

BAB Uji Normalitas

Uji distribusi normal adalah uji untuk mengukur apakah data kita memiliki distribusi normal sehingga dapat dipakai dalam statistik parametrik (*statistik inferensial*). Cara yang biasa dipakai untuk menghitung masalah ini adalah **Chi Square**. Tapi karena tes ini memiliki kelemahan, maka yang kita pakai adalah **Kolmogorov-Smirnov**. Kedua tes dinamakan masuk dalam kategori **Goodness Of Fit Tes**.

Januar: Makanan lagi ini? Ayo Jelaskan apa yang kau maksud dengan fitness tes ini.

Bukan Fitness Tes, tapi Goodness Of Fit Tes. Artinya, uji apakah **data empirik** yang kamu dapatkan dari lapangan itu sesuai dengan **distribusi teoritik** tertentu. Dalam kasus ini, distribusi normal. Dengan kata lain, apakah datamu itu dari populasi yang berdistribusi normal.

Januar: Mengapa kita harus, ngetes normalitas segala?

Pertama, Tes-tes parametrik itu dibangun dari distribusi normal, kau lihat tabel t-tes misalnya, pembuatannya itu mengacu pada tabel normalitas. Kedua, kita bisa berasumsi bahwa sampel kita benar-bener mewakili populasi. sehingga hasil penelitian kita bisa digeneralisasikan pada populasi. Bukankah dalam pandangan statistik itu sifat dan karakteristik populasi adalah terdistribusi secara normal.

Terus, bagaimana kalau kita langsung meneliti populasi secara langsung. Misalnya Hubungan Antara Independensi Anak yang Jarang Mandi di Fakultas Psikologi UGM Dengan Kreativitas. Populasinya khan cuma tiga. Aku, kamu, dan Sony '93. Apakah harus di tes normal segala?.

Mbuh!

Chi-Square

Filosofi mengapa *Chi-Square* kok bisa dikatakan *Goodness Of Fit Tes*, adalah begini:

Aku punya uang seratus rupiah. Tak lempar seratus kali, sisi A keluar sebanyak 35 kali, sisi B keluar sebanyak 65 kali. Apakah koinku dapat dikatakan seimbang..maksud'e koinku gak penceng?.

Macam Data	Kemunculan Sisi Koin	Total
Data Teoritik	A = 50	100
	B = 50	
Data Observasi	A = 35	100
	B = 65	

Uji Chi Square
 Sig. p>0,05 : Tidak Ada Beda
 Sig. p<0.05 : Ada Beda

Kalau hasilnya tidak ada perbedaan, maka dapat dikatakan bahwa koin kita setimbang.

Kita Lihat dulu Data teoritik kurve Normal. Kurve normal punya 6 Standar Deviasi (sd). Masing-masing sd luasnya seperti ini.

Sd	Distribusi Kurve Normal			Data Hasil
	Kategori	%	Frek.	Frek.
-3	1 - 10	2%	2	5
-2	11 - 15	14%	14	15
-1	16 - 20	34%	34	20
1	21 - 30	34%	34	38
2	31 - 35	14%	14	12
3	36 - 40	2%	2	10

Sig. p>0,05 :
Tidak Ada
Beda
Sig. p<0.05 :
Ada Beda

Kolmogorov -Smirnov

Chi Square membandingkan distribusi teoritik dan distribusi empirik (observasi) berdasarkan **kategori-kategori**, kalau KS berdasarkan **frekuensi kumulatif**. Jadi yang dibandingkan adalah frekuensi kumulatif distribusi teoritik dengan frekuensi kumulatif distribusi empirik.

Data	frekuensi	Frekuensi kumulatif
2	5	5
3	2	7
5	3	10
7	5	15
Total	15	

Frekuensi Kumulatif
Frekuensi kumulatif adalah penjumlahan frekuensi per-baris hingga ke bawah

Lihat tabel dibawah ini. Ini contoh-contohan uji normalitas data. Ada 6 pembagian.

Uji KS outputnya adalah D. Kalau t-tes khan t. Kalau korelasi khan r. D itu didapatkan dari distribusi kumulatif teoritik dikurangi distribusi kumulatif empirik. Tapi tidak semua yang diambil. Hanya satu yang diambil yaitu yang selisihnya terbesar.

Hasil ini lalu dibandingkan dengan tabel D.

	Sebaran Normal						Total
	1	2	3	4	5	6	
Frekuensi teoririk	1	2	3	3	2	1	12
Distribusi kumulatif (teoritik).....A	1/12	3/12	6/12	9/12	11/12	12/12	
Frekuensi empirik	1	3	2	3	1	2	12
Distribusi kumulatif (empirik).....B	1/12	4/12	6/12	9/12	10/12	12/12	
A - B	0	-1/12	0	0	1/12	0	

Cara Membaca Angka

One-Sample Kolmogorov-Smirnov Test

		VAR00001
N		12
Normal Parameters ^{a,b}	Mean	3,0833
	Std. Deviation	1,37895
Most Extreme Differences	Absolute	,164
	Positive	,117
	Negative	-,164
Kolmogorov-Smirnov Z		,567
Asymp. Sig. (2-tailed)		,905

a. Test distribution is Normal.

b. Calculated from data.

Test Distribution is Normal artinya, yang diuji itu distribusi normal, bukan distribusi eksponen, atau poisson

Dari sini dapat dikatakan bahwa data anda berdistribusi normal. $D = 0,164$ ($p > 0,05$).
or
Beberapa orang ada yang menjadikan acuan signifikansi adalah Z. dan biasanya mereka menulis $Z = 0,567$ ($p > 0,05$)

Negative

Pengurangan yang menghasilkan angka negatif terbesar

Positive

Pengurangan yang menghasilkan angka negatif terbesar

Absolut (D)

Dari perbandingan antara negatif dan positif, yang terbesar lah yang dimasukkan sebagai absolut. Dalam kasus ini $D = 0,164$. Jika D anda lebih kecil dari tabel, maka data anda NORMAL.

Deviasi Distribusi Normal

Jika Z anda di bawah 1,97 maka dapat dikatakan tidak ada perbedaan antara distribusi teoritik dan distribusi empirik..data anda NORMAL !

Menampilkan Uji Kolmogorov-Smirnov

Cara Pertama

The screenshot shows the SPSS 'One-Sample Kolmogorov-Smirnov Test' dialog box. The 'Test Variable List' contains several variables including 'Age in years [age]'. In the 'Test Distribution' section, the 'Normal' checkbox is selected. A callout box with an arrow points to this section, containing the text 'Pilih distribusi normal'.

Cara Kedua

1. Pilih **Descriptive Statistics** → **Explore**

The screenshot shows the SPSS 'Data Editor' with a table of data. The 'Analyze' menu is open, and 'Explore...' is selected. The 'Explore' dialog box is open, showing 'jenis kelamin [gen]' in the 'Factor List' and 'depresi [depresi]' and 'makna hidup [makn]' in the 'Dependent List'. The 'Plots' button is highlighted.

	depresi	makna
1	13,00	96,00
2	9,00	108,00
3	15,00	111,00
4	10,00	114,00
5	11,00	112,00
6	4,00	108,00
7	4,00	88,00
8	10,00	115,00

2. Masukkan variabel yang hendak di uji pada kotak **Dependen**.

3. Tekan tombol **Plots**.

4. Beri tanda pada **Normality Plot With Test**

The screenshot shows the 'Explore: Plots' dialog box. The 'Normality plots with tests' checkbox is checked. Other options include 'Factor levels together', 'Dependents together', 'None', 'Stem-and-leaf', and 'Histogram'.

Output Kolmogorov Smirnov Cara Kedua

Descriptives

			Statistic	Std. Error
depresi	Mean		9,5732	,77137
	95% Confidence Interval for Mean	Lower Bound	8,0384	
		Upper Bound	11,1080	
	5% Trimmed Mean		9,1287	
	Median		9,0000	
	Variance		48,791	
	Std. Deviation		6,98505	
	Minimum		,00	
	Maximum		29,00	
	Range		29,00	
	Interquartile Range		9,2500	
	Skewness		,834	,266
	Kurtosis		,357	,526
	makna hidup	Mean		105,5366
95% Confidence Interval for Mean		Lower Bound	103,2176	
		Upper Bound	107,8556	
5% Trimmed Mean			105,6233	
Median			106,0000	
Variance			111,388	
Std. Deviation			10,55403	
Minimum			75,00	
Maximum			131,00	
Range			56,00	
Interquartile Range			13,0000	
Skewness			-,199	,266
Kurtosis			,639	,526

Perhatian

Karena ada koreksi Lilliefors, maka harga tes ini jadi mahal. Peluang tidak normal, lebih besar di sini. Jadi enakan pakai cara pertama saja.

Contoh :

Data Depresi milik Hendro. Jika dihitung dengan cara pertama sig.-nya 0,273 $p > 0,05 \rightarrow$ Normal

tetapi jika dihitung dengan cara kedua sig.-nya 0,015 $p < 0,05 \rightarrow$ Tak Normal

Mengapa ?

Cara pertama adalah uji *Kolmogorov-Smirnov Plus*. Soalnya ada tambahan Koreksi Lilliefors segala.

Apa Itu Koreksi Lilliefors?

Tuhan belum mengizinkan aku untuk menjawabnya.

Berkatalah hanya pada apa yang anda ketahui saja

Anonim

Tests of Normality

	Kolmogorov-Smirnov ^a			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
depresi	,110	82	,015	,937	82	,001
makna hidup	,088	82	,177	,982	82	,310

a. Lilliefors Significance Correction

Arti

Statistic

Adalah nilai D, seperti yang saya jelaskan sebelumnya.

Sig.

Adalah perbedaan antara distribusi teoritik dan distribusi empirik.

Keunggulan Kolmogorov Smirnov (KS) dibanding (Chi Square)

1. CS memerlukan data yang terkelompokkan, KS tidak memerlukannya.
2. CS tidak bisa untuk sampel kecil, sementara KS bisa.
Bayangkan jika data anda berjumlah 5, sedangkan anda harus membuat 6 kategori sd, Cs tidak bisa digunakan bukan?
3. Oleh karena data Chi Square adalah bersifat kategorik. Maka ada data yang terbuang maknanya. Misalkan kategori 11-15. Anda membuat angka 15 marah-marah. Ia merasa rugi karena dibulatkan ke bawah, padahal kurang satu no dia masuk kategori 16-20. Dan anda membuat angka 11 untung, karena ia dibulatkan ke atas, dan disamakan dengan angka di atasnya yaitu 12,13, 14 dan 15.
4. KS lebih fleksibel dibanding CS.
KS dapat mengestimasi variasi sd, sedangkan CS, sd nya sama, karena dibagi secara seimbang.
(info lebih lengkap baca buku *Non Parametrik Statistical Inference*, Gibbons, 1971)

Apa yang harus dilakukan jika sebaran data tidak normal

1. Kita transformasikan data kita dalam bentuk yang lain (*remedies for non normal*). Ada banyak cara mentransformasikan, tetapi cara yang sering dipakai adalah transformasi dalam bentuk akar kuadrat, arcsin, dan log 10. Lihat buku "*Multivariate Data Analysis*" karangan Hair dkk. (1995)
2. Jika cara 1 tidak bisa, tambah jumlah sampel penelitian, hingga katakanlah 100 sampel.
3. Jika tidak bisa juga, buang subjek yang teridentifikasi sebagai *outliers*. Lihat **Bab Outliers**
4. Jika tidak bisa...Relakan, data anda memang 'gak normal. Gunakan statistik non parametrik.

Transformasi Data

1. Tekan Menu Transform
2. Kemudian pilih Compute.
Lalu muncul seperti yang ini..

Macam-macam rumus yang tersedia di SPSS
Pada contoh ini, SQRT (*square root*) adalah akar kuadrat. Numexpr...di sini adalah variabel anda yg hendak ditransformasi.
Klik tanda panah, di atasnya untuk memilih fungsi ini.

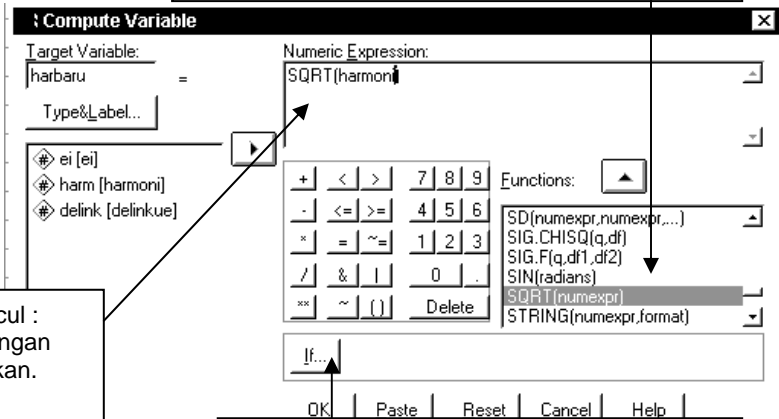
Nama Baru
Nama baru bagi variabel yang ditransformasikan

Pilih variabel yang hendak ditransformasikan

Ketika fungsi sudah dipilih, akan muncul : SQRT (?), tanda tanya ini ada ganti dengan variabel yang hendak di transformasikan. Kemudian jadi..SQRT (harmoni)

3. Klik OK jika sudah selesai.
4. Ulangi Uji Normalitas sekali lagi pada data anda.
5. Semoga menjadi normal.

Rumus Buatn Sendiri
Anda juga bisa membuat rumus dengan cara sendiri. Untuk mengerti arti lambang2 di sana. Klik kanan pada lambang itu. help akan muncul



Tambahan tentang Normalitas (Bagi yang ingin mendalami saja)

Satu istilah yang ngetrend dalam Kurve Normal adalah *Skewness* dan *Kurtosis*. *Skewness* berkaitan dengan lebar kurve, sedangkan *kurtosis* dengan tinggi kurve. Jika data terlihat sebarannya normal, tapi kalau nilai kurtosisnya besar (alias salah satu kategori terlalu tinggi) ya nggak normal. Dua nilai ini harus diperhatikan...

Nilai Kritis (Z) = $Skewness / \sqrt{6/N}$. Z tidak boleh lebih dari 2,58 (sig. 1%) dan 1,96 (sig. 5%).
Untuk Kurtosis juga lho..rumusnya sama.