


Nama: AUREL REGINA NIM: 065002300023	 Praktikum Probabilitas dan Statistika	MODUL 8 Nama Dosen: Dedy Sugiarto
Hari/Tanggal: Rabu, 22 May 2024		Nama Asisten Labratorium: 1. Kharisma Maulida Saara (064002200024) 2. Tarum Widyasti Pertiwi (064002200027)

Uji Kebebasan & Uji Kenormalan

1. Teori Singkat

Uji Kebebasan (*Independence Test*):

Uji ini dapat digunakan untuk melihat hubungan antar dua peubah yang umumnya bersifat kategorik. Cara kerja uji ini didasarkan pada tabel tabulasi frekuensi secara silang (cross tabulation) dari dua peubah.

Tabel tersebut disebut juga *tabel $r \times c$* (r silang c), dimana tabel tersebut mempunyai r baris dan c kolom. Total baris dan total kolom dalam tabel disebut "*Frekuensi Marjinal*".

Karakteristiknya :

1. Ukuran sampel grand total telah ditentukan.
1. Sampel berasal dari satu populasi
2. Hipotesis :
 H_0 : Dua variabel dari klasifikasi adalah independent (bebas)
 H_1 : Kedua variabel tidak independent (saling bergantung)

Langkah-langkah pengujian :

1. Tentukan H_0 dan H_1
1. Tentukan taraf nyata α
2. Tentukan Wilayah kritiknya, yaitu tolak H_0 jika $\chi^2_{hit} > \chi^2_{tabel}$

(χ^2_{tabel} dapat dilihat pada tabel A.6 pada buku Walpole hal 472) dengan v atau derajat bebas sebesar $(r-1) \cdot (c-1)$

4. Lakukan perhitungan untuk χ^2 dengan :

Hitung frekuensi harapan :

$$e_{ij} = \frac{(\text{Total kolom ybs}) \times (\text{Total baris ybs})}{(\text{Grand total } n)}$$

$$\chi^2 = \sum_{i=1}^r \sum_{j=1}^c \frac{(o_{ij} - e_{ij})^2}{e_{ij}}$$

Catatan :

Frekwensi (sel) harapan biasa dibulatkan ke integer terdekat atau satu desimal.

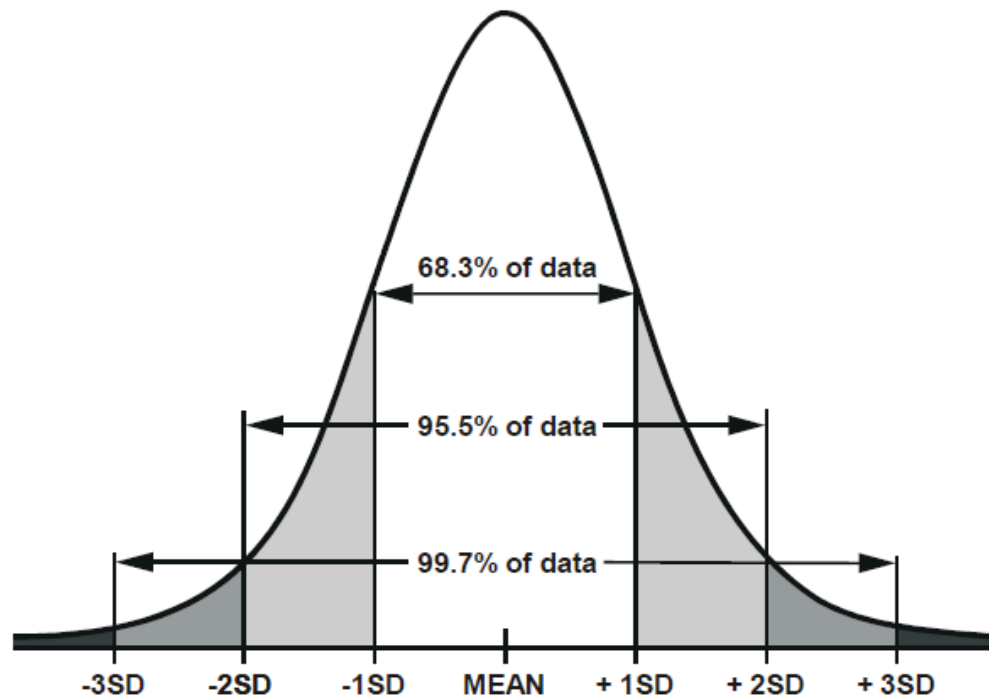
Terdapat “rule of five” dimana frekuensi dalam tiap sel minimum harus 5, jika kurang, maka 2 atau lebih sampel / data harus digabung sedemikian rupa sehingga tidak ada e yang dibawah 5.

5. Hitunglah :
6. Bandingkan χ^2_{hit} dengan χ^2_{tabel} , kemudian buat kesimpulannya.

Uji Kenormalan (*Normality Test*):

Peubah acak kontinu seperti tinggi badan, denyut jantung, waktu tempuh, umur sebuah merek bohlam dapat memiliki bentuk distribusi atau sebaran peluang yang berbeda-beda diantaranya adalah distribusi normal, eksponensial atau distribusi weibull. Sebaran peluang kontinu yang cukup penting dalam ilmu statistika adalah sebaran/distribusi peluang normal dengan kurvan yang berbentuk lonceng atau disebut bell-shaped distribution. Untuk mengetahui apakah suatu populasi mengikuti sebaran normal atau tidak dapat digunakan salah satu uji kesesuaian distribusi (Goodness of Fit) yaitu menggunakan uji Kolmogorov Smirnov.

Areas under the normal curve that lie between 1, 2, and 3 standard deviations on each side of the mean



Gambar distribusi Normal (berbentuk seperti lonceng)

2. Alat dan Bahan

Hardware : Laptop/PC Software: Jupyter Notebook

3. Elemen Kompetensi

- a. Latihan pertama – Distribusi Binomial
 1. Buka note baru pada Jupyter Notebook
 2. Implementasi manual rumus distribusi binomial

Latihan

1. Seseorang ingin menguji apakah kecukupan tidur mempengaruhi (ada hubungan) dengan kekuatan gowes sepeda seseorang. Didapat data survey sebagai berikut:

		Kecukupan Tidur			
		Kelebihan	Cukup	Kurang	Kurang sekali
Kemampuan gowes	35 km	8	22	15	5
	25 km	10	28	20	7
	15 km	12	30	20	8

Dengan uji kebebasan, apakah kecukupan tidur mempengaruhi kekuatan gowes seseorang dengan taraf nyata 1%.

R

```
df=read.delim("clipboard")
> # 1. convert the data as a table
> dt <- as.table(as.matrix(df))
> dt
> chisq <- chisq.test(df)
> chisq
> chisq$observed
> # Expected counts
> round(chisq$expected,2)
> # printing the p-value
> chisq$p.value
```

Output:

```
> df=read.delim("clipboard")
warning message:
In read.table(file = file, header = header, sep = sep, quote = quote, :
incomplete final line found by readTableHeader on 'clipboard'
> # 1. convert the data as a table
> dt <- as.table(as.matrix(df))
> dt
      kelebihan cukup kurang kurang.sekali
A           8    22    15           5
B          10    28    20           7
C          12    30    20           8
> chisq <- chisq.test(df)
> chisq

      Pearson's Chi-squared test

data:  df
X-squared = 0.18734, df = 6, p-value = 0.9999
> chisq$observed
      kelebihan cukup kurang kurang.sekali
[1,]          8    22    15           5
[2,]         10    28    20           7
[3,]         12    30    20           8
> # Expected counts
> round(chisq$expected,2)
      kelebihan cukup kurang kurang.sekali
[1,]      8.11 21.62 14.86           5.41
[2,]     10.54 28.11 19.32           7.03
[3,]     11.35 30.27 20.81           7.57
> # printing the p-value
> chisq$p.value
[1] 0.9998723
> |
```

[deskripsi] (minimal 4 baris)

menampilkan data dengan clipboard , mendapat $df=6$, $\chi^2=0.18734$, p-value data yg diharapkan tidak terlalu signifikan data awal, Karena nilai p-value yang didapatkan adalah 0.9999, yang jauh lebih besar dari tingkat signifikansi (0.01), maka kita gagal menolak hipotesis nol (H_0).

Python

```
import numpy as np
from scipy.stats import chi2_contingency

# Data dari tabel
data = np.array([[8, 22, 15, 5],
                 [10, 28, 20, 7],
                 [12, 30, 20, 8]])

# Menghitung chi-kuadrat
chi2, p, dof, expected = chi2_contingency(data)

# Menampilkan hasil
print(f"Chi-Square Statistic: {chi2}")
print(f"P-value: {p}")
print(f"Degrees of Freedom: {dof}")
print("Expected Frequencies:")
print(expected)

# Menentukan keputusan
alpha = 0.01
if p < alpha:
    print("Tolak H0: Terdapat hubungan antara kecukupan tidur dan kekuatan gowes.")
else:
    print("Gagal Tolak H0: Tidak terdapat hubungan antara kecukupan tidur dan kekuatan gowes.")
```

Output:

```

: import numpy as np
  from scipy.stats import chi2_contingency

# Data dari tabel
data = np.array([[8, 22, 15, 5],
                 [10, 28, 20, 7],
                 [12, 30, 20, 8]])

# Menghitung chi-kuadrat
chi2, p, dof, expected = chi2_contingency(data)

# Menampilkan hasil
print(f"Chi-Square Statistic: {chi2}")
print(f"P-value: {p}")
print(f"Degrees of Freedom: {dof}")
print("Expected Frequencies:")
print(expected)

# Menentukan keputusan
alpha = 0.01
if p < alpha:
    print("Tolak H0: Terdapat hubungan antara kecukupan tidur dan kekuatan gowes.")
else:
    print("Gagal Tolak H0: Tidak terdapat hubungan antara kecukupan tidur dan kekuatan gowes.")

```

Chi-Square Statistic: 0.1873409923409925
P-value: 0.999872291854867
Degrees of Freedom: 6
Expected Frequencies:
[[8.10810811 21.62162162 14.86486486 5.40540541]
[10.54054054 28.10810811 19.32432432 7.02702703]
[11.35135135 30.27027027 20.81081081 7.56756757]]
Gagal Tolak H0: Tidak terdapat hubungan antara kecukupan tidur dan kekuatan gowes.

[deskripsi] (minimal 4 baris)

didapatkan chi square 0.18734, df 6, dan berbagai nilai expected frequency, nilai p 0.999872291854867, yang jauh lebih besar dari tingkat signifikansi umum ($\alpha = 0.01$), maka kita gagal menolak hipotesis nol (H_0). tidak ada bukti yang cukup untuk menyatakan bahwa ada hubungan yang signifikan antara kecukupan tidur dan kekuatan gowes.

Excel

prob8_Aurel												
File Edit View Insert Format Data Tools Extensions Help												
Q Menu 100% \$ % .0 .00 123 Default... 11 B I A												
A	B	C	D	E	F	G	H	I	J	K	L	M
		Kecukupan Tidur								Total		
		Kelebihan		Cukup	Kurang		Kurang Sekali					
Kemampuan Gowes	11 km	8		22	15		5			50		
			8.10811		21.62162		14.86486		5.40541			
	15 km	10		28	20		7			65		nama: Aurel Regina
			10.54054		28.10811		19.32432		7.02703			
	20 km	12		30	20		8			70		
			11.35135		30.27027		20.81081		7.56757			
Total		30		80	55		20			185		
Kesimpulan :												
chi-kuadrat :		16.81										
STATISTIK UJI :		0.187341	Karena nilai statistika uji 0.187341 < dari nilai chi-kuadrat tabel 16.81 maka dapat disimpulkan terima H_0 yg berarti tidak terdapat hubungan signifikan antara kecukupan tidur dan kemampuan gowes.									

[deskripsi] (minimal 4 baris)

dengan menggunakan rumus excel nilai statistik uji chi-kuadrat sebesar 0.187341 jauh lebih kecil dari nilai chi-kuadrat tabel sebesar 16.81, maka kita gagal

menolak hipotesis nol (H_0). tidak terdapat hubungan yang signifikan antara kecukupan tidur dan kemampuan gowes.

2. Dari suatu autopsi diketahui berat otak 15 orang dewasa penderita penyakit tertentu sebagai berikut:

Berat Otak (gram)				
1348	1140	1086	1039	920
1233	1146	1002	1012	904
1255	1168	1016	1001	973

Berdasarkan data di atas, ujlilah apakah distribusi frekuensi mengikuti distribusi normal atau tidak ? ($\alpha = 5\%$)

Jawab:

1. H_0 : Populasi data berdistribusi normal
2. H_1 : Populasi data tidak berdistribusi normal
3. $\alpha = 0.05$
4. Daerah kritis: Tolak H_0 jika $D_{\text{maks}} > D_{\text{tabel}} (0.338)$ dengan $n = 15$.

Excel:

prob8_Aurel									
File Edit View Insert Format Data Tools Extensions Help									
Q Menus									
17 Nama : Aurel Regina									
A	B	C	D	E	F	G	H	I	J
	Xi	z	Ft(xi)	Fs(xi)	Ft(xi) - Fs(xi)	Nilai Max			
	904	-1.39	0.0824	0.0667	0.0158	0.1666			
	920	-1.26	0.1030	0.1333	0.0303				
	973	-0.85	0.1968	0.2000	0.0032				
	1001	-0.64	0.2625	0.2667	0.0042				
	1002	-0.63	0.2650	0.3333	0.0683				
	1012	-0.55	0.2911	0.4000	0.1089		Nama : Aurel Regina		
	1016	-0.52	0.3018	0.4667	0.1649				
	1039	-0.34	0.3667	0.5333	0.1666				
	1086	0.02	0.5097	0.6000	0.0903				
	1140	0.44	0.6713	0.6667	0.0047		Kesimpulan : Karena $0.1666 < 0.338$, maka terima H_0 dan disimpulkan bahwa data berdistribusi normal		
	1146	0.49	0.6880	0.7333	0.0453				
	1168	0.66	0.7457	0.8000	0.0543				
	1233	1.17	0.8781	0.8667	0.0115	Average :	1082.87		
	1255	1.34	0.9093	0.9333	0.0240	Stdev :	128.79		
	1348	2.06	0.9802	1.0000	0.0198				

[deskripsi] (minimal 4 baris)

didapatkan average 1082.87, standar deviasi 128.79, nilai maksimum dari selisih antara distribusi, nilai max 0.1666. Nilai ini lebih kecil dari nilai kritis 0.338. Oleh karena itu, kita menerima hipotesis nol (H_0) yang menyatakan bahwa data berdistribusi normal.

Python:

```
import numpy as np
from scipy import stats

# Data sampel nilai dari 15 mahasiswa
data = [904, 920, 973, 1001, 1002, 1002, 1012, 1016, 1039, 1086, 1140, 1146, 1168, 1233, 1255, 1348]

# Melakukan Uji Kolmogorov-Smirnov
stat, p_value = stats.kstest(data, 'norm', args=(np.mean(data), np.std(data, ddof=1)))

print("Statistik Uji Kolmogorov-Smirnov:", stat)
print("p-value:", p_value)

# Menentukan apakah data berdistribusi normal atau tidak
alpha = 0.05
if p_value > alpha:
    print("Data berdistribusi normal (gagal menolak H0)")
else:
    print("Data tidak berdistribusi normal (menolak H0)")
```

Output:

```
import numpy as np
from scipy import stats

# Data sampel nilai dari 15 mahasiswa
data = [904, 920, 973, 1001, 1002, 1002, 1012, 1016, 1039, 1086, 1140, 1146, 1168, 1233, 1255, 1348]

# Melakukan Uji Kolmogorov-Smirnov
stat, p_value = stats.kstest(data, 'norm', args=(np.mean(data), np.std(data, ddof=1)))

print("Statistik Uji Kolmogorov-Smirnov:", stat)
print("p-value:", p_value)

# Menentukan apakah data berdistribusi normal atau tidak
alpha = 0.05
if p_value > alpha:
    print("Data berdistribusi normal (gagal menolak H0)")
else:
    print("Data tidak berdistribusi normal (menolak H0)")

Statistik Uji Kolmogorov-Smirnov: 0.18805925481308217
p-value: 0.5606487063640894
Data berdistribusi normal (gagal menolak H0)
```

[deskripsi] (minimal 4 baris)

Didapatkan nilai statistik uji kolmo adalah 0.1880 dan p-value adalah 0.5606. Karena p-value lebih besar dari alpha (0.05), kita gagal menolak hipotesis nol (H_0). Dengan demikian, data berdistribusi normal.

TUGAS

- Seorang peneliti ahli gizi sedang melakukan penelitian dan ingin meneliti apakah ada pengaruh (hubungan dependent) antara pendapatan dan kualitas bahan makanan yang dikonsumsi oleh konsumen tersebut. Untuk itu maka diadakan penyelidikan terhadap 100 sampel individu dan diperoleh data sebagai berikut :
Ujilah data diatas dengan taraf nyata 5%

		Pendapatan			
		Tinggi	Sedang	Rendah	Jumlah
Mutu Bahan Makanan	Baik	14	6	9	29
	Cukup	10	16	10	36
	Jelek	2	13	20	35
Jumlah		26	35	39	100

Excel

prob8_Aurel										
File Edit View Insert Format Data Tools Extensions Help										
Q Menus 100% 123 Default...										
D20	A	B	C	D	E	F	G	H	I	J
1										
2										
3										
4										
5										
6										
7										
8										
9										
10										
11										
12										
13										
14										
15										

R

```
df=read.delim("clipboard")
# 1. convert the data as a table
dt <- as.table(as.matrix(df))
dt
chisq <- chisq.test(df)
chisq
chisq$observed
# Expected counts
round(chisq$expected,2)
# printing the p-value
chisq$p.value
```

Output:

```
> df=read.delim("clipboard")
Warning message:
In read.table(file = file, header = header, sep = sep, quote = quote, :
  incomplete final line found by readTableHeader on 'clipboard'
> # 1. convert the data as a table
> dt <- as.table(as.matrix(df))
> dt
      tinggi sedang rendah
A         14         6      9
B         10        16     10
C          2        13     20
> chisq <- chisq.test(df)
> chisq

        Pearson's Chi-squared test

data:  df
X-squared = 18.367, df = 4, p-value = 0.001046

> chisq$observed
      tinggi sedang rendah
[1,]      14         6      9
[2,]      10        16     10
[3,]       2        13     20
> # Expected counts
> round(chisq$expected,2)
      tinggi sedang rendah
[1,]   7.54  10.15  11.31
[2,]   9.36  12.60  14.04
[3,]   9.10  12.25  13.65
> # printing the p-value
> chisq$p.value
[1] 0.001046273
> |
```

[deskripsi] (minimal 4 baris)

Didapat derajat kebebasan (df) sebesar 4, nilai X-squared sebesar 18.367 dengan Nilai p-value yang diperoleh adalah 0.001046 lebih kecil daripada 0,05. Ini berarti bahwa menolak (H_0). p-value data yang diharapkan tidak terlalu signifikan dari awal, berarti data tidak saling berhubungan.

Python

```
import numpy as np
from scipy.stats import chi2_contingency

# Data dari tabel
data = np.array([[14, 6, 9],
                 [10, 16, 10],
                 [2, 13, 20]])

# Menghitung chi-kuadrat
chi2, p, dof, expected = chi2_contingency(data)

# Menampilkan hasil
print(f"Chi-Square Statistic: {chi2}")
print(f"P-value: {p}")
print(f"Degrees of Freedom: {dof}")
print("Expected Frequencies:")
```

```
print(expected)

# Menentukan keputusan
alpha = 0.05
if p < alpha:
    print("Tolak H0: Terdapat pengaruh (hubungan dependent) antara
pendapatan dan kualitas bahan makanan yang dikonsumsi.")
else:
    print("Gagal Tolak H0: Tidak terdapat pengaruh (hubungan dependent)
antara pendapatan dan kualitas bahan makanan yang dikonsumsi.")
```

Output:

```
import numpy as np
from scipy.stats import chi2_contingency

# Data dari tabel
data = np.array([[14, 6, 9],
                 [10, 16, 10],
                 [2, 13, 20]])

# Menghitung chi-kuadrat
chi2, p, dof, expected = chi2_contingency(data)

# Menampilkan hasil
print(f"Chi-Square Statistic: {chi2}")
print(f"P-value: {p}")
print(f"Degrees of Freedom: {dof}")
print("Expected Frequencies:")
print(expected)

# Menentukan keputusan
alpha = 0.05
if p < alpha:
    print("Tolak H0: Terdapat pengaruh (hubungan dependent) antara pendapatan dan kualitas bahan makanan yang dikonsumsi.")
else:
    print("Gagal Tolak H0: Tidak terdapat pengaruh (hubungan dependent) antara pendapatan dan kualitas bahan makanan yang dikonsumsi.")
```

Chi-Square Statistic: 18.366532937961512
P-value: 0.001046273329595159
Degrees of Freedom: 4
Expected Frequencies:
[[7.54 10.15 11.31]
 [9.36 12.6 14.04]
 [9.1 12.25 13.65]]
Tolak H0: Terdapat pengaruh (hubungan dependent) antara pendapatan dan kualitas bahan makanan yang dikonsumsi.

[deskripsi] (minimal 4 baris)

didapatkan nilai P-value (0.00104627) p-value tidak terlalu signifikan dari p-value awal, sehingga akan menolak H_0 karena p-value lebih kecil dari taraf nyata 0.05. Ini

```
import numpy as np
from scipy import stats

# Data sampel nilai dari 15 mahasiswa
```

```
data = [12, 25, 45, 67, 43, 33, 24, 45, 34, 11, 8, 34, 67, 99, 22]

# Melakukan Uji Kolmogorov-Smirnov
stat, p_value = stats.kstest(data, 'norm', args=(np.mean(data), np.std(data, ddof=1)))
print("Statistik Uji Kolmogorov-Smirnov:", stat)
print("p-value:", p_value)

# Menentukan apakah data berdistribusi normal atau tidak
alpha = 0.05
if p_value > alpha:
    print("Data berdistribusi normal (gagal menolak H0)")
else:
    print("Data tidak berdistribusi normal (menolak H0)")
```

Output:

```
import numpy as np
from scipy import stats

# Data sampel nilai dari 15 mahasiswa
data = [12, 25, 45, 67, 43, 33, 24, 45, 34, 11, 8, 34, 67, 99, 22]

# Melakukan Uji Kolmogorov-Smirnov
stat, p_value = stats.kstest(data, 'norm', args=(np.mean(data), np.std(data, ddof=1)))
print("Statistik Uji Kolmogorov-Smirnov:", stat)
print("p-value:", p_value)

# Menentukan apakah data berdistribusi normal atau tidak
alpha = 0.05
if p_value > alpha:
    print("Data berdistribusi normal (gagal menolak H0)")
else:
    print("Data tidak berdistribusi normal (menolak H0)")

Statistik Uji Kolmogorov-Smirnov: 0.18719291305040942
p-value: 0.6043842954060732
Data berdistribusi normal (gagal menolak H0)
```

[deskripsi] (minimal 4 baris)

didapatkan statistika uji kolmo 0.187192 ,p-value 0.604384 lebih besar dari alpha (0.05), kita gagal menolak hipotesis nol (H_0), yang berarti data berdistribusi normal. Jika p-value kurang dari atau sama dengan alpha, kita menolak H_0 , yang berarti data tidak berdistribusi normal.

4. File Praktikum

Github Repository:

<https://github.com/aurelregina/probabilitas8prak.git>

5. Kesimpulan

- a. Dalam pengerjaan praktikum Statistika, ... Uji Kebebasan (*Independence Test*) Uji ini digunakan untuk melihat hubungan antar dua peubah yang umumnya bersifat kategorik, pengujian kebebasan dapat menggunakan Excel, Python, serta Rstudio. Sedangkan uji kenormalan Untuk mengetahui apakah suatu populasi mengikuti sebaran normal atau tidak dapat digunakan salah satu uji kesesuaian distribusi (Goodness of Fit) yaitu menggunakan uji Kolmogorov Smirnov, pengujian hanya menggunakan Excel dan Python saja
- b. Kita juga dapat mengetahui... kita gagal menolak hipotesis nol (H_0), yang berarti data berdistribusi normal. Jika p-value kurang dari atau sama dengan alpha, kita menolak H_0 , yang berarti data tidak berdistribusi normal.

6. Cek List (✓)

No	Elemen Kompetensi	Penyelesaian	
		Selesai	Tidak Selesai
1.	Latihan	...✓	
2.	Tugas	...✓	

7. Formulir Umpan Balik

No	Elemen Kompetensi	Waktu Pengerjaan	Kriteria
1.	Latihan	20... Menit	menarik...
2.	Tugas	20... Menit	menarik...

Keterangan:

1. Menarik
2. Baik
3. Cukup
4. Kurang