maka akan semakin kecil kemungkinan untuk menolak hipotesis nol atau nilai-p cenderung lebih besar dari tingkat signifikansi α yang dipilih. Terlihat bahwa untuk ukuran sampel $n = 100$, belum ada pasangan sampel yang mengakibatkan hipotesis nol ditolak, sedangkan untuk ukuran sampel $n = 500$, pada saat $\sigma_2 = 1, 1.2, 1.4$, seperti yang kita harapkan hipotesis nol sudah ditolak. Demikian juga, seperti yang kita harapkan untuk semakin besar ukuran sampel n maka akan semakin banyak σ_2 di antara antara $\sigma_2 = 1$ sampai dengan $\sigma_2 = 5$ yang ditolak. Untuk hipotesis nol yang ditolak berarti bahwa koefisien variasi antara kedua sampel berbeda secara signifikan.

Mengingat data inflasi bulanan ketiga komoditas mempunyai cukup banyak *outlier*, dapat juga digunakan MAD (*Median Absolut Deviation*) sebagai pengganti simpangan baku dan Median sebagai pengganti mean dalam estimasi koefisien variasi yang mempunyai sifat *robust*. Pemilihan statistik yang diusulkan tetap memperhatikan lama waktu yang diperlukan untuk menghitung nilai-p dengan pendekatan bootstrap dan mempertimbangkan kemungkinan median yang bernilai nol sehingga koefisien variasi menjadi tidak terdefiniskan.

Tabel 3. Hasil simulasi berdasarkan prosedur studi simulasi S1.

σ_2	$n = 50$	$n=100$	$n=500$	$n=1000$	$n=5000$
1	0.1809	0.0616	0	0	0
1.2	0.2461	0.0997	0.0001	0	0
1.4	0.3104	0.1555	0.0014	0	0
1.6	0.3672	0.2337	0.0065	0.0002	0
1.8	0.4256	0.2988	0.0212	0.003	0
2	0.4827	0.3679	0.0967	0.0087	0
2.2	0.4728	0.4857	0.1694	0.0717	0
2.4	0.4736	0.4478	0.3192	0.1574	0.0049
2.6	0.5144	0.4599	0.4108	0.3261	0.0694
2.8	0.5818	0.4423	0.4715	0.4628	0.3539
3	0.5497	0.5248	0.5177	0.5247	0.5085
3.2	0.5198	0.5169	0.4872	0.4292	0.4004
3.4	0.5130	0.496	0.4802	0.4241	0.1231
3.6	0.5228	0.4686	0.4274	0.3588	0.0426
3.8	0.4980	0.4540	0.4267	0.2699	0.0229
4	0.4734	0.4573	0.2353	0.1811	0.0043
4.2	0.4676	0.4723	0.2677	0.1492	0.0008
4.4	0.5586	0.5022	0.2849	0.1079	0.0003
4.6	0.5356	0.4684	0.1825	0.0897	0.0001
4.8	0.5242	0.4823	0.2046	0.0752	0
5	0.4393	0.4935	0.1652	0.0526	0

5. Kesimpulan

Dalam makalah ini telah dijelaskan bagaimana melakukan pengujian hipotesis untuk menguji apakah koefisien variasi antara 2 sampel sama atau berbeda secara signifikan. Di samping itu, berdasarkan data nyata yang dimiliki, koefisien variasi antara 2 sampel dari inflasi bulanan dari ketiga komoditas yang menjadi perhatian yaitu komoditas beras, cabe merah dan bawang putih dapat dikatakan tidak berbeda secara signifikan. Studi simulasi digunakan untuk memberikan gambaran pada sampel yang seperti apa mengakibatkan koefisien variasi antara kedua sampel berbeda secara signifikan.

6. Daftar Pustaka

- [1] Amiri, S., Zwanzig, S. (2010). An Improvement of the Nonparametric Bootstrap Test for the Comparison of the Coefficient of Variations. *Communications in Statistics - Simulation and Computation*, 39, 1726-1734
- [2] Amiri, S., Zwanzig, S. (2011). Assessing the Coefficient of Variations of Chemical Data using Bootstrap Method. *Journal of Chemometrics*.
- [3] Harinaldi, 2005, *Prinsip-prinsip Statistik untuk Teknik dan Sains*, Penerbit Erlangga, Jakarta.
- [4] Setiawan, Adi, 2002, *Analisis Data Statistika*, FSM Universitas Kristen Satya Wacana, Salatiga.