

Statistik Robust Berdasarkan Desil dengan Penerapan pada Data Wisatawan Mancanegara Provinsi Jawa Barat Tahun 2015-2019

Dio Indra Rukmana*

Prodi Statistika, Fakultas Matematika dan Ilmu Pengetahuan Alam,
Universitas Islam Bandung, Indonesia.

* dioindrarukmana@gmail.com

Abstract. (1) Parameter is a term that describes a quantity related to the population. Parameters can be interpreted as a value or condition that is used as a measure against other values or conditions. In statistics, one of the most important aspects to describe the distribution of data is the central value of the observed data. (2) The concentration size reflects where the data center is located, while the size of the spread reflects the variation or how spread the data is. Among the centrally used statistical measures are mean, median and mode. (3) The distribution size statistics include range, variance and standard deviation. In a statistical analysis, we often find outliers or outliers. (4) Outlier is an observed value or a collection of observed values that is separate from most of the majority observed values. If the data contains outliers then there are concentration size statistics or spread sizes whose values will be affected. (5) But there are also statistics whose value is not affected even though the data contains outliers. Such statistics are known as robust statistics. In this thesis, a new type of robust statistic is discussed which is formed using the decile concept. A decile is 9 values that divide the data into 10 equal parts. (6) There are 4 (four) robust statistics formed, namely decile average, decile standard deviation, decile variation coefficient and decile standard error. Robust statistics that are formed based on this decile are then applied to the Data of Foreign Tourists of the Answer Province for 2015-2019. (7) With data obtained from the Central Statistics Agency of West Java Province. The results of the application show that the decile version of robust statistics has a smaller statistical value. This is because there are outliers of large values in the data which cause ordinary statistical values to tend to be carried away by these outliers.

Keywords: Decile Average, Decile Standard Deviation, Decile Variation Coefficient, Decile Standard Error.

Abstrak. (1) Parameter adalah istilah yang mendeskripsikan besaran yang berkaitan dengan populasi. Parameter dapat diartikan sebagai suatu nilai atau kondisi yang dijadikan sebagai tolak ukur terhadap nilai atau kondisi yang lainnya. Dalam statistika salah satu aspek yang paling penting untuk menggambarkan distribusi data adalah nilai pusat data pengamatan. (2) Ukuran pemusatan mencerminkan dimana pusat data berada, sedangkan ukuran penyebaran mencerminkan variasi atau seberapa menyebarnya data itu. Di antara statistik ukuran pemusatan yang biasa digunakan adalah rata-rata, median dan modus. (3) Adapun statistik ukuran penyebaran diantaranya rentang, variansi dan simpangan baku. Dalam suatu analisis statistika sering kali kita mendapati data pencilan atau *outlier*. (4) *Outlier* adalah adalah nilai pengamatan atau kumpulan nilai pengamatan yang terpisah dari kebanyakan nilai pengamatan

mayoritas. Jika data mengandung pencilan maka ada statistik ukuran pemusatan atau ukuran penyebaran yang nilainya akan terpengaruh. (5) Tetapi ada juga statistik yang nilainya tidak terpengaruh meskipun datanya mengandung pencilan. Statistik yang demikian disebut sebagai statistik yang robust. Dalam skripsi ini dibahas tentang tipe baru statistik robust yang dibentuk dengan menggunakan konsep desil. Desil adalah 9 buah nilai yang membagi data menjadi 10 bagian sama banyak. (6) Terdapat 4 (empat) statistik robust yang dibentuk, yaitu rata-rata desil, simpangan baku desil, koefisien variasi desil dan standar error desil. Statistik robust yang dibentuk berdasarkan desil ini kemudian diterapkan pada Data Wisatawan Mancanegara Provinsi Jawabart Tahun 2015-2019. (7) Dengan data yang diperoleh dari Badan Pusat Statistik Provinsi Jawa Barat. Hasil penerapan menunjukkan bahwa statistik robust versi desil mempunyai nilai statistika yang lebih kecil. Hal ini dikarenakan adanya pencilan nilai besar dalam data yang menyebabkan nilai statistik biasa cenderung terbawa oleh nilai pencilan tersebut.

Kata Kunci: Rata-rata desil, Simpangan Baku Desil, Koefisien Variasi Desil, Standar error Desil.

1. Pendahuluan

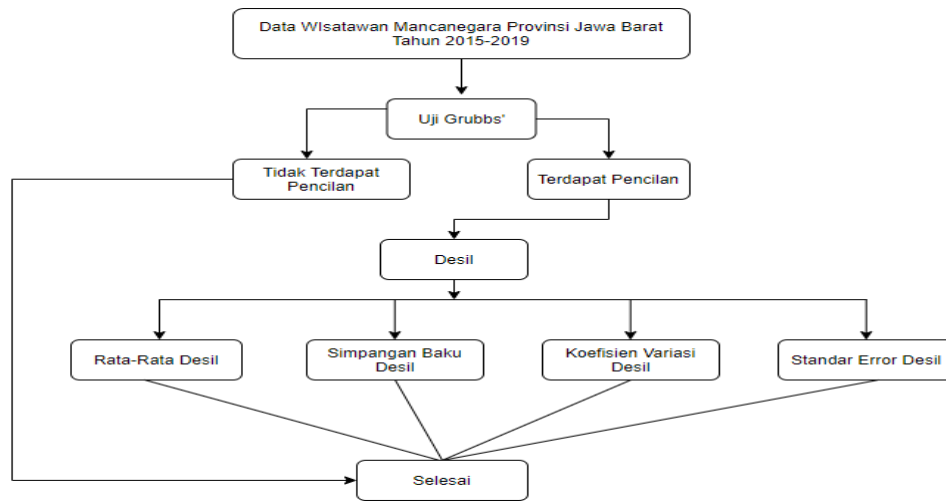
Statistika sangat penting untuk bidang pariwisata, karena dengan statistik sektor pariwisata menjadi dapat mengetahui pertumbuhan bidang pariwisata pada suatu daerah serta untuk mengetahui kecenderungan (tren) yang terjadi dalam pariwisata pada jangka waktu tertentu. Dengan data yang didapatkan pengelola atau pemerintah dapat menentukan strategi dalam pengembangan, perbaikan danantisipasi untuk tahun selanjutnya. Menurut data pengunjung pariwisata mancanegara Provinsi Jawa Barat di Badan Pusat Statistika Provinsi Jawa Barat, wisatawan mancanegara mengalami peningkatan antara tahun 2015-2016 namun selama tahun 2016-2019 mengalami penurunan setiap tahunnya. Dengan data yang diperoleh dari Badan Pusat Statistik Provinsi Jawa Barat. Menurut Badan Pusat Statistika data dalam bidang pariwisata bisa dikategorikan menjadi dua kelompok yaitu (1) data statistik kunjungan wisatawan mancanegara dan (2) statistik hotel dan akomodasi lainnya di Indonesia (Badan Pusat Statistika Provinsi Jawa Barat). Dalam statistik kunjungan wisatawan mancanegara terdapat dua klasifikasi yaitu wisatawan dan pelancong. Berdasarkan penjelasan statistik *robust* berdasarkan desil yang disampaikan diatas. Dalam skripsi ini akan dilakukan penerapan statistik robust berdasarkan desil untuk data pariwisata. Data yang digunakan adalah data kunjungan wisatawan mancanegara Provinsi Jawa Barat tahun 2015-2019.

Adapun tujuan untuk penelitian ini adalah sebagai berikut:

1. Untuk mengetahui Berapakah nilai rata-rata wisatawan mancanegara Provinsi Jawa Barat tahun 2015-2019 menggunakan teori rata-rata desil.
2. Untuk mengetahui Berapakah nilai simpangan baku wisatawan mancanegara Provinsi Jawa Barat 2015-2019 menggunakan teori simpangan baku desil.
3. Untuk mengetahui Berapakah nilai koefisien variasi wisatawan mancanegara Provinsi Jawa Barat 2015-2019 menggunakan teori koefisien variasi desil.
4. Untuk mengetahui Berapakah nilai standar error wisatawan mancanegara Provinsi Jawa Barat 2015-2019 menggunakan teori standar error desil.
5. Untuk mengetahui Berapakah perbandingan antara statistika biasa dengan statistik desil wisatawan Provinsi Jawa Barat 2015-2019

2. Landasan Teori

Dalam penulisan materi pokok dari skripsi ini diperlukan beberapa teori-teori yang mendukung, yang menjadi uraian pokok pada bab ini. Teori yang dipakai dalam skripsi ini adalah *Grubbs' Test*, rata-rata desil, simpangan baku desil, koefisien variasi desil, standar error desil dan bootstrap.



Gambar 1. Data Wisatawan Mancanegara Provinsi Jawa Barat tahun 2015-2019

1. Pertama lakukan pengujian outlier menggunakan uji Grubbs'. Jika data mempunyai pencilan maka dapat dilanjutkan kepada perhitungan selanjutnya. Jika data tidak mempunyai pencilan maka tidak dapat dilanjutkan pada tahap berikutnya.
2. Hitung nilai desil dari data wisatawan mancanegara Provinsi Jawa Barat
3. Setelah didapatkan nilai desil selanjutnya hitung nilai rata-rata desil
4. Selanjutnya hitung nilai simpangan baku desil menggunakan
5. Setelah didapat simpangan baku desil selanjutnya hitung nilai koefisien variasi desil
6. Lalu terakhir hitung nilai standar error

Grubbs' test digunakan untuk mendeteksi pencilan tunggal dalam satu set data univariat yang mengikuti distribusi yang mendekati normal [6]. Jika kita mencurigai kemungkinan terdapat lebih dari satu outlier, maka disarankan untuk menggunakan uji yang lainnya seperti box`plot.

3. Hasil Penelitian dan Pembahasan

Bahan yang digunakan adalah data wisatawan mancanegara yang berkunjung ke Provinsi Jawa Barat dari tahun 2015-2019. Data tersebut berasal dari website Badan Pusat Statistika Provinsi Jawa Barat. Untuk perhitungan data penulis menggunakan Microsoft Excell sebagai software pendukung.

Tabel 1. Jumlah Wisatawan Mancanegara Provinsi Jawa Barat Tahun 2015-2019 (Jiwa)

Bulan	Tahun				
	2015	2016	2017	2018	2019
Januari	10453	11065	8614	11600	12529
Februari	13138	8497	13410	12302	15172
Maret	15224	15971	17439	15793	16440
April	16978	30922	16189	14248	14830
Mei	18902	16841	15452	10571	8168

Juni	15423	9055	8360	6493	8881
Juli	6688	9499	12017	12814	12645
Agustus	10409	12663	15189	13766	14129
September	10652	15141	14120	13399	13028
Oktober	10755	17444	14151	13918	13569
November	14951	12876	15541	14364	14715
Desember	17067	22410	18031	17375	15159
Total	160640	182384	168513	156643	159265

sumber: Badan Pusat Statistika Provinsi Jawa Barat tahun 2019.

Pada bagian ini akan dibahas metode atau langkah-langkah dalam menghitung rata-rata desil, simpangan baku desil, koefisien variasi desil dan standar error desil wisatawan mancanegara Provinsi Jawa Barat Tahun 2015-2019.

Uji Grubbs'

Akan dilakukan uji *Grubbs'* untuk mengetahui apakah terdapat pencilan atau tidak pada data wisatawan mancanegara Provinsi Jawa Barat tahun 2015-2019 dengan hipotesis sebagai berikut:

H_0 : Tidak terdapat pencilan pada data wisatawan mancanegara Provinsi Jawa Barat tahun 2015-2019.

H_1 : Terdapat tepat satu pencilan pada data wisatawan mancanegara Provinsi Jawa Barat tahun 2015-2019.

Diperoleh nilai $\max|x_i - \bar{x}|$ sebesar 17131,25 lalu didapatkan nilai G sebagai berikut :

$$G = \frac{\max|x_i - \bar{x}|}{s}$$

$$= \frac{|30922 - 13790,75|}{3877,16}$$

$$= 4,419$$

Dengan taraf signifikansi $\alpha = 0,05$ maka uji hipotesis dua sisi menggunakan $\alpha/2n = 0,05/120 = 0.00042$ sehingga nilai tabel $t_{(\frac{\alpha}{2n}, n-2)} = t_{(0,00042, 58)} = 3.525$ dan titik kritisnya

$$\frac{(n-1)}{\sqrt{n}} \sqrt{\frac{\left(t_{\frac{\alpha}{2n}, n-2}\right)^2}{n-2 + \left(t_{\frac{\alpha}{2n}, n-2}\right)^2}} = \frac{59}{\sqrt{60}} \sqrt{\frac{3,525^2}{58 + 3,525^2}} = 3.200. \text{ Sehingga } H_0 \text{ ditolak karena statistik uji } G > \text{ titik}$$

kritis, yakni $4,419 > 3,200$. Dengan demikian terdapat pencilan dalam data wisatawan mancanegara Provinsi Jawa Barat tahun 2015-2019.

Rata-Rata Desil

$$\text{Letak } D_i = \frac{i(n+1)}{10}$$

$$\text{Letak } D_1 = \frac{1(60+1)}{10} = \frac{61}{10} = 6,1$$

$$D_1 = x_6 + 0,1 (x_7 - x_6)$$

$$D_1 = 8614 + 0,1 (8881 - 8614)$$

$$D_1 = 8640,7$$

Dengan menggunakan rumus yang sama akan dihitung nilai D_1, D_2, \dots, D_9 .

Tabel 2. Hasil Perhitungan Desil

No.	Desil	Nilai Desil	No.	Desil	Nilai Desil	No.	Desil	Nilai Desil
1	D ₁	8640,7	4	D ₄	13.072	7	D ₇	15.213,5
2	D ₂	10.587,2	5	D ₅	14.019	8	D ₈	16.145,4
3	D ₃	12.370,1	6	D ₆	14.784	9	D ₉	17.432,6

Kemudian dengan menggunakan nilai desil dapat dihitung nilai rata-rata sebagai berikut:

$$\begin{aligned}
 DM &= \frac{D_1 + D_2 + \dots + D_9}{9} \\
 &= \frac{(8640,7 + 10.587,2 + \dots + 17.432,6)}{9} \\
 &= \frac{122.264,5}{9} \\
 &= 13.584,9444
 \end{aligned}$$

Maka nilai rata-rata desil (DM) untuk data wisatawan mancanegara Provinsi Jawa Barat tahun 2015-2019 adalah sebesar 13.584,9444.

Simpangan Baku Desil, Koefisien Variasi desil dan Standar Error Desil

Sebelum menghitung nilai simpangan baku desil, koefisien variasi desil dan standar error desil kita perlu menghitung nilai $(x_i - DM)^2$ untuk perhitungan D_i^* menggunakan nilai rata-rata desil sebelumnya. perhitungan $(x_i - DM)^2$ sebagai berikut :

$$\begin{aligned}
 x_1 &= (x_1 - DM)^2 \\
 &= (6493 - 13.584,9444)^2 \\
 &= 50.295.675,37
 \end{aligned}$$

Maka nilai $(x_1 - DM)^2$ untuk x_i adalah 50.295.675,37. Seterusnya hingga nilai x_{60} seperti pada tabel sebagai berikut :

Tabel 3. Hasil Perhitungan $(x_i - DM)^2$

i	x	$(x_i - DM)^2$	i	x	$(x_i - DM)^2$	i	x	$(x_i - DM)^2$
1	6493	50295675,37	21	12663	849981,4767	41	15172	2518745,477
2	6688	47567842,06	22	12814	594355,2679	42	15189	2572994,368
3	8168	29343286,63	23	12876	502602,1623	43	15224	2686503,26

4	8360	27300043,98	24	13028	310187,0647	44	15423	3378448,389
5	8497	25887178,22	25	13138	199759,2967	45	15452	3485896,613
6	8614	24710288,23	26	13399	34575,31989	46	15541	3826153,51
7	8881	22127092,92	27	13410	30605,54309	47	15793	4875509,533
8	9055	20520396,27	28	13569	254,2238914	48	15971	5693261,326
9	9499	16694941,64	29	13766	32781,13029	49	16189	6781105,568
10	10409	10086622,83	30	13918	110926,0327	50	16440	8151342,479
11	10453	9809075,725	31	14120	286284,4951	51	16841	10601898,07
12	10571	9083860,846	32	14129	295996,4959	52	16978	11512826,3
13	10652	8602162,853	33	14151	320418,9423	53	17067	12124711,2
14	10755	8008585,307	34	14248	439642,7287	54	17375	14364521,45
15	11065	6350119,779	35	14364	606927,6279	55	17439	14853744,57
16	11600	3940004,271	36	14715	1277025,659	56	17444	14892310,12
17	12017	2458449,641	37	14830	1550163,447	57	18031	19767410,4
18	12302	1645946,333	38	14951	1866107,902	58	18902	28271080,25
19	12529	1115018,576	39	15141	2421309,03	59	22410	77881606,34
20	12645	883495,4751	40	15159	2477651,032	60	30922	300573496,9

Setelah didapatkan nilai seperti pada tabel diatas selanjutnya kita urutkan data tersebut dari yang terkecil ke terbesar lalu kita dapat mencari D_i^* sebagai berikut :

Tabel 4. Hasil Pengurutan Data $(x_i - DM)^2$

x	$(x_i - DM)^2$	x	$(x_i - DM)^2$	x	$(x_i - DM)^2$
1	254,2238914	21	1866107,902	41	10086622,83
2	30605,54309	22	2421309,03	42	10601898,07
3	32781,13029	23	2458449,641	43	11512826,3
4	34575,31989	24	2477651,032	44	12124711,2
5	110926,0327	25	2518745,477	45	14364521,45

6	199759,2967	26	2572994,368	46	14853744,57
7	286284,4951	27	2686503,26	47	14892310,12
8	295996,4959	28	3378448,389	48	16694941,64
9	310187,0647	29	3485896,613	49	19767410,4
10	320418,9423	30	3826153,51	50	20520396,27
11	439642,7287	31	3940004,271	51	22127092,92
12	502602,1623	32	4875509,533	52	24710288,23
13	594355,2679	33	5693261,326	53	25887178,22
14	606927,6279	34	6350119,779	54	27300043,98
15	849981,4767	35	6781105,568	55	28271080,25
16	883495,4751	36	8008585,307	56	29343286,63
17	1115018,576	37	8151342,479	57	47567842,06
18	1277025,659	38	8602162,853	58	50295675,37
19	1550163,447	39	9083860,846	59	77881606,34
20	1645946,333	40	9809075,725	60	300573496,9

$$\text{Letak } D_i^* = \frac{i(n+1)}{10}$$

$$\text{Letak } D_1^* = \frac{1(60+1)}{10} = \frac{61}{10} = 6,1$$

$$D_1^* = x_6 + 0,1 (x_7 - x_6)$$

$$D_1^* = 199759,2967 + 0,1 (286284,4951 - 199759,2967)$$

$$D_1^* = 208411,8165$$

Maka nilai D_1^* sebesar 208.411,8165 dan seterusnya hingga didapat nilai D_9^*

Tabel 5. Hasil Perhitungan $D_1^*, D_2^*, \dots, D_9^*$

D_1^*	208411,8165
D_2^*	520952,7834
D_3^*	1358966,995
D_4^*	2494088,81
D_5^*	3883078,891
D_6^*	8094239,61
D_7^*	11239547,83
D_8^*	19152916,65
D_9^*	28173976,63
Total	75126180,01

Menghitung nilai simpangan baku desil sebagai berikut :

$$\begin{aligned}
 SD_{DM} &= \sqrt{\frac{D_1^* + D_2^* + \dots + D_9^*}{9}} \\
 &= \sqrt{\frac{208411,8165 + 520952,7834 + \dots + 28173976,63}{9}} \\
 &= 2889,179
 \end{aligned}$$

Setelah mendapatkan nilai simpangan baku desil (SD_{DM}) maka selanjutnya akan dihitung nilai koefisien variasi desil

$$\begin{aligned}
 CD &= \frac{SD_{DM}}{DM} \\
 &= \frac{2889,179}{13.584,9444} \\
 &= 0,212675
 \end{aligned}$$

Maka nilai koefisien variasi desil untuk data wisatawan mancanegara Provinsi Jawa Barat tahun 2015-2019 adalah sebesar 0,212675. Selanjut menghitung SE_{DM}

$$\begin{aligned}
 SE_{DM} &= \frac{SD_{DM}}{\sqrt{n}} \\
 &= \frac{2889,179}{\sqrt{60}} \\
 &= \frac{2889,179}{7,745967} \\
 &= 372,9914
 \end{aligned}$$

Tabel 6. Perbandingan statistika biasa dengan statistik desil

Statistik	Versi Biasa	Versi Statistik Robust
Rata-rata	13.790,75	13.584,94
Simpangan Baku	3.877,160	2.889,179
Koefisien Variasi	0,28114208	0,212675
Standar Error	500,5392	372,991

4. Kesimpulan

Kesimpulan dari penulisan skripsi ini adalah :

1. Pada data wisatawan mancanegara Provinsi Jawa Barat tahun 2015-2019 memiliki nilai rata-rata desil sebesar 13.585 jiwa.
2. Pada data wisatawan mancanegara Provinsi Jawa Barat tahun 2015-2019 memiliki nilai simpangan baku desil sebesar 2.890.
3. Pada data wisatawan mancanegara Provinsi Jawa Barat tahun 2015-2019 memiliki nilai koefisien variasi desil sebesar 0,212675.
4. Pada data wisatawan mancanegara Provinsi Jawa Barat tahun 2015-2019 memiliki nilai standar error desil sebesar 373.
5. Dengan melihat perbandingan antara statistika biasa dengan statistika desil terlihat bahwa kecenderungannya versi desil mempunyai nilai statistik yang lebih kecil. Hal ini dikarenakan adanya pencilan nilai besar dalam data yang menyebabkan nilai statistik versi biasa cenderung terbawa oleh nilai pencilan tersebut.

5. Saran

Saran Teoritis

Saran teoritisnya untuk mengembangkan statistik robust yang lain selain berdasarkan desil.

Saran Praktis

Manfaat statistik robust berdasarkan desil untuk berbagai kepentingan analisis data yang kemungkinan mengandung pencilan. Banyak sekali data yang mungkin mengandung pencilan selain data wisatawan mancanegara yang datang ke Jawa Barat.

Daftar Pustaka

- [1] Badan Pusat Statistika Provinsi Jawa Barat.2019. Tabel Jumlah Wisatawan Mancanegara Provinsi Jawa Barat. Bandung.
- [2] Ating Somantri dan Sambas Ali Muhidin (2006). Aplikasi Statistika Dalam Penelitian. Bandung : Pustaka Setia.
- [3] Prof. Dr. Vincent Gaspersz (1989). Statistika Untuk Fakultas Ekonomi Dan Fakultas Ilmu – Ilmu Sosial. Bandung : Armico.
- [4] Grubbs, Frank (1969). “Procedures for Detecting Outlying Observations in Samples”.
- [5] Stefansky, W. (1972). “Rejecting Outliers In Factorial Design”.
- [6] Sohel Rana, Md.Siraj-Ud-Doulah, Habshah Midi, A.H.M.R Imon.(2012). Decile Mean: A New Robust Measure of Central Tendency. Department of Mathematics, Faculty of Science, University Putra Malaysia. Malaysia.
- [7] Md. Siraj-Ud-Doulah.(2018).Alternative Measures of Standard Deviation Coefficient of Variation and Standard Error. Department of Statistiks, Begum Rokeya University, Rangpur, Bangladesh. Bangladesh.
- [8] www.itl.nist.gov. Grubbs’ Test For Outlier. Diakses pada tanggal 10 Juni 2020, dari <https://www.itl.nist.gov/div898/handbook/eda/section3/eda35h1.htm>.

- [9] www.jabar.bps.go.id. Konsep dan Definisi Statistik Kunjungan Wisatawan mancanegara.
- [10] <https://jabar.bps.go.id/subject/16/pariwisata.html#subjekViewTab1>.