HUKUM LOTKA
Mengenai Produktifitas Pengarang

Oleh: B. Mustafa
mus@ipb.ac.id dan mustafa.mustari@gmail.com

Abstrak:
Terdapat hubungan matematis antara jumlah pengarang dengan jumlah karya yang dihasilkan. Fenomena ini sesungguhnya telah diteliti oleh Lotka pada tahun 1926, namun baru pada tahun 1949 dinyatakan sebagai **hukum Lotka tentang produktifitas pengarang**. Hukum Lotka termasuk salah satu objek kajian utama dalam ilmu bibliometrika. Menurut Lotka jika ada seratus orang yang menghasilkan 1 karya, maka akan ada seperempat dari 100 orang yang menghasilkan 2 karya, akan ada sepersembilan dari 100 orang yang akan menghasilkan 3 karya, dan akan ada seperenambelas dari 100 orang yang menghasilkan 4 karya, demikian seterusnya. Ini yang disebut hukum kuadrat terbalik dari Lotka mengenai produktifitas pengarang. Sejak tahun 1973 Hukum Lotka makin banyak dikembangkan oleh para peneliti. Dalam tulisan ini juga diuraikan teknik penghitungan dengan metoda statistik untuk melihat hubungan nyata antara jumlah pengarang dan jumlah karya yang dihasilkan. Selain itu dibahas cara pengujian **Kolmogorov-Smirnov (Uji K-S)** menggunakan program otomasi langsung secara online via internet. Uji K-S adalah suatu uji dalam kelompok Statistik Nonparametrik untuk melihat derajat keeratan dua set data atau dua sebaran (*Goodness of fit*).

Kata kunci: Lotka's law of authors productity, bibliometrics, Kolmogorov-Smirnov Test.

Alfred J. Lotka (1880-1949) seorang statistikawan yang juga seorang ahli kependudukan pada tahun 1926 melakukan penelitian mengenai produktifitas pengarang dalam menghasilkan karya ilmiah. Hasil penelitiannya kemudian diterbitkan dalam *Journal of the Washington Academy of Science,* dengan judul *The frequency distribution of scientific productivity*.

Ia menghitung jumlah nama pengarang perseorangan (pengarang badan korporasi diabaikan) yang terdapat dalam *Chemical Abstracts* antara tahun 1907 sampai 1916. Nama yang diamati hanya

pengarang yang nama keluarganya berawalan A dan B, sehingga didapatkan 6891 nama. Selain itu diteliti juga nama-nama pengarang dari jurnal Anerbach's Geschictstafeln der Physik hanya untuk tahun 1900. Kali ini semua abjad diambil sehingga didapatkan 1325 nama. Jika ada karya yang pengarangnya lebih dari satu, maka yang diambil hanya satu pengarang yaitu pengarang utama atau pengarang pertama atau pengarang "senior".

Menentukan satu pengarang dari satu karya yang banyak pengarangnya merupakan suatu hal yang penting dalam analisis hukum Lotka. Untuk itu biasa digunakan salah satu dari tiga cara yaitu ***adjusted count, complete count*** atau ***straight count***. Lotka dalam penelitiannya menggunakan prinsip ***straight count*** atau biasa juga disebut ***senior count*** atau ***primary count***. Cara menentukan pengarang dengan teknik adjusted count dan complete count banyak juga digunakan oleh para ahli lain. Pada **Lampiran 1** tulisan ini dijelaskan ketiga cara menentukan jumlah karangan setiap pengarang pada dokumen yang dikarang bersama lebih dari satu pengarang.

Data jumlah pengarang dan jumlah karyanya kemudian oleh Lotka dipetakan pada suatu grafik koordinat xy dengan skala logaritma. Hasilnya adalah titik-titik yang dibentuk dengan dengan sumbu x adalah jumlah pengarang dan sumbu y adalah jumlah karyanya. Ketika diamati maka titik-titik yang tersebut tersebar sekitar suatu garis lurus dengan sudut kemiringan dua. Dengan kata lain ada hubungan terbalik antara jumlah karya yang dihasilkan dengan jumlah penulisnya.

Rumus umum yang menunjukkan hubungan antara jumlah pengarang (sumbu y) yang menghasilkan sejumlah karya tertentu (sumbu x) kemudian disebut sebagai hukum kuadrat terbalik, yang model matematikanya seperti berikut:

$$\mathbf{Y = k / X^n}$$

Untuk n = 2 berlaku hukum kuadrat terbalik produktifitas ilmiah sebagai berikut:

$$\mathbf{Y_1 = k/1^2}$$
$$\mathbf{Y_2 = k/2^2}$$
$$\mathbf{Y_3 = k/3^2}$$
$$\mathbf{\ldots}$$
$$\mathbf{Y_n = k/n^2}$$

Jika dijumlahkan hasilnya sebagai berikut:

$$\mathbf{\Sigma\, y_n = k\,(1/1^2 + 1/2^2 + 1/3^2 + \ldots 1/n^2)}$$
$$\mathbf{= k\,(\Sigma\, 1/x^2),}$$ ingat bahwa deret harmonik $\mathbf{\Sigma\, 1/x^2 = \Pi^2/6}$ sehingga
$$\mathbf{= k\,\Pi^2/6}$$
$$\mathbf{6\,\Sigma\, y_n = k\,\Pi^2}$$
$$\mathbf{k = 6\,\Sigma\, y_n/\Pi^2,}$$ ingat juga bahwa $\mathbf{\Pi = 3.1416}$ dan $\mathbf{\Sigma\, y_n \approx 1,}$ maka
$$\mathbf{k = 6 \times 1/\Pi^2 = 6/(3.1416)^2 = 6/9.87 = 0{,}6079}$$ atau **60.79** persen

Jadi menurut dalil kuadrat terbalik, proporsi dari penulis yang mempunyai kontribusi satu karya adalah sekitar 60 persen.

Disimpulkan juga bahwa:

- Jumlah pengarang yang menghasilkan 2 karya adalah seperempat dari yang membuat satu karya,
- Jumlah pengarang yang menghasilkan 3 karya adalah sepersembilan dari yang membuat satu karya, dan seterusnya,
- Jumlah pengarang yang membuat n karya adalah seper-n pangkat dua dari yang membuat satu karya.

Dalam penelitian Lotka terungkap bahwa:

- Hanya ada beberapa orang pengarang yang membuat karya dalam jumlah banyak
- Banyak pengarang yang hanya menghasilkan satu atau dua karya dalam suatu bidang
- Pada data subjek Fisika, ada 59.2 persen dari total 1325 pengarang yang mengarang hanya satu karya
- Pada data subjek Kimia, untuk abjad A ada 57.7 persen dari 1543 pengarang yang mengarang satu karya dan untuk abjad B ada 57.98 persen dari 5348 pengarang yang mengarang satu karya. Sedangkan untuk data gabungan ada 57.9 persen dari 6891 pengarang yang mengarang satu karya.
- Untuk subjek Kimia, konstanta c = 0,5669 dan n = 1.888
- Sedangkan subjek Fisika, konstanta c = 0,6079 dan n = 2.02.

Lebih lengkap hasil pengamatan dan perhitungan Lotka dimuat dalam **Lampiran 2**.

Artikel Lotka yang diterbitkan tahun 1926 baru mulai dikutip pada tahun 1941 dan dalil sebaran ini baru disebut sebagai Hukum atau Dalil Lotka pada tahun 1949. Price mengembangkan Dalil Lotka dan menyatakan bahwa ada 50 persen dari publikasi ilmiah ditulis oleh 60 persen pengarang. Juga ditemukan bahwa rata-rata ilmuwan menghasilkan tiga karya selama hidupnya.

Percobaan untuk menguji penerapan Dalil Lotka pada disiplin ilmu lain baru banyak dilakukan sejak tahun 1973. Ternyata terungkap bahwa tidak selamanya Dalil Lotka bisa diterapkan. Pada disiplin ilmu tertentu atau cara pengumpulan data menentukan apakah Dalil Lotka bisa digunakan atau tidak. Tahun 1974 Voos menyatakan pada subjek *information science*, rasionya adalah $1/n^{3.5}$.

Untuk menguji apakah dalil Lotka dapat digunakan pada sekelompok data tertentu biasanya digunakan uji K-S atau uji **Kolmogorov-Smirnov (*the Kolmogorov-Smirnov Test*)**. Salah satu instrumen uji statistik dengan metode nonparametrik adalah uji K-S ini. Instrumen uji statistik digunakan untuk mengetahui perbedaan yang nyata (signifikan) antara distribusi frekuensi pengamatan dengan distribusi frekuensi teoritis. Uji K-S juga merupakan ukuran ketepatan (***goodness of fit***) suatu distribusi frekuensi teoritis (frekuensi harapan), seperti pada pengujian ***Chi-Square*** yang digunakan pada metode statistika parametrik. Seperti diketahui bahwa uji statistik parametrik biasa digunakan pada set data yang dapat diasumsikan terdistribusi atau menyebar secara normal. Pada data dari populasi yang sangat kecil biasa digunakan uji statistik nonparametrik. Karena asumsi bahwa objek yang digunakan sebagai sampel berasal dari populasi yang terdistribusi normal sangat sulit dilakukan pada populasi yang anggotanya sedikit.

Uji Kolmogorov-Smirnov (Uji K-S)

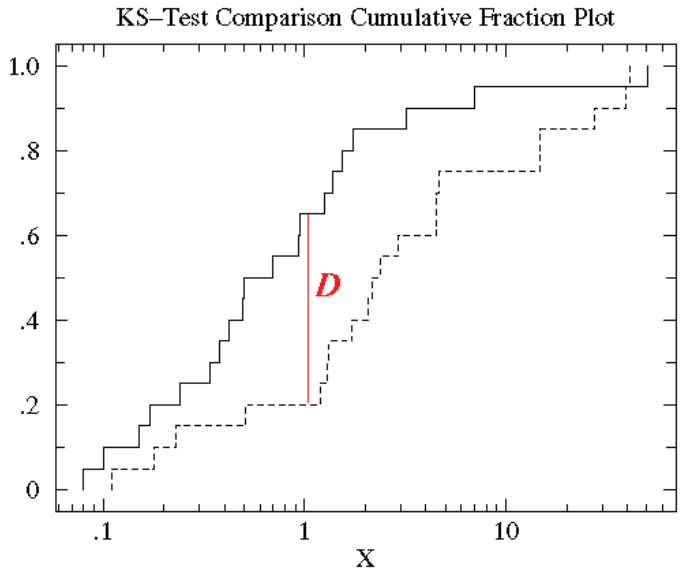
$Dmaks = |F_o(x) - S_n(x)|$

$F_o(x)$ = fungsi frekuensi kumulatif secara teoritis (harapan)

$S_n(x)$ = fungsi frekuensi kumulatif pengamatan.

Nilai Dmaks adalah deviasi absolut (mutlak) tertinggi, berupa selisih positif tertinggi antara frekuensi harapan dan frekuensi pengamatan. Nilai ini untuk membuat dugaan mengenai keeratan antara distribusi frekuensi pengamatan dengan frekuensi teoritis. Peluang distribusi Dmaks tidak tergantung pada banyaknya sampel (n) yang digunakan dan tidak tergantung dari distribusi frekuensi teoritis (harapan)nya.

Untuk memberi gambaran jelas, ilustrasi secara grafik nilai Dmaks dapat dilihat pada gambar berikut:



D pada gambar diatas menunjukkan jarak terbesar antara distribusi set data.

Pada tingkat kepercayaan 0.01, Statistik K-S = 1.63/√N, dimana N adalah jumlah total pengarang. Jika Dmaks > K-S, dimana Dmaks adalah nilai mutlak (positif) selisih nilai pengamatan dan nilai perhitungan teoritis, maka ini berarti sebaran contoh pengamatan tidak tepat dengan sebaran teoritis. (Lihat tabel uji K-S selengkapnya pada **Lampiran 3**)

Dalam kasus subjek Kimia yang diteliti Lotka:

K-S = 1.63 /√ 6891 = 0.0195

Sedangkan Dmaks= 0.0287 > dari nilai K-S, maka sebaran tidak tepat menggambarkan sebaran teoritis.

Dalam kasus subjek Fisika yang diteliti Lotka:

K-S = 1.63 /√1325 = 0.0448

Sehingga Dmaks=0.0253 < dari nilai K-S, maka sebaran tepat menggambarkan sebaran teoritis.

Perhatikan tabel berikut yang merupakan perhitungan oleh Lotka untuk Subjek Fisika

Jumlah Karya (x)	Hasil Perhitungan (f'x)	Hasil Pengamatan (fx)	$F_0(x)$ >f'x	$S_n(x)$ >fx	D_{maks} $\lvert F_0(x)-S_n(x)\rvert$
1	0,6079	0,5917	0,6079	0,5917	0,0162
2	0,1520	0,1540	0,7599	0,7457	0,0142
3	0,0675	0,0958	0,8274	0,8415	0,0141
4	0,0380	0,0377	0,8654	0,8792	0,0138
5	0,0243	0,0249	0,8897	0,9041	0,0144
6	0,0169	0,0211	0,9066	0,9252	0,0186 ◄
.	.	.	.	.	.

Dmaks = 0,0186 N = 1325
Pada tingkat kepercayaan 0,05 maka statistik K-S = 1.36/√N =1,36/√1325 = 0,0374

Dmaks = 0.0186 < K-S = 0,0374, maka tidak ada perbedaan yang signifikan antara rumus kuadrat terbalik dari Lotka dengan data hasil pengamatan. Ini berarti data tersebut sesuai dengan batasan produktifitas ilmiah Lotka berdasarkan rumus kuadrat terbalik.

Untuk subjek Kimia, perhatikan tabel berikut:

Jumlah Karya (x)	Hasil Pengamatan (fx)	$S_n(x)$ >fx	Hasil Perhitungan (f'x)	$F_0(x)$ >f'x	D_{maks} $\lvert F_0(x)-S_n(x)\rvert$
1	0,5792	0,5792	0,6079	0,6079	0,0287 ◄
2	0,1537	0,7329	0,1520	0,7599	0,0270
3	0,0715	0,8044	0,0675	0,8274	0,0230
4	0,0416	0,8460	0,0380	0,8654	0,0194
5	0,0267	0,8727	0,0243	0,8897	0,0170
6	0,0190	0,8917	0,0169	0,9066	0,0149
7	0,0164	0,9081	0,0124	0,9190	0,0109
8	0,0123	0,9204	0,0095	0,9285	0,0081
9	0,0093	0,9297	0,0075	0,9360	0,0063
10	0,0094	0,9391	0,0061	0,9421	0,0030
.	.	.	.	.	.

Dmaks = 0,0287
Pada derajat kepercayaan 0,01 statistik K-S = 1,63/√6891=0,0195
Dmaks= 0,0287 > K-S yang hanya 0,0195, jadi data dari Chemical Abstracts tidak sesuai dengan hukum Lotka.

Sebagai perbandingan perhatikan hasil penelitian lain yang lebih baru yang diringkas dalam tabel berikut ini.

Penelitian Hukum Lotka bidang *Management Information System* (MIS) pada 10 jurnal

Jumlah Karya (x)	Jumlah Pengarang	Persentase Pengarang	$S_n(x)$	$F_0(x)$	D_{maks} $\lvert F_0(x)-S_n(x)\rvert$
1	439	73,91	0,7391	0,6079	0,1312
2	91	15,32	0,8923	0,7599	0,1324◄
3	30	5,05	0,9428	0,8274	0,1154
4	14	2,36	0,9664	0,8654	0,1010
5	7	1,18	0,9782	0,8897	0,0885
6	7	1,18	0,9900	0,9066	0,0834
7	3	0,51	0,9951	0,9190	0,0760
8	0	0	0,9951	0,9285	0,0666
9	1	0,17	0,9968	0,9360	0,0608
10	0	0	0,9968	0,9421	0,0547
11	0	0	0,9986	0,9472	0,0496
12	1	0,17	0,9985	0,9514	0,0471
13	1	0,17	1,0002	0,9549	0,0453
	594	100,02			

Dmaks =0,1324. Statistik K-S pada tingkat kepercayaan 0.01 = $1.63/\sqrt{594}$= 0,0669
Data MIS tidak sesuai dengan Hukum Lotka. Namun pada tingkat kepercayan 0.05 sesuai dengan Hukum Lotka. Perhatikan tabel berikut:

Jumlah Karya (x)	Jumlah Pengarang	Persentase Pengarang	$S_n(x)$	$F_0(x)$	D_{maks} $\lvert F_0(x)-S_n(x)\rvert$
1	439	73,91	0,7391	0,7775	0,0384◄
2	91	15,32	0,8923	0,9005	0,0082
3	30	5,05	0,9428	0,9423	0,0005
4	14	2,36	0,9664	0,9618	0,0046
5	7	1,18	0,9782	0,9726	0,0056
6	7	1,18	0,9900	0,9792	0,0008
7	3	0,51	0,9951	0,9836	0,0015
8	0	0	0,9951	0,9867	0,0084
9	1	0,17	0,9968	0,9890	0,0078
10	0	0	0,9968	0,9907	0,0061
11	0	0	0,9986	0,9920	0,0048
12	1	0,17	0,9985	0,9930	0,0055
13	1	0,17	1,0002	0,9938	0,0064
	594	100,02			

N=594, Statitik K-S = 0.0669 < Dmaks-0,0384. Jadi data MIS sesuai dengan hukum Lotka.

Dalam era internet ini, kita semakin dipermudah. Termasuk dalam menghitung nilai-nilai pada uji K-S. Kini tersedia via internet program untuk menghitung nilai Dmaks dalam uji K-S. Selain itu beberapa nilai dari tiap set data (dataset) yang dimasukkan langsung dapat dihitung dan

ditampilkan misalnya **mean, median, standar deviasi** dan sebagainya. Kita hanya perlu mengakses situs program Uji K-S di http://**www.physics.csbsju.edu/stats/KS-test.n.plot_form.html** untuk mendapatkan *template* penghitungan nilai uji K-S seperti pada **Lampiran 4**. Semoga situs itu selalu tersedia, sebagaimana ketika terakhir diakses pada tanggal 2 Mei 2009. Untuk mendapatkan nilai dari Uji K-S, kita hanya perlu memasukkan nilai-nilai dari dua set data pengamatan, lalu mengklik tombol "**Calculate Please**", dan program segera mulai menghitung. Sesaat kemudian beragam nilai hasil perhitungan dari dua set data yang dimasukkan ditampilkan seperti pada **Lampiran 5**.

Kesimpulan

Terdapat hubungan matematis antara jumlah pengarang dengan jumlah karya yang dihasilkan oleh para pengarang tersebut. Fenomena ini diteliti oleh Lotka pada tahun 1926, namun baru beberapa puluh tahun kemudian para ahli menyatakannya sebagai **hukum Lotka tentang produktifitas pengarang**. Hukum Lotka, sebagaimana Hukum Bradford mengenai penyebaran artikel ilmiah pada berbagai jurnal, Hukum Zipf mengenai frekuensi kata, termasuk objek kajian utama dalam ilmu bibliometrika. Menurut Lotka jika ada seratus orang yang menghasilkan 1 karya, maka akan ada seperempat dari 100 orang yang menghasilkan 2 karya, dan akan ada sepersembilan dari seratus orang yang akan menghasilkan tiga karya, demikian seterusnya, sehingga rumus umumnya adalah $\mathbf{Y_n = k/n^2}$. Ini yang disebut hukum kuadrat terbalik dari Lotka mengenai produktifitas pengarang.

Untuk mengetahui bahwa perhitungan dengan rumus dari Hukum Lotka berlaku, dua set data pengamatan perlu diuji menggunakan uji statistik nonparamterik. Uji yang dianggap paling cocok untuk mengukur derajat keeratan dua set data atau dua sebaran (*Goodness of fit*) adalah Uji Kolmogorv-Smirnov (Uji K-S). Kini dengan bantuan fasilitas sistem internet, setiap orang dapat dengan mudah melakukan uji K-S, karena program aplikasi otomatisnya sudah tersedia dan dapat diakses dan dijalankan dengan mudah, yang diperlukan hanyalah akses ke internet.

Kajian Hukum Lotka mengenai produktifitas pengarang ini diharapkan dapat memperluas wawasan pustakawan Indonesia, khususnya mengenai kajian Bibliometrika, untuk melakukan kajian-kajian lain demi pengembangan ilmu perpustakaan pada umumnya di Indonesia.

Daftar Bacaan

Diodato, Virgil. 1994. *Dictionary of Bibliometrics*. New York: The Haworth Press. 185p.

Egghe, L. 2005. *Relations between the continuous and the discrete Lotka power function*. Journal of the American Society for Information Science and Technology, 56 (7): 664–668.

Lotka, A.J. 1926. *The frequency distribution of scientific productivity*. Journal of the Washington Academy of Sciences, 16: 317-323.

Newby, Gregory B, Greenberg, Jane and Jones, Paul. 2003. *Open source software development and Lotka's law: bibliometric patterns in programming*. Source Journal of the American Society for Information Science and Technology, 54(2), January: 169 – 178.

Nicholls, Paul Travis. 1986. *Empirical validation of Lotka's law*. Information Processing and Management: an International Journal, 22(5), September: 417-419.

Nicholls, Paul Travis. 1989. *Bibliometric modeling processes and the empirical validity of Lotka's law*. Journal of the American Society for Information Science, 40(6): 379-85.

Pao, M.L. 1985. *Lotka's Law: A testing procedure*. Information Processing & Management, 21(4): 305-320.

Saam, N.J, Reiter, L. 1999. *Lotka's law reconsidered : The evolution of publication and citation distributions in scientific fields*. Scientometrics, 44 (2): 135-155.

Schorr, A.E. 1975a. *Lotka's Law and map librarianship*. Journal of the American Society for Information Science, 26(5): 189-190.

Schorr, A.E. 1975b. *Lotka's Law and the history of legal medicine*. Research Librarianship, 30: 205-209.

Sengupta, I.N. 1992. *Bibliometrics, informetrics, scientometrics and librametrics: An overview*. Libri, 42(2): 75-98.

Voos, Henry. 1974. *Lotka and Information Science*. Journal of the American Society for Information Science. July-August: 270-272 dalam Augmenting Information Seeking on the World Wide Web using Collaborative Filtering Techniques. Diakses tanggal 11 April 2007. Http://www.gslis.utexas.edu/~donturn/research/augmentis.html

Lampiran 1

Teknik menghitung jumlah karangan seseorang pada dokumen yang dikarang lebih dari satu orang

Misalnya ada empat dokumen yang dikarang oleh lebih dari satu orang pengarang seperti berikut:

Dokumen 1 : oleh pengarang A, B dan C
Dokumen 2 : oleh pengarang B
Dokumen 3 : oleh pengarang C dan A
Dokumen 4 : oleh pengarang B dan D

Penghitungan dengan cara *Adjusted Count:*

Dokumen 1 dikarang oleh tiga orang, sehingga pengarang A mendapat nilai 1/3, B mendapat nilai 1/3 dan pengarang C mendapat nilai 1/3.
Dokumen 2 dikarang oleh satu orang, sehingga pengarang B mendapat nilai 1.
Dokumen 3 dikarang oleh dua orang, sehingga pengarang C dan pengarang A masing-masing mendapat nilai ½.
Dokumen 4 dikarang oleh dua orang, sehingga pengarang B dan pengarang D masing-masing mendapat nilai ½.

Sehingga secara total pengarang A mendapat nilai = 1/3 + ½ = 5/6
Pengarang B mendapat nilai = 1/3 + 1 + ½ = 1 5/6
Pengarang C mendapat nilai = 1/3 + ½ + = 5/6
Pengarang D mendapat nilai = ½

Perhitungan dengan cara *Complete Count*:

Pengarang A mengarang dua kali (dokumen 1 dan dokumen 3) sehingga mendapat nilai 2.
Pengarang B mengarang tiga kali (dokumen 1, 2 dan dokumen 3) sehingga mendapat nilai 3.
Pengarang C mengarang dua kali (dokumen 1 dan dokumen 3) sehingga mendapat nilai 2.
Pengarang D mengarang satu kali (dokumen 4) sehingga mendapat nilai 1.

Perhitungan dengan cara *Straight Count (Senior Count, Primary Count)*:

Pengarang A menjadi pengarang pertama pada dokumen 1 sehingga mendapat nilai 1.
Pengarang B menjadi pengarang pertama pada dokumen 2 dan 4 sehingga mendapat nilai 2.
Pengarang C menjadi pengarang pertama pada dokumen 3, sehingga mendapat nilai 1.
Pengarang D tidak pernah menjadi pengarang pertama dari ke 4 dokumen, sehingga tidak mendapat nilai.

Sumber: Virgil Diodato. *Dictionary of Bibliometrics*. New York: The Haworth Press, 1994.

Lampiran 2 Tabel Lengkap Pengamatan dan Perhitungan Menurut Hukum Lotka

Lampiran 3

Table *Kolmogorov-Smirnov Test*

(If calculated ratio is greater than value shown, then reject the null hypothesis at the chosen level of confidence.)

SAMPLE SIZE (N)	LEVEL OF SIGNIFICANCE FOR D = MAXIMUM [$F_0(X) - S_n(X)$]				
	.20	.15	.10	.05	.01
1	.900	.925	.950	.975	.995
2	.684	.726	.776	.842	.929
3	.565	.597	.642	.708	.828
4	.494	.525	.564	.624	.733
5	.446	.474	.510	.565	.669
6	.410	.436	.470	.521	.618
7	.381	.405	.438	.486	.577
8	.358	.381	.411	.457	.543
9	.339	.360	.388	.432	.514
10	.322	.342	.368	.410	.490
11	.307	.326	.352	.391	.468
12	.295	.313	.338	.375	.450
13	.284	.302	.325	.361	.433
14	.274	.292	.314	.349	.418
15	.266	.283	.304	.338	.404
16	.258	.274	.295	.328	.392
17	.250	.266	.286	.318	.381
18	.244	.259	.278	.309	.371
19	.237	.252	.272	.301	.363
20	.231	.246	.264	.294	.356
25	.210	.220	.240	.270	.320
30	.190	.200	.220	.240	.290
35	.180	.190	.210	.230	.270
OVER 35	$\frac{1.07}{\sqrt{N}}$	$\frac{1.14}{\sqrt{N}}$	$\frac{1.22}{\sqrt{N}}$	$\frac{1.36}{\sqrt{N}}$	$\frac{1.63}{\sqrt{N}}$

Lampiran 4

Program Online untuk Menghitung Nilai K-S
(http://www.physics.csbsju.edu/stats/KS-test.n.plot_form.html)

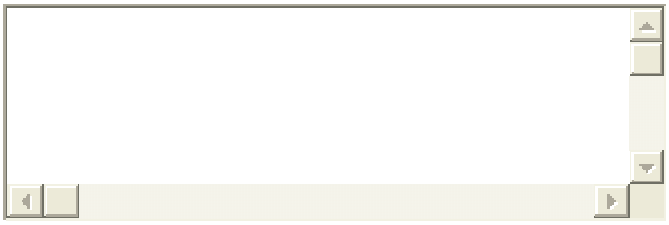
Nilai K-S dapat langsung diketahui setelah memasukkan **dataset1** dan **dataset2**, kemudian mengklik **Calculate please**.

KS-test Data Entry

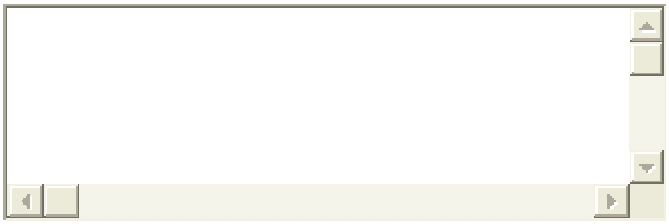
Use the below form to enter your data for a Kolmogorov-Smirnov test. The KS-test seeks differences between your two datasets; it is non-parametric and distribution free. Reject the null hypothesis of no difference between your datasets if *P* is "small". In addition this page reports if your datasets seem to have normal or lognormal distribution. This may allow you to use other tests like the *t*-test.

For each dataset, enter your data into the given box separating each datum from its neighbor with tabs, commas, or spaces. Very commonly you will already have the data in your computer in some format. You should be able to just copy and paste that data into the appropriate area. This KS-test form is designed to handle datasets with between 10 and 1024 items in each dataset.

Dataset 1:



Dataset 2:



Calculate Please

Lampiran 5

Hasil perhitungan nilai K-S dan nilai lain, ketika data pada tabel 3 dihitung menggunakan program penghitungan nilai K-S secara online.

KS Test: Results
Kolmogorov-Smirnov Comparison of Two Data Sets

The results of a Kolmogorov-Smirnov test performed at 01:37 on 9-APR-2007
The maximum difference between the cumulative distributions, *D*, is: 0.3846 with a corresponding *P* of: 0.226

Data Set 1:

13 data points were entered
Mean = 7.000
95% confidence interval for actual Mean: 4.647 thru 9.353
Standard Deviation = 3.89
High = 13.0 Low = 1.00
Third Quartile = 10.5 First Quartile = 3.50
Median = 7.000
Average Absolute Deviation from Median = 3.23
KS finds the data is consistent with a normal distribution: *P*= 1.00 where the normal distribution has mean= 6.998 and sdev= 4.389
KS finds the data is consistent with a log normal distribution: *P*= 0.72 where the log normal distribution has geometric mean= 5.292 and multiplicative sdev= 2.718

Items in Data Set 1:

1.00 2.00 3.00 4.00 5.00 6.00 7.00 8.00 9.00 10.0 11.0 12.0 13.0

Data Set 2:

13 data points were entered
Mean = 45.69
95% confidence interval for actual Mean: -27.30 thru 118.7
Standard Deviation = 121.
High = 439. Low = 0.00
Third Quartile = 22.0 First Quartile = 0.500
Median = 3.000
Average Absolute Deviation from Median = 45.0
John Tukey defined data points as *outliers* if they are 1.5*IQR above the third quartile or below the first quartile. Following Tukey, the following data points are outliers: 439. 91.0
KS says it's unlikely this data is normally distributed: *P*= 0.00 where the normal distribution has mean= 110.9 and sdev= 194.6

Items in Data Set 2:

0.00 0.00 0.00 1.00 1.00 1.00 3.00 7.00 7.00 14.0 30.0 91.0 439.