Pengujian Kesamaan Distribusi Multivariat dengan Menggunakan Uji Nonparametrik Data Depth pada Nilai UNBK Sekolah Menengah Atas Kota Sukabumi

Similarity Test for Multiple Multivariate Distributions Using Nonparametric test with Data Depth Applied in UNBK Score In Senior High School of Sukabumi

[1]Siti Sarah Sobariah Lestari, [2].Suwanda, , [3]Abdul Kudus
[1,2,3]Prodi Statistika, Fakultas Matematika dan Ilmu Pengetahuan Alam, Universitas Islam Bandung, Jl. Tamansari No.1 Bandung 40116
email:[1]sobariahara11696@gmai.com [2]idris100358@gmail.com [3] akudus69@unisba.ac.id

Abstract. A nonparametric test for multivariate analysis have been developed by using component wise signs or rank. However these method give some lack of result or in other words it is not affine invariant. Chenouri S and G.Christopher (2012) developed Kruskal-Wallis test by using statistical depth function (Mahalanobis depth and halfspace) in determinig depth rank that computes the depth on distance of the data to the data center that provides an affine invariant result in testing similarity of multivariate distribution. In this paper we are going to apply Mahalanobis depth methods to compute rank of multivariate data and nonparametric test (Kruskal-Wallis) based on data depth for testing the similarity of some multivariate distribution between two samples of UNBK Score in one of senior high school in Sukabumi city at 2015/2016 school year with 2017/2018 school year. The data sets contains by three variables ($X_1 =$ Bahasa, $X_2 =$ Math, and $X_3 =$ English). So that the value of statistic test H is 123,50027 and after comparison with ${\chi^2}_{1(0,05)} = 3.8415$ we can conclude that distribution of two samples are different.

Keywords: Multivariat Nonparametric Test, Data Depth, Kruskal-Wallis Test.

Abstrak. Pengembangan terhadap pengujian nonparametrik pada analisis multivariat telah banyak dikembangkan dengan menghitung *component wise* berdasarkan tanda atau ranking. Namun demikian masih terdapat kekurangan pada metode yang telah ditawarkan, salah satunya hasil analisis yang tidak *affine invariant*. Chenouri S dan G. Christopher (2012) mengembangkan pengujian Kruskal-Wallis dengan memanfaatkan *statistical depth function* (*Mahalanobis depth* dan *halfspace method*) dalam penentuan ranking yang menghitung kedalaman pada jarak data terhadap pusat data yang mampu memberikan hasil pengujian yang *affine invariant* dalam melakukan pengujian kesamaan beberapa distribusi multivariat. Dalam skripsi ini akan dilihat penerapan metode *Mahalanobis depth* untuk menentukan ranking dari data multivariat dan uji statistik nonparametrik Kruskal-Wallis yang berdasarkan kepada *data depth* untuk menguji apakah terdapat perbedaan distribusi pada nilai UNBK pada salah satu SMAN Kota Sukabumi tahun pelajaran 2015/2016 dengan tahun pelajaran 2017/2018 dimana dari data nilai UNBK tersebut diambil nilai siswa pada program IPA dengan membandingkan tiga mata pelajaran yang menjadi mata pelajaran wajib dalam UNBK yakni X_1 = Bahasa Indonesia, X_2 = Matematika, dan X_3 = Bahasa Inggris. Sehingga diperoleh nilai statistik uji H sebesar 123,50027 dengan nilai kritis ${\chi^2}_{1(0,05)} = 3.8415$, maka disimpulkan bahwa dua buah distribusi mulrivariat nilai UNBK tersebut adalah berbeda.

Kata Kunci: Uji Multivariat Nonparametrik, *Data Depth,* Uji Kruskal-Wallis.

A. Pendahuluan

Analisis multivariat memiliki peran penting dalam perkembangan ilmu statistik. Analisis ini bertujuan untuk meminimalisir penggunaan hanya satu variabel dalam melakukan pengukuran. Akan tetapi, analisis ini sangat bergantung kepada asumsi normal multivariat. Sedangkan pada praktiknya asumsi ini cukup sulit terpenuhi sehingga apabila asumsi tersebut tidak terpenuhi harus dilakukan pengujian nonparametrik pada data multivariat. Pengembangan terhadap pengujian nonparametrik pada data multivariat telah banyak dikembangkan salah satunya dengan menghitung rank dengan component-wise. Namun demikian metode ini tidak memberikan hasil yang *affine invariant*.

Chenouri S dan G. Christopher (2012) mengembangkan pengujian nonparametrik pada data multivariat dengan menentukan ranking berdasarkan data *depth* dengan perhitungan *statistical depth function* seperti *Mahalanobis depth* atau *halfspace* yang mana perhitungan tersebut memberikan hasil affine invariant. Dalam makalah ini akan digunakan perhitungan *statistical depth function* yakni *Mahalanobis depth* hal ini berdasarkan penelitian yang dilakukan oleh Zuo (2000) bawasanya *Mahalanobis depth* merupakan salah satu *statistical depth function* yang memenuhi kriteria untuk digunakan dalam perhitungan nilai *depth*.

Pendekatan data *depth* dapat mengidentifikasi letak data dengan titik pusat sehingga skala multivariat dapat dikarakterisasi dengan baik oleh data *depth*. Liu dan Singh menggunakan konsep data *depth* untuk melakukan perluasan uji Kruskal-Wallis univariat ke model multivariat. Dengan mengadaptasi konsep tersebut dalam kumpulan data multivariat, dapat memperluas aplikasi analisis multivariat, karena dapat membuka kemungkinan metode nonparametrik yang akan digunakan dalam analisis data multivariat.

Berdasarkan latar belakang yang telah diuraikan, maka perumusan masalah dalam penelitian ini sebagai berikut: “Bagaimana prosedur penentuan nilai rank pada data multivariat dan prosedur perhitungan statistik uji Kruskal-Wallis dengan perhitungan *Mahalanobis depth* pada kasus data tertentu?”. Selanjutnya, tujuan dalam penelitian ini diuraikan dalam pokok-pokok sbb.

1. Untuk mengetahui prosedur penentuan nilai rank pada data multivariat dengan menggunakan perhitungan *Mahalanobis depth*.
2. Untuk mengetahui prosedur perhiutngan statistik uji Kruskal-Wallis dengan menggunakan nilai rank dari *Mahalanobis depth* dan penerapannya pada data nilai UNBK salah satu SMA di Kota Sukabumi.

B. Landasan Teori

Analisis multivariat muncul ketika peneliti memiliki n buah amatan dan setiap amatan dilakukan pengukuran p buah peubah (Hajarisman, 2008). Gugus data dalam multivariat berbentuk matriks dengan baris ke-i $(i = 1,2,\dots,n)$, sedangkan kolom ke-j merupakan peubah ke-$j(j = 1,2,\dots,3)$ dengan struktur matriks sebagai berikut :

$$\begin{bmatrix} x_{11} & \cdots & x_{1p} \\ \vdots & \ddots & \vdots \\ x_{n1} & \cdots & x_{np} \end{bmatrix}$$

Katakanlah $X \sim F(x)$ dimana X mengikuti distribusi $F(x)$ sehingga penaksir untuk rata-rata dan variansnya adalah sebagai berikut

Penaksir rata-rata :

$$\hat{\mu} = \frac{1}{n}\sum_{i=1}^{n} x_i = \begin{pmatrix} \bar{x}_1 \\ \bar{x}_2 \\ \vdots \\ \bar{x}_p \end{pmatrix} \tag{1}$$

Penaksir Varians :

$$\hat{\Sigma} = \frac{1}{n-1}\sum (x_i - \bar{x})(x_i - \bar{x})' \tag{2}$$

Chenouri dan Christopher (2012), membuat tes multivariat yang asimtotik dengan distribusi bebas bersyarat multivariat berdasarkan data *depth*. Selain itu, uji yang dikembangkan dibuat agar hasil yang diberikan *affine equivalen*. Dalam pengaturan standar, statistik uji pada metode ini memiliki distribusi asimtotik chi-square dengan derajat bebas $t-1$. Dimana metode ini dikombiasikan dengan Sebuah teknik grafis untuk membandingkan dua distribusi yang dikenal sebagai depth-depth plot, atau DD-plot, diperkenalkan oleh Liu et al. (1999). Metode ini memiliki beberapa keuntungan. Diantaranya, memungkinkan peneliti untuk memilih *depth function* dengan menggabungkan konsep *equivariant*, *robustness* dan komputasi. Dengan kata lain, dapat menyesuaikan berdasarkan konteksnya. Kedua, metode ini didasarkan pada peringkat multidimensi pada data dengan menggunakan perhitungan kedalaman data ke X_i dengan pusat data. Sehingga Perankingan Statistik memberikan keuntungan sebab mereka berdimensi tunggal. Penggunaan peringkat ini sangat berguna jika data analisis diolah dengan teknik pereduksi dimensi seperti PCA.

Fungsi *depth* menjadi fungsi nonnegatif debatasi dari bentuk $D: \mathbb{R}^p \times F \rightarrow \mathbb{R}$ untuk setiap vektor acak $\boldsymbol{Y}$ dimana F_Y menunjukkan distribusi Y. Fungsi *depth* D akan dikatakan *affine invariant* jika $D\left(A_x + b, F_{A_x+b}\right) = D\,(x, F_x)$ berlaku untuk vektor acak sembarang $\boldsymbol{X}$ dalam $\mathbb{R}^p$ untuk setiap $d \times d$ merupakan matriks nonsingular A dan setiap vektor $\boldsymbol{x}$ dan $\boldsymbol{b}$ di $\mathbb{R}^p$(Zuo, 2000).

Jarak Mahalanobis sebagai fungsi *depth* merupakan fungsi yang paling umum yang terkait dengan teori normal multivariat. Mengikuti Rao (1988), untuk matriks definit $d \times d$ positif $\sum$, kita definisikan jarak Mahalanobis dari x ke μ dengan :

$$JM_i = (x_i - \mu)' \textstyle\sum^{-1} (x_i - \mu) \qquad (3)$$

Mahalanobis depth dirumuskan dengan :

$$D_i = \frac{1}{1+JM_i} \qquad (4)$$

Apabila apabila vektor rata-rata μ dan matiks kovarians Σ tidak diketahui maka ditaksir dari data sampel yaitu $\hat{\mu}$ dan $\hat{\Sigma}$. Sebuah titik pengamatan dengan nilai Mahalanobis *depth* mendekati 1 menunjukkan bahwa pengamatan tersebut mendekati pusat *data cloud* (μ).

Pada bahasan ini akan dijelaskan perluasan metode Kruskal-Wallis untuk pengujian kesamaan beberapa distribusi multivariat berdasarkan data depth yang rumusan hipotesisnya sebagai berikut :

$$H_0 : F_1(x) = F_2(x) = \cdots = F_t(x) \text{ untuk semua } x \qquad (5)$$

Secara umum, data sampel yang dibangun dengan konsep multivariat, disajikan pada Tabel 2.1. Dibawah H_0 benar kita akan menentukan distribusinya secara umum dengan F. $\widehat{F_j}$ menunjukan distribusi empiris sampel yang berisi data ke-j. $X_j = \left\{X_{1j \ldots\ldots} X_{n_j j}\right\}$ maka, $\hat{F} = n^{-1} \sum_{j=1}^{t} n_j \widehat{F_j}$ dimana $n = \sum n_j$ dan $X = \bigcup_{j=1}^{t} X_j$ sehingga, $D_{ij} = \left(X_{ij,} \hat{F}\right)$ menunjukkan depth dari X_{ij} untuk $i = 1,2, \ldots, n_j$ dan $j = 1,2, \ldots, t$ dengan menggunakan distribusi $\hat{F}$.

Tabel. 1 Data berasal dari t distirbusi $d - variate$

Treatment			
1	2	…	t
x_{11}	x_{12}	…	x_{1t}
x_{21}	x_{22}	…	x_{2t}
$\vdots$	$\vdots$	$\ddots$	$\vdots$
$x_{n_1 1}$	$x_{n_2 2}$	…	$x_{n_t t}$

Kita tentukan R_{ij} sebagai ranking dari X_{ij} yang dihubungkan dengan fungsi depth $D_{ij} = \left(., \hat{F}\right)$, sehingga struktur rank nya sebagai berikut : (3)

Tabel. 2 Struktur Rank Dari Depth Berdasarkan Sampel

Sampel ke-			
1	2	...	t
R_{11}	R_{12}	...	R_{1t}
R_{21}	R_{22}	...	R_{2t}
$\vdots$	$\vdots$	$\ddots$	$\vdots$
$R_{n_1 1}$	$R_{n_2 2}$	...	$R_{n_t t}$

Dari susunan data dalam Tabel 2 diperoleh nilai $R_{.j}$ sebagai berikut

$$R_{.j} = \sum_{i=1}^{n_j} R_{ij} \quad (6)$$

dan

$$\bar{R}_{.j} = \frac{R_{.j}}{n_j} \quad (7)$$

sehingga dapat ditunjukkan dengan mudah bahwa

$$E_{H_0}(R_{.j}) = \frac{n_j(n+1)}{2} \quad (8)$$

$$Var_{H_0}(R_{.j}) = \frac{n_j(n-n_j)(n+1)}{12} \quad (9)$$

$$Cov_{H_0}(R_{.j}, R_{.j'}) = \frac{-n_j n_{j'}(n+1)}{12} \quad (10)$$

Kita misalkan,

$$K_j = \frac{\left[\bar{R}_{.j} - E_{H_0}(\bar{R}_{.j})\right]^2}{Var_{H_0}(\bar{R}_{.j})} \quad (11)$$

Sama halnya dengan uji Kruskal-Wallis pada data univariat, maka nilai dari rata-rata yang diboboti dari K_j yakni

$$K = \sum_{j=1}^{t} \left(1 - \frac{n_j}{n}\right) K_j \quad (12)$$

Perlu dicatat bahwa dibawah hipotesis nol benar statistik K merupakan variabel acak bebas distribusi. Akan tetapi, distribusi ini tidak cukup sensitif dalam hal mendeteksi pergeseran lokasi (lihat Chenouri (2004), Liu dan Singh (2006) dan Chenouri et al (2011)). Untuk mengatasi hal tersebut, Chenouri dan Christopher (2012) mengubah metode perankingan yang lebih sensitif melalui sampel yang dipisah berdasarkan distribusi empiris masing-masing sampel. Perhatikan fungsi *depth* $D_{ij}(k) = \left(X_{ij}, \hat{F}_k\right)$ dan peringkat $R_{ij}(k)$ untuk $i = 1,2,\dots,n_j$ dan j, $k = 1,2,\dots,t$, seperti yang dijelaskan di atas. Tabel 3 merangkum peringkat $R_{ij}(k)$, untuk setiap $k = 1,2,\dots,t$.

Tabel. 3 Struktur Depth rank berdasarkan *k* sample

Sampel ke-			
1	2	...	t
$R_{11}(k)$	$R_{12}(k)$	...	$R_{1t}(k)$
$R_{21}(k)$	$R_{22}(k)$	...	$R_{2t}(k)$
$\vdots$	$\vdots$	$\ddots$	$\vdots$
$R_{n_1 1}(k)$	$R_{n_2 2}(k)$	...	$R_{n_t t}(k)$

Dengan demikian, akan lebih masuk akal untuk membangun statistik uji $H(k)$, sebab rank $R_{ij}(k)$ memiliki kemampuan untuk membedakan observasi sampel k yang berada di sampel j dimana $j \neq k$. Katakanlah, berdasarkan peringkat $R_{ij}(k)$ dengan cara yang mirip dengan konstruksi *K* seperti dalam Persamaan (12) maka, dikumpulkan statistik-statistik uji $H(1),\dots,H(t)$ bersama-sama sehingga membentuk sebuah statistik uji keseluruhan untuk hipotesis nol bahwa semua sampel memiliki distribusi yang sama. Maka dapat didefinisikan :

$$H_j(k) = \frac{12 n_j}{(n-n_j)(n+1)} \left(\bar{R}_{.j}(k) - \frac{n+1}{2}\right)^2 \quad (13)$$

Berdasarkan pada persamaan (13) maka,

$$H(k) = \sum_{j=1}^{t} \left(1 - \frac{n_j}{n}\right) H_j(k) \quad (14)$$

Persamaan (14) dapat dipertimbangkan sebagai statistik uji namun tentu hal ini bergantung pada pemilihan *k*. Untuk mengatasi hal tersebut maka statistik uji yang sejenis dengan statistik uji *K* adalah

$$H = \frac{1}{t} \sum_{k=1}^{t} H(k) \quad (15)$$

Berdasarkan statistik uji *H* Perasamaan (15), diperoleh kriteria uji yaitu tolak H_0 jika nilai statistik uji H sangat besar.

Untuk menentukan nilai kritis, berdasarkan hasil simulasi Chenouri dan Christopher (2012) pada studi Monte Carlo diperoleh bahwa untuk ukuran sampel pada kelompok sampel, jika salah satu kelompok sampel memiliki ukuran sampel ≤ 50 maka nilai kritis diperoleh dengan uji permutasi. Sedangkan jika salah satu kelompok sampel memiliki ukuran sampel ≥ 50 menggunakan pendekatan ke distribusi χ^2 dengan derajat bebas $t-1$.

C. Hasil Penelitian dan Pembahasan

Menghitung Jarak Mahalanobis

Dengan bantuan Matlab 2014 diperoleh nilai jarak Mahalanobis dan Mahalanobis *depth* dari setiap responden untuk kelompok nilai UNBK tahun ajaran 2015/2016 dan kelompok nilai UNBK tahun ajaran 2017/2018 terhadap distribusi $\hat{F}(1)$ dan $\hat{F}(2)$. Selanjutnya akan ditulis dengan $\hat{F}(k)$. Dimana kelompok 1 adalah kelompok nilai UNBK tahun ajaran 2015/2016 dan kelompok 2 adalah kelompok nilai UNBK tahun ajaran 2017/2018. Hasil output dari Matlab 2014 untuk perhitungan nilai Mahalanobis *depth* disajikan pada Tabel 4.

Ranking Berdasarkan Nilai Data *Depth*

Untuk menentukan ranking berdasarkan nilai mahalanobis *depth*, maka akan kita gabungkan nilai mahalanobis depth dari kedua sampel berdasarkan pendekatan distribusi $\hat{F}(k)$ nya, kemudian diurutkan. Dimana, rank pertama merupakan nilai *depth* yang paling kecil, menunjukkan bahwa nilai *depth* tersebut merupakan titik yang paling jauh dari pusat dan nilai rank terbesar merupakan nilai *depth* yang paling besar, menunjukkan nilai *depth* tersebut merupakan titik yang paling mendekati pusat. Dengan bantuan Matlab 2014 maka diperoleh nilai ranking dari mahalanobis *depth* pada Tabel 5.

Tabel. 4 Nilai Mahalanobis *Depth*

Mahalanobis *depth* berdasarkan distribusi $\hat{F}(1)$ $k=1$				Mahalanobis *depth* berdasarkan distribusi $\hat{F}(2)$ $k=2$			
No Res.	Kelompok 1 $j=1$	No Res.	Kelompok 2 $j=2$	No Res	Kelompok 1 $j=1$	No Res.	Kelompok 2 $j=2$
1	0,3144	1	0,2352	1	0,2228	1	0,5989
2	0,2965	2	0,1007	2	0,2741	2	0,3042
3	0,2592	3	0,0891	3	0,1604	3	0,2180
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
291	0,2217	291	0,2211	291	0,1780	291	0,7187
		⋮	⋮			⋮	⋮
		337	0,0674			337	0,1964

Tabel. 5 Rank Mahalanobis *Depth*

Rank Mahalanobis *depth* berdasarkan distribusi $\hat{F}(1)$ $k = 1$				Rank Mahalanobis *depth* berdasarkan distribusi $\hat{F}(2)$ $k = 2$			
No Res.	Kelompok 1 $j = 1$	No Res.	Kelompok 2 $j = 2$	No Res	Kelompok 1 $j = 1$	No Res.	Kelompok 2 $j = 2$
1	446	1	368	1	255	1	574
2	430	2	111	2	346	2	387
3	386	3	75	3	148	3	249
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
291	348	291	347	291	170	291	601
		⋮	⋮			⋮	⋮
		337	27			337	208
$\boldsymbol{R_{.1}(1)}$	**122464**	$\boldsymbol{R_{.2}(1)}$	**75042**	$\boldsymbol{R_{.1}(2)}$	**75361**	$\boldsymbol{R_{.2}(2)}$	**122145**
$\boldsymbol{\bar{R}_{.1}(1)}$	**420.8385**	$\boldsymbol{\bar{R}_{.2}(1)}$	**222.6766**	$\boldsymbol{\bar{R}_{.1}(2)}$	**258.9725**	$\boldsymbol{\bar{R}_{.2}(2)}$	**362.4481**

Setelah dilakukan perankingan, kemudian dihitung jumlah dan rata-rata rankingnya. Berdasarkan Tabel 5 di atas, diperoleh nilai $\bar{R}_{.1}(1)$ dan $\bar{R}_{.2}(1)$ terhadap distribusi $\hat{F}(1)$ sebesar 420.8385 dan 222.6766 kemudian untuk nilai $\bar{R}_{.1}(2)$ dan $\bar{R}_{.2}(2)$ terhadap distribusi $\hat{F}(2)$ sebesar 258.9725 dan 362.4481. Kemudian akan digunakan dalam perhitungan statistik $H_j(k)$.

Menghitung Statistik Uji Kruskal-Wallis Berdasarkan Data *Depth*

$H_0: F_1(x) = F_2(x)$:tidak terdapat perbedaan distribusi antara nilai UNBK tahun pelajaran 2015/2016 dengan nilai UNBK tahun pelajaran 2017/2018.

H_1: $F_1(x) \neq F_2(x)$:terdapat perbedaan distribusi antara nilai UNBK tahun pelajaran 2015/2016 dengan nilai UNBK tahun pelajaran 2017/2018

Menghitung Statistik Uji

Dalam perhitungan statistik uji, akan digunakan peramaan (21) untuk memperoleh nilai statistik uji H yang kemudian akan dibandingkan dengan nilai kritis. Namun sebelum mendapatkan nilai statistik uji H akan dihitng terlebih dahulu nilai statistik uji $H_j(k)$ dengan persamaan (19) sehingga akan diperoleh statistik uji $H_1(1)$ dan $H_2(1)$ untuk memperoleh nilai statistik uji $H(1)$ yang didasarkan kepada jarak x_i ke $\hat{\mu}_1$ dan statistik uji $H_1(2)$ dan $H_2(2)$ untuk memperoleh nilai statistik uji $H(2)$ yang didasarkan kepada jarak x_i ke $\hat{\mu}_2$. Dengan perhitungan menggunakan Matlab 2014 diperoleh nilai statistik uji $H_1(1)$,$H_2(1)$, $H_1(2)$, dan $H_2(2)$ seperti pada Tabel 6.

Tabel 6 Nilai statistik $H_j(k)$ berdasarkan distribusi $\hat{F}(1)$ dan $\hat{F}(2)$

Nilai statistik $H_j(k)$ Kelompok 1 dan 2 thd distribusi $\hat{F}(1)$		Nilai statistik $H_j(k)$ Kelompok 1 dan 2 thd distribusi $\hat{F}(2)$	
$H_1(1)$	186.28365	$H_1(2)$	50.79371
$H_2(1)$	186.28344	$H_2(2)$	50.79375

Berdasarkan nilai pada Tabel 6 maka kita dapat menghitung nilai statistik uji $H(1)$ dan statistik uji $H(2)$ dengan persamaan (20) sehingga diperoleh nilai sebagai berikut :

$$H(1) = \sum_{j=1}^{2}\left(1 - \frac{n_j}{n}\right) H_j(1)$$
$$= 199{,}92695$$

$$H(2) = \sum_{j=1}^{2}\left(1 - \frac{n_j}{n}\right) H_j(2)$$
$$= 47{,}07360$$

Sehingga, nilai statistk uji H dapat dihitung yakni

$$H = \frac{H(1) + H(2)}{2} = 123{,}50027$$

Menghitung Nilai Kritis

Mengkuti Chenouri dan Christopher (2012) berdasarkan hasil simulasi yang dilakukan, maka nilai kritis dibagi menjadi dua kriteria yakni saat sampel lebih besar dari 50 maka, digunakan nilai kritis dengan chi-kuadrat. Sedangkan untuk sampel lebih kecil dari 50 maka, digunakan nilai kritis dengan cara permutasi. Dalam skripsi ini ukuran sampel lebih besar dari 50. Sehingga nilai kritis yang digunakan adalah nilai Chi-kuadrat. Dengan taraf nyata 0,05 maka, ${\chi^2}_{1(0{,}05)} = 3.8415$

Kriteria Uji

Kriteria Pengujian adalah bentuk pembuatan keputusan dalam menerima atau menolak hipotesis nol (H_0) dengan cara membandingkan nilai kritis dengan nilai uji statistiknya, sesuai dengan bentuk pengujiannya. Dalam kasus ini Penolakan H_0 terjadi jika nilai uji statistiknya lebih besar dari ${\chi^2}_{1(0{,}05)}$. Atau nilai statistik uji berada di luar nilai kritis (Sudjana, 2008). Seperti diketahui nilai statistik uji sebesar $123{,}50027$ dan nilai kritis sebesar 3,8415.

sehingga

$$H > {\chi^2}_{1(0{,}05)}$$

Atau

$$123{,}50027 > 3{,}8415$$

Sehingga, H_0 ditolak.

Kesimpulan

Terdapat perbedaan distribusi nilai UNBK tahun pelajaran 2015/2016 dengan distribusi nilai UNBK tahun pelajaran 2017/2018.

D. Kesimpulan

Berdasarkan hasil analisis dan penelitian, perhitungan rank berdasarkan *Mahalanobis depth* disimpulkan bahwa,

1. Berdasarkan prosedur penentuan nilai rank pada data multivariat dua sampel nilai UNBK SMAN XXX dengan menggunakan perhitungan Mahalanobis depth maka diperoleh nilai $\bar{R}_{.1}(1)$ dan $\bar{R}_{.2}(1)$ terhadap distribusi $\hat{F}(1)$ sebesar 420.8385 dan 222.6766 kemudian untuk nilai $\bar{R}_{.1}(2)$ dan $\bar{R}_{.2}(2)$ terhadap distribusi $\hat{F}(2)$ sebesar 258.9725 dan 362.4481. Prosedur perhitungan rank berdasarkan Mahalanobis depth untuk data multivariat cukup membantu dan cenderung lebih mudah dilakukan untuk data multivariat. Selain itu hasil yang diberikan cukup representatif atau invariant sebab, setiap obesrvasi dihitung jaraknya terhadap pusat yang mana dilakukan berdasarkan distribusi $\hat{F}(k)$ untuk $k = 1,2,..t$. Kemudahan perhitungan ini dapat dibantu dengan beberapa software seperti Matlab atau Excel.
2. Berdasarkan prosedur perhitungan statistik uji Kruskal-Wallis dengan menggunakan nilai rank dari perthitungan Mahalanobis depth diperoleh nilai statistik uji $H(1)$ dan statistik uji $H(2)$ sebesar $199{,}92695$ dan $47{,}07360$ dan

nilai statistik uji H sebesar 123,50027. Hasil ini menunjukkan terdapat perbedaan distribusi dari dua sampel nilai UNBK pada SMA XXX.

Daftar Pustaka

Chenouri, S., Small, C. G., and Farrar, T. J. (2011). Data depth-based nonparametric scale tests. Canad. J. Statist., 39:356–369

Chenouri, S. and Farrar, T. J. (2011) A Two-sample Nonparametric Multivariate Scale Test based on Data Depth Canad. J. Statist.

Chenouri, S., and Small, C. G. (2012) A nonparametric multivariate multisample test based on data depth Electronic Journal of Statistics,. 06:760-782

Hajarisman Nusar (2008). Statistika Multivariat. Seri buku Ajar Universitas Islam Bandung

Junaidi (2010) Statistik Uji Kruskall-Wallis Repository Fakultas Ekonimo Universeitas Jambi

Karl Mosler (2010) Depth Statistic.dalam C. Becker et al. (eds.), Robustness and Complex Data Structures, Springer-Verlag Berlin Heidelberg

Liu, R. Y., Parelius, J. M., and Singh, K. (1999). Multivariate analysis by data depth: Descriptive statistics, graphics and inference. Ann. Statist., 27:783–858.

Yanti Tety Sofia (2010) Perluasan Uji Kruskal Wallis untuk Data Multivariat Statistika Unisba, Vol. 10 No. 1, 43 – 49

Zhang C, Lin, and Wu J (2009) Nonparametric tests for the general multivariate multi-sample problem Journal of Nonparametric Statistics Vol. 21, No. 7, October 2009, 877–888

Zuo, Y. and Serfling, R. (2000). General notions of statistical depth function. Ann. Statist., 28:461–482.