

Normalitas Data Menggunakan Uji Kolmogorov-Smirnov dan Saphiro-Wilk

(Studi kasus penghasilan orang tua mahasiswa Prodi Pendidikan Matematika Unismuh Makassar)

Andi Quraisy

Universitas Muhammadiyah Makassar

Kampus FKIP UNISMUH Makassar Jln. Sultan Alauddin No. 259, Makassar, 90221

Email : andiquraisy@unismuh.ac.id



©2020 J-HESTFDI DPD Sulawesi Barat

Ini adalah artikel dengan akses terbuka dibawah lisensi CC BY-NC-4.0

(<https://creativecommons.org/licenses/by-nc/4.0/>).

ABSTRACT

This research is a nonparametric study using the Kolmogorov-Smirnov and Saphiro-Wilk methods or research to test data normality. This study aims to determine a description of the income of the parents of students in Mathematics Education Study Program UNISMUH Makassar. This research data is primary data obtained directly from students. The research was analyzed descriptively and used nonparametric analysis methods. The results of the normality test using the Kolmogorov-Smirnov and Saphiro-Wilk tests, it was found that the p-value was smaller than α so that the income data of the student's parents was not normal.

Keywords: Data Normality, Kolmogorov-smirnov method, Saphiro-Wilk method

ABSTRAK

Penelitian ini merupakan salah satu penelitian nonparametrik dengan menggunakan metode Kolmogorov-smirnov dan saphiro-wilk atau penelitian untuk menguji kenormalan data .penelitian ini bertujuan untuk mengetahui gambaran tentang penghasilan orangtua mahasiswa Prodi Pendidikan Matematika UNISMUH Makassar. Data penelitian ini merupakan data primer yang diperoleh langsung dari mahasiswa.Penelitian dianalisis secara deskriptif dan menggunakan metode analisis nonparametrik. Hasil uji normalitas dengan menggunakan uji Kolmogorov-smirnov dan Saphiro-Wilk diperoleh nilai p-value lebih kecil dari α sehingga data penghasilan orangtua mahasiswa tersebut tidak normal.

Kata Kunci: Normalitas data, Metode Kolmogorov-Smirnov, Metode Saphiro-Wilk

PENDAHULUAN

Seperti yang diketahui bersama, statistik dan statistika merupakan dua kata yang berbeda dan mempunyai arti yang berbeda pula. Statistik sendiri adalah data yang berupa angka yang dikumpulkan, digolongkan sehingga dapat

memberi informasi yang berarti sedangkan statistika merupakan ilmu yang mempelajari cara mengumpulkan, menganalisis, menginterpretasi, dan menyimpulkan data yang telah diperoleh. Statistika sendiri dapat dibedakan menjadi statistika parametrik dan statistika non parameterik.Menurut Tiro, M.

Arif (2015) menyatakan bahwa dalam melakukan penaksiran parameter atau pengujian hipotesis, sering diperlukan asumsi tentang sebaran populasi, seperti asumsi kenormalan populasi atau asumsi sebaran khusus yang lainnya. Dalam kasus ini, statistika yang digunakan disebut statistika parametrik, sedangkan statistika yang longgar asumsinya tentang sebaran populasi disebut statistika non parametrik.

Salah satu syarat utama dalam analisis statistika parametrik adalah terpenuhinya kenormalan data. Menurut Lukiastruti & Hamdani (2012) menyatakan bahwa persyaratan dan asumsi statistika parametrik antara lain (a). sampel data yang diperoleh berasal dari populasi yang memiliki pola distribusi data yang normal, (b). data yang dijadikan sampel diperoleh dengan penentuan secara acak, (c). skala pengukuran harus bersifat kontinu, (d). nilai variansi harus sama.

Menurut Purwanto (2011) menyatakan bahwa statistika parametrik adalah statistika yang digunakan apabila berbagai asumsi yang dituntut terpenuhi. Pengolahan data meliputi dua kegiatan yaitu pengujian asumsi dan pengujian hipotesis. Pengujian asumsi dilakukan sebelum dilakukan pengujian hipotesis. Jika dalam pengujian asumsi menunjukkan bahwa asumsi yang dituntut dapat dipenuhi oleh data maka pengujian hipotesisnya menggunakan statistika parametrik.

Menurut Siregar, Syofiyani (2015) menyatakan bahwa statistika parametrik adalah statistika yang mempertimbangkan jenis/sebaran data yang berdistribusi normal dan memiliki variansi homogen. Pada umumnya, data yang digunakan ini bersifat interval atau rasio.

Menurut Nurgiyantoro, dkk. (2012) menyatakan bahwa data – data berskala interval sebagai hasil pengukuran pada umumnya mengikuti asumsi distribusi normal. Namun bahwa suatu data ternyata tidak mengikuti asumsi itu bukanlah hal mustahil. Untuk mengetahui kepastian sebaran data yang diperoleh, haruslah dilakukan uji normalitas terhadap data yang bersangkutan.

Pengujian untuk membuktikan normal atau tidaknya suatu data dapat dilakukan dengan menggunakan analisis *kolmogoroff – smirnov* dan *Shapiro – wilk*. Pembahasan analisis uji normalitas dengan menggunakan uji

kolmogoroff – smirnov dan *Shapiro – wilk* ini pada umumnya dapat ditemukan pada literatur – literatur dengan pokok bahasan statistika nonparametrik.

Uji statistik parametrik adalah suatu pengujian yang modelnya menerapkan adanya asumsi tertentu tentang parameter populasi yang merupakan sumber sampel penelitian. Selain itu, tes statistik parametrik menuntut skor – skor yang dianalisis harus merupakan hasil suatu pengukuran yang sedikitnya berkekuatan sebagai skala interval. Uji asumsi normalitas yang lebih kompleks dan lengkap juga disebut dengan uji kesesuaian model dimaksudkan untuk menguji apakah model yang diusulkan memiliki kesesuaian dengan data atau tidak. Suatu model dikatakan *fit* apabila matriks korelasi sampel tidak jauh berbeda dengan matriks korelasi estimasi (Riadi, Edi ; 2016).

Menurut Sugiyono (2015) menyatakan bahwa bila suatu data tidak normal, maka statistik parametrik tidak dapat digunakan, untuk itu perlu digunakan statistik nonparametrik. Tetapi perlu diketahui bahwa hal yang menyebabkan tidak normal, misalnya terdapat kesalahan instrument dan pengumpulan data maka dapat mengakibatkan data yang diperoleh menjadi tidak normal. Menurut Sujarweni, Wiratna (2016) menyatakan bahwa uji normalitas data sebaiknya dilakukan sebelum data dioleh berdasarkan model – model penelitian.

Normalitas data dapat dilihat dengan menggunakan uji normalitas Kolmogorov–smirnov. Uji data berdistribusi normal bisa dilakukan beberapa prosedur (Santoso, Singgih (2015)) yaitu melakukan metode statistik tertentu seperti uji Kolmogorov–Smirnov, uji Shapiro–Wilk dan sebagainya, dan membuat grafik dengan prosedur tertentu dan mengamati pola plot atau grafik tersebut.

Dalam statistik, uji Kolmogorov-Smirnov (uji K-S atau uji KS) adalah tes nonparametrik dari persamaan kontinu, distribusi probabilitas satu dimensi yang dapat digunakan untuk membandingkan sebuah sampel dengan distribusi probabilitas referensi (uji K-S satu sampel), atau untuk membandingkan dua buah sampel (uji K-S dua sampel). Dinamai setelah Andrey Kolmogorov dan Nikolai Smirnov. Tes Kolmogorov-Smirnov dapat berfungsi sebagai uji goodness of fit, dalam kasus khusus pengujian normalitas distribusi, sampel

distandarisasi dan dibandingkan dengan distribusi normal standar. Ini setara dengan menetapkan mean dan varians dari distribusi referensi yang sama dengan estimasi sampel (Wikipedia, 2020).

Uji ini digunakan untuk menguji *goodness of fit* antara distribusi sampel dan distribusi lainnya serta membandingkan serangkain data pada sampel terhadap distribusi normal serangkaian nilai dengan *mean* dan standar deviasi yang sama (Siregar, 2015).

Uji Kolmogorof-Smirnof dilakukan dengan membandingkan *Dhit* dan *Dtabel*.

$$Dhitung = \max |F0(x) - Sn(x)|$$

Keterangan,

$F0(x)$ = distribusi frekuensi kumulatif teoritis

$Sn(x)$ = distribusi frekuensi kumulatif skor observasi (Purwanto, 2011). Menurut Real-Statistic (2019) menyatakan bahwa Uji Kolmogorov-Smirnov satu sampel digunakan untuk menguji apakah sampel berasal dari distribusi tertentu. Kita dapat menggunakan prosedur ini untuk menentukan apakah sampel berasal dari populasi yang terdistribusi normal.

Menurut Statistikian (2013) Uji Shapiro Wilk adalah sebuah metode atau rumus perhitungan sebaran data yang dibuat oleh shapiro dan wilk. Metode shapiro wilk adalah metode uji normalitas yang efektif dan valid digunakan untuk sampel berjumlah kecil.

Menurut Razali, N.M & Wah, Y.B. (2011) menyatakan bahwa uji Shapiro dan Wilk awalya dibatasi untuk ukuran sampel yang kurang dari 50. Uji ini merupakan uji pertama yang mampu mendeteksi kenormalan data berdasarkan skewness dan kurtosis atau keduanya. Berikut perumusan untuk Shapiro – Wilk :

$$W = \frac{\left(\sum_{i=1}^n a_i y_i \right)^2}{\sum_{i=1}^n \left(y_i - \bar{y} \right)^2}$$

Keterangan,

y_i = nilai data atau urutan data ke – i

\bar{y} = rata – rata sampel

$$a_i = (a_1, \dots, a_n) = \frac{m^T V^{-1}}{(m^T V^{-1} V^{-1} m)^{1/2}} \text{ dan}$$

$m = (m_1, \dots, m_n)^T$ adalah nilai yang diharapkan dari urutan statistik yang independen dan didistribusikan secara identik. variabel acak yang diambil dari distribusi normal standar dan V adalah matriks kovarians dari statistik orde tersebut.

METODE

Penelitian ini termasuk penelitian kuantitatif dengan menggunakan tehnik analisis non-parametrik dengan uji Kolmogorof-smirnov dan uji Shapiro-Wilk. Penelitian ini dilaksanakan di Universitas Muhammadiyah Makassar khususnya di Prodi Pendidikan Matematika dengan jumlah sampel yaitu sebanyak 136 orang. Data yang digunakan merupakan data primer yang diambil langsung dari sampel yang dipilih secara acak dan menjadi objek penelitian. Data yang diperoleh dari penelitian ini selanjutnya dianalisis dengan menggunakan teknik statistik deskriptif dan analisis non-parametrik.

Analisis statistik deskriptif digunakan untuk mengetahui gambaran awal dari data yang diperoleh. Analisis deskriptif ditampilkan dalam bentuk rata-rata, standar deviasi, skor maksimum, skor minimum, variansi, serta berupa tabel dan diagram batang. Sedangkan analisis non-parametrik digunakan untuk mengetahui kenormalan data.

HASIL DAN PEMBAHASAN

Analisis Deskriptif

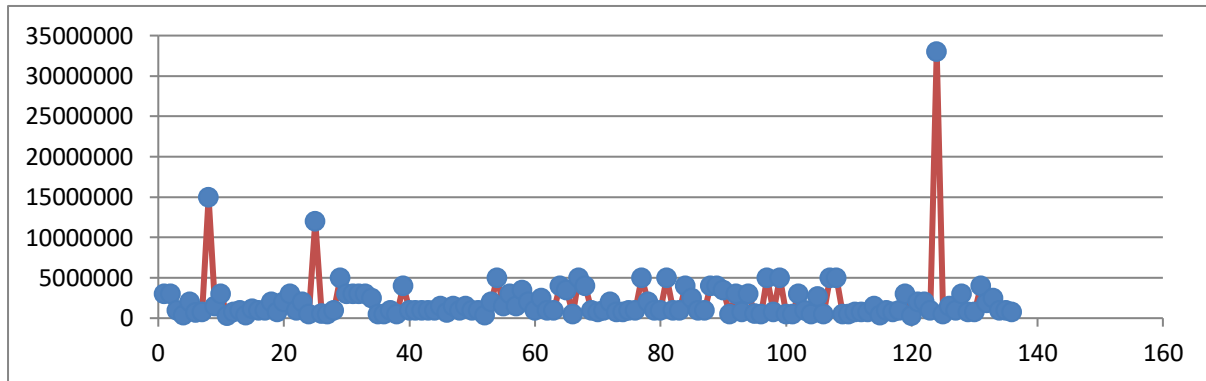
Adapun gambaran awal dari hasil analisis deskriptif penghasilan orangtua mahasiswa sebagai berikut :

Tabel 1. Hasil analisis deskriptif

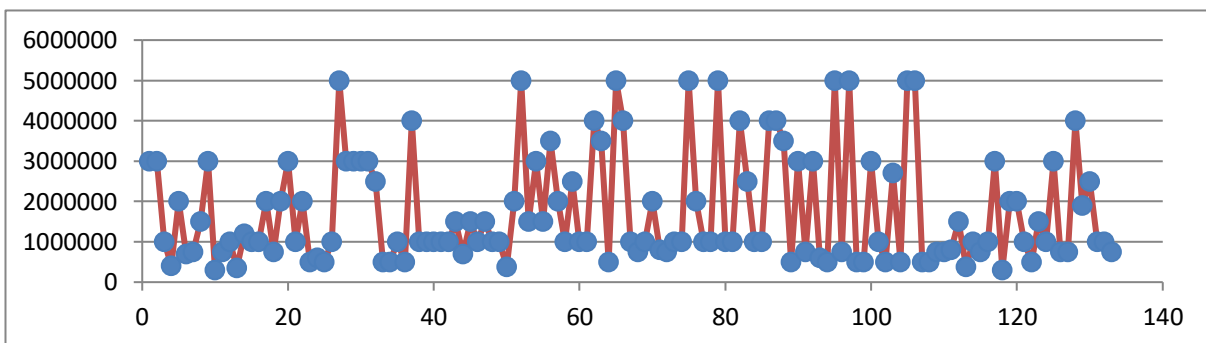
<i>Deskriptif</i>	
Mean	1743233.083
Standard Error Standard	117093.3333
Deviation	1350386.196
Sample Variance	1.82354E+12
Kurtosis	0.147091614

Skewness	1.120607179
Range	4700000
Minimum	300000
Maximum	5000000
Sum	231850000
Count	133

sebesar Rp.1.350.386,-. Adapun nilai minimum dan maksimumnya yaitu berada diangka Rp.300.000,- dan Rp.5.000.000,- dengan rentang Rp.4.700.000,-. Dari analisis deskriptif tersebut hanya digunakan 133 data dari 136 data yang telah dipilih. Hal ini disebabkan karena terdapat 3 data yang sangat jauh melenceng dari rata-rata sehingga dapat mengakibatkan standar deviasi dari data penghasilan orangtua mahasiswa sangat besar dan dapat menampilkan pola data yang tidak baik. Berikut ditampilkan perbedaan grafik data dengan 136 jumlah sampel dan 133 jumlah sampel



Grafik 1. Penghasilan orang tua 136 sampel



Grafik 2. Penghasilan orang tua 133 sampel

Hasil Uji Kenormalan Data

Dari hasil analisis data dengan menggunakan uji *kolmogorof-smirnov* dan uji *saphiro-wilk* untuk data penghasilan orang tua mahasiswa diperoleh hasil sebagai berikut :

Tabel 2. Hasil uji normalitas

	K-S test statistic (D)	Shapiro-Wilk Test
value	0.27437	0.834435
p - value	0.00001	0.0000

Hipoetesis :
H0 : data berdistribusi normal

H1 : data tidak berdistribusi normal

Tingkat kesalahan dalam pengambilan keputusan ditetapkan $\alpha = 5\% = 0,05$. Dengan kriteria pengambilan dengan menggunakan sig. α atau p-value, apabila sig. $\alpha < \alpha$ maka tolak H₀. Hasil analisis data di atas diperoleh bahwa nilai p-value uji kolmogorof smirnof dan saphiro-wilk yaitu 0,00001 dan 0,0000 dimana masing – masing dari nilai ini lebih kecil dari α data penghasilan orangtua mahasiswa tidak berdistribusi normal.

SIMPULAN DAN SARAN

Dari hasil analisis yang telah dilakukan, dapat diambil kesimpulan bahwa dari total 136 jumlah sampel terdapat 3 sampel yang tidak dianalisis. Hal tersebut diakibatkan karena ke 3 data tersebut merupakan data pencilan dan dapat mempengaruhi penyebaran data terhadap rata – rata yang sangat besar sehingga mengakibatkan hasil dari standar deviasinya terlalu besar. Sedangkan untuk hasil analisis uji normalitas diperoleh bahwa baik uji kolmogorof-smirnov maupun uji sapiro-wilk menghasilkan nilai p-value yang lebih kecil dari α sehingga dapat disimpulkan data penghasilan orangtua mahasiswa tersebut tidak normal.

DAFTAR RUJUKAN

- Lukiastuti, Fitri & Hamdani, Muliawan.(2012). *Statistika Non Parametris*. Yogyakarta : CAPS.
- Nurgiantoro, dkk.(2012). *Statistika Terapan untuk Penelitian Ilmu-ilmu Sosial*. Yogyakarta : Gadjah Mada University Press.
- Purwanto.(2011). *Statistik untuk Penelitian*. Yogyakarta : Pustaka Pelajar.
- Razali, N.M & Wah, Y.B. (2011). *Power Comparisons Saphiro Wilk, Kolmogorov – Smirnov, Lilliefors and Anderson Darling Test*. Jurnal of Statistical modeling and analytics Vol.2.No.1, 21 -33, 2011.
- Real-Statistic. (2019). *Kolomogorov – Smirnov Test For Normality*. Diakses 2020.<http://www.real-statistics.com/tests-normality-and-symmetry/statistical-tests-normality-symmetry/kolmogorov-smirnov-test/>
- Riadi, Edi. (2016). *Statistika Penelitian (Analisis Manual dan IBM SPSS)*. Yogyakarta : CV. ANDI OFFSET.
- Santoso, Singgih, (2015). *Statistik NonParametrik*. Jakarta : PT. Elex Media Komputindo.
- Siregar, Syofian. (2015). *Statistik Terapan untuk Perguruan Tinggi*. Jakarta : Prenadamedia Group.
- Sugiyono.(2015). *Statistika Non Parametris untuk Penelitian*. Bandung : ALFABETA.
- Sujarweni, Wiratna. (2016). *Kupas Tuntas Penelitian Akuntansi dengan SPSS*. Yogyakarta : PUSTAKA BARU PRESS.
- Statistikian.(24 Januari 2013). *Pengertian dan Rumus Uji Saphiro Wilk – Cara*
- Hitung*.Diakses 2020.<https://www.statistikian.com/2013/01/saphiro-wilk.html>.
- Tiro, M. Arif.(2015). *Dasar – Dasar Statistika*. Makassar : Andira Publisher.
- Wikipedia.(2020). *Kolmogorov – Smirnov Test*. Diakses 2020.https://en.wikipedia.org/wiki/Kolmogorov%E2%80%93Smirnov_test