



PAPER – OPEN ACCESS

Perbandingan Keakuratan Dari Model Tabel Distribusi Frekuensi Berkelompok Antara Metode Sturges Dan Metode Scott

Author : Wanda Maya Sari
DOI : 10.32734/st.v1i1.182
Electronic ISSN : 2654-7094
Print ISSN : 2654-7086

Volume 1 Issue 1 – 2018 TALENTA Conference Series: Science & Technology (ST)



This work is licensed under a [Creative Commons Attribution-NonCommercial 4.0 International License](https://creativecommons.org/licenses/by-nc/4.0/).

Published under licence by TALENTA Publisher, Universitas Sumatera Utara



Perbandingan Keakuratan dari Model Tabel Distribusi Frekuensi Berkelompok Antara Metode Sturges dan Metode Scott

Wanda Maya Sari^a, Open Darnius^a, dan Pasukat Sembiring^a

Fakultas Matematika dan Ilmu Pengetahuan Alam, Universitas Sumatera Utara, Medan 20155, Indonesia

wandamayasaril@gmail.com

Abstrak

Distribusi Frekuensi berkelompok adalah penyusunan urutan data tunggal ke dalam tabel kelas-kelas interval, untuk kemudian ditentukan frekuensinya. Penyusunan ini berguna untuk menyelidiki kumpulan data sehingga dapat membawa informasi dengan cara yang pintas. Metode Sturges adalah aturan yang paling dikenal untuk menentukan jumlah kelas interval dalam distribusi frekuensi, namun terdapat kelemahan di dalamnya sehingga pada penelitian ini dilakukan analisis perbandingan keakuratan model tabulasi antara Metode Sturges dengan Metode Scott yang merupakan aturan untuk menentukan panjang kelas dalam distribusi frekuensi. Perbandingan dilihat dari hasil perhitungan selisih ukuran pemusatan, ukuran lokasi, serta dispersiantara kedua model terhadap ukuran dari data simulasinya (data aktual), kemudian dianalisis perbandingannya melalui jenis distribusi data, yang dalam penelitian ini menggunakan data berdistribusi normal dan eksponensial. Hasil perbandingan keakuratan dari masing-masing tabel distribusi frekuensi dari data normal maupun data eksponensial, menunjukkan persentasi keakuratan dari tabel distribusi frekuensi dengan metode scott lebih tinggi dibandingkan dengan metode sturges, hal ini dikarenakan, pada tabel distribusi frekuensi dengan metode sturges, terdapat data yang tidak masuk ke dalam tabel distribusi frekuensi diakibatkan nilai panjang kelas interval tidak dapat mencakup pengelompokkan sampai data tertinggi.

Kata Kunci: Distribusi Frekuensi Kelompok; Kelas Interval; Metode Sturges; Metode Scott; Data Simulasi.

1. Pendahuluan

Untuk menyajikan data kuantitatif menjadi bentuk yang baik dapat disusun dengan cara distribusi frekuensi berkelompok, di mana data yang sudah berkelompok dimasukkan ke dalam rentang ukuran yang sama atau dikenal dengan istilah kelas interval, yang biasa disimbolkan dengan K . Metode yang paling dikenal dalam menentukan banyak kelas interval (K) adalah aturan yang dirancang oleh ^[1]. Salah satu kritikan oleh ^[2] yaitu aturan sturges mengandung beberapa kelemahan, di mana untuk jumlah data (n) yang terlampau besar akan memberikan hasil yang tidak baik. Adapun dalam jurnalnya, ^[3] menyatakan bahwa rumus jumlah kelas interval sturges mempunyai kemungkinan dapat bertahan selama ini, karena untuk n sedang (kurang dari 200) dapat memberikan hasil histogram yang layak. Hyndman mengajukan salah satu rumus dalam penentuan panjang kelas interval, yaitu aturan ^[4] yang dipublikasikan dalam jurnalnya yang berjudul *–On Optimal and Data-Based Histograms*

¹ Sturges, H. (1926). The Choice of A Class Interval. Journal American Statistical Association. 21.65-66

² Dajan, A. nto. (1986). Pengantar Metode Statistik 2. LP3ES. Jakarta.

³ Hyndman, Rob J. (1995). The Problem with Sturges's Rule for Constructing Histograms. A shortnotes. Unpublished Article.

⁴ Scott, David W. (1979). On Optimal Data-Based Histograms. Biometrika Trust. 3.605-610

Jika dilakukan perhitungan nilai kriteria ukuran pemusatan, lokasi, dan dispersi pada data berkelompok dalam tabel distribusi frekuensi akan terjadi perbedaan perhitungan dengan data tunggalnya (sebelum dikelompokkan). Perbedaan

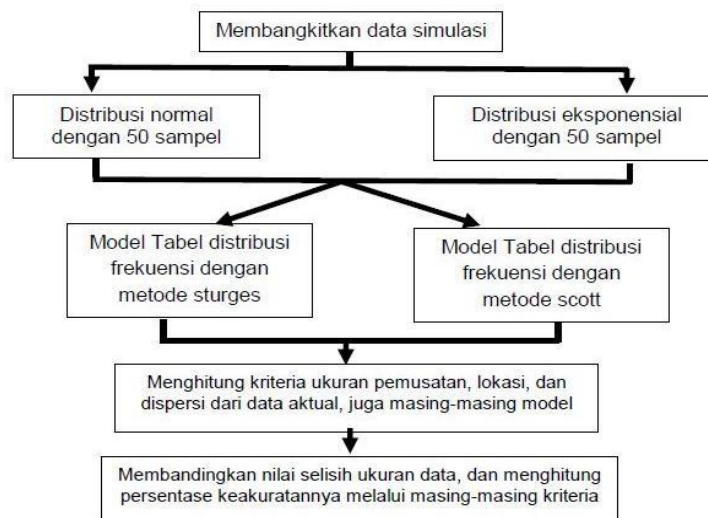
demikian dikatakan sebagai selisih akibat pengelompokan (*grouping error*) [2]. Melalui permasalahan tersebut, dalam penelitian ini dilakukan analisa perbandingan keakuratan dari model tabel distribusi frekuensi berkelompok antara penggunaan metode sturges dan metode scott terhadap nilai selisih pengelompokan yang berdasarkan nilai kriteria rata-rata, median, kuartil 1, kuartil 3, standar deviasi, dan varians data.

2. Tujuan

Mengetahui perbandingan keakuratan model tabel distribusi frekuensi berkelompok yang dapat mendekati data aktual, yaitu antara model yang menggunakan metode sturges dan metode scott, sehingga selisih akibat pengelompokan (*grouping error*) yang terjadi dalam tabel distribusi frekuensi semakin kecil.

3. Metode

Model tabel distribusi frekuensi yang diteliti dalam penelitian ini, diperoleh dari hasil program yang dirancang menggunakan bantuan *software* Matlab Guide, dan untuk membandingkan keakuratan model tabel distribusi frekuensi dilakukan dengan analisis perbandingan selisih hasil simulasi perhitungan ukuran pemusatan, lokasi, dan dispersi data menggunakan kode program yang dirancang dengan *software* Matlab Guide. Adapun data yang digunakan adalah data kontinu dari hasil data simulasi yang dibangkitkan melalui program R. Berikut tahapan analisis penelitian ini:



Gambar 1: Tahapa analisis penelitian

4. Hasil Penelitian

4.1. Metode Sturges

Metode sturges adalah aturan untuk menentukan seberapa luas dipilih data dalam suatu grafik batang pada histogram. Aturan ini dipublikasikan pada jurnal *American Statistical Association* (ASA) pada tahun 1926. Sturges menganggap jumlah data dalam sebuah interval adalah dari koefisien binomial, dan sturges mengasumsikan bahwa satu kumpulan data terdistribusi baik didekati dengan distribusi binomial dengan probabilitas 0,5 (memberikan distribusi simetris). Pada penurunan aturan tersebut digunakan identitas ekspansi binomial dan sifat logaritma, sehingga aturan sturges adalah

$$K \approx 1 + 3,322 \log n$$

4.2. Metode Scott

Metode scott adalah sebuah aturan yang memberikan sebuah pilihan dari nilai lebar (panjang) kelas dalam sebuah histogram dengan ukuran data kontinu. Dalam jurnal yang dipublikasikan Biometrika Trust, Aturan ini dideskripsikan berasal dari perhitungan ketidaksesuaian dalam integrasi rata-rata kesalahan kuadrat (*Integrated Mean Squared Error*) pada lebar kelas interval. Aturan tersebut didefinisikan sebagai berikut:

$$p = 3,4908 sn^{1/3}$$

4.3. Proses Pembangkitan Data Simulasi dan Perhitungan Ukuran Pemusatan Data, Lokasi, dan Dispersi Data

Dalam penelitian ini, data simulasi yang dibangkitkan menggunakan program R, dengan data berdistribusi normal dan berdistribusi eksponensial. Dengan ukuran sampel dari 20, 25, 30, ..., 265 data (50 data simulasi) dan dengan parameter rata-rata (μ) adalah 50 dan standar deviasi (σ) adalah 5. Fungsi perintah untuk membangkitkan data simulasi berdistribusi normal adalah:

$$rnorm(n, \mu, \sigma)$$

dan fungsi perintah untuk membangkitkan data eksponensial adalah:

$$rexp(n, \frac{1}{\mu})$$

Masing-masing hasil data simulasi dari program yang dijalankan, disimpan dalam variabel $X_1, X_2, X_3, \dots, X_{50}$ untuk data berdistribusi normal, dan variabel $Y_1, Y_2, Y_3, \dots, Y_{50}$ untuk data berdistribusi eksponensial. Data tersebut merupakan data tunggal sehingga dalam penelitian ini diolah menjadi distribusi frekuensi berkelompok. Dari data tersebut dilakukan perhitungan kriteria ukuran pemusatan data yaitu rata-rata dan median, lalu ukuran lokasi yang dihitung adalah kuartil 1 dan kuartil 3, serta ukuran dispersi yang hitung adalah varians dan standar deviasi. Perhitungan dilakukan dengan merancang program dalam Matlab. Hasil perhitungan kriteria ukuran dari data simulasi pada variabel X_i ditunjukkan pada Gambar 2, dan dilakukan perhitungan yang sama untuk semua data simulasi berdistribusi normal dan berdistribusi eksponensial.



Parameter	Nilai
UKURAN DATA TUNGGAL	
Rata-Rata Tunggal =	51.764
Median Tunggal =	52.1017
Kuartil 1 Tunggal =	47.5424
Kuartil 3 Tunggal =	55.8975
Varians Tunggal =	26.3355
Standar Deviasi Tunggal =	5.1318

Gambar 2: Hasil Perhitungan Ukuran Pemusatan, Lokasi, Dispersi Data X_i

Untuk semua data simulasi dilakukan tabulasi menjadi tabel distribusi frekuensi, masing-masingnya dengan menggunakan metode sturges dan metode scott. Pembentukan tabel distribusi frekuensi tersebut didapatkan dari hasil program aplikasi yang dirancang pada Matlab. Hasil keluaran program untuk data tunggal pada sampel X_i menjadi model tabel distribusi frekuensi dengan metode sturges, ditunjukkan pada Gambar 3. Untuk model dengan metode scott, ditunjukkan pada Gambar 4.

Dari Gambar 3, terlihat bahwa model tabel distribusi frekuensi data normal dengan metode sturges tidak mencakup keseluruhan data, di mana hanya 19 data dari 20 data, sedangkan model menggunakan metode scott pada Gambar 4, dapat mencakup keseluruhan data. Untuk data eksponensial terlihat hal yang sama pula, yaitu pada model sturges, tabel distribusi frekuensi tidak mencakup keseluruhan data.

Tabel Distribusi Frekuensi
dengan Metode Sturges

Data: 43.72728 43.72919 50.40886 50.43863 60.24156 53.15687 52.81946 56.92664 54.08441 55.55451 51.79754 45.36

Proses:

Hasil:

Parameter	Nilai
Banyak Data (n)	20
Min Data	43.7273
Max Data	60.2416
Rentang Data (Range)	16.5143
Banyak Kelas (k) = $1 + 3.322 \times \log n$	5.322
Banyak Kelas (k) dibulatkan	6
Panjang Kelas = R/k	2.7524
NILAI FREKUENSI	
1 43.7273 - 46.4796	4
2 46.4797 - 49.232	2
3 49.232 - 51.9844	4
4 51.9844 - 54.7368	4
5 54.7368 - 57.4892	4
6 57.4892 - 60.2416	1
Jumlah =	19

Gambar 3: Tabel Distribusi Frekuensi Sturges Sampel X_i

Tabel Distribusi Frekuensi
dengan Metode Scott

Data: 43.72728 43.72019 50.40886 50.43863 60.24156 53.15687 52.81046 56.02664 54.08441 55.55451 51.70754 45.36

Proses:

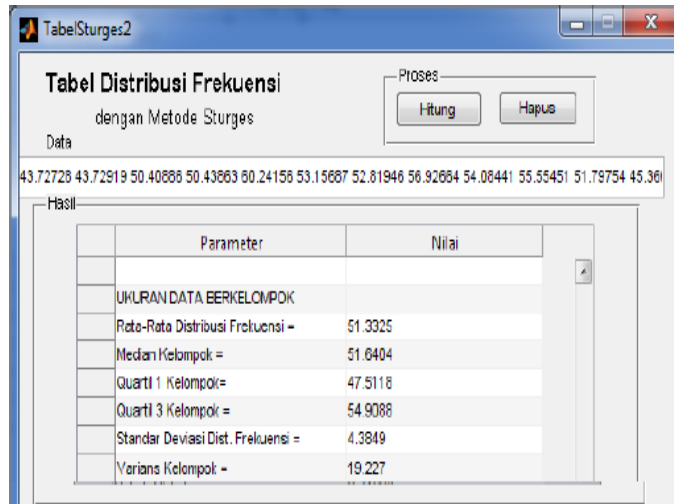
Hasil:

Parameter	Nilai
Banyak Data (n)	20
Min Data	43.7273
Max Data	60.2416
Rentang Data (Range)	16.5143
Panjang Kelas = $3.5 \times n^{-1/3}$	6.5996
Banyak Kelas (k) = R/p	2.5023
Banyak Kelas dibulatkan	3
NILAI FREKUENSI	
1 43.7273 - 50.3269	6
2 50.3269 - 56.9265	9
3 56.9265 - 63.5261	5
Jumlah =	20

Gambar 4: Tabel Distribusi Frekuensi Scott Sampel X_i

4.4. Perhitungan Ukuran Pemusatan, Lokasi, dan Dispersi Data pada Tabel Distribusi Frekuensi

Masing-masing model tabel distribusi frekuensi dengan data normal dan eksponensial, dilakukan perhitungan kriteria ukuran data seperti pada data tunggalnya, yaitu rata-rata, median, kuartil 1, kuartil 3, varians, dan standar deviasi. Perhitungan dilakukan dengan merancang program pada Matlab. Hasil perhitungan kriteria ukuran data dari tabel distribusi frekuensi sturges pada sampel X_i ditunjukkan pada Gambar 5.



Gambar 5: Perhitungan Ukuran Data Berkelompok dengan Metode Sturges dari Sampel X_i

4.5. Analisis Perbandingan Keakuratan Model Tabel Distribusi Frekuensi

Perbandingan keakuratan didapatkan dengan membandingkan selisih kriteria ukuran data yang telah dihitung sebelumnya pada data tunggal dan data berkelompok dalam masing-masing tabel distribusi frekuensi. Untuk data berdistribusi normal, perbandingan selisih kriteria ukuran data pada masing-masing sampel data simulasi ditunjukkan pada Tabel 1, Tabel 2, dan Tabel 3.

Berdasarkan hasil perbandingan dalam 50 percobaan simulasi dengan ukuran sampel 20 sampai 265, Untuk data berdistribusi normal, model tabel distribusi frekuensi menggunakan metode scott lebih akurat dibandingkan dengan metode sturges, persentasi perbandingan hasil simulasi menunjukkan 58% nilai rata-rata, 58% nilai Q_1 , 52% nilai Q_3 , 86% nilai varians dan 84% nilai standar deviasi lebih akurat ke metode scott. Hanya pada nilai median 58% lebih akurat menggunakan metode sturges dibandingkan metode scott.

Untuk data berdistribusi eksponensial, model tabel distribusi frekuensi menggunakan metode scott juga lebih akurat dibandingkan dengan metode sturges, persentasi hasil simulasi menunjukkan 74% nilai median, 54% nilai Q_1 , 64% nilai Q_3 , 100% nilai varians dan 98% nilai standar deviasi lebih akurat dibandingkan dengan metode sturges. Hanya pada nilai rata-rata 62% lebih akurat menggunakan metode sturges dibandingkan metode scott.

Tabel 1: Perbandingan Selisih Nilai Rata-Rata dan Nilai Median dari Data Normal

Ukuran Sampel	Rata-Rata		Keterangan Lebih Baik	Median		Keterangan Lebih Baik
	Sturges	Scott		Sturges	Scott	
20	0,4315	1,5327	Sturges	0,4614	1,1583	Sturges
25	0,3203	0,3710	Sturges	0,0095	0,5645	Sturges
30	0,4227	0,3133	Scott	0,2516	0,3083	Sturges
35	0,0374	0,0504	Sturges	0,0326	0,1440	Sturges

40	0,0471	0,2697	Sturges	0,7503	1,0614	Sturges
45	0,2992	0,0403	Scott	0,2504	0,1470	Scott
50	0,3614	0,1284	Scott	0,5692	0,4433	Scott
55	0,0137	0,1104	Sturges	0,1259	0,0189	Scott
60	0,2406	0,1898	Scott	0,9653	0,9672	Sturges
65	0,2033	0,3770	Sturges	0,2366	0,0530	Scott
70	0,2887	0,2328	Scott	0,7150	0,1625	Scott
75	0,1850	0,1755	Scott	0,1617	0,0323	Scott
80	0,2438	0,0721	Scott	0,5532	0,3951	Scott
85	0,1918	0,1010	Scott	0,1080	0,2488	Sturges
90	0,2052	0,2054	Sturges	0,1127	0,3789	Sturges
95	0,0092	0,0776	Sturges	0,2088	0,0142	Scott
100	0,2242	0,0351	Scott	0,1578	0,1324	Scott
105	0,3457	0,0338	Scott	0,5833	0,1382	Scott
110	0,0953	0,0709	Scott	0,0477	0,3062	Sturges
115	0,1362	0,0228	Scott	0,0959	0,2220	Sturges
120	0,1081	0,0119	Scott	0,1762	0,1085	Scott
125	0,1063	0,1302	Sturges	0,2285	0,3605	Sturges
130	0,0633	0,0623	Scott	0,0159	0,1448	Sturges
135	0,1698	0,2560	Sturges	0,1563	0,2742	Sturges
140	0,1103	0,0252	Scott	0,3038	0,2521	Scott
145	0,1635	0,1689	Sturges	0,0227	0,3105	Sturges
150	0,1520	0,0439	Scott	0,2386	0,0238	Scott
155	0,0502	0,0916	Sturges	0,0228	0,1279	Sturges
160	0,1409	0,0111	Scott	0,0526	0,2832	Sturges
165	0,1705	0,0669	Scott	0,2663	0,2634	Scott
170	0,1270	0,0252	Scott	0,2853	0,0701	Scott
175	0,0105	0,0696	Sturges	0,0855	0,2846	Sturges
180	0,0255	0,0456	Sturges	0,5645	0,2918	Scott
185	0,0874	0,1061	Sturges	0,2663	0,1269	Scott
190	0,0478	0,0889	Sturges	0,0897	0,0957	Sturges
195	0,0229	0,0147	Scott	0,0422	0,1010	Sturges
200	0,1162	0,0941	Scott	0,0361	0,0359	Scott
205	0,1790	0,0730	Scott	0,0603	0,2571	Sturges
210	0,1040	0,1103	Sturges	0,0878	0,1355	Sturges
215	0,0102	0,0241	Sturges	0,0094	0,0189	Sturges
220	0,0922	0,0891	Scott	0,1074	0,2271	Sturges
225	0,1194	0,0035	Scott	0,0301	0,0657	Sturges
230	0,0631	0,1721	Sturges	0,0547	0,0614	Sturges
235	0,0152	0,0904	Sturges	0,2105	0,2537	Sturges
240	0,0605	0,0123	Scott	0,0427	0,0034	Scott
245	0,1811	0,0530	Scott	0,0159	0,0248	Sturges
250	0,0042	0,0010	Scott	0,0764	0,0622	Scott
255	0,0559	0,0295	Scott	0,0392	0,0089	Scott
260	0,0044	0,1461	Sturges	0,0667	0,0928	Sturges
265	0,0724	0,0710	Scott	0,1224	0,2648	Sturges

Tabel 2: Perbandingan Selisih Nilai Q_1 dan Nilai Q_3 dari Data Normal

Ukuran Sampel	Q_1		Keterangan Lebih Baik	Q_3		Keterangan Lebih Baik
	Sturges	Scott		Sturges	Scott	
20	0,0306	1,6846	Sturges	0,9887	1,0290	Sturges
25	0,2028	0,0539	Scott	1,6858	1,7686	Sturges
30	0,7585	0,0227	Scott	1,0285	0,0095	Scott
35	0,1256	0,0736	Scott	0,3198	0,3113	Scott
40	0,4763	0,1324	Scott	0,2876	0,1910	Scott
45	0,2051	0,0209	Scott	0,6902	0,3170	Scott
50	0,6016	0,1264	Scott	0,2269	0,3269	Sturges
55	0,2260	0,3951	Sturges	0,1254	0,0054	Scott
60	0,0132	0,2828	Sturges	0,1600	0,0857	Scott
65	0,2265	0,1292	Scott	1,3346	0,0862	Scott
70	0,5916	0,0067	Scott	0,0027	0,4633	Sturges
75	0,0188	0,1248	Sturges	0,2697	0,7672	Sturges
80	0,5334	0,0242	Scott	0,3943	0,3988	Sturges
85	0,4372	0,3784	Scott	0,0279	0,0942	Sturges
90	0,2210	0,0057	Scott	0,3138	0,2469	Scott
95	0,1004	0,4668	Sturges	0,0153	0,1641	Sturges
100	0,1781	0,1646	Scott	0,1860	0,1333	Scott
105	0,2799	0,1408	Scott	0,2883	0,0346	Scott
110	0,0081	0,0302	Sturges	0,1727	0,1108	Scott
115	0,0042	0,4582	Sturges	0,1435	0,6158	Sturges
120	0,0265	0,4029	Sturges	0,0529	0,4279	Sturges
125	0,0831	0,1554	Sturges	0,2243	0,3204	Sturges
130	0,1309	0,2254	Sturges	0,0295	0,0015	Scott
135	0,2095	0,4144	Sturges	0,4179	0,1853	Scott
140	0,1159	0,1355	Sturges	0,0641	0,1954	Sturges
145	0,1461	0,2670	Sturges	0,1750	0,2101	Sturges
150	0,1738	0,1221	Scott	0,1621	0,2781	Sturges
155	0,3549	0,2418	Scott	0,0378	0,6370	Sturges
160	0,5268	0,4050	Scott	0,1556	0,1697	Sturges
165	0,3016	0,0170	Scott	0,1517	0,0554	Scott
170	0,3229	0,0114	Scott	0,2531	0,1226	Scott
175	0,0357	0,1535	Sturges	2,0093	0,0849	Scott
180	0,0007	0,3586	Sturges	0,4270	0,0249	Scott
185	0,0786	0,1900	Sturges	0,1572	0,6163	Sturges
190	0,0673	0,1425	Sturges	0,1468	0,0779	Scott
195	0,0235	0,0465	Sturges	0,2615	0,2862	Sturges
200	0,2566	0,1461	Scott	0,2965	0,1722	Scott
205	0,2838	0,1571	Scott	0,4092	0,5205	Sturges
210	0,1570	1,1527	Sturges	0,1265	0,0414	Scott
215	0,0002	0,0218	Sturges	0,1072	0,0232	Scott
220	0,4739	0,1213	Scott	0,2943	0,3749	Sturges
225	0,2427	0,0967	Scott	0,0078	0,0697	Sturges
230	0,3679	0,1021	Scott	0,1132	0,0496	Scott
235	0,3166	0,1855	Scott	0,3059	0,0534	Scott

240	0,1163	0,0514	Scott	0,4079	0,2791	Scott
245	0,2076	0,1059	Scott	0,0023	0,1191	Sturges
250	0,0293	0,0145	Scott	0,3644	0,0941	Scott
255	0,0351	0,0230	Scott	0,1432	0,2392	Sturges
260	0,1755	0,0835	Scott	0,0036	0,0377	Sturges
265	0,1167	0,2244	Sturges	0,3079	0,0164	Scott

Tabel 3: Perbandingan Selisih Nilai s^2 dan Nilaisdari Data Normal

Ukuran Sampel	S^2		Keterangan Lebih Baik	Q		Keterangan Lebih Baik
	Sturges	Scott		Sturges	Scott	
20	7,1086	1,2341	Scott	0,7469	0,1217	Scott
25	0,2261	2,3715	Sturges	0,0244	0,2648	Sturges
30	6,7034	1,8445	Scott	0,6758	0,1770	Scott
35	2,908	0,1917	Scott	0,3571	0,0226	Scott
40	2,3953	1,2888	Scott	0,2546	0,1316	Scott
45	3,0642	0,4661	Scott	0,3224	0,0476	Scott
50	5,3537	1,1112	Scott	0,4505	0,0907	Scott
55	2,0173	1,1287	Scott	0,2328	0,1287	Scott
60	2,7623	1,2306	Scott	0,3090	0,1350	Scott
65	2,3792	0,3502	Scott	0,2566	0,0366	Scott
70	4,3682	0,8986	Scott	0,4518	0,0896	Scott
75	1,7041	0,8573	Scott	0,1594	0,0785	Scott
80	2,7936	1,7000	Scott	0,2846	0,1655	Scott
85	3,3657	0,4931	Scott	0,4084	0,0567	Scott
90	1,8099	0,9036	Scott	0,1728	0,0856	Scott
95	0,9699	0,1981	Scott	0,0942	0,0191	Scott
100	1,0269	0,4023	Scott	0,1033	0,0402	Scott
105	3,0713	0,1133	Scott	0,3072	0,0110	Scott
110	1,7748	0,4346	Scott	0,1765	0,0423	Scott
115	1,1856	0,1873	Scott	0,1159	0,0181	Scott
120	1,3858	0,8044	Scott	0,1449	0,0822	Scott
125	2,3845	0,2952	Scott	0,2157	0,0262	Scott
130	3,1367	0,1571	Scott	0,3072	0,0149	Scott
135	1,3228	0,0758	Scott	0,1389	0,0078	Scott
140	2,4324	0,1799	Scott	0,2447	0,0177	Scott
145	1,6566	1,0130	Scott	0,1726	0,1026	Scott
150	1,1941	0,7325	Scott	0,129	0,0774	Scott
155	0,4247	0,0569	Scott	0,0406	0,0055	Scott
160	0,3781	0,4169	Sturges	0,0395	0,0436	Sturges
165	0,3567	0,0876	Scott	0,0383	0,0094	Scott
170	0,2796	0,2171	Scott	0,0256	0,4801	Sturges
175	1,4250	0,8186	Scott	0,1306	0,0754	Scott
180	0,6593	0,0795	Scott	0,0669	0,0080	Scott
185	0,4946	0,5053	Sturges	0,0560	0,0565	Sturges
190	0,0445	0,1466	Sturges	0,0046	0,0152	Sturges
195	0,8824	0,3872	Scott	0,0891	0,0389	Scott
200	1,0659	1,9690	Sturges	0,1071	0,1921	Sturges
205	0,1604	0,3616	Sturges	0,0172	0,0385	Sturges

210	4,2286	0,0128	Scott	0,0315	0,0492	Sturges
215	4,2043	0,5076	Scott	0,0317	0,0013	Scott
220	0,5581	0,4734	Scott	0,0571	0,0485	Scott
225	1,1915	1,0189	Scott	0,1092	0,0918	Scott
230	0,8685	0,1688	Scott	0,0906	0,0174	Scott
235	0,6737	0,2018	Scott	0,0687	0,0207	Scott
240	1,2715	0,5777	Scott	0,1344	0,0598	Scott
245	0,4008	0,0509	Scott	0,0382	0,0054	Scott
250	0,3811	0,3797	Sturges	0,0374	0,0373	Scott
255	0,9062	0,6121	Scott	0,0870	0,0589	Scott
260	0,4415	0,0871	Scott	0,0452	0,0089	Scott
265	0,9140	0,5796	Scott	0,0919	0,0575	Scott

Pada percobaan simulasi dengan ukuran sampel 200 sampai 445, didapatkan perbandingan keakuratan tabel distribusi frekuensi dengan metode scott lebih akurat dari metode sturges. Untuk data normal, persentase perbandingannya adalah 70% nilai rata-rata, 56% nilai median, 64% nilai μ_1 , 52% nilai μ_3 , 86% nilai s , dan 86% nilai s^2 . Untuk data eksponensial, persentase perbandingannya adalah 66% nilai rata-rata, 52% nilai median, 86% nilai μ_1 , 60% nilai μ_3 , 100% nilai s , dan 100% nilai s^2 . Hasil perbandingan menunjukkan keakuratan yang lebih tinggi dibandingkan percobaan simulasi dengan ukuran sampel 20 sampai 265.

5. Kesimpulan

Dari hasil persentase perbandingan keakuratan antara model tabel distribusi frekuensi menggunakan metode sturges dan metode scott, dapat diambil kesimpulan bahwa keakuratan model scott lebih baik dari model sturges pada data normal, dan terjadi peningkatan persentase keakuratan pada data eksponensial. Pada percobaan simulasi dengan ukuran sampel yang lebih tinggi yaitu ukuran sampel 200 sampai 445, terjadi pula peningkatan persentase keakuratan, dengan model scott tetap lebih baik dari model sturges. Pada penelitian ini, nilai pengukuran dari model sturges kurang akurat dibandingkan dengan model scott, dikarenakan adanya data yang tidak masuk dalam tabel distribusi frekuensi, diakibatkan nilai panjang kelas interval tidak mencakup pengelompokkan sampai data tertinggi.

Referensi

- [1] Sturges, H. (1926). The Choice of A Class Interval. Journal American Statistical Association. 21.65-66
- [2] Dajan, A.nto. (1986). Pengantar Metode Statistik 2. LP3ES. Jakarta.
- [3] Hyndman, Rob J. (1995). The Problem with Sturges's Rule for Constructing Histograms. A shortnotes. Unpublished Article.
- [4] Scott, David W. (1979). On Optimal Data-Based Histograms. Biometrika Trust. 3.605-610