BAB 6

UJI NORMALITAS POPULASI

Uji Normalitas dengan Uji Kolmogorov-Smirnov

Uji Kolmogorov-Smirnov dapat digunakan untuk menguji suatu asumsi apakah suatu data sampel berasal dari populasi yang berdistribusi normal atau tidak. Pada pembahasan Bab 5 telah dibahas mengenai distribusi sampling dari rata-rata $\bar{X}$. Apabila data sampel berasal dari populasi yang berdistribusi normal, maka distribusi sampling dari rata-rata $\bar{X}$ juga mengikuti distribusi normal. Asumsi normalitas memiliki peranan penting dalam uji-uji parametrik, seperti uji beda rata-rata dari dua populasi dengan uji t dan analisis varians. Hal ini karena uji-uji parametrik akan bekerja dengan baik ketika asumsi normalitas dipenuhi. Conover (1999:115) menyatakan sebagai berikut.

"Most parametric methods are based on the normality assumption because the theory behind the test can be worked out with the normal population distribution. The resulting procedures are efficient and powerful procedures for normally distributed data. Other parametric procedures have been developed assuming the population has other distributions, such as the exponential, Weibull, and soon".

Pada uji Kolmogorov-Smirnov, hipotesis nol menyatakan data yang diteliti berasal dari populasi yang berdistribusi normal, sedangkan hipotesis alternatif menyatakan data yang diteliti tidak berasal dari populasi yang berdistribusi normal. Andaikan $X_1, X_2, X_3, \ldots, X_k$ merupakan nilai-nilai pada sampel acak (*random sample*). Misalkan $f(X_i)$ menyatakan probabilitas dari nilai X_i, sedangkan $F(X_i) = f(X \leq X_i)$ menyatakan probabilitas kumulatif dari nilai X_i, di mana $i = 1,2,3, \ldots, k$. Selanjutnya andaikan Z_i merupakan nilai normal (sampel) terstandarisasi dari hasil transformasi nilai X_i dan $F(Z_i) = f(Z \leq Z_i)$ menyatakan probabilitas kumulatif dari nilai normal Z_i terstandarisasi. Nilai normal Z_i terstandarisasi merupakan hasil transformasi dari nilai X_i yang dihitung dengan rumus sebagai berikut.

$$Z_i = \frac{X_i - \bar{X}}{s}, i = 1,2,3, \ldots, k$$

Perhatikan bahwa $\bar{X}$ merupakan rata-rata sampel sebagai estimasi dari rata-rata populasi μ, sedangkan s merupakan standar deviasi sampel sebagai estimasi dari standar deviasi populasi σ. Misalkan D_i menyatakan nilai mutlak dari selisih antara $F(Z_i)$ dan $F(X_i)$, yakni

$$\mathrm{D_i} = |F(Z_i) - F(X_i)|, i = 1,2,3, \ldots, k.$$

Nilai D_i paling besar (*maximum*) atau D_{max} merupakan nilai statistik dari uji Kolmogorov-Smirnov. Nilai statistik dari uji Kolmogorov-Smirnov (D_{max}) kemudian dibandingkan dengan nilai kritis berdasarkan tabel distribusi Kolmogorov-Smirnov untuk pengambilan keputusan terhadap hipotesis. Berikut aturan pengambilan keputusan terhadap hipotesis berdasarkan uji Kolmogorov-Smirnov.

$$Jika\ D_{max} \leq nilai\ kritis, maka\ H_0\ diterima\ dan\ H_1\ ditolak.$$
$$Jika\ D_{max} > nilai\ kritis, maka\ H_0\ ditolak\ dan\ H_1\ diterima.$$

Tabel 6.1 merupakan tabel distribusi Kolmogorov-Smirnov. Pengambilan keputusan terhadap hipotesis juga dapat dilakukan dengan membandingkan nilai probabilitas (*p-value*) dari uji Kolmogorov-Smirnov terhadap tingkat signifikansi α (*significance level*). Berikut aturan pengambilan keputusan berdasarkan pendekatan nilai probabilitas.

$$Jika\ nilai\ probabilitas \geq tingkat\ signifikansi, maka\ H_0\ diterima\ dan\ H_1\ ditolak.$$
$$Jika\ nilai\ probabilitas < tingkat\ signifikansi, maka\ H_0\ ditolak\ dan\ H_1\ diterima.$$

Tabel 6.1

n	$\alpha = 0,2$	$\alpha = 0,1$	$\alpha = 0,05$	$\alpha = 0,02$	$\alpha = 0,01$
1	0,900	0,950	0,975	0,990	0,995
2	0,684	0,776	0,842	0,900	0,929
3	0,565	0,636	0,708	0,785	0,829
4	0,493	0,565	0,624	0,689	0,734
5	0,447	0,509	0,563	0,627	0,669
6	0,410	0,468	0,519	0,577	0,617
7	0,381	0,436	0,483	0,538	0,576
8	0,359	0,410	0,454	0,507	0,542
9	0,339	0,387	0,430	0,480	0,513
10	0,323	0,369	0,409	0,457	0,486
11	0,308	0,352	0,391	0,437	0,468
12	0,296	0,338	0,375	0,419	0,449
13	0,285	0,325	0,361	0,404	0,432
14	0,275	0,314	0,349	0,390	0,418
15	0,266	0,304	0,338	0,377	0,404
16	0,258	0,295	0,327	0,366	0,392
17	0,250	0,286	0,318	0,355	0,381
18	0,244	0,279	0,309	0,346	0,371

Contoh Kasus Uji Normalitas Populasi dengan Uji Kolmogorov-Smirnov (Contoh Perhitungan)

Misalkan seorang mahasiswa semester 8 sedang menyusun tugas akhir dan baru saja mengumpulkan data sampel mengenai nilai ujian matematika kelas 6 SD sebanyak 16 siswa. Berikut data yang telah dikumpulkan oleh mahasiswa tersebut.

Tabel 6.2 (Data Fiktif)

Nomor	Nama	Nilai	Nomor	Nama	Nilai	Nomor	Nama	Nilai
1	A	40	7	H	70	13	N	80
2	B	50	8	I	70	14	O	90
3	C	50	9	J	70	15	P	90
4	D	60	10	K	70	16	Q	100
5	F	60	11	L	80			
6	G	60	12	M	80			

Berikut akan digunakan pendekatan uji Kolmogorov-Smirnov untuk menguji hipotesis apakah data tersebut ditarik dari populasi yang berdistribusi normal atau tidak (misalkan tingkat signifikansi yang digunakan $\alpha = 5\%$). Perhitungan akan dilakukan secara manual.

➔Menghitung nilai rata-rata $(\bar{X})$ dan standar deviasi (s).

Tabel 6.3

No.	X	Frekuensi	$f(X)$	$F(X)$	Z	$F(Z)$	$D=\lvert F(Z)-F(X)\rvert$
1	40	1	0,0625	0,0625	-1,83712	0,033096276	0,029403724
2	50	2	0,125	0,1875	-1,22474	0,110335658	0,077164342
3	60	3	0,1875	0,375	-0,61237	0,270145667	0,104854333
4	70	4	0,25	0,625	0	0,5	**0,125**
5	80	3	0,1875	0,8125	0,612372	0,729854333	0,082645667
6	90	2	0,125	0,9375	1,224745	0,889664342	0,047835658
7	100	1	0,0625	1	1,837117	0,966903724	0,033096276

Berdasarkan Tabel 6.3, berikut akan dihitung nilai rata-rata hitung $(\bar{X})$ dan standar deviasi (s).

$$\bar{X} = \frac{\sum X}{n}$$

$$\bar{X} = \frac{(40 \times 1) + (50 \times 2) + (60 \times 3) + (70 \times 4) + (80 \times 3) + (90 \times 2) + (100 \times 1)}{16}$$

$$\bar{X} = 70$$

$$s = \sqrt{\frac{\sum (X - \bar{X})^2}{n - 1}}$$

$$s = \sqrt{\frac{4000}{15}}$$

$$s = 16{,}330.$$

➔Menghitung probabilitas dari X_i atau $f(X_i)$.

Setelah diperoleh $\bar{X} = 70$ dan $s = 16{,}330$, selanjutnya akan dihitung probabilitas dari X_i atau $f(X_i)$. Probabilitas untuk nilai $X = 40$ atau $f(40)$ adalah $\frac{1}{16} = 0{,}0625$, probabilitas untuk nilai $X = 50$ atau $f(50)$ adalah $\frac{2}{16} = 0{,}125$, probabilitas untuk nilai $X = 70$ atau $f(70)$ adalah $\frac{4}{16} = 0{,}25$, dan seterusnya.

➔Menghitung probabilitas kumulatif dari X_i atau $F(X_i) = f(X \leq X_i)$.

Nilai dari $F(40) = 0{,}0625$, nilai dari $F(50) = f(X \leq 50) = f(40) + f(50) = 0{,}0625 + 0{,}125 = 0{,}1875$, nilai dari $F(60) = f(X \leq 60) = f(40) + f(50) + f(60) = 0{,}375$, dan seterusnya.

➔Mentransformasi nilai X_i menjadi nilai normal Z_i terstandarisasi.

Selanjutnya mentransformasi nilai X_i ke dalam nilai normal Z_i terstandarisasi yang dihitung dengan rumus

$$Z_i = \frac{X_i - \bar{X}}{s}.$$

Untuk $X = 40$, maka

$$Z(X = 40) = \frac{40 - 70}{16{,}330} = -1{,}837.$$

Untuk $X = 50$, maka

$$Z(X = 50) = \frac{50 - 70}{16{,}330} = -1{,}2247,$$

dan seterusnya.

➔ Menghitung probabilitas kumulatif dari Z_i atau $F(Z_i) = f(Z \leq Z_i)$.

Setelah diperoleh nilai-nilai normal terstandarisasi, maka akan dihitung probabilitas kumulatif dari nilai-nilai normal terstandarisasi tersebut. Probabilitas kumulatif dari $Z = -1{,}837$ atau $f(Z \leq -1{,}837)$ berdasarkan tabel distribusi normal kumulatif adalah $0{,}033$, probabilitas kumulatif dari $Z = 0{,}61$ atau $f(Z \leq 0{,}61)$ berdasarkan tabel distribusi normal kumulatif adalah $0{,}729$, dan seterusnya.

➔ Menghitung nilai mutlak dari selisih antara $F(Z_i)$ dan $F(X_i)$.

Selanjutnya menghitung nilai mutlak dari selisih antara $F(Z_i)$ dan $F(X_i)$.

$$\mathrm{D_i} = |F(Z_i) - F(X_i)|.$$

Nilai D untuk $X = 40$ adalah $|0{,}033 - 0{,}0625| = 0{,}0295$, nilai D untuk $X = 50$ adalah $|0{,}110 - 0{,}1875| = 0{,}077$, dan seterusnya.

➔ Menghitung nilai statistik dari uji Kolmogorov-Smirnov (D_{max}).

Nilai statistik dari uji Kolmogorov-Smirnov merupakan nilai D yang paling besar atau maksimum. Berdasarkan Tabel 6.3, nilai D terbesar adalah $0{,}125$, sehingga nilai statistik dari uji Kolmogorov-Smirnov adalah $0{,}125$ atau $D_{max} = 0{,}125$.

➔ Menghitung nilai kritis Kolmogorov-Smirnov.

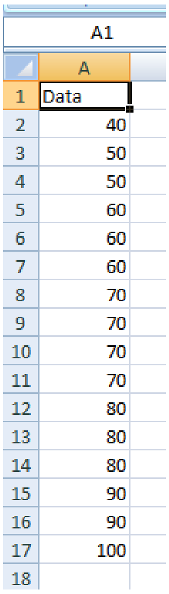
Nilai kritis Kolmogorov-Smirnov pada tingkat signifikansi 5% dan jumlah elemen sampel 16 berdasarkan tabel distribusi Kolmogorov-Smirnov adalah 0,327.

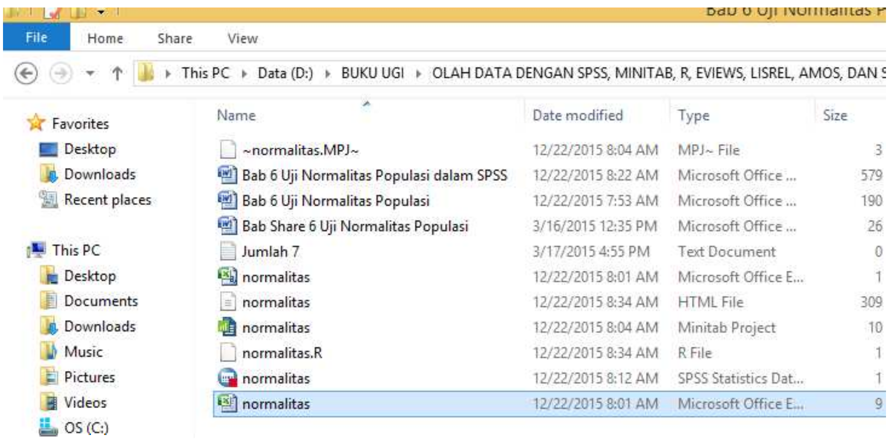
➔ Pengambilan keputusan terhadap hipotesis.

Perhatikan bahwa karena nilai statistik dari uji Kolmogorov-Smirnov (0,125) lebih kecil dibandingkan nilai kritis Kolmogorov-Smirnov (0,327), maka hipotesis nol diterima dan hipotesis alternatif ditolak, sehingga asumsi mengenai data nilai ujian matematika kelas 6 SD ditarik dari populasi yang berdistribusi normal dapat diterima pada tingkat signifikansi 5%.

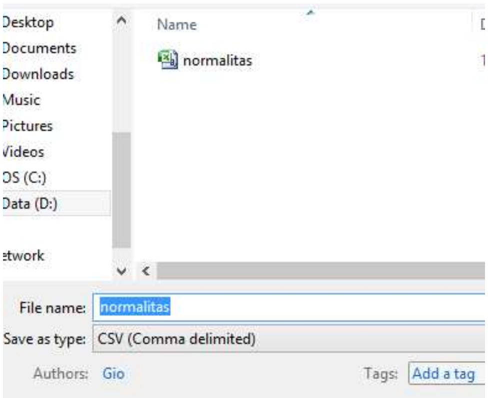
Penyelesaian dalam R untuk Uji Normalitas Populasi dengan Uji Kolmogorov-Smirnov

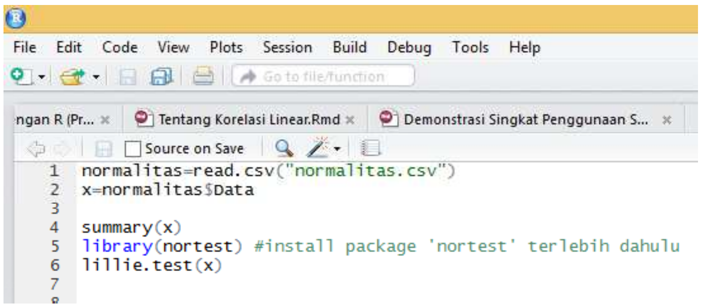
Data terlebih dahulu dibuat dalam *Microsoft Excel* (Gambar 6.1) dan disimpan dengan format tipe **.csv** (Gambar 6.2 dan Gambar 6.3). Ketik kode R seperti pada Gambar 6.4. Kemudian *Compile* dan pilih HTML (Gambar 6.5). Hasilnya seperti pada Gambar 6.6.



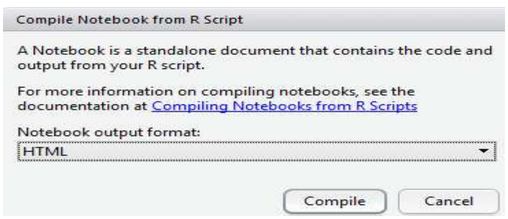


Gambar 6.1 **Gambar 6.2**





Gambar 6.3 **Gambar 6.4**



Gambar 6.5

```
normalitas=read.csv("normalitas.csv")
x=normalitas$Data

summary(x)
```

```
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##      40      60      70      70      80     100
```

```
lillie.test(x)
```

```
##
##  Lilliefors (Kolmogorov-Smirnov) normality test
##
## data:  x
## D = 0.125, p-value = 0.7235
```

Gambar 6.6

Pada Gambar 6.6, terlihat bahwa nilai statistik dari uji Kolmogorov-Smirnov (D) 0,125, lebih kecil dibandingkan nilai kritis Kolmogorov-Smirnov 0,327, maka hipotesis nol diterima dan hipotesis alternatif ditolak, sehingga asumsi mengenai data nilai ujian matematika kelas 6 SD ditarik dari populasi yang berdistribusi normal dapat diterima pada tingkat signifikansi 5%.

Perhatikan juga bahwa nilai probabilitas atau *p-value* adalah 0,7235. Karena nilai probabilitas, yakni 0,7235, lebih besar dibandingkan tingkat signifikansi, yakni 0,05, maka hipotesis nol diterima, dan hipotesis alternatif ditolak. Hal ini berarti asumsi mengenai data nilai ujian matematika kelas 6 SD ditarik dari populasi yang berdistribusi normal dapat diterima pada tingkat signifikansi 5%.

Pada Gambar 6.4, *package* **nortest** diaktifkan dengan maksud untuk menggunakan fungsi **lillie.test**. Fungsi **lillie.test** digunakan untuk menghitung nilai statistik dari uji Kolmogorov-Smirnov, dan probabilitasnya.

Uji Normalitas Populasi dengan Uji Jarque-Bera (Contoh Perhitungan dan Penyelesaian dalam R)

Berdasarkan data pada Tabel 6.2, berikut akan digunakan pendekatan uji Jarque-Bera (JB) untuk menguji hipotesis apakah data tersebut ditarik dari populasi yang berdistribusi normal atau tidak (misalkan tingkat signifikansi yang digunakan $\alpha = 5\%$). Perhitungan akan dilakukan secara manual. Nilai statistik dari uji JB dihitung dengan rumus sebagai berikut (Gujarati, 2003:148).

$$JB = n\left[\frac{S^2}{6} + \frac{(K-3)^2}{24}\right]$$

Perhatikan bahwa n menyatakan banyaknya elemen dalam sampel, S menyatakan kemiringan atau *skewness*, dan K menyatakan kurtosis. Untuk variabel yang terdistribusi secara normal, $S = 0$ dan $K = 3$. Oleh karena itu, uji normalitas JB merupakan suatu uji dari hipotesis gabungan (*joint hypothesis*), yakni S dan K masing-masing bernilai 0 dan 3. Dalam hal ini, nilai statistik dari uji JB diharapkan 0 (Gujarati, 2003:148).
Untuk kemiringan dan kurtosis dihitung dengan rumus sebagai berikut.

$$Kemiringan = \frac{\frac{1}{n}\sum(X-\bar{X})^3}{\left(\frac{1}{n}\sum(X-\bar{X})^2\right)^{3/2}}$$

$$Kurtosis = \frac{\frac{1}{n}\sum(X-\bar{X})^4}{\left(\frac{1}{n}\sum(X-\bar{X})^2\right)^{2}}$$

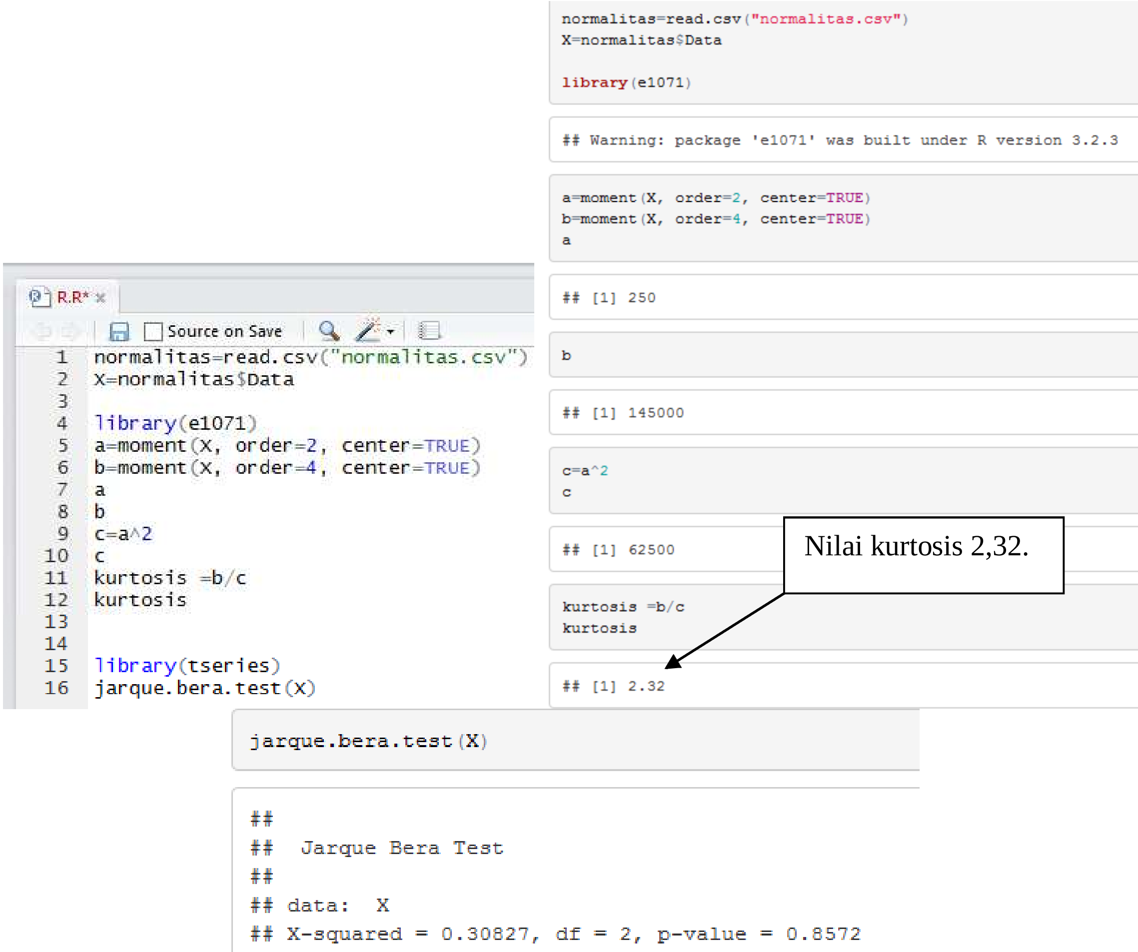
Tabel 6.4

	X	$(X-\bar{X})^2$	$(X-\bar{X})^3$	$(X-\bar{X})^4$
	40	900	-27000	810000
	50	400	-8000	160000
	50	400	-8000	160000
	60	100	-1000	10000
	60	100	-1000	10000
	60	100	-1000	10000
	70	0	0	0
	70	0	0	0
	70	0	0	0
	70	0	0	0
	80	100	1000	10000
	80	100	1000	10000
	80	100	1000	10000
	90	400	8000	160000
	90	400	8000	160000
	100	900	27000	810000
Jumlah	1120	4000	0	2320000
Rata-Rata	70	250	0	145000
Standar Deviasi	16.32993	296.6479395	10708.25227	268179.0447

$$Kemiringan = \frac{\frac{1}{n}\sum(X-\bar{X})^3}{\left(\frac{1}{n}\sum(X-\bar{X})^2\right)^{3/2}} = \frac{0}{\left(\frac{1}{n}\sum(X-\bar{X})^2\right)^{3/2}} = 0$$

$$Kurtosis = \frac{\frac{1}{n}\sum(X-\bar{X})^4}{\left(\frac{1}{n}\sum(X-\bar{X})^2\right)^{2}} = \frac{\frac{1}{16}(2320000)}{\left(\frac{1}{16}(4000)\right)^{2}} = \frac{145000}{62500} = 2,32$$

Gambar 6.7 menyajikan hasil perhitungan kurtosis. Berdasarkan Gambar 6.7, nilai dari kurtosis adalah 2,32.



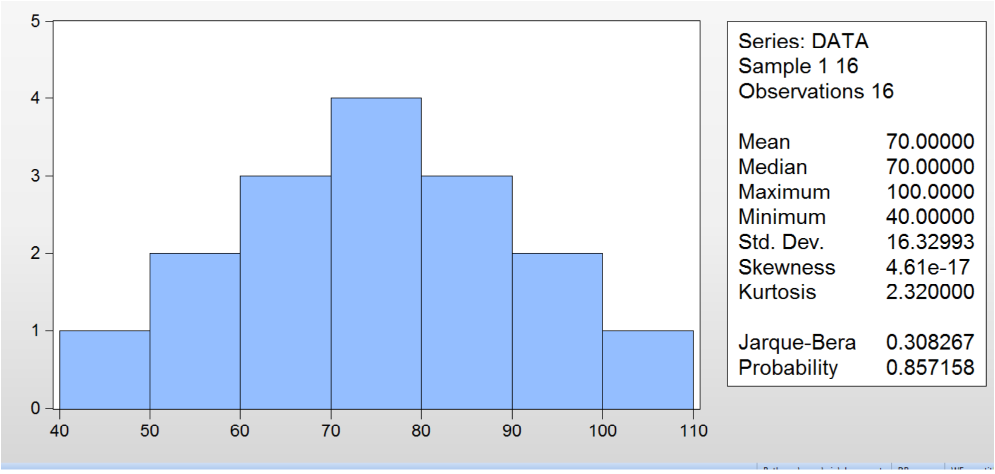
Gambar 6.7

Diketahui nilai kemiringan adalah 0 dan nilai kurtosis adalah 2,32. Sehingga nilai statistik dari uji JB dihitung sebagai berikut.

$$JB = n\left[\frac{S^2}{6} + \frac{(K-3)^2}{24}\right] = 16\left[\frac{0^2}{6} + \frac{(2{,}32-3)^2}{24}\right]$$

$$JB = 0{,}308267$$

Gambar 6.8 ditampilkan hasil perhitungan nilai statistik dari uji JB berdasarkan *software* EViews. Untuk hasil perhitungan nilai statistik dari uji JB berdasarkan R, disajikan pada Gambar 6.7 (*X-squared* = 0,30827).



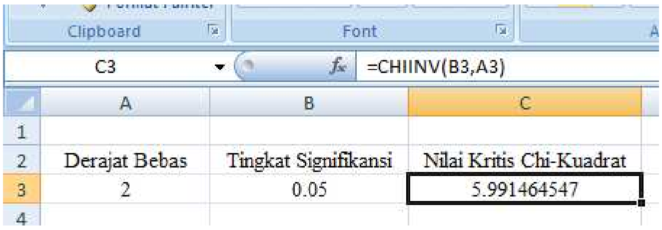
Gambar 6.8

Pengambilan keputusan terhadap hipotesis, dapat dilakukan dengan membandingkan nilai statisik dari uji Jarque-Bera terhadap nilai kritis chi-kuadrat χ^2_{kritis}. Statistik dari uji Jarque-Bera berdistribusi sampling chi-kuadrat dengan derajat bebas 2 untuk ukuran sampel yang besar. Gujarati (2003:148) menyatakan sebagai berikut.

"*Under the null hypothesis that the residuals are normally distributed, Jarque and Bera showed that asymptotically (i.e., in large samples) the JB statistic given in (5.12.1)* ***follows the chi-square distribution with 2 df***."

Berikut aturan pengambilan keputusan terhadap hipotesis.

$$Jika\ nilai\ statistik\ JB \leq \chi^2_{kritis}, H_0\ diterima\ dan\ H_1\ ditolak.$$
$$Jika\ nilai\ statistik\ JB > \chi^2_{kritis}, H_0\ ditolak\ dan\ H_1\ diterima.$$



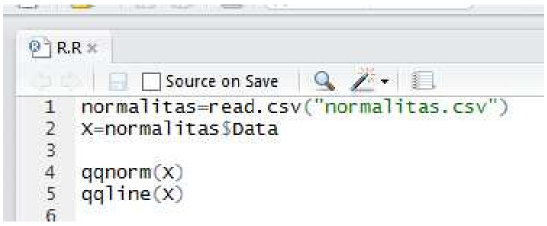
Gambar 6.9 Menghitung Nilai Kritis Chi-kuadrat dengan *Microsoft Excel*

Berdasarkan Gambar 6.9, diketahui nilai kritis chi-kuadrat bernilai 5,991. Karena nilai statisik dari uji Jarque-Bera, yakni 0,308, lebih kecil dibandingkan nilai kritis chi-kuadrat, yakni 5,991, maka hipotesis nol diterima dan hipotesis alternatif ditolak, sehingga asumsi mengenai data nilai ujian matematika kelas 6 SD ditarik dari populasi yang berdistribusi normal dapat diterima pada tingkat signifikansi 5%.

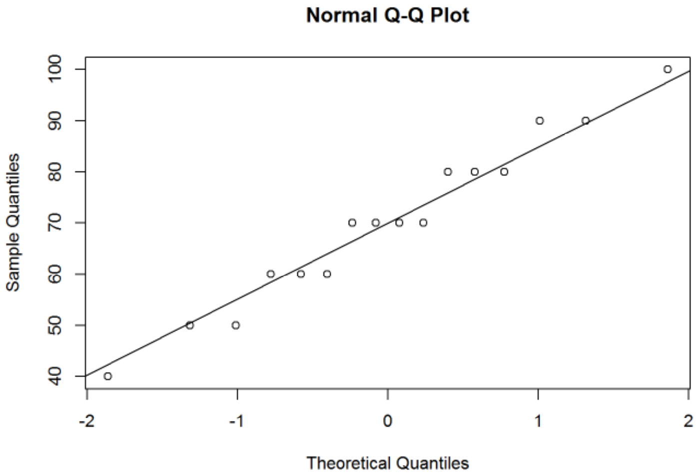
Perhatikan juga bahwa nilai probabilitas atau *p-value* adalah 0,8572 (lihat Gambar 6.7). Karena nilai probabilitas, yakni 0,8572, lebih besar dibandingkan tingkat signifikansi, yakni 0,05, maka hipotesis nol diterima, dan hipotesis alternatif ditolak. Hal ini berarti asumsi mengenai data nilai ujian matematika kelas 6 SD ditarik dari populasi yang berdistribusi normal dapat diterima pada tingkat signifikansi 5%.

Uji Normalitas Populasi dengan Quantile-Quantile Plot (Q-Q Plot)

Untuk menguji asumsi normalitas juga dapat digunakan pendekatan analisis grafik, yakni *Q-Q (quantile-quantile) plot*. Pada pendekatan *Q-Q plot*, jika titik-titik (*dots*) menyebar jauh (menyebar berliku-liku pada garis diagonal seperti ular) dari garis diagonal, maka diindikasi asumsi normalitas tidak dipenuhi. Jika titik-titik menyebar sangat dekat pada garis diagonal, maka asumsi normalitas dipenuhi. Ilustrasi dalam R diperlihatkan pada Gambar 6.10 dan Gambar 6.11.



Gambar 6.10



Gambar 6.11

Berdasarkan Gambar 6.11, titik-titik (*dots*) menyebar cukup dekat dari garis diagonal, maka maka asumsi normalitas dipenuhi.

Referensi

1. Conover, W.J. 1999. *Practical Nonparametric Statistics*, 3rd *Edition*. New York: John Wiley & Sons, Inc.

2. Field, A. 2009. *Discovering Statistics Using SPSS, 3rd Edition*. London: Sage.

3. Gio, P.U. dan E. Rosmaini, 2015. Belajar Olah Data dengan SPSS, Minitab, R, Microsoft Excel, EViews, LISREL, AMOS, dan SmartPLS. USUpress.

4. Mann, P. S. dan C.J. Lacke. 2011. *Introductory Statistics, International Student Version, 7th Edition*, Asia: John Wiley & Sons, Inc.

5. Montgomery, D.C. dan G.C. Runger. 2011. *Applied Statistics and Probability for Engineers, 5^{th} Edition*. United States of America: John Wiley & Sons, Inc.

6. http://www.r-tutor.com/elementary-statistics/numerical-measures/skewness

7. http://www.r-tutor.com/elementary-statistics/numerical-measures/moment

8. http://stats.stackexchange.com/questions/130368/why-do-i-get-this-p-value-doing-the-jarque-bera-test-in-r

9. http://www.inside-r.org/packages/cran/tseries/docs/jarque.bera.test

10. https://cran.r-project.org/web/packages/nortest/nortest.pdf

11. https://cran.r-project.org/web/packages/e1071/e1071.pdf

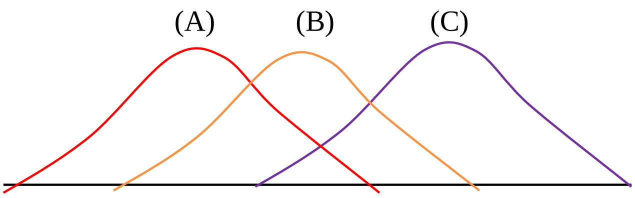
12. https://cran.r-project.org/web/packages/tseries/tseries.pdf

BAB 7

UJI KESAMAAN VARIANS POPULASI

Uji Kesaman Varians Populasi dengan Uji Levene

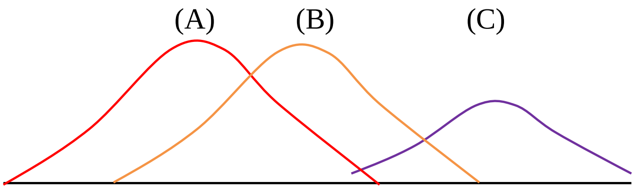
Uji Levene merupakan salah satu uji dalam statistika yang dapat digunakan untuk menguji kesamaan varians dari dua atau lebih populasi. Selain uji Levene, dapat juga digunakan uji F, uji Hartley, dan uji Bartlett untuk menguji kesamaan varians populasi. Varians populasi dilambangkan dengan σ^2, sedangkan varians sampel dilambangkan dengan s^2.



$$\sigma_A^2 = \sigma_B^2 = \sigma_C^2$$

Gambar 7.1

Pada Gambar 7.1, varians dari populasi A, B, dan C adalah sama, namun rata-ratanya berbeda. Pada Gambar 7.2, varians dari populasi A dan B sama, namun berbeda dengan C.



Gambar 7.2

Pada uji Levene, hipotesis nol menyatakan tidak terdapat perbedaan varians di antara populasi, sedangkan hipotesis alternatif menyatakan terdapat paling tidak sepasang varians populasi yang berbeda. Field (2009:150) menyatakan sebagai berikut.

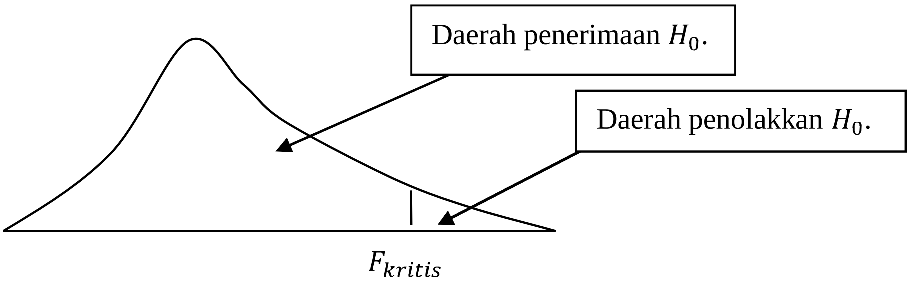
"*Levene's test tests null hypothesis that the variances in different groups are equal (i.e. the difference between the variances is zero).*"

Untuk pengambilan keputusan terhadap hipotesis dapat dilakukan dengan membandingkan nilai statistik dari uji Levene (L) terhadap nilai kritis dari tabel distribusi F (F_{kritis}). Berikut aturan pengambilan keputusan terhadap hipotesis berdasarkan uji Levene.

$$Jika\ L \leq nilai\ kritis\ F, maka\ H_0\ diterima\ dan\ H_1\ ditolak.$$
$$Jika\ L > nilai\ kritis\ F, maka\ H_0\ ditolak\ dan\ H_1\ diterima.$$

Pengambilan keputusan terhadap hipotesis juga dapat dilakukan dengan membandingkan nilai probabilitas dari uji Levene terhadap tingkat signifikansi α (*significance level*).

$$Jika\ nilai\ probabilitas \geq tingkat\ signifikansi, maka\ H_0\ diterima\ dan\ H_1\ ditolak.$$
$$Jika\ nilai\ probabilitas < tingkat\ signifikansi, maka\ H_0\ ditolak\ dan\ H_1\ diterima.$$



Contoh Kasus Uji Kesamaan Varians Populasi dengan Uji Levene (Contoh Perhitungan)

Misalkan diberikan data mengenai nilai ujian matematika kelas 1,2, dan 3 SMA (Tabel 7.1). Berdasarkan data pada Tabel 7.1, X menyatakan nilai ujian matematika siswa kelas 1 SMA, Y menyatakan nilai ujian matematika siswa kelas 2 SMA, dan Z menyatakan nilai ujian siswa kelas 3 SMA. Berikut akan digunakan pendekatan uji Levene untuk menguji apakah asumsi populasi X, Y, dan Z memiliki varians yang sama (secara statistika), dapat diterima atau tidak, pada tingkat signifikansi 5%.

Tabel 7.1 (Data Fiktif)

Nilai Ujian Matematika		
X	Y	Z
70	80	70
80	85	87
87	70	90
77	77	77
80	85	76
	60	87
	80	

Tabel 7.2 menyajikan proses perhitungan untuk memperoleh nilai statistik dari uji Levene (L).

Tabel 7.2

	X	Y	Z	$a = \lvert X - \bar{X} \rvert$	$b = \lvert Y - \bar{Y} \rvert$	$c = \lvert Z - \bar{Z} \rvert$
	70	80	70	8,8	3,28571429	11,16666667
	80	85	87	1,2	8,28571429	5,833333333
	87	70	90	8,2	6,71428571	8,833333333
	77	77	77	1,8	0,28571429	4,166666667
	80	85	76	1,2	8,28571429	5,166666667
		60	87		16,7142857	5,833333333
		80			3,28571429	
Jumlah	394	537	487	21,2	46,8571429	41
Rata-rata	78,8	76,71429	81,16667	4,24	6,69387755	6,833333333

	$d = (a - \bar{a})^2$	$e = (b - \bar{b})^2$	$f = (c - \bar{c})^2$
	20,7936	11,61557684	18,77777778
	9,2416	2,53394419	1
	15,6816	0,000416493	4
	5,9536	41,06455643	7,111111111
	9,2416	2,53394419	2,777777778
		100,4085798	1
		11,61557684	
Jumlah	60,912	169,7725948	34,66666667
Rata-rata			

➔Menghitung rata-rata gabungan dari data a, b, dan c.

$$\bar{X}_{a,b,c} = \frac{\sum a + \sum b + \sum c}{n_a + n_b + n_c}$$

$$\bar{X}_{a,b,c} = \frac{21{,}2 + 46{,}8571429 + 41}{5 + 7 + 6}$$

$$\bar{X}_{a,b,c} = 6{,}05873.$$

➔Menghitung nilai statistik dari uji Levene (L).

$$L = \frac{\frac{n_a\left(\bar{X}_a - \bar{X}_{a,b,c}\right)^2 + n_b\left(\bar{X}_b - \bar{X}_{a,b,c}\right)^2 + n_c\left(\bar{X}_c - \bar{X}_{a,b,c}\right)^2}{(k-1)}}{\frac{(\sum d + \sum e + \sum f)}{(N-k)}}.$$

$$n_a\left(\bar{X}_a - \bar{X}_{a,b,c}\right)^2 = (5)(4{,}24 - 6{,}05873)^2 = 16{,}5389$$

$$n_b\left(\bar{X}_b - \bar{X}_{a,b,c}\right)^2 = (7)(6.69387755 - 6{,}05873)^2 = 2{,}823885$$

$$n_c\left(\bar{X}_c - \bar{X}_{a,b,c}\right)^2 = (6)(6.833333333 - 6{,}05873)^2 = 3{,}60006$$

$$L = \frac{\frac{16{,}5389 + 2{,}823885 + 3{,}60006}{3 - 1}}{\frac{60{,}912 + 169{,}7725948 + 34{,}66667}{18 - 3}}$$

$$L = \frac{\frac{22{,}96284}{2}}{\frac{265{,}3513}{15}}$$

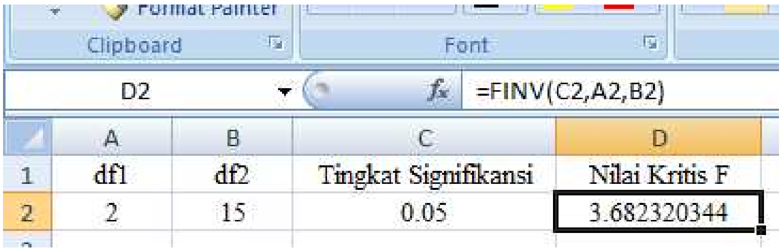
$$L = 0{,}64903148.$$

➔Menghitung nilai kritis F.

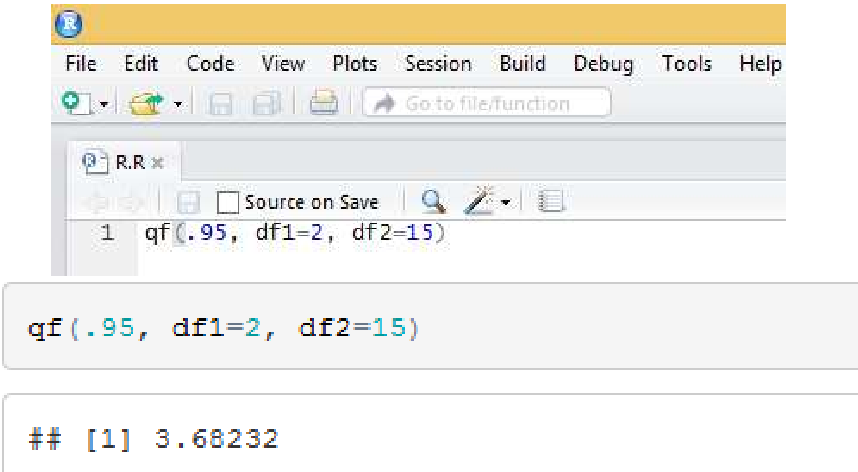
Berikut rumus untuk menghitung nilai derajat bebas pembilang dan derajat bebas penyebut.

$$Derajat\ bebas\ pembilang = k - 1.$$
$$Derajat\ bebas\ penyebut = N - k.$$

Perhatikan bahwa k menyatakan banyaknya sampel, sedangkan N merupakan jumlah elemen atau pengamatan dari seluruh sampel. Diketahui nilai k adalah 3, sedangkan nilai N adalah 18 ($n_1 + n_2 + n_3 = 5 + 7 + 6 = 18$). Diketahui tingkat signifikansi yang digunakan adalah 5%, sehingga nilai kritis F dengan derajat bebas pembilang $3 - 1 = 2$, derajat bebas penyebut $18 - 3 = 15$, dan tingkat signifikansi 5% adalah 3,68.



Gambar 7.3 Menentukan Nilai Kritis F dengan *Microsoft Excel*



Gambar 7.4 Menentukan Nilai Kritis F dengan R

➔Pengambilan keputusan terhadap hipotesis.

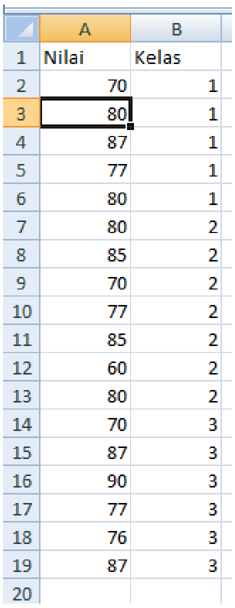
Berikut aturan pengambilan keputusan terhadap hipotesis berdasarkan uji Levene.

$$Jika\ L \leq nilai\ kritis\ F, maka\ H_0\ diterima\ dan\ H_1\ ditolak.$$
$$Jika\ L > nilai\ kritis\ F, maka\ H_0\ ditolak\ dan\ H_1\ diterima.$$

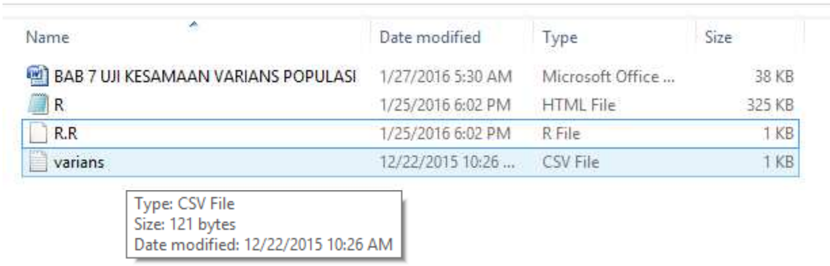
Perhatikan bahwa karena nilai statistik dari uji Levene, yakni 0,649, lebih kecil dibandingkan nilai kritis F, yakni 3,68, maka hipotesis nol diterima dan hipotesis alternatif ditolak, sehingga asumsi bahwa sampel X, Y, dan Z berasal dari populasi-populasi yang memiliki varians populasi yang sama, dapat diterima pada tingkat signifikansi 5%.

Penyelesaian dalam R untuk Uji Kesamaan Varians Populasi dengan Uji Levene

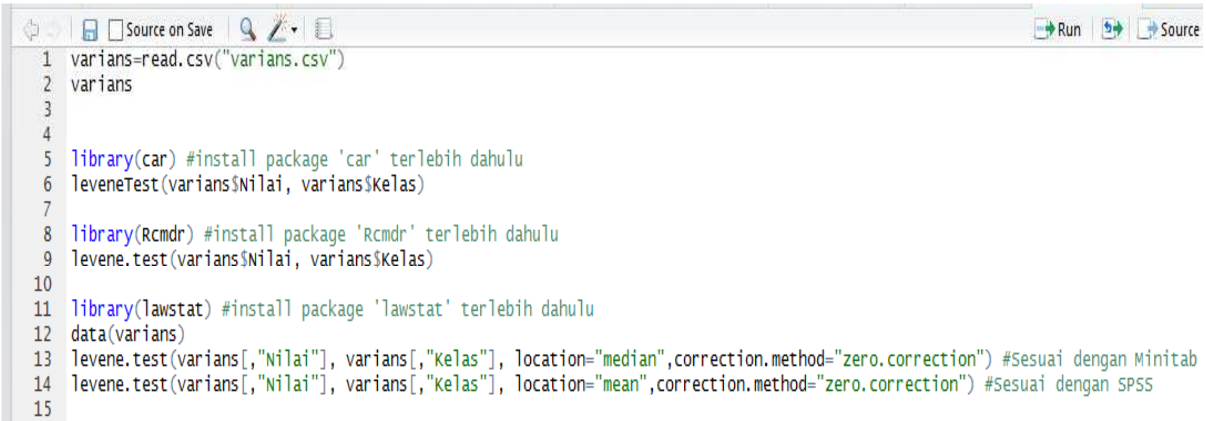
Data terlebih dahulu dibuat dalam *Microsoft Excel* (Gambar 7.5) dan disimpan dengan format tipe **.csv** (Gambar 7.6). Ketik kode R seperti pada Gambar 7.7. Kemudian *Compile* dan pilih HTML (Gambar 7.8). Hasilnya seperti pada Gambar 7.9 dan Gambar 7.10.



	A	B
1	Nilai	Kelas
2	70	1
3	80	1
4	87	1
5	77	1
6	80	1
7	80	2
8	85	2
9	70	2
10	77	2
11	85	2
12	60	2
13	80	2
14	70	3
15	87	3
16	90	3
17	77	3
18	76	3
19	87	3
20		



Gambar 7.5 **Gambar 7.6**



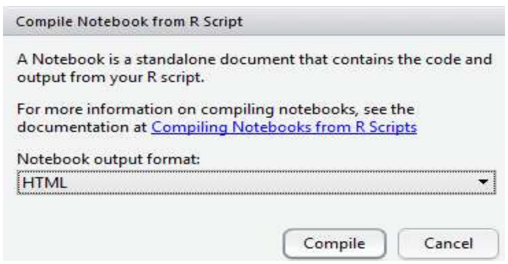

```
varians=read.csv("varians.csv")
varians


library(car) #install package 'car' terlebih dahulu
leveneTest(varians$Nilai, varians$Kelas)

library(Rcmdr) #install package 'Rcmdr' terlebih dahulu
levene.test(varians$Nilai, varians$Kelas)

library(lawstat) #install package 'lawstat' terlebih dahulu
data(varians)
levene.test(varians[,"Nilai"], varians[,"Kelas"], location="median",correction.method="zero.correction") #Sesuai dengan Minitab
levene.test(varians[,"Nilai"], varians[,"Kelas"], location="mean",correction.method="zero.correction") #Sesuai dengan SPSS
```

Gambar 7.7



Gambar 7.8

```
varians=read.csv("varians.csv")
varians
```

```
##    Nilai Kelas
## 1     70     1
## 2     80     1
## 3     87     1
## 4     77     1
## 5     80     1
## 6     80     2
## 7     85     2
## 8     70     2
## 9     77     2
## 10    85     2
## 11    60     2
## 12    80     2
## 13    70     3
## 14    87     3
## 15    90     3
## 16    77     3
## 17    76     3
## 18    87     3
```

```
## Levene's Test for Homogeneity of Variance (center = median)
##       Df F value Pr(>F)
## group  2  0.4267 0.6604
##       15
```

Gambar 7.9

```
levene.test(varians[,"Nilai"], varians[,"Kelas"], location="median",correction.method="zero.correction") #Sesuai dengan Minitab
```

```
##
##  modified robust Brown-Forsythe Levene-type test based on the
##  absolute deviations from the median with modified structural zero
##  removal method and correction factor
##
## data:  varians[, "Nilai"]
## Test Statistic = 0.4372, p-value = 0.6557
```

```
levene.test(varians[,"Nilai"], varians[,"Kelas"], location="mean",correction.method="zero.correction") #Sesuai dengan SPSS
```

```
##
##  classical Levene's test based on the absolute deviations from the
##  mean ( zero.correction not applied because the location is not
##  set to median )
##
## data:  varians[, "Nilai"]
## Test Statistic = 0.64903, p-value = 0.5366
```

Gambar 7.10

Perhatikan Gambar 7.10. Nilai statistik dari uji Levene dengan pendekatan *Location = "median"* adalah 0,4372, yang mana hasil ini sama dengan hasil Minitab. Namun nilai statistik dari uji Levene dengan pendekatan *Location = "mean"* adalah 0,649, yang mana hasil ini sama dengan hasil SPSS.

Diketahui juga berdasarkan Gambar 7.10 nilai probabilitas (*p-value*) adalah 0,5366, yakni lebih besar dibandingkan tingkat signifikansi 0,05, maka hipotesis nol diterima dan hipotesis alternatif ditolak, sehingga asumsi bahwa sampel X, Y, dan Z berasal dari populasi-populasi yang memiliki varians populasi yang sama, dapat diterima pada tingkat signifikansi 5%.

Contoh Kasus 2, Uji Kesamaan Varians Populasi dengan Uji Levene (Contoh Perhitungan dan Penyelesaian dengan R)

Misalkan diberikan data mengenai nilai ujian matematika kelas 1 dan 2 SMA (Tabel 7.3). Berdasarkan data pada Tabel 7.3, X menyatakan nilai ujian matematika siswa kelas 1 SMA, dan Y menyatakan nilai ujian matematika siswa kelas 2 SMA. Berikut akan digunakan pendekatan uji Levene untuk menguji apakah asumsi populasi X, Y, dan Z memiliki varians yang sama, dapat diterima atau tidak, pada tingkat signifikansi 5%.

Tabel 7.3 (Data Fiktif)

X	Y
30	10
40	20
50	30
60	40
70	50
80	60
90	70

Perhatikan bahwa sudah bisa diduga atau ditebak bahwa hipotesis nol diterima, yakni sampel X dan sampel Y ditarik dari populasi-populasi yang memiliki varians (*variance*) yang sama. Hal ini dikarenakan nilai nilai varians dari X dan Y **bernilai sama,** yakni 466,67 (lihat Gambar 7.11).

```
R.R*
Source on Save
Simpan=read.csv("varians2.csv",header=TRUE, sep=",") #membaca data
Simpan

library(doBy)
summaryBy(Nilai ~ Kelas, data = Simpan, FUN = function(x)
{ c(ratarata = mean(x), varians = var(x), jumlah=sum(x)   ) } )
```

```
Simpan=read.csv("varians2.csv",header=TRUE, sep=",") #membaca data
Simpan
```

```
##    Nilai Kelas
## 1     30     1
## 2     40     1
## 3     50     1
## 4     60     1
## 5     70     1
## 6     80     1
## 7     90     1
## 8     10     2
## 9     20     2
## 10    30     2
## 11    40     2
## 12    50     2
## 13    60     2
## 14    70     2
```

```
library(doBy)
```

```
## Loading required package: survival
```

```
summaryBy(Nilai ~ Kelas, data = Simpan, FUN = function(x)
{ c(ratarata = mean(x), varians = var(x), jumlah=sum(x)   ) } )
```

```
##   Kelas Nilai.ratarata Nilai.varians Nilai.jumlah
## 1     1             60      466.6667          420
## 2     2             40      466.6667          280
```

Gambar 7.11

Berdasarkan Gambar 7.11, diketahui nilai varians (*variance*) dari sampel X dan sampel Y, masing-masing adalah 466,6667. Tabel 7.4 menyajikan proses perhitungan untuk memperoleh nilai statistik dari uji Levene (L).

Tabel 7.4

	X	Y	$a = \|X - \bar{X}\|$	$b = \|Y - \bar{Y}\|$	$c = (a - \bar{a})^2$	$d = \left(b - \bar{b}\right)^2$
	30	10	30	30	165,3061	165,3061
	40	20	20	20	8,163265	8,163265
	50	30	10	10	51,02041	51,02041
	60	40	0	0	293,8776	293,8776
	70	50	10	10	51,02041	51,02041
	80	60	20	20	8,163265	8,163265
	90	70	30	30	165,3061	165,3061
Rata-Rata	60	40	17,14285714	17,14286	106,1224	106,1224
Jumlah	420	280	120	120	742,8571	742,8571

➔Menghitung rata-rata gabungan dari data a dan b.

$$\bar{X}_{a,b} = \frac{\sum a + \sum b}{n_a + n_b} = \frac{120 + 120}{7 + 7} = 17{,}14285714.$$

➔Menghitung nilai statistik dari uji Levene (L).

$$L = \frac{\dfrac{n_a\left(\bar{X}_a - \bar{X}_{a,b}\right)^2 + n_b\left(\bar{X}_b - \bar{X}_{a,b}\right)^2}{(k-1)}}{\dfrac{(\sum c + \sum d)}{(N-k)}}.$$

$$n_a\left(\bar{X}_a - \bar{X}_{a,b}\right)^2 = (7)(17{,}1428 - 17{,}1428)^2 = 0$$

$$n_b\left(\bar{X}_b - \bar{X}_{a,b}\right) = (7)(17{,}1428 - 17{,}1428)^2 = 0$$

$$L = \frac{\dfrac{0 + 0}{2 - 1}}{\dfrac{742{,}8571 + 742{,}8571}{14 - 2}}$$

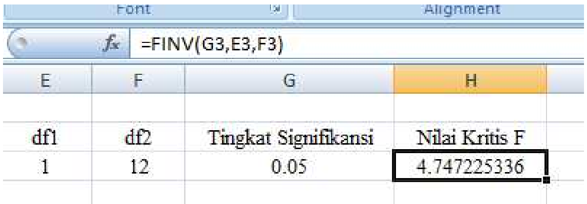
$$L = 0.$$

➔Menghitung nilai kritis F.

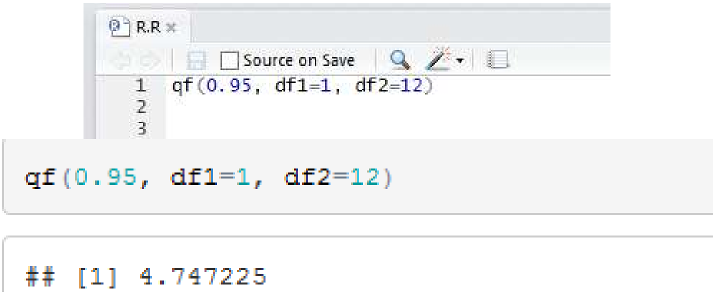
Berikut rumus untuk menghitung nilai derajat bebas pembilang dan derajat bebas penyebut.

$$Derajat\ bebas\ pembilang = k - 1.$$
$$Derajat\ bebas\ penyebut = N - k.$$

Perhatikan bahwa k menyatakan banyaknya sampel, sedangkan N merupakan jumlah elemen atau pengamatan dari seluruh sampel. Diketahui nilai k adalah 2, sedangkan nilai N adalah 14 $(n_1 + n_2 = 7 + 7 = 14)$. Diketahui tingkat signifikansi yang digunakan adalah 5%, sehingga nilai kritis F dengan derajat bebas pembilang $2 - 1 = 1$, derajat bebas penyebut $14 - 2 = 12$, dan tingkat signifikansi 5% adalah 4,747.



Gambar 7.12   Menentukan Nilai Kritis F dengan *Microsoft Excel*



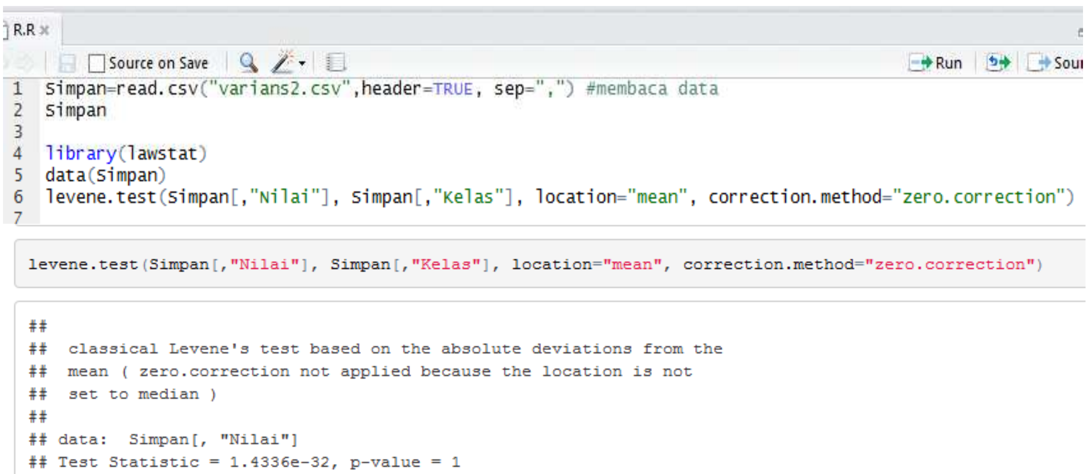
Gambar 7.13   Menentukan Nilai Kritis F dengan R

➔Pengambilan keputusan terhadap hipotesis.

Berikut aturan pengambilan keputusan terhadap hipotesis berdasarkan uji Levene.

$$Jika\ L \leq nilai\ kritis\ F, maka\ H_0\ diterima\ dan\ H_1\ ditolak.$$
$$Jika\ L > nilai\ kritis\ F, maka\ H_0\ ditolak\ dan\ H_1\ diterima.$$

Perhatikan bahwa karena nilai statistik dari uji Levene, yakni 0, lebih kecil dibandingkan nilai kritis F, yakni 4,747, maka hipotesis nol diterima dan hipotesis alternatif ditolak, sehingga asumsi bahwa sampel X dan sampel Y berasal dari populasi-populasi yang memiliki varians populasi yang sama, dapat diterima pada tingkat signifikansi 5%. Gambar 7.14 menyajikan hasil penyelesaian dengan R.



Gambar 7.14

Perhatikan Gambar 7.14. Nilai statistik dari uji Levene dengan pendekatan *Location* = *"mean"* adalah $1{,}4336 \times 10^{-32} = 0{,}0000000$ ..... Diketahui juga berdasarkan Gambar 7.14 nilai probabilitas (*p-value*) adalah 1, yakni lebih besar dibandingkan tingkat signifikansi 0,05, maka hipotesis nol diterima dan hipotesis alternatif ditolak, sehingga asumsi bahwa sampel X dan sampel Y berasal dari populasi-populasi yang memiliki varians populasi yang sama, dapat diterima pada tingkat signifikansi 5%.

.

Referensi

1. Field, A. 2009. *Discovering Statistics Using SPSS, 3rd Edition*. London: Sage.

2. Gamst, G., L.S. Meyers dan A.J. Guarino. 2008. *Analysis of Variance Designs*. New York: Cambridge University Press.

3. Gio, P.U. dan E. Rosmaini, 2015. Belajar Olah Data dengan SPSS, Minitab, R, Microsoft Excel, EViews, LISREL, AMOS, dan SmartPLS. USUpress.

4. Ott, R.L. dan M. Longnecker. 2001. *An Introduction to Statistical Methods and Data Analysis, 5th Edition*. United States of America: Duxbury.

5. https://cran.r-project.org/web/packages/lawstat/lawstat.pdf

6. https://cran.r-project.org/web/packages/doBy/doBy.pdf

7. https://cran.r-project.org/web/packages/car/car.pdf

8. https://cran.r-project.org/web/packages/Rcmdr/index.html