PEMILIHAN DISTRIBUSI PROBABILITAS PADA ANALISA HUJAN DENGAN METODE *GOODNESS OF FIT TEST*

Togani Cahyadi Upomo[1], **Rini Kusumawardani**[2]
[1)] Jurusan Teknik Sipil, Fakultas Teknik, Universitas Negeri Semarang (UNNES)
Kampus Unnes Gd E4, Sekaran, Gunungpati, Semarang 50229, email: togani.cahyadi@mail.unnes.ac.id
[2)] Jurusan Teknik Sipil, Fakultas Teknik, Universitas Negeri Semarang (UNNES)
Kampus Unnes Gd E4, Sekaran, Gunungpati, Semarang 50229, email: rini.kusumawardani@mail.unnes.ac.id

Abstract: *Rainfall event is a stochastic process, so to explain and analyze this processes the probability theory and frequency analysisare used. There are four types of probability distributions.They are normal, log normal, log Pearson III and Gumbel. To find the best probabilities distribution, it will used goodness of fit test. The tests consist of chi-square and smirnov-kolmogorov. Results of the chi-square test for normal distribution, log normal and log Pearson III was 0.200, while for the Gumbel distribution was 2.333. Results of Smirnov Kolmogorov test for normal distribution D = 0.1554, log-normal distribution D = 0.1103, log Pearson III distribution D = 0.1177 and Gumbel distribution D = 0.095. All of the distribution can be accepted with a confidence level of 95%, but the best distribution is log normal distribution.*

***Keywords** :goodness of fit test, chi square, smirnov kolmogorov, probability distribution, frequency analysis*

Abstrak: Kejadian hujan merupakan proses stokastik, sehingga untuk keperluan analisa dan menjelaskan proses stokastik tersebut digunakan teori probabilitas dan analisa frekuensi. Terdapat empat jenis distribusi probabilitas yaitu distribusi normal, log normal, log pearson III dan gumbel. Untuk mencari distribusi probabilitas terbaik maka akan digunakan pengujian metode goodness of fit test. Pengujian tersebut meliputi uji chi-kuadrat dan uji smirnov kolmogorov. Hasil pengujian chi kuadrat untuk distribusi normal, log normal dan log pearson III adalah 0.200, sedangkan untuk distribusi gumbel 2.333. Hasil pengujian smirnov kolmogorov untuk distribusi normal dengan nilai D = 0.1554, distribusi log normal dengan nilai D = 0.1103, distribusi log pearson III dengan nilai D = 0.1177 dan distribusi gumbel dengan nilai D = 0.095. Seluruh distribusi dapat diterima dengan tingkat kepercayaan 95%, tetapi distribusi terbaik adalah distribusi log normal.

Kata kunci : goodness of fit test, chi kuadrat, smirnov kolmogorov, distribusi probabilitas, analisa frekuensi

PENDAHULUAN

Banjir atau kekeringan akan mengakibatkan dampak negatif bagi kehidupan. Curah hujan yang sangat tinggi akan mengakibatkan banjir dan sebaliknya, jika tidak ada hujan akan mengakibatkan kekeringan. Kejadian hujan merupakan proses stokastik, sehingga untuk keperluan analisa dan menjelaskan proses stokastik tersebut digunakan teori probabilitas dan analisa frekuensi.

Terdapat empat distribusi probabilitas yang cukup dikenal dalam ilmu hidrologi, yaitu : distribusi normal, distribusi log-normal, distribusi log-pearson III dan distribusi gumbel. Untuk mendapatkan model terbaik perlu dilakukan pengujian terhadap masing-masing model tersebut. Pengujian distribusi probabilitas dapat menggunakan metode goodness of fit test, yaitu uji chi-kuadrat dan uji smirnov-kolmogorov.

Penelitian ini merupakan studi literatur tentang cara analisa distribusi probabilitas serta penggunaan metode goodness of fit test yang meliputi uji chi-kuadrat dan smirnov-kolmogorov dalam menentukan distribusi probabilitas yang tepat.

ANALISA STATISTIK DASAR

Terdapat beberapa parameter penting dalam analisa statistik, meliputi rerata, deviasi standar, koefisien varian, koefisien kemencengan dan koefisien kurtosis.

Menurut Triatmodjo (2008), tidak semua variat dari variabel hidrologi sama dengan nilai reratanya, tetapi ada yang lebih besar atau lebih kecil. Besarnya derajad sebaran variat di sekitar nilai reratanya disebut varian (*variance*) atau penyebaran dispersi (*dispersion*). Penyebaran data dapat diukur dengan deviasi standar (*standard deviation*) dan varian.

Nilai rerata dapat dihitung dengan persamaan berikut :

$$\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i \qquad (1)$$

dengan $\bar{x}$ = rerata, x_i= variabel random dan n = jumlah data.

Deviasi standar dapat digunakan untuk mengetahui variabilitas dari distribusi. Semakin besar deviasi standar maka akan semakin besar penyebaran dari distribusi.Deviasi standar dapat dihitung dengan persamaan berikut :

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

Koefisien varian adalah nilai perbandingan antara deviasi standar dan nilai rerata, yang mempunyai bentuk:

$$C_v = \frac{s}{t} \qquad (3)$$

Kemencengan (*skewness*) merupakan derajad ketidaksimetrisan atau dapat juga didefinisikan sebagai penyimpangan kesimetrisan dari suatu distribusi. Jika suatu kurva frekuensi dari suatu distribusi memiliki ekor kurva yang lebih panjang ke arah sisi kanan dibandingkan ke arah sisi kiri dari nilai maksimum tengah, maka distribusi ini dikenal dengan nama distribusi miring ke kanan, atau memiliki kemencengan positif. Untuk kondisi kebalikannya, distribusinya dikenal sebagai distribusi miring ke kiri atau memiliki kemencengan negatif. Untuk mengetahui derajad ketidaksimetrisan (*assymetry*) dari suatu bentuk distribusi. Kemencengan diberikan oleh bentuk berikut:

$$a = \frac{n}{(n-1)(n-2)}\sum_{i=1}^{n}(x_i - \bar{x})^3 \qquad (4)$$

$$C_s = \frac{a}{s^3} \qquad (5)$$

Kurtosis adalah derajat ketinggian puncak atau keruncingan suatu distribusi. Sebuah distribusi yang mempunyai puncak yang relatif tinggi disebut leptokurtik, sementara kurva yang memiliki puncak datar atau rata disebut platikurtik sedangkan kurva dengan puncak yang tidak terlalu runcing ataupun terlalu datar disebut mesokurtik.Koefisien kurtosis diberikan oleh persamaan berikut:

$$C_k = \frac{n}{(n-1)(n-2)(n-3)s^4}\sum_{i=1}^{n}(x_i - \bar{x})^4 \qquad (6)$$

KALA ULANG

Menurut Triatmodjo (2008), periode ulang (return period) didefinisikan sebagai waktu hipotetik dimana debit atau hujan dengan suatu besaran tertentu (XT) akan disamai atau dilampaui sekali dalam jangka waktu tersebut. Berdasarkan data debit atau hujan untuk beberapa tahun pengamatan dapat diperkirakan debit/hujan yang diharapkan disamai atau dilampaui satu kali dalam T tahun; dan debit/hujan tersebut dikenal sebagai debit/hujan dengan periode ulang T tahun atau debit/hujan T Tahunan.

Untuk mencari probabilitas dapat menggunakan persamaan weibull, yaitu :

$$P = \frac{m}{n+1} \qquad (7)$$

sedangkan periode ulang dapat dicari dengan persamaan

$$T_r = \frac{1}{P} \qquad (8)$$

dengan m = nomor urut peringkat data setelah diurutkan dari besar ke kecil, n = banyaknya data atau jumlah kejadian, P = probabilitas, Tr = periode ulang.

DISTRIBUSI PROBABILITAS

Distribusi Normal

Perhitungan dengan distribusi normal secara praktis dapat didekati dengan persamaan sebagai berikut :

$$X_T = \bar{x} + z.s \quad (9)$$

dengan :X_T=perkiraan nilai yang diharapkan terjadi dengan periode ulang T-tahunan,$\bar{X}$=nilai rata-rata hitung variat, s = deviasi standar nilai variat,z=faktor frekuensi dari distribusi normal (tabel z untuk distribusi normal), merupakan fungsi dari peluang atau periode ulang dan tipe model matematik distribusi peluang yang digunakan untuk analisis peluang.

Distribusi Log Normal

Jika Y = log X, maka perhitungan dengan distribusi normal secara praktis dapat didekati dengan persamaan sebagai berikut :

$$Y_T = \bar{Y} + z.s \quad (10)$$

dengan Y_T=perkiraan nilai yang diharapkan terjadi dengan periode ulang T-tahunan,$\bar{Y}$=nilai rata-rata hitung variat, s =deviasi standar nilai variat,z=faktor frekuensi, merupakan fungsi dari peluang atau periode ulang dan tipe model matematik distribusi peluang yang digunakan untuk analisis peluang.

Distribusi Log-Pearson III

Jika Y = log X, maka perhitungan dengan distribusi normal secara praktis dapat didekati dengan persamaan sebagai berikut :

$$Y_T = \bar{Y} + K_T s \quad (11)$$

dengan Y_T=perkiraan nilai yang diharapkan terjadi dengan periode ulang T-tahunan,$\bar{Y}$=nilai rata-rata hitung variat, s =deviasi standar nilai variat,K_T=faktor frekuensi (tabel nilai K_T untuk distribusi log pearson III), nilai K_T ini tergantung dari koefisien kemencengan (*skewness*) dan probabilitasnya.

Distribusi Gumbel

Perhitungan curah hujan rencana menurut metode Gumbel, mempunyai perumusan sebagai berikut:

$$X_T = \bar{X} + s.K \quad (12)$$

dengan X_T = perkiraan nilai yang diharapkan terjadi dengan periode ulang T-tahunan,$\bar{X}$ = nilai rata-rata hitung variat,s = deviasi standar nilai variat,K = faktor frekuensi, merupakan fungsi dari peluang atau periode ulang dan tipe model matematik distribusi peluang yang digunakan untuk analisis peluang.

Faktor probabilitas K untuk harga-harga ekstrim Gumbel dapat dinyatakan dengan persamaan sebagai berikut :

$$K = \frac{Y_{T_r} - Y_n}{S_n} \quad (13)$$

denganYn =*reduced mean*yang tergantung jumlah sampel/data n, Sn =*reduced standard deviation*yang juga tergantung pada jumlah sampel/data n, Y_{Tr}=*reduced variate*, yang dapat dihitung dengan persamaan sebagai berikut :

$$Y_{T_r} = -\ln\left(-\ln\frac{T_r - 1}{T_r}\right) \quad (14)$$

Dengan Tr = kala ulang.

PENGUJIAN DISTRIBUSI

Uji Chi-Kuadrat (X^2)

Uji chi kuadrat merupakan pengujian terhadap perbedaan antara data sampel dan distribusi probabilitas. Uji chi kuadrat dapat dihitung dengan persamaan berikut :

$$X^2 = \sum_{i=1}^{N} \frac{(\mathrm{Oi} - \mathrm{Ei})^2}{\mathrm{Ei}} \quad (15)$$

Dengan X^2 = Nilai chi-kuadrat terhitung, Ei = Frekuensi yang diharapkan sesuai pembagian kelasnya, Oi= frekuensi yang terbaca pada kelas yang sama, N=Jumlah sub kelompok dalam satu grup (jumlah kelas). Nilai Ei dapat dicari dengan persamaan berikut :

$$Ei = \frac{n}{N} \quad (16)$$

dengan n = jumlah data, N= jumlah kelas.

Derajat kebebasan dapat dihitung dengan persamaan :

$$dk = K - (\alpha+1) \quad (17)$$

dengan dk=derajad kebebasan, K=banyaknya kelas, α=jumlah parameter. Menurut McCuen (2003), jika nilai rerata dan deviasi standar digunakan dalam perhitungan, maka terdapat dua parameter. Sehingga nilai α untuk uji chi-kuadrat adalah 2. Tetapi jika nilai rerata dan deviasi standar didapatkan dari penelitian atau data sebelumnya maka nilai α untuk uji chi-kuadrat adalah 0.

Menurut Meylan dkk (2011), pada masing-masing kelas, jumlah data minimum adalah 5. Sehingga untuk menentukan jumlah kelas (k) dapat dihitung dengan persamaan berikut ini :

$$k = \frac{n}{5} \quad (18)$$

dengan n = jumlah data.

Pengujian chi-kuadrat selanjutnya membandingkan antara chi-kuadrat yang didapatkan dengan chi kritik. Nilai chi kritik tergantung dari derajad kebebasan (dk) dan tingkat signifikansinya.

Uji Smirnov-Kolmogorov

Kelemahan dari uji chi-kuadrat adalah jumlah sampel yang kecil, karena paling tidak pada masing-masing kelas harus mempunyai frekuensi 5 atau lebih. Uji smirnov-kolmogorov dapat digunakan untuk menguji sampel yang kecil.

Pada uji smirnov-kolmogorov akan dihitung nilai D, yaitu perbedaan maksimum antara fungsi kumulatif sampel dan fungsi probabilitas kumulatif. Nilai D tersebut selanjutnya dibandingkan dengan nilai Dα. Distribusi probabilitas akan diterima jika nilai D lebih kecil dari Dα.

HASIL PEMBAHASAN

Data Hujan

Data hujan untuk analisa menggunakan data hujan di stasiun hujan simongan dengan data hujan selama 15 tahun. Data hujan dapat dilihat pada tabel berikut ini :

Tabel 1 . Data Hujan

Tahun	Curah hujan (mm)
2000	203
2001	147
2002	84
2003	122
2004	163
2005	121
2006	198
2007	162
2008	169
2009	216
2010	110
2011	83
2012	115
2013	111
2014	125

Sumber : BMKG

Analisa Statistik Dasar

Perhitungan statistik dasar meliputi rerata, deviasi standar, koefisien kurtosis dan koefisien kemencengan (*skewness*). Jumlah data (n) adalah sebesar 15. Perhitungan dilakukan pada data yang belum di logaritma maupun data yang telah di logaritma. Hasil perhitungan dapat dilihat pada tabel berikut ini :

Tabel 2. Hasil Perhitungan Statistik Dasar

Data Sebelum di Logaritma	
Rerata	141.933
Deviasi standar	41.932
Koefisien kurtosis	2.700
Koefisien skewness	0.385
Koefisien varian	0.295
Data Setelah di Logaritma	
Rerata	2.134
Deviasi standar	0.130
Koefisien kurtosis	2.744
Koefisien skewness	-0.078
Koefisien varian	0.061

Uji Chi-Kuadrat

Pada pengujian chi-kuadrat data dibuat menjadi beberapa kelas. Pada masing-masing kelas minimal mempunyai frekuensi 5. Oleh karena jumlah data 15, maka sesuai dengan persamaan 17, data dibagi menjadi 3 kelas.

Nilai dk dengan persamaan 16 dapat dicari. Untuk jumlah parameter (α) = 2, yaitu rerata dan deviasi standar dan terdapat 3 kelas maka nilai dk nya adalah 0. Karena pembagian 3 kelas akan mengakibatkan nilai derajad kebebasan 0, maka untuk analisa dibagi menjadi 4 kelas. Sehingga nilai dk adalah 1. Hal inilah yang menjadikan kelemahan uji chi-kuadrat dengan data yang sedikit.

Pembagian kelas berdasarkan rentang probabilitasnya, yaitu 0 < P ≤ 0.25, 0.25 < P ≤ 0.50, 0.50 < P ≤ 0.75, 0.75 < P ≤ 0.999. Uji chi-kuadrat untuk masing-masing distribusi yaitu distribusi normal, log-normal, log pearson III dan gumbel adalah sebagia berikut :

1. Distribusi normal

   Perhitungan untuk distribusi normal dibagi menjadi 4 kelas. Batas probabilitas tiap-tiap kelas adalah 0.25, 0.50, 0.75 dan 0.999. Pada batas tersebut selanjutnya dianalisa nilai yang diharapkan.

Perhitungan nilai yang diharapkan pada distribusi normal menggunakan persamaan 9. Nilai rerata $\bar{X}$=141.933, deviasi standar s=41.932, serta nilai z yang didapatkan dari tabel z pada distribusi normal. Maka nilai yang diharapkan dari distribusi normal dapat dilihat pada Tabel 3.

Tabel 3. Nilai yang Diharapkan pada Distribusi Normal

Batas F(z)	P(z) = 1- F(z)	z	Nilai yang diharapkan
0.250	0.750	0.68	170.448
0.500	0.500	0.00	141.933
0.750	0.250	-0.67	113.838
0.999	0.001	-3.10	11.942

Setelah didapatkan nilai yang diharapkan, maka dilakukan perhitungan frekuensi yang terjadi (Oi) dengan batasan nilai yang diharapkan. Hasil perhitungan frekuensi yang terjadi (Oi) dapat dilihat pada Tabel 4.

Tabel 4. Frekuensi yang terjadi (Oi) pada Distribusi Normal

Kelas	Nilai yang diharapkan	Frekuensi yang terjadi (Oi)
0<P≤0.25	170.448	3
0.25<P≤0.50	141.933	4
0.50<P≤0.75	113.838	4
0.75<P≤0.999	11.9420	4

Nilai frekuensi yang diharapkan (Ei), dihitung dengan persamaan 16 adalah 3.75. Chi-kuadrat dapat dihitung dengan persamaan 15 dan disajikan dalam tabel berikut:

Tabel 5. Nilai Chi-Kuadrat Distribusi Normal

Kelas	Ei	Oi	$(Ei-Oi)^2/Ei$
0<P≤0.25	3.75	3	0.15
0.25<P≤0.50	3.75	4	0.167
0.50<P≤0.75	3.75	4	0.167
0.75<P≤0.999	3.75	4	0.167
Chi-Kuadrat			0.200

Dengan tingkat kepercayaan 95%, tingkat signifikansi 5% serta derajat kebebasan dk =1, didapatkan nilai chi kritik adalah sebesar 3.841 (Tabel Chi-Kritik). Sehingga dapat disimpulkan dengan tingkat kepercayaan 95% dan kesalahan 5% distribusi normal dapat diterima karena nilai chi-kuadrat lebih kecil dari chi kritik (0.200 < 3.841).

2. Distribusi log normal

Tahap perhitungan chi kuadrat untuk distribusi log normal, sama seperti dengan distribusi normal. Perbedaannya adalah data terlebih dahulu dilogaritmakan.

Perhitungan statistika dasar setelah data dilogaritmakan menghasilkan nilai rerata $\bar{X}$=2.134 dan deviasi standar s=0.130.

Hasil perhitungan chi kuadrat untuk distribusi log normal disajikan dalam tabel 6, tabel 7 dan tabel 8 berikut ini.

Tabel 6. Nilai yang Diharapkan pada Distribusi Log Normal

Batas F(z)	P(z)=1-F(z)	z	Log X	Nilai yang diharapkan
0.250	0.750	0.68	2.222	166.95
0.500	0.500	0.00	2.134	136.19
0.750	0.250	-0.67	2.047	111.44
0.999	0.001	-3.10	1.731	53.83

Tabel 7. Frekuensi yang terjadi (Oi) pada Distribusi Log Normal

Kelas	Nilai yang diharapkan	Frekuensi yang terjadi (Oi)
0<P≤0.25	166.95	4
0.25<P≤0.50	136.19	3
0.50<P≤0.75	111.44	4
0.75<P≤0.999	53.83	4

Tabel 8. Nilai Chi-Kuadrat Distribusi Log Normal

Kelas	Ei	Oi	$(Ei-Oi)^2/Ei$
0<P≤0.25	3.75	4	0.167
0.25<P≤0.50	3.75	3	0.15
0.50<P≤0.75	3.75	4	0.167
0.75<P≤0.999	3.75	4	0.167
Chi-Kuadrat			0.200

Hasil perhitungan chi kuadratuntukdistribusi log normaldapat disimpulkan dengan tingkat kepercayaan 95% dan kesalahan 5% distribusi log normal dapat diterima karena nilai chi-kuadrat lebih kecil dari chi kritik (0.200 < 3.841).

3. Distribusi log pearson III

Tahap perhitungan chi kuadrat untuk distribusi log pearson III, sama seperti dengan distribusi log normal. Perbedaannya adalah faktor frekuensi (K).

Perhitungan nilai yang diharapkan pada distribusi normal menggunakan persamaan 11. Nilai rerata $\bar{X}$=2.134, deviasi standar s=0.130, koefisien kemencengan (*skewness*) = -0.078 ≈ -0.1, serta nilai K_T yang didapatkan dari tabel nilai K_T untuk distribusi Log Pearson III. Maka nilai yang diharapkan dari distribusi normal dapat dilihat pada Tabel 9.

Hasil perhitungan chi kuadrat untuk distribusi log pearson III disajikan dalam tabel 9, tabel 10 dan tabel 11 berikut ini.

Tabel 9. Nilai yang Diharapkan pada Distribusi Log Pearson III

Batas F(z)	P(z)=1-F(z)	K	Log X	Nilai yang diharapkan
0.250	0.750	0.69	2.224	167.51
0.500	0.500	0.00	2.134	136.19
0.750	0.250	-0.67	2.046	111.30
0.999	0.001	-3.23	1.714	51.727

Tabel 10. Frekuensi yang terjadi (Oi) pada Distribusi Log Pearson III

Kelas	Nilai yang diharapkan	Frekuensi yang terjadi (Oi)
0<P≤0.25	167.51	4
0.25<P≤0.50	136.19	3
0.50<P≤0.75	111.30	4
0.75<P≤0.999	51.727	4

Tabel 11. Nilai Chi-Kuadrat Distribusi Log Pearson III

Kelas	Ei	Oi	$(Ei-Oi)^2/Ei$
0<P≤0.25	3.75	4	0.167
0.25<P≤0.50	3.75	3	0.15
0.50<P≤0.75	3.75	4	0.167
0.75<P≤0.999	3.75	4	0.167
Chi-Kuadrat			0.200

Hasil perhitungan chi kuadratuntukdistribusi log pearson III dapat disimpulkan dengan tingkat kepercayaan 95% dan kesalahan 5% distribusi log pearson III dapat diterima karena nilai chi-kuadrat lebih kecil dari chi kritik (0.200 < 3.841).

4. Distribusi Gumbel

Pada distribusi gumbel, pembagian kelas sama dengan distribusi normal, log normal maupun log pearson III. Batas probabilitas tiap-tiap kelas adalah 0.25, 0.50, 0.75 dan 0.999. Dengan menggunakan persamaan 8, didapatkan kala ulang 4 tahun, 2 tahun, 1.333 tahun dan 1.001 tahun.

Pada batas tersebut selanjutnya dianalisa nilai yang diharapkan. Untuk menentukan nilai yang diharapkan pada distribusi gumbel, maka terlebih dahulu mencari nilai Y_{TR} dengan persamaan 14 dan K dengan persamaan 13. Untuk nilai Yn dan Sn tergantung dari jumlah data dan didapatkan dari tabel nilai Yn dan Sn untuk distribusi gumbel. Nilai yang diharapkan dicari dengan persamaan 12.

Untuk data sebanyak 15, maka nilai Yn=0.5128 dan Sn=1.0206. Sedangkan dari perhitungan statistik nilai rerata $\bar{X}$=141.933, deviasi standar s=41.932. Perhitungan chi kuadrat untuk distribusi gumbel dapat dilihat pada tabel 12, tabel 13 dan tabel 14.

Tabel 12. Nilai yang Diharapkan pada Distribusi Gumbel

Batas F(z)	T_R	Y_{TR}	K	Nilai yang diharapkan
0.250	4	1.24	0.718	172.054
0.500	2	0.37	-0.143	135.923
0.750	1.333	-0.33	-0.822	107.444
0.999	1.001	-1.93	-2.396	41.459

Tabel 13. Frekuensi yang terjadi (Oi) pada Distribusi Gumbel

Kelas	Nilai yang diharapkan	Frekuensi yang terjadi (Oi)
0<P≤0.25	172.054	3
0.25<P≤0.50	135.923	4
0.50<P≤0.75	107.444	6
0.75<P≤0.999	41.459	2

Tabel 14. Nilai Chi-Kuadrat Distribusi Gumbel

Kelas	Ei	Oi	$(Ei-Oi)^2/Ei$
0<P≤0.25	3.75	3	0.15
0.25<P≤0.50	3.75	4	0.0167
0.50<P≤0.75	3.75	6	1.35
0.75<P≤0.999	3.75	2	0.817
Chi-Kuadrat			2.333

Hasil perhitungan chi kuadratuntukdistribusigumbel dapat disimpulkan dengan tingkat kepercayaan 95% dan kesalahan 5% distribusi gumbel dapat diterima karena nilai chi-kuadrat lebih kecil dari chi kritik (2.333< 3.841).

Uji Smirnov-Kolmogorov

Pada pengujian smirnov-kolmogorov, data terlebih dahulu dan diurutkan, lalu dihitung probabilitasnya dengan persamaan 7. Tiap data selanjutnya dihitung probabilitas secara teoritis pada masing-masing distribusi. Selisih maksimum probabilitas yang terjadi merupakan nilai D yang dicari. Nilai D ini dibandingkan dengan nilai Dα yang didapatkan dari tabel nilai kritis untuk pengujian smirnov-kolmogorov. Distribusi probabilitas akan diterima jika nilai D lebih kecil dari Dα.

1. Distribusi Normal

   Uji smirnov kolmogorov dapat dilihat pada tabel 15.

Tabel 15. Penentuan Nilai D pada Distribusi Normal

Curah Hujan (mm)	P=m/N+1	z	P(z)=1-F(z)	F(z)=1-P(z)	F(z)-P
216	0.0625	1.766	0.9608	0.0392	0.0233
203	0.1250	1.456	0.9265	0.0735	0.0515
198	0.1875	1.337	0.9082	0.0918	0.0957
169	0.2500	0.645	0.7389	0.2611	0.0111
163	0.3125	0.502	0.6915	0.3085	0.0040
162	0.3750	0.478	0.6772	0.3228	0.0522
147	0.4375	0.120	0.5478	0.4522	0.0147
125	0.5000	-0.403	0.3446	0.6554	**0.1554**
122	0.5625	-0.475	0.3192	0.6808	0.1183
121	0.6250	-0.499	0.3121	0.6879	0.0629
115	0.6875	-0.642	0.2611	0.7389	0.0514
111	0.7500	-0.737	0.2327	0.7673	0.0173
110	0.8125	-0.761	0.2236	0.7764	0.0361
84	0.8750	-1.381	0.1922	0.8078	0.0672
83	0.9375	-1.405	0.08	0.92	0.0175

Perhitungan dimulai dengan mengurutkan data dari terbesar ke terkecil dan menghitung probabilitas masing-masing data dengan persamaan 7. Selanjutnya dihitung nilai z dari persamaan 9. Dari nilai z akan diketahui nilai P(z) yang selanjutnya dapat digunakan untuk menghitung F(z). Selisih maksimum antara F(z) dan P adalah nilai D yang dicari. Nilai Dα dengan jumlah data 15 dan tingkat signifikansi 5% didapatkan nilai 0.338. Hasil perhitungan menunjukkan nilai D adalah sebesar 0.1554 dan lebih kecil dari Dα, sehingga distribusi normal bisa diterima.

2. Distribusi Log Normal

   Prosedur perhitungan untuk distribusi log normal sama dengan distribusi normal. Perbedaannya adalah data pada distribusi log normal dilogaritmakan terlebih dahulu.

Tabel 16. Penentuan Nilai D pada Distribusi Log Normal

Log (Xi)	P=m/N+1	z	P(z)=1-F(z)	F(z)=1-P(z)	F(z)-P
2.3345	0.0625	1.540	0.9382	0.0618	0.0007
2.3075	0.1250	1.332	0.9082	0.0918	0.0332
2.2967	0.1875	1.249	0.8944	0.1056	0.0819
2.2279	0.2500	0.720	0.7642	0.2358	0.0142
2.2122	0.3125	0.600	0.7257	0.2743	0.0382
2.2095	0.3750	0.579	0.719	0.281	0.094
2.1673	0.4375	0.255	0.5981	0.4019	0.0356
2.0969	0.5000	-0.286	0.3897	0.6103	**0.1103**
2.0864	0.5625	-0.367	0.3557	0.6443	0.0818
2.0828	0.6250	-0.395	0.3483	0.6517	0.0267
2.0607	0.6875	-0.564	0.2877	0.7123	0.0248
2.0453	0.7500	-0.683	0.2483	0.7517	0.0017
2.0414	0.8125	-0.713	0.2389	0.7611	0.0514
1.9243	0.8750	-1.613	0.0537	0.9463	0.0713
1.9191	0.9375	-1.653	0.0495	0.9505	0.013

Nilai D didapatkan sebesar 0.1103 dan lebih kecil dari Dα=0.338. Jadi dengan tingkat kepercayaan 95% distribusi log normal bisa diterima.

3. Distribusi Log Pearson III

Prosedur perhitungan untuk distribusi log pearson III mirip dengan distribusi log normal. Perbedaannya terletak pada tabel standar untuk mencari nilai K yang tergantung terhadap nilai koefisien kemencengan.

Nilai koefisien kemencengan (*skewness*) = -0.078 ≈ -0.1, maka probabilitas teoritis dapat dicari.

Tabel 17. Penentuan Nilai D pada Distribusi Log Pearson III

Log (Xi)	P=m/N+1	K	F(z)=1-P(z)	F(z)-P
2.3345	0.0625	1.540	0.0609	0.0015
2.3075	0.1250	1.332	0.0909	0.0340
2.2967	0.1875	1.249	0.1048	0.0826
2.2279	0.2500	0.720	0.2404	0.0095
2.2122	0.3125	0.600	0.2794	0.0330
2.2095	0.3750	0.579	0.2860	0.0889
2.1673	0.4375	0.255	0.4054	0.0320
2.0969	0.5000	-0.286	0.6177	**0.1177**
2.0864	0.5625	-0.367	0.6473	0.0848
2.0828	0.6250	-0.395	0.6573	0.0325
2.0607	0.6875	-0.564	0.7162	0.0287
2.0453	0.7500	-0.683	0.7527	0.0027
2.0414	0.8125	-0.713	0.7620	0.0504
1.9243	0.8750	-1.613	0.9422	0.0672
1.9191	0.9375	-1.653	0.9475	0.0100

Nilai D didapatkan sebesar 0.1177 dan lebih kecil dari Dα=0.338. Jadi dengan tingkat kepercayaan 95% distribusi log normal bisa diterima.

4. Distribusi Gumbel

Pada distribusi gumbel untuk uji smirnov-kolmogorov, nilai K yang dicari sama dengan nilai z. Selanjutnya nilai K tersebut dengan persamaan 13 dan 14 untuk mencari nilai Y_{TR} dan T_R. Hasil perhitungan dapat dilihat pada tabel 18.

Tabel 18. Penentuan Nilai D pada Distribusi Gumbel

Curah Hujan (mm)	P=m/N+1	Y_T	T_R	F(z)=1/T_R	F(z)-P
216	0.0625	2.315	10.638	10.638	0.031
203	0.1250	1.999	7.893	7.893	0.001
198	0.1875	1.877	7.049	7.049	0.045
169	0.2500	1.171	3.752	3.752	0.016
163	0.3125	1.025	3.318	3.318	0.011
162	0.3750	1.001	3.252	3.252	0.067
147	0.4375	0.636	2.433	2.433	0.026
125	0.5000	0.100	1.680	1.680	0.095
122	0.5625	0.027	1.607	1.607	0.059
121	0.6250	0.003	1.585	1.585	0.009
115	0.6875	-0.142	1.461	1.461	0.003
111	0.7500	-0.240	1.389	1.389	0.030
110	0.8125	-0.264	1.373	1.373	0.084
84	0.8750	-0.897	1.094	1.094	0.038
83	0.9375	-0.921	1.088	1.088	0.018

Nilai D didapatkan sebesar 0.095 dan lebih kecil dari Dα=0.338. Jadi dengan tingkat kepercayaan 95% distribusi log normal bisa diterima.

Pemilihan Distribusi Probabilitas

Untuk pemilihan distribusi probabilitas yang bisa digunakan adalah dengan membandingkan nilai chi kuadrat dan nilai D yang didapatkan untuk masing-masing distribusi.

Tabel 19. Pembandingan Nilai Chi Kuadrat dan D pada masing-masing Distribusi Probabilitas

Jenis Distribusi Probabilitas	Chi kuadrat (X^2)	D
Normal	0.200	0.1554
Log Normal	0.200	0.1103
Log Pearson III	0.200	0.1177
Gumbel	2.333	0.095

Berdasarkan hasil pengujian chi kuadrat distribusi terbaik adalah normal, log normal dan log pearson III sedangkan dari uji smirnov kolmogorov hasil terbaik adalah gumbel karena mempunyai nilai terkecil.

Jika membandingkan kedua pengujian tersebut, maka distribusi log normal adalah yang terbaik. Selain memberikan nilai chi kuadrat terkecil juga nilai D yang masih mendekati dengan nilai D pada distribusi gumbel.

KESIMPULAN DAN SARAN

Kesimpulan

Berdasarkan hasil pembahasan, maka dapat disimpulkan sebagai berikut :

1. Pengujian chi kuadrat memiliki kelemahan jika jumlah data sedikit karena untuk masing-masing kelas dibutuhkan frekuensi minimal 5.
2. Pengujian smirnov kolmogorov mempunyai keunggulan apabila datanya sedikit karena tiap data yang ada diuji terhadap nilai standar pada masing-masing distribusi probabilitas.
3. Hasil pengujian menunjukkan nilai chi kuadrat terkecil adalah pada distribusi normal, log normal dan log pearson III yaitu 0.200. Sedangkan pada pengujian smirnov kolmogorov nilai D terkecil adalah terdapat pada distribusi gumbel yaitu 0.095. Tetapi seluruh distribusi masih bisa diterima dengan tingkat kepercayaan 95%.

Saran

Untuk penentuan distribusi probabilitas pada analisa hujan perlu dilakukan dua pengujian, karena masing-masing pengujian mempunyai kelemahan. Hasil terbaik adalah dengan membandingkan hasil pada masing-masing pengujian.

DAFTAR PUSTAKA

Buku

Das M.M., Saikia M.D., 2009. *Hydrology*. New Delhi: PHI Learning Private Limited.

Haan, 1977. *Statistical Methods in Hydrology*. Iowa: The Iowa State Universities Press.

Mc Cuen, 2003. *Modelling Hydrologic Change Statistical Methods*. Florida : CRC Press.

Triatmodjo, 2008. Hidrologi Terapan. Yogyakarta: Beta Offset Yogyakarta.

Soemarto, 1999. Hidrologi Teknik. Jakarta: Erlangga.

Sri Harto, 2000. Hidrologi. Yogyakarta: Nafiri Offset.

Subramanya, 2013. *Engineering Hydrology*. New Delhi: McGrawHill.

Suripin, 2004. Sistem Drainase Perkotaan yang Berkelanjutan. Yogyakarta: Andi Offset.

Jurnal

Kang H.M. and Yusof F., 2013.*Determination of Best-fit Distribution and Rainfall Events in Damansara and Kelantan, Malaysia.*Matematika. Hal: 43-52.

Khudri M.M. and Sadia F., 2013. *Determination of the Best Fit Probability Distribution for Annual Extreme Precipitation in Bangladesh*.*European Journal of Scientific Research*.103 (3), Hal: 391-404.