

PEMROGRAMAN R UNTUK DISTRIBUSI GAMMA-NORMAL

SKRIPSI

**Sebagai Salah Satu Syarat Memperoleh Gelar
Sarjana Matematika**



Oleh

**DELIA PARAMITHA
NIM 08011181419066**

**FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM
UNIVERSITAS SRIWIJAYA
APRIL 2018**

LEMBAR PENGESAHAN

PEMROGRAMAN R UNTUK DISTRIBUSI GAMMA-NORMAL

SKRIPSI

**Sebagai Salah Satu Syarat Memperoleh Gelar
Sarjana Matematika**

Oleh

**DELIA PARAMITHA
NIM 08011181419066**

Indralaya, April 2018

Pembimbing Kedua

Pembimbing Utama



**Novi Rustiana Dewi, M.Si
NIP.19701113 199603 2 002**



**Dr. Ir. Herlina Hanum, M.Si
NIP.19650108 199003 2 007**



**Drs.Sugandi Yahdin, M.M
NIP. 19580727 198603 1003**

HALAMAN PERSEMBAHAN

MOTTO

“Once you choose hope, anything’s possible”

“banyak kegagalan dalam hidup ini dikarenakan orang-orang tidak menyadari betapa dekatnya mereka dengan keberhasilan saat mereka menyerah”

*Sesungguhnya sesudah kesulitan itu ada kemudahan, maka apabila kamu telah selesai (dari suatu urusan), kerjakanlah dengan sungguh-sungguh (urusan) yang lain, dan hanya kepada Allah lah hendaknya kamu berharap.
(QS Al-Insyirah ayat 6-8)*

Skripsi ini Kupersembahkan untuk :

- Allah SWT
- Rasulullah SAW
- Kedua Orang Tuaku
- Kakak & adikku tersayang
- Keluarga Besarku
- Sahabat-Sahabatku
- Almamaterku

KATA PENGANTAR

Assalammu'alaikum wr. wb

Puji syukur penulis panjatkan kepada Allah SWT karena atas rahmat dan berkat-Nya yang luar biasa sehingga penyelesaian skripsi yang berjudul **“Pemrograman R untuk Distribusi Gamma-Normal”** dapat berjalan dengan baik dan selesai pada waktunya. Skripsi ini sebagai salah satu syarat untuk memperoleh gelar Sarjana Sains bidang studi Matematika di Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Sriwijaya.

Dengan segala hormat dan kerendahan hati mengucapkan terima kasih yang sedalam-dalamnya kepada Kedua Orang Tua, yaitu Bapak **Tony Kosim** dan Ibu **Dahlia Wati** yang telah menuntun, mendidik, mengajari, menasehati, memberi semangat, dan tidak lelah untuk selalu berdoa yang terbaik untuk anaknya. Penulisan skripsi ini tidak lepas dari bantuan berbagai pihak, untuk itu penulis menyampaikan terima kasih kepada:

1. Bapak **Drs. Sugandi Yahdin, M.Si** : Ketua Jurusan Matematika atas bimbingan yang telah diberikan selama penulis belajar di Jurusan Matematika Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Sriwijaya.
2. Ibu **Des Alwine Zayanti, M.Si** : Sekretaris Jurusan Matematika Fakultas Matematika dan Ilmu Pengetahuan Alam (FMIPA) Universitas Sriwijaya.
3. Bapak **Dr. Ngudiantoro** : Dosen Pembimbing akademik yang telah memberikan bimbingan selama penulis berstatus sebagai mahasiswa di Jurusan Matematika FMIPA Universitas Sriwijaya.

4. Ibu **Dr. Ir. Herlina Hanum, M.Si** : Pembimbing Utama yang telah bersedia meluangkan banyak waktu, pikiran, tenaga, dan motivasi sehingga penulis dapat menyelesaikan skripsi ini dengan baik.
5. Ibu **Novi Rustiana Dewi, M.Si** : Pembimbing kedua yang telah bersedia meluangkan banyak waktu, pikiran, tenaga dan motivasi sehingga penulis dapat menyelesaikan skripsi ini dengan baik.
6. Ibu **Irmeilyana, M.Si**, Ibu **Des Alwine Zayanti, M.Si**, dan Bapak **Alfensi Faruk, M.Sc** : Dosen Penguji yang telah bersedia meluangkan waktu dalam memberikan tanggapan, kritik dan saran yang bermanfaat dalam perbaikan dan penyelesaian skripsi ini.
7. **Seluruh Dosen di Jurusan Matematika Fakultas Matematika dan Ilmu Pengetahuan Alam** yang telah memberikan ilmu yang bermanfaat, nasehat serta bimbingan selama penulis menjalani perkuliahan.
8. Pak **Irwansyah** dan Ibu **Hamidah** : Tenaga kependidikan jurusan Matematika Fakultas dan Ilmu Pengetahuan Alam yang telah membantu penulis selama masa perkuliahan.
9. Bapak **Tony Kosim** dan Ibu **Dahlia Wati** serta Adik-Adikku **M Burlian Ardhi, Dimas Wicaksono** dan seluruh keluarga besar yang telah memberikan semangat, cinta dan motivasi agar penulis dapat menyelesaikan skripsi secepat mungkin serta selalu optimis menggapai cita-cita..
10. Sahabat tercinta, **Siti Julaeha, Annisa Larashati, Nur Atika, Desti Destiansari, Istinabiyah, Rizki Utami, Dhea Adistia, Yulia Saputri, Gusti Rahmawati, Reni Meilina Sari, Idham Fahrul, Junaidi Mangku Pratama dan Fachri**

Kusriadi yang telah sabar mendengar keluhan, membantu suka dan cita, dan memberikan semangat serta motivasi untuk menyelesaikan skripsi ini. Semoga kita semua sukses ke depannya.

11. Teman-teman seperjuangan, **Putri Asia, Liani, Lady, Elka, Haliza, Vinny, Elsia, Yiyi, Oci** dan semua angkatan **2014** yang tidak bisa disebutkan semuanya, penulis ucapkan banyak terima kasih atas bantuan dan semangat yang ditujukan kepada penulis.
12. Kakak-kakak tingkat angkatan **2011, 2012, dan 2013**, serta adik-adik tingkat istimewa **Novika, Elen, Elsa, Vidya, Feren, Annisa, Indah Meiliana** dan semua angkatan **2015, 2016, dan 2017**.
13. Semua pihak yang telah membantu penulis dalam menyelesaikan skripsi yang tidak dapat disebutkan satu persatu.

Terimakasih atas semua dukungan, semoga Allah Subhanahuwata'ala membalas semua kebaikan yang diberikan kepada penulis dengan rahmat dan karunia-Nya. Penulis mengharapkan kritik dan saran untuk meningkatkan kualitas dari skripsi ini dan semoga dapat bermanfaat bagi semua yang membacanya.

Wassalammu'alaikum wr. wb

Indralaya, Maret 2018

Penulis

R PROGRAMMING FOR GAMMA-NORMAL DISTRIBUTION

By:

DELIA PARAMITHA
NIM 08011181419066

ABSTRACT

The Gamma-Normal Distribution (G-N) is a new distribution which is in the form of two continuous distributions of gamma and normal with the formation of transformer-transformer distributions (T-X). The G-N distribution is consist of G-N distributions of two parameters (gamma-standard normal) with $\mu = 0$ and $\sigma = 1$ and four parameters (gamma-normal). Software R is software that is open source and suitable for the distribution of G-N. This study discussed the development of R programmers for special functions of G-N distribution. The program created is the program 1) parameter estimation that begins by determining the initial value, 2) pdf, 3) cumulative function, 4) the quantile function and 5) the generation of randomly distributed G-N variables. Testing the program on Breaking stress of carbon fibers data gives almost similar results for the analysis of SAS software on the data. For the data obtained the value of AIC, MAPE and p-value KS is 175.8405, 6.057551%, 0.9915 for two parameters and 178.9064, 4.855483%, 0.9915 for four parameters G-N distribution.

Keywords: T-X distribution, G-N distribution, and R programming.

PEMROGRAMAN R UNTUK DISTRIBUSI GAMMA-NORMAL

Oleh:

DELIA PARAMITHA

NIM 08011181419066

ABSTRAK

Distribusi Gamma-Normal (G-N) merupakan distribusi baru yang di bentuk dari dua distribusi kontinu yaitu gamma dan normal dengan pembentukan distribusi *transformed-transformer* ($T-X$). Distribusi G-N terbagi menjadi distribusi G-N dua parameter (gamma-normal standar) dengan $\mu=0$ dan $\sigma=1$ dan empat parameter (gamma-normal). *Software* R adalah *software* yang bersifat *open source* dan cocok digunakan untuk distribusi G-N. Pada penelitian ini dibahas mengenai pembuatan pemrogram R untuk fungsi-fungsi khusus distribusi G-N. Program yang dibuat adalah program 1) pendugaan parameter yang diawali dengan menentukan nilai awal, 2) fkp, 3) fungsi kumulatif, 4) fungsi kuantil dan 5) pembangkitan peubah acak berdistribusi G-N. Uji coba program pada data *Breaking stress of carbon fibers* memberikan hasil yang hampir sama dengan hasil analisis pada *software* SAS pada data tersebut. Untuk data tersebut diperoleh nilai AIC, MAPE, p-value KS masing-masing adalah 175,8405, 6,057551%, 0,9915 untuk dua parameter dan 178,9064, 4,855483%, 0,9915 untuk empat parameter distribusi G-N.

Kata kunci: distribusi T-X, distribusi G-N, dan pemrograman R.

DAFTAR ISI

Halaman

HALAMAN JUDUL	i
LEMBAR PENGESAHAN	Error! Bookmark not defined.
HALAMAN PERSEMBAHAN	ii
KATA PENGANTAR.....	iv
ABSTRACT	vii
ABSTRAK	viii
DAFTAR ISI.....	ix
DAFTAR TABEL	xiii
DAFTAR GAMBAR.....	xiv
DAFTAR LAMPIRAN	xvi
BAB I PENDAHULUAN.....	1
1.1. Latar Belakang	1
1.2. Perumusan Masalah.....	3
1.3. Pembatasan Masalah	3
1.4. Tujuan.....	4
1.5. Manfaat.....	4
BAB II TINJAUAN PUSTAKA.....	5
2.1. Ruang Lingkup Statistika	5

2.1.1. Statistika Deskriptif	5
2.1.2. Statistika Inferensia.....	5
2.2. Distribusi Peluang Kontinu	5
2.2.1. Distribusi Gamma.....	6
2.2.2. Distribusi Normal	6
2.3. Distribusi G-N	7
2.3.1. Fungsi kuantil distribusi Gamma-Normal	8
2.3.2. Pendugaan Parameter.....	8
2.4. Uji Kesesuaian Distribusi.....	9
2.4.1. Plot Kuantil.....	9
2.4.2. AIC.....	10
2.4.3. MAPE.....	10
2.4.4. Uji KS	11
2.5. Sejarah dan Pengenalan <i>Software R</i>	12
2.5.1. Kelebihan dan Kelemahan <i>Software R</i>	12
2.5.2. R Studio.....	13
2.6. Definisi-Definisi dalam <i>Software R</i>	14
2.7. <i>Package</i> dan <i>Library</i>	16
2.8. Skema Kerja R.....	17

2.9.	Mendefinisikan fungsi dalam R	18
2.10.	Fungsi Pembangkit Data Peubah Acak	19
BAB III METODELOGI PENELITIAN.....		21
3.1.	Tempat.....	21
3.2.	Waktu	21
3.3.	Jenis dan Sumber Data	22
3.4.	Metode Penelitian.....	22
BAB IV HASIL DAN PEMBAHASAN		26
4.1.	Penginstalan dan Pemanggilan <i>Package</i>	26
4.2.	Pemrograman Input Data.....	28
4.3.	Pemrograman Menentukan Nilai Awal	29
4.4.	Pemrograman Fungsi Pendugaan Parameter Distribusi G-N	32
4.5.	Pemrograman FKP Distribusi G-N	36
4.6.	Pemrograman Fungsi Kumulatif Distribusi G-N	38
4.7.	Pemrograman Fungsi Kuantil Distribusi G-N.....	41
4.8.	Pemrograman Pembangkitan Peubah Acak Distribusi G-N.....	43
4.9.	Pemrograman Uji Kesesuaian Data Uji Coba Program pada Distribusi G-N	45
4.9.1.	MAPE	45
4.9.2.	Uji KS	47

BAB V KESIMPULAN DAN SARAN	49
5.1. Kesimpulan.....	49
5.2. Saran.....	51
DAFTAR PUSTAKA.....	52
LAMPIRAN.....	54

DAFTAR TABEL

	Halaman
Tabel 2.1. Distribusi gamma dan normal pada R	20
Tabel 3.1. Jadwal kegiatan penelitian dan penulisan skripsi	21
Tabel 4.1. Pendugaan parameter dengan <i>software R</i>	35
Tabel 4.2 Tabel hasil pemrograman <i>software R</i> dan <i>software SAS</i>	48

DAFTAR GAMBAR

	Halaman
Gambar 4.1 <i>Package</i> yang telah di instal dan diaktifkan.....	26
Gambar 4.2. <i>Package</i> yang telah dipanggil pada <i>R script</i>	26
Gambar 4.3 Pemrograman untuk menginput data yang tersimpan pada Ms.Excel	28
Gambar 4.4. Output pemrograman input data.....	28
Gambar 4.5. Pemrograman dalam mencari nilai awal	29
Gambar 4.6. Output nilai awal parameter μ dan σ	31
Gambar 4.7. Output nilai awal parameter α dan β	31
Gambar 4.8. Pemrograman fungsi pendugaan parameter distribusi G-N	33
Gambar 4.9. Output dari fungsi pendugaan dua parameter	34
Gambar 4.10. Output dari fungsi pendugaan parameter	35
Gambar 4.11. Pemrograman fkp distribusi G-N	36
Gambar 4.12. Pemrograman membuat histogram fkp distribusi G-N	37
Gambar 4.13. Histogram fkp distribusi G-N.....	38
Gambar 4.14. Pemrograman fungsi kumulatif Distribusi G-N.....	39
Gambar 4.15. Pemrograman plot fungsi kumulatif distribusi G-N.....	40
Gambar 4.16. Plot fungsi kumulatif distribusi G-N.....	40
Gambar 4.17. Pemrograman fungsi kuantil Distribusi G-N	41

Gambar 4.18. Pemrograman plot fungsi kuantil G-N.....	42
Gambar 4.19. Plot fungsi kuantil G-N	43
Gambar 4.20. Pemrograman pembangkitan bilangan acak Distribusi G-N.....	44
Gambar 4.21. Pemrograman nilai MAPE	45
Gambar 4.22. Pemrograman nilai MAPE	46
Gambar 4.23. Output nilai MAPE.....	46
Gambar 4.24. Pemrograman nilai KS	47
Gambar 4.25 Output pemrograman nilai KS	47

DAFTAR LAMPIRAN

	Halaman
Lampiran 1. Data <i>breaking stress of carbon fibers</i>	55
Lampiran 2. Estimasi parameter (standar eror parameter) pada data <i>breaking stress of carbon fibers</i> dengan <i>software SAS</i>	56
Lampiran 3. Output nilai fkp distribusi G-N.....	57
Lampiran 4. Output nilai fungsi kumulatif distribusi G-N	58
Lampiran 5. Output nilai fungsi kuantil distribusi G-N.....	59
Lampiran 6. Pemrograman fungsi-fungsi khusus distribusi G-N	60

BAB I

PENDAHULUAN

1.1. Latar Belakang.

Distribusi peluang baru sangat berpengaruh terhadap statistika inferensia karena jenis prosedur pengolahan data yang diterapkan. Sebuah metode dilakukan untuk menghasilkan distribusi baru yaitu dengan pembentukan distribusi *transformed-transformer* ($T-X$). Distribusi $T-X$ pertama kali dipopulerkan oleh Alzaatreh *et al.*(2013), yang menjelaskan bahwa variabel acak X sebagai “transformator” digunakan untuk mengubah variabel acak T yang “ditransformasikan”.

Pembentukan distribusi baru dapat dilakukan dengan beberapa distribusi peluang kontinu. Alzaatreh *et al.* (2014) menggunakan distribusi peluang kontinu gamma dan normal dalam pembentukan distribusi $T-X$ yang dimana distribusi gamma sebagai variabel acak T dan distribusi normal sebagai variabel acak X , sehingga dinamakan distribusi Gamma-Normal (G-N).

Distribusi G-N merupakan distribusi baru sehingga diperlukan software yang cocok dalam mengolah datanya. Perkembangan teknologi informasi berdampak positif pada berkembangnya berbagai perangkat lunak pengolah data. Secara umum ada dua macam kategori *software* statistik untuk keperluan analisis data, yaitu *software* statistik yang bersifat *komersil* dan *software* statistik yang bersifat *open source* atau *freeware*. Hardhono dan Sitanggung (2014) menjelaskan *Software R* adalah suatu *software* yang masuk dalam kategori *open source* yang artinya bahwa *source program*-nya dapat didistribusikan secara terbuka dan gratis serta produk perangkat lunak yang dihasilkan boleh dipakai secara terbuka oleh siapa saja serta

software R bersifat *multiplatform*, yang artinya dapat di *install* dan digunakan baik pada system operasi Windows, Unix/Linux maupun pada Macintosh. *Software* R dapat melakukan analisis dengan menuliskan fungsi yang diinginkan pada *package* yang telah tersedia atau membuat sendiri fungsi dengan menggunakan rumus dari data yang akan dianalisis, sehingga *software* R sangat cocok digunakan untuk menganalisis distribusi baru yang fungsi-fungsinya tidak tersedia dalam *software* melainkan harus dibuat sendiri.

Penelitian sebelumnya, aplikasi dari distribusi G-N telah dilakukan oleh Alzaatreh *et al.* (2014) dengan menggunakan *software* SAS dan *PROC NLMIXED* (*Procedure Non Linier Mixed*) sebagai model dalam SAS yang digunakan untuk mencari penduga parameter G-N dua parameter (gamma-normal standar) dengan parameter $\mu=0$ dan $\sigma=1$ dan G-N empat parameter (gamma-normal) yang semua parameternya tidak ditetapkan sebelumnya. Duningan (2011), menjelaskan *PROC NLMIXED* merupakan prosedur pada *software* SAS yang digunakan untuk menganalisis model regresi nonlinier yang mengandung lebih dari satu parameter yaitu parameter tetap dan acak.

Penelitian ini, akan dibuat pemrograman untuk mencari nilai kemungkinan maksimum dengan menggunakan metode *Maximum Likelihood Estimation* (MLE) yang diperkirakan menghasilkan nilai yang hampir sama dan membuat fungsi (*function*) peluang distribusi G-N dengan menggunakan *software* Rstudio dikarenakan *software* R dinilai *open source* serta memiliki kapasitas memori yang lebih kecil dibandingkan dengan *software* SAS yang bersifat *komersil* serta memiliki kapasitas memori yang lebih besar. *Script* dengan program R ini juga memanfaatkan

beberapa *package* yang harus diinstall terlebih dahulu diantaranya *package* *bbmle*, *stats4*, *MASS*, *zipfR*, dan *numDeriv*. Data yang digunakan sebagai uji coba pemrograman yaitu data *breaking stress of carbon fibers* yang merupakan data tekanan pecahan serat karbon (Alzaatreh *et al.* 2014). Versi perangkat lunak yang digunakan adalah R Studio dengan versi 3.4.2. Uji kesesuaian distribusi yang digunakan adalah dengan plot kuantil, uji KS (*Kolmogorov-Smirnov*), MAPE (*Mean Absolute Percentage Error*) dan AIC (*Akaike Information Criterion*).

1.2. Perumusan Masalah

Permasalahan dalam penelitian ini adalah :

1. Bagaimana membuat pemrograman *function* suatu fungsi khusus distribusi G-N dan memperoleh nilainya pada *software* RStudio dari :
 - a. Fungsi kemungkinan maksimum (*likelihood*)
 - b. Fungsi kepekatan peluang (*fkp*)
 - c. Fungsi kumulatif
 - d. Fungsi kuantil
 - e. Pembangkitan data acak
2. Bagaimana uji kesesuaian data uji coba pemrograman pada distribusi G-N dengan menggunakan *software* R?

1.3. Pembatasan Masalah

Penelitian ini dibatasi pada pemrograman fungsi-fungsi untuk menghitung peluang distribusi G-N pada data dengan pemodelan satu peubah pada *software* R

dengan mencocokkan hasil dari salah satu data penelitian Alzaatreh *et al.* (2014) sebagai acuan dari penelitian ini.

1.4. Tujuan

Tujuan penelitian ini adalah untuk :

1. Menghasilkan pemrograman *function* suatu fungsi khusus distribusi G-N dan memperoleh nilainya pada *software* RStudio dari :
 - a. Fungsi kemungkinan maksimum (*likelihood*)
 - b. Fungsi kepekatan peluang (fkp)
 - c. Fungsi distribusi kumulatif
 - d. Fungsi kuantil
 - e. Pembangkitan data acakuntuk distribusi G-N pada program R.
2. Mengetahui uji kesesuaian data uji coba pemrograman pada distribusi G-N dengan menggunakan *software* R.

1.5. Manfaat

Pada hasil penelitian ini diharapkan dapat dijadikan sebagai bahan diskusi ilmiah untuk mengembangkan pengaplikasian *software* R. Untuk Statistikawan penelitian ini sangat berguna karena memudahkan menganalisis data jika dalam suatu data ditemukan data yang berdistribusi G-N.

DAFTAR PUSTAKA

- Alzaatreh, A., Famoye, F., and Lee, C. 2013. A New Method for Generating Families of Continuous Distributions. *Metron* 71: 63-79.
- Alzaatreh, A., Famoye, F., and Lee, C. 2014. The Gamma-Normal Distribution: Properties and Application. *Computational Statistics and Data Analysis* 69 : 67-80.
- Asdi, Yudiantri. 2017. *Pengenalan Software R*. matematika. fmipa.unand.ac.id/images/bahan-seminar/pakyudi.pdf (diakses pada 29 Oktober 2017).
- Budiharto, W., dan Rachmawati, R. N. 2013. *Pengantar Praktis Pemograman R untuk Ilmu Komputer*. Jakarta Barat : Halaman Moeka Publishing.
- Burnham, K. P., and Anderson, D. R. 2002. *Model Selection and Multimodel Inference*. New York: Springer-Verlag.
- Duningan, Keith. 2011. PROC NLMIXED for Basic Non-Linear Regression. Paper SA-10-2011.
- Hardhono, A. P., dan Sitanggang, I. S. 2014. *Pengenalan dan Instalasi Perangkat Lunak dan Lingkungan Pemrograman R*. <http://www.pustaka.ut.ac.id/lib/wp-content/uploads/pdfmk/SATS4111-M1.pdf> (diakses pada 20 November 2017).
- Hassani, H., and Silva, E. S. 2015. A Kolmogorov-Smirnov Based Test for Comparing the Predictive Accuracy of Two Sets of Forecasts. *Econometrics*, 3, 590-609.
- Herrhyanto, N., dan Gantini, T. 2000. *Pengantar Statistika Matematis*. Bandung: Yrama Widya.
- Hosmer, D. W., and Lemeshow, S. 2000. *Applied Logistic Regression*. New York: Jhon Wiley and Sons.
- Mallor, Nualart dan Omey. 2009. *An Introduction to Statistical Modelling of Extreme Value Application to Calculate Extreme Wind Speeds*. Hogeschool University Briscel.

- Nurlaila, D., Kusnandar, D., dan Sulistianingsih, E . (2013). Perbandingan Metode *Maximum Likelihood Estimation* (MLE) dan Metode Bayes dalam Pendugaan Parameter Distribusi Eksponensial. *Bimaster*, volume 3, 51-5.
- Sauddin, A. 2014. Pengenalan R Programming. *Jurnal MSA*, 2(2), 51-59.
- Suhartono. 2009. *Analisis Data Statistik Dengan R*. Surabaya : Graha Ilmu.
- Tirta, I. M. 2014. *Analisis Data dengan R*. Jember: UNEJ.
- Tirta, I. M. 2015. *Buku Panduan Program Staistika R*. Jember: UNEJ.
- Wibisono, Yusuf. 2005. *Metode Statistik*. Yogyakarta : Gadjah Mada University Press.
- Yudistira, I. G. A. A. 2013. Desain Animasi Statistika Berbasis Bahasa Pemograman R. *E-Journal Widya Eksata*, Vol.1(1).
- Yudistira, I. G. A. A. 2005. Pengenalan Bahasa R Untuk Analisis Data dan Graphik.