

UJI CHI SQUARE

MODUL PERKULIAHAN SESI 9



Disusun oleh:

TIM DOSEN

Pelaksana Akademik Mata Kuliah Umum (PAMU)

Universitas Esa Unggul

Jakarta Barat

2019

UJI CHI SQUARE

Topik

1. Pendahuluan
2. Uji Independensi
3. Uji *Goodness of Fit*
4. Latihan soal

PENDAHULUAN

Pengujian dengan menggunakan Chi-Square diterapkan pada kasus dimana akan diuji apakah frekuensi data yang diamati (frekuensi/data observasi) sama atau tidak dengan frekuensi harapan atau frekuensi secara teoritis. Chi-Square disebut juga dengan Kai Kuadrat. Chi Square adalah salah satu jenis uji komparatif yang dilakukan pada dua variabel, di mana skala data kedua variabel adalah nominal. Apabila dari 2 variabel, ada 1 variabel dengan skala nominal maka dilakukan uji chi square dengan merujuk bahwa harus digunakan uji pada derajat yang terendah.

Nilai dari frekuensi observasi adalah suatu nilai yang diperoleh dari hasil percobaan sedangkan nilai frekuensi harapan (ekspektasi) adalah nilai yang diperoleh dari hasil perhitungan secara teoritis. Untuk selanjutnya, frekuensi observasi dinotasikan dengan (o) dan frekuensi harapan dinotasikan dengan (e).

Nilai χ^2 adalah nilai kuadrat karena itu nilai χ^2 selalu positif. Bentuk distribusi χ^2 tergantung dari derajat kebebasan. Untuk lebih jelasnya, akan diilustrasikan cara membaca tabel χ^2 yang ada pada Lampiran. Misalkan diberikan derajat kebebasan $db = 5$ dengan $\alpha = 0,01$. Dengan membaca tabel χ^2 pada lampiran, diperoleh nilai $\chi^2 = 15,086$.

Dalam pengujian Chi square, hal yang dapat diuji antara lain adalah uji independensi, uji *Goodness of Fit*, uji homogenitas, dan uji varians. Dalam buku ini, yang dibahas adalah mengenai uji independensi dan uji *goodness of fit*.

Uji independensi adalah uji untuk menentukan apakah antara variabel independen dan variabel dependennya terdapat perbedaan (hubungan) yang nyata atau tidak. Misalnya, kita ingin mengamati apakah terdapat perbedaan yang nyata antara pendidikan dengan pekerjaan, maka uji yang tepat dilakukan adalah uji independensi dengan Chi square.

Uji *Goodness of Fit* (kecocokan) adalah uji untuk menentukan apakah sebuah populasi mengikuti distribusi tertentu atau tidak. Misalnya, kita ingin mengetahui apakah populasi yang diamati berdistribusi normal atau tidak, atau mungkin populasi yang diamati ternyata berdistribusi poisson, dan seterusnya. Sehingga uji yang digunakan adalah uji *Goodness of Fit* dengan Chi square.

Untuk lebih memahami bagaimana penerapan uji independensi dan uji *goodness of fit*, secara lanjut dibahas pada subbab berikutnya, dan dilengkapi dengan beberapa contoh-contoh soalnya.

Uji Independensi

Seperti yang telah disebutkan di pendahuluan, Uji independensi adalah uji yang dilakukan untuk mengetahui apakah terdapat hubungan antara variabel independen dengan variabel dependen. Sebelum membahas contoh soal, alangkah baiknya kita mengetahui terlebih dahulu syarat uji independensi.

Syarat Uji independensi

Ada beberapa syarat yang harus dipenuhi jika akan melakukan pengujian dengan Chi Square. Berikut dijelaskan syarat-syarat yang harus dipenuhi, diantaranya:

- Tidak ada cell dengan nilai frekuensi kenyataan atau disebut juga *Actual Count* (F_o) sebesar 0 (Nol).
- Apabila bentuk tabel kontingensi 2×2 , maka tidak boleh ada 1 cell saja yang memiliki frekuensi harapan atau disebut juga *expected count* (" F_h ") kurang dari 5.
- Apabila bentuk tabel lebih dari 2×2 , misal 2×3 , maka jumlah cell dengan frekuensi harapan yang kurang dari 5 tidak boleh lebih dari 20%.

Jenis Uji Chi Square

Ada beberapa rumus yang digunakan untuk menyelesaikan suatu pengujian Chi Square. Seperti rumus koreksi yates, Fisher Exact Test, dan Pearson Chi Square. Berikut rincian penggunaan rumus-rumusnya.

- Jika tabel kontingensi berbentuk 2×2 , maka rumus yang digunakan adalah "koreksi yates".
- Apabila tabel kontingensi 2×2 , tetapi cell dengan frekuensi harapan kurang dari 5, maka rumus harus diganti dengan rumus "Fisher Exact Test".
- Rumus untuk tabel kontingensi lebih dari 2×2 , rumus yang digunakan adalah "Pearson Chi-Square",

KOREKSI YATES

$$\chi^2 = \frac{N(ad - bc)^2}{(a + b)(c + d)(a + c)(b + d)}$$

Dimana:

a, b, c, d = cell dari hasil persilangan dua variabel.

N = banyaknya sampel

FISHER EXACT TEST

$$\chi^2 = \frac{(a+b)!(c+d)!(a+c)!(b+d)!}{N! a! b! c! d!}$$

dimana:

a, b, c, d = cell dari hasil persilangan dua variabel.

N = banyaknya sampel

PEARSON CHI SQUARE

$$\chi^2 = \sum \frac{(f_o - f_e)^2}{f_e}$$

dimana :

χ^2 = Chi Square

f_o = Frekuensi Observasi

f_e = Frekuensi Ekspektasi

Untuk memahami apa itu “cell”, perhatikan tabel 2 dibawah ini:

Pendidikan	Pekerjaan		Total
	1	2	
1	a	b	$a + b$
2	c	d	$c + d$
3	e	f	$e + f$
Total	$a + c + e$	$b + d + f$	N

Tabel Kontingensi Chi Square

Tabel 2, terdiri dari 6 cell, yaitu cell a, b, c, d, e dan f.

PROSEDUR UJI CHI SQUARE.

1. Perumusan Hipotesis

$$H_0: \chi = 0, .$$

$$H_a: \chi \neq 0,$$

2. Menetapkan taraf nyata

Menentukan nilai kesalahan = α . Setelah α ditetapkan selanjutnya menghitung nilai dari χ^2_{tabel} dengan menggunakan tabel χ^2 yang ada di Lampiran, dengan

$$db = (r - 1)(c - 1)$$

dimana:

db = derajat kebebasan

r = jumlah baris

c = jumlah kolom

3. Menghitung nilai χ^2_{hitung}

Rumus yang digunakan untuk menghitung nilai dari χ^2_{hitung} adalah sebagai berikut.

$$\chi^2_{hitung} = \sum \frac{(f_o - f_e)^2}{f_e}$$

dimana:

f_o = nilai frekuensi observasi

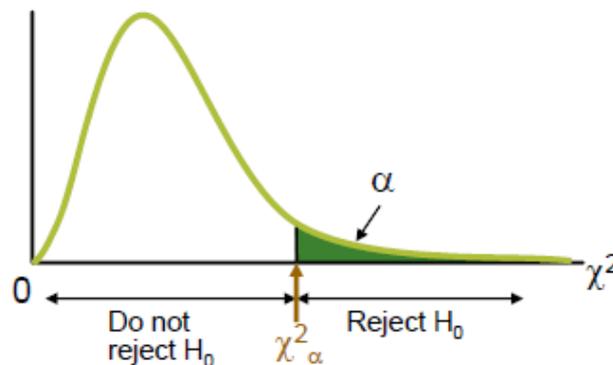
f_e = nilai frekuensi harapan

4. Penarikan keputusan dan kesimpulan

Kriteria keputusan dari pengujian Chi square adalah sebagai berikut.

Jika $\chi^2_{hitung} \leq \chi^2_{tabel}$, maka H_0 diterima.

Jika $\chi^2_{hitung} > \chi^2_{tabel}$, maka H_0 ditolak.



Contoh 1.

Sampel 500 mahasiswa yang berpartisipasi dalam studi yang dirancang untuk mengevaluasi tingkat pengetahuan mahasiswa dari suatu perguruan tinggi tentang penyakit umum tertentu. Tabel berikut menunjukkan mahasiswa yang telah diklasifikasikan berdasarkan bidang studi utama dan tingkat pengetahuan kelompok penyakit.

Jurusan	Pengetahuan tentang kelompok penyakit		
	Baik	Lemah	Total
Kesehatan	31	91	122
Lainnya	19	359	378
Total	50	450	500

Apakah data ini menunjukkan bahwa ada hubungan antara pengetahuan tentang kelompok penyakit dan jurusan, dari sampel 500 mahasiswa tersebut? (Gunakan $\alpha = 0,05$).

Penyelesaian:

Diketahui

$n = 500$

$\alpha = 5\%$

Tabel kontingensi berukuran 2×2 . Sehingga nilai $r = 1$ dan $c = 1$.

Prosedur pengujian hipotesis

1) Perumusan Hipotesis

$H_0: \chi = 0$, Tidak terdapat hubungan yang signifikan antara pengetahuan kelompok penyakit dengan jurusan.

$H_a: \chi \neq 0$, Terdapat hubungan yang signifikan antara pengetahuan kelompok penyakit dengan jurusan.

2) Menetapkan taraf nyata

nilai dari taraf nyata $\alpha = 0,05$. derajat kebebasannya adalah

$$db = (r - 1)(c - 1) = (2 - 1)(2 - 1) = 1$$

Selanjutnya, dilihat nilai dari χ_{tabel}^2 dengan nilai $\alpha = 0,05$ dan $db = 1$, sehingga diperoleh

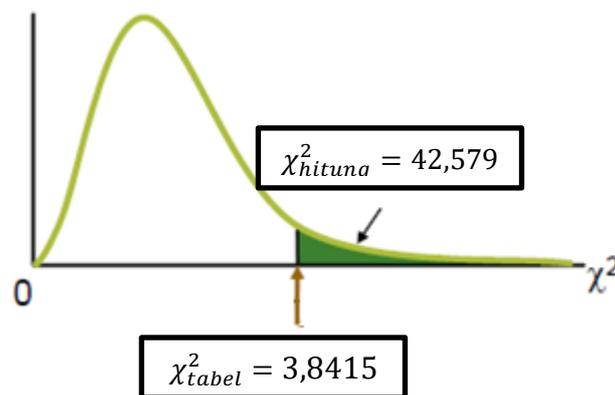
$$\chi_{tabel}^2 = 3,8415$$

3) Menghitung nilai χ_{hitung}^2

Ukuran tabel kontingensinya adalah 2×2 . sehingga rumus yang digunakan untuk menghitung nilai dari χ_{hitung}^2 adalah rumus Korensi Yates.

$$\begin{aligned}\chi_{hitung}^2 &= \frac{N(ad - bc)^2}{(a + b)(c + d)(a + c)(b + d)} \\ \chi_{hitung}^2 &= \frac{500(31 \times 359 - 91 \times 19)^2}{(31 + 91)(19 + 359)(31 + 19)(91 + 359)} \\ \chi_{hitung}^2 &= \frac{500 \times (9400)^2}{122 \times 378 \times 50 \times 450} = 42,579\end{aligned}$$

4) Penarikan keputusan dan kesimpulan



Dari poin 2) diketahui nilai dari $\chi_{tabel}^2 = 3,8415$ dan dari poin 3) diketahui nilai $\chi_{hitung}^2 = 42,579$. Berdasarkan kriteria keputusan, jika nilai dari $\chi_{hitung}^2 > \chi_{tabel}^2$ maka H_0 ditolak. Artinya Terdapat hubungan yang signifikan antara pengetahuan kelompok penyakit dengan jurusan dengan tingkat keyakinan 95%.

Contoh 2.

Dari 100 karyawan di PT XYZ, 60 adalah pria dan 40 adalah wanita. Dari 60 orang pria ternyata 10 menyukai pakaian warna merah muda, 20 menyukai warna putih dan 30 menyukai warna biru. Sedangkan dari 40 orang karyawan wanita, 20 menyukai warna merah muda, 10 menyukai warna putih dan 10 menyukai warna biru. Dengan tingkat kepercayaan 95% apakah antara pemilihan warna dengan jenis kelamin berbeda secara nyata?

Penyelesaian.

Sebelum melakukan pengujian hipotesis, terlebih dahulu dibuat tabel kontingensi berdasarkan informasi yang ada di contoh 4.

	Pria	Wanita	Jumlah
Merah muda	10	20	30
Putih	20	10	30
Biru	30	10	40
Jumlah	60	40	100

Tabel kontingensi

Berdasarkan Tabel 8, dapat dilihat bahwa ukuran dari tabel kontingensi tersebut adalah 3×2 . Artinya terdapat 3 baris dan 2 kolom. Selanjutnya, kita akan melakukan prosedur pengujian hipotesis.

1) Perumusan hipotesis

$H_0: \chi = 0$, Antara pemilihan warna dengan jenis kelamin tidak berbeda secara nyata

$H_a: \chi \neq 0$, Antara pemilihan warna dengan jenis kelamin berbeda secara nyata

2) Menetapkan taraf nyata

Dari soal diketahui bahwa nilai $\alpha = 5\%$

Selanjutnya dihitung nilai dari χ^2_{tabel} dengan $db = (r - 1)(c - 1) = (3 - 1)(2 - 1) = 2$ dan menggunakan tabel χ^2 pada Lampiran, sehingga diperoleh

$$\chi^2_{tabel} = 5,991$$

3) Menghitung nilai χ^2_{hitung}

Pertama, kita akan menghitung nilai dari f_o dari setiap cellnya, yaitu

Cell	f_o
a	10
b	20
c	20
d	10
e	30
f	10

dengan

a = Pria yang menyukai warna merah muda

b = Wanita yang menyukai warna merah muda

c = Pria yang menyukai warna putih

d = Wanita yang menyukai warna putih

e = Pria yang menyukai warna biru

f = Wanita yang menyukai warna biru

Selanjutnya dihitung nilai dari f_e dengan rumus:

$$F_e = \frac{\text{Jumlah baris}}{\text{Jumlah total}} \times \text{Jumlah kolom}$$

sehingga diperoleh:

$$F_{e \text{ cell } a} = \frac{30}{100} \times 60 = 18$$

$$f_{e \text{ cell } b} = \frac{30}{100} \times 40 = 12$$

$$f_{e \text{ cell } c} = \frac{30}{100} \times 60 = 18$$

$$f_{e \text{ cell } d} = \frac{30}{100} \times 40 = 12$$

$$f_{e \text{ cell } e} = \frac{40}{100} \times 60 = 24$$

$$f_{e \text{ cell } f} = \frac{40}{100} \times 40 = 16$$

Selanjutnya nilai dari f_o dan f_e dari setiap cell disajikan dalam satu tabel

Cell	f_o	f_e
a	10	18
b	20	12
c	20	18
d	10	12
e	30	24
f	10	16

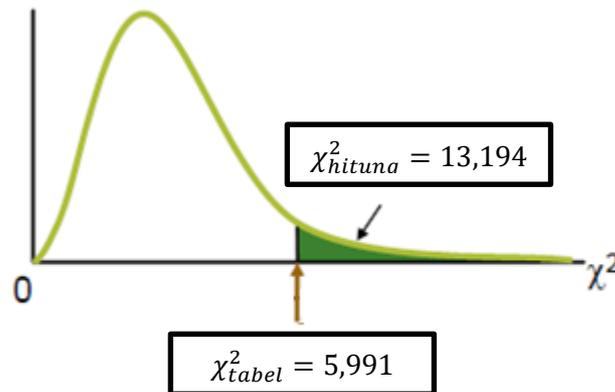
Selanjutnya dihitung nilai dari $f_e, f_o - f_e, (f_o - f_e)^2$, dan $\frac{(f_o - f_e)^2}{f_e}$ serta dinyatakan dalam satu tabel sebagai berikut.

Cell	f_o	f_e	$f_o - f_e$	$(f_o - f_e)^2$	$\frac{(f_o - f_e)^2}{f_e}$
a	10	18	-8	64	3,556
b	20	12	8	64	5,333
c	20	18	2	4	0,222
d	10	12	-2	4	0,333
e	30	24	6	36	1,5
f	10	16	-6	36	2,25

Kemudian dihitung nilai dari χ^2_{hitung} dengan menggunakan rumus berikut.

$$\chi^2_{hitung} = \sum \frac{(f_o - f_e)^2}{f_e} = 3,556 + 5,333 + 0,222 + 0,333 + 1,5 + 2,25 = 13,194$$

4) Penarikan keputusan dan kesimpulan.



Dari poin 2) diketahui nilai dari $\chi^2_{tabel} = 5,991$ dan dari poin 3) diketahui nilai $\chi^2_{hitung} = 13,194$. Berdasarkan kriteria keputusan, jika nilai dari $\chi^2_{hitung} > \chi^2_{tabel}$ maka H_0 ditolak, artinya antara pemilihan warna dengan jenis kelamin berbeda secara nyata.

CONTOH 3.

Diketahui data 60 responden. Dari ke 60 responden tersebut, ada responden yang mempunyai pekerjaan 1 dan 2 serta pendidikan 1 dan 2. Misalkan pekerjaan 1 adalah pegawai negeri dan pekerjaan 2 adalah pegawai swasta. Serta pendidikan 1 adalah lulusan SMA, pendidikan 2 adalah lulusan sarjana, dan pekerjaan 3 adalah lulusan magister. Data dari ke 60 responden tersebut, disajikan dalam Tabel berikut.

Responden	Pendidikan	Pekerjaan	Responden	Pendidikan	Pekerjaan
1	1	1	31	2	2
2	2	2	32	2	1
3	1	2	33	2	1
4	2	2	34	1	1
5	1	2	35	2	2
6	3	2	36	1	1
7	2	2	37	3	2
8	1	2	38	2	2
9	2	2	39	2	1
10	1	2	40	3	2
11	1	2	41	1	1
12	3	1	42	3	2

Responden	Pendidikan	Pekerjaan	Responden	Pendidikan	Pekerjaan
13	3	1	43	1	1
14	2	1	44	2	2
15	1	2	45	1	1
16	3	2	46	3	1
17	2	2	47	3	2
18	2	2	48	2	1
19	1	1	49	3	2
20	2	2	50	2	1
21	3	1	51	2	1
22	1	1	52	2	2
23	3	2	53	3	2
24	1	2	54	1	1
25	3	1	55	2	2
26	2	2	56	2	2
27	1	2	57	1	1
28	1	2	58	3	1
29	2	2	59	2	1
30	1	1	60	3	1

Tabel Data responden

Dengan menggunakan $\alpha = 0,05$, ujilah apakah terdapat hubungan antara pendidikan dengan pekerjaan.

Penyelesaian:

Pertama yang dilakukan adalah membuat tabel kontingensi seperti disajikan dalam Tabel berikut.

Pendidikan	Pekerjaan		Total
	1	2	
1	11	9	20
2	8	16	24
3	7	9	16
Total	26	34	60

Tabel kontingensi

Dari Tabel tersebut, dapat diketahui bahwa tabel kontingensi yang diperoleh adalah tabel kontingensi 3 x 2. Selanjutnya, dilakukan prosedur pengujian hipotesis.

Prosedur pengujian hipotesis:

- 1) Perumusan hipotesis

$H_0: \chi = 0$, Tidak terdapat hubungan antara pendidikan dengan pekerjaan.

$H_a: \chi \neq 0$, Terdapat hubungan antara pendidikan dengan pekerjaan.

2) Menetapkan taraf nyata

Dari soal diketahui bahwa nilai $\alpha = 5\%$

Selanjutnya dihitung nilai dari χ^2_{tabel} dengan $db = (r - 1)(c - 1) = (3 - 1)(2 - 1) = 2$ dan menggunakan tabel χ^2 pada Lampiran, sehingga diperoleh

$$\chi^2_{tabel} = 5,991$$

3) Menghitung nilai χ^2_{hitung}

Pertama, kita akan menghitung nilai dari f_o dari setiap cellnya, yaitu

Cell	f_o
<i>a</i>	11
<i>b</i>	9
<i>c</i>	8
<i>d</i>	16
<i>e</i>	7
<i>f</i>	9

Tabel Nilai f_o dari setiap cell

Langkah berikutnya kita hitung nilai frekuensi harapan (f_e) dari setiap cellnya. Rumus menghitung frekuensi harapan adalah sebagai berikut:

$$F_e = \frac{\text{Jumlah baris}}{\text{Jumlah total}} \times \text{Jumlah kolom}$$

sehingga diperoleh:

$$F_e \text{ cell a} = (20/60) \times 26 = 8,667$$

$$F_e \text{ cell b} = (20/60) \times 34 = 11,333$$

$$F_e \text{ cell c} = (24/60) \times 26 = 10,400$$

$$F_e \text{ cell d} = (24/60) \times 34 = 13,600$$

$$F_e \text{ cell e} = (16/60) \times 26 = 6,933$$

$$F_e \text{ cell f} = (16/60) \times 34 = 9,067$$

Setelah diperoleh nilai dari f_o dan f_e selanjutnya disajikan dalam satu tabel sebagai berikut.

Cell	f_o	f_e
a	11	8,667
b	9	11,333
c	8	10,400
d	16	13,600
e	7	6,933
f	9	9,067

Tabel Nilai f_o dan f_e dari setiap cell

Untuk memudahkan perhitungan nilai Chi Square, semua data dimasukkan dalam tabel, dengan kolom-kolomnya adalah nilai dari *cell*, f_o , f_e , $f_o - f_e$, $(f_o - f_e)^2$, dan $\frac{(f_o - f_e)^2}{f_e}$. Sehingga diperoleh tabel sebagai berikut.

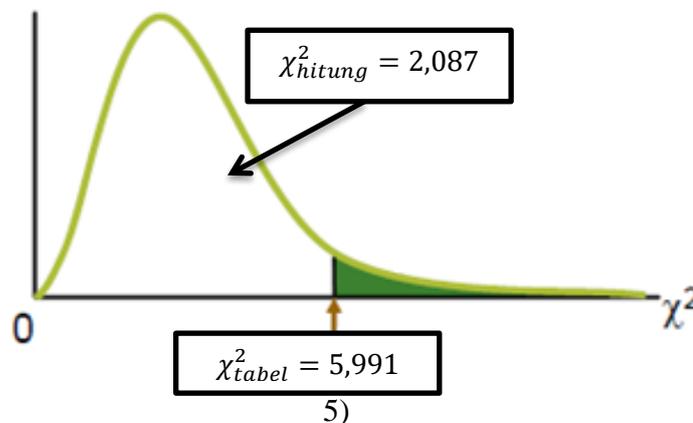
<i>Cell</i>	f_o	f_e	$f_o - f_e$	$(f_o - f_e)^2$	$\frac{(f_o - f_e)^2}{f_e}$
a	11	8,667	2,333	5,444	0,628
b	9	11,333	-2,333	5,444	0,480
c	8	10,400	-2,400	5,760	0,554
d	16	13,600	2,400	5,760	0,424
e	7	6,933	0,067	0,004	0,001
f	9	9,067	-0,067	0,004	0,000

Tabel 7. Tabel perhitungan Chi Square

Sehingga nilai dari Chi Square adalah sebagai berikut.

$$\chi^2_{hitung} = \sum \frac{(f_o - f_e)^2}{f_e} = 0,628 + 0,480 + 0,554 + 0,424 + 0,001 + 0,000 = 2,087$$

4) Penarikan keputusan dan kesimpulan.



Dari poin 2) diketahui nilai dari $\chi^2_{tabel} = 5,991$ dan dari poin 3) diketahui nilai $\chi^2_{hitung} = 2,087$. Berdasarkan kriteria keputusan, jika nilai dari $\chi^2_{hitung} > \chi^2_{tabel}$ maka H_0 diterima. Artinya Tidak terdapat hubungan antara pendidikan dengan pekerjaan dengan tingkat kepercayaan 95%.

Goodness of Fit Test (Uji kecocokan)

Uji kecocokan atau *goodness of fit test* menentukan apakah sebuah populasi mengikuti distribusi tertentu. Chi-Square Goodness of Fit dapat digunakan ketika bertemu dengan kondisi sebagai berikut:

- Metode sample yang digunakan adalah simple random sampling
- Variabel yang digunakan adalah kategorikal
- Nilai yang diharapkan pada sampel yang diobservasi minimal 5 dalam setiap level variabel

Prosedur pengujian Goodness of Fit

1. Perumusan Hipotesis

$H_0: \chi = 0$, Data berdistribusi tertentu.

$H_a: \chi \neq 0$, Data tidak berdistribusi tertentu.

2. Menetapkan taraf nyata

Menentukan nilai kesalahan = α . Setelah α ditetapkan selanjutnya menghitung nilai dari χ^2_{tabel} dengan menggunakan tabel χ^2 yang ada di Lampiran, dengan

$$db = r - 1$$

dimana:

db = derajat kebebasan

r = jumlah baris

3. Menghitung nilai χ^2_{hitung}

Rumus yang digunakan untuk menghitung nilai dari χ^2_{hitung} adalah sebagai berikut.

$$\chi^2_{hitung} = \sum \frac{(f_o - f_e)^2}{f_e}$$

dimana:

f_o = nilai frekuensi observasi

f_e = nilai frekuensi harapan

4. Penarikan keputusan

Kriteria keputusan dari pengujian Chi square adalah sebagai berikut.

Jika $\chi^2_{hitung} \leq \chi^2_{tabel}$, maka H_0 diterima.

Jika $\chi^2_{hitung} > \chi^2_{tabel}$, maka H_0 ditolak.

Contoh 4.

Berikut diberikan data hasil survey 1000 perokok terhadap 5 merek rokok yang mereka pilih.

Preferensi merk rokok	Jumlah konsumen
A	210
B	310
C	170
D	85
E	225
Jumlah	1000

- a. Ujilah apakah preferensi konsumen dalam memilih merek rokok sama pada $\alpha = 5\%$.
- b. Ujilah apakah yang memilih rokok merek A 20%, merek B 30%, merek C 15%, merek D 10% dan merek E 25% pada alpha 5%?

Penyelesaian.

Kasus a.

- 1) Perumusan Hipotesis

H_0 : Distribusi jumlah konsumen yang diamati untuk setiap merk rokok tidak berbeda secara signifikan dari jumlah konsumen yang sama di setiap merk rokok.

H_a : Distribusi jumlah konsumen yang diamati untuk setiap merk rokok berbeda secara signifikan dari jumlah konsumen yang sama di setiap merk rokok

atau dapat dituliskan sebagai berikut

$$H_0: P_A = P_B = P_C = P_D = P_E$$

$$H_a: P_A \neq P_B \neq P_C \neq P_D \neq P_E$$

- 2) Menetapkan taraf nyata

Dari soal diketahui bahwa nilai $\alpha = 5\%$

Selanjutnya dihitung nilai dari χ^2_{tabel} dengan $db = r - 1 = 5 - 1 = 4$ dan menggunakan tabel χ^2 pada Lampiran, sehingga diperoleh

$$\chi^2_{tabel} = 9,488$$

- 3) Menghitung nilai χ^2_{hitung}

Pertama, kita akan menghitung nilai dari f_o dari setiap cellnya, yaitu

Cell	f_o
A	210
B	310
C	170
D	85
E	225

Selanjutnya dihitung nilai dari f_e . Karena $P_A = P_B = P_C = P_D = P_E$ maka nilai f_e untuk setiap cellnya sama, yaitu nilai rata-rata dari jumlah konsumennya, sehingga diperoleh

$$F_e = \frac{1000}{5} = 200$$

Selanjutnya nilai dari f_o dan f_e dari setiap cell disajikan dalam satu tabel

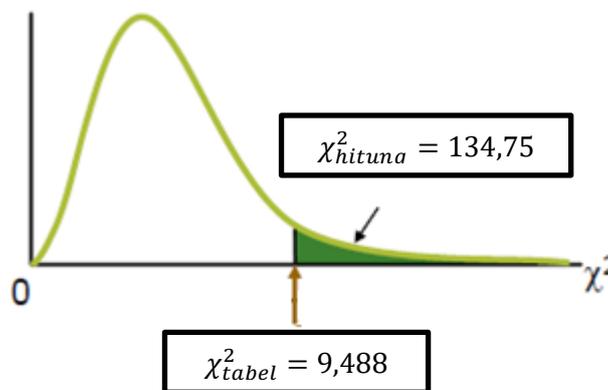
Cell	f_o	f_e
A	210	200
B	310	200
C	170	200
D	85	200
E	225	200

Selanjutnya dihitung nilai dari $f_e, f_o - f_e, (f_o - f_e)^2$, dan $\frac{(f_o - f_e)^2}{f_e}$ serta dinyatakan dalam satu tabel sebagai berikut.

	f_o	f_e	$f_o - f_e$	$(f_o - f_e)^2$	$\frac{(f_o - f_e)^2}{f_e}$
A	210	200	10	100	0,5
B	310	200	110	12100	60,5
C	170	200	-30	900	4,5
D	85	200	-115	13225	66,125
E	225	200	25	625	3,125

$$\chi^2_{hitung} = \sum \frac{(f_o - f_e)^2}{f_e} = 0,5 + 60,5 + 4,5 + 66,125 + 3,125 = 134,75$$

4) Penarikan keputusan dan kesimpulan



Dari poin 2) diketahui nilai dari $\chi^2_{tabel} = 9,488$ dan dari poin 3) diketahui nilai

$\chi^2_{hitung} = 134,75$. Berdasarkan kriteria keputusan, jika nilai dari $\chi^2_{hitung} > \chi^2_{tabel}$ maka H_0 ditolak. Artinya distribusi jumlah konsumen yang diamati untuk setiap merk rokok berbeda secara signifikan dari jumlah konsumen yang sama di setiap merk rokok dengan tingkat kepercayaan 95%.

Kasus b.

1) Perumusan Hipotesis

$$H_0: P_A = 0,2; P_B = 0,3; P_C = 0,15; P_D = 0,1; P_E = 0,25$$

$$H_0: P_A \neq 0,2; P_B \neq 0,3; P_C \neq 0,15; P_D \neq 0,1; P_E \neq 0,25$$

2) Menetapkan taraf nyata

Dari soal diketahui bahwa nilai $\alpha = 5\%$

Selanjutnya dihitung nilai dari χ^2_{tabel} dengan $db = r - 1 = 5 - 1 = 4$ dan menggunakan tabel χ^2 pada Lampiran, sehingga diperoleh

$$\chi^2_{tabel} = 9,488$$

3) Menghitung nilai χ^2_{hitung}

Pertama, kita akan menghitung nilai dari f_o dari setiap cellnya, yaitu

Cell	f_o
A	210
B	310
C	170
D	85
E	225

Selanjutnya dihitung nilai dari f_e . Karena $P_A \neq P_B \neq P_C \neq P_D \neq P_E$ maka nilai f_e untuk setiap cellnya tidak sama.

$$f_e \text{ cell A} = 0,2 \times 1000 = 200$$

$$f_e \text{ cell B} = 0,3 \times 1000 = 300$$

$$f_e \text{ cell C} = 0,15 \times 1000 = 150$$

$$f_e \text{ cell D} = 0,1 \times 1000 = 100$$

$$f_e \text{ cell E} = 0,25 \times 1000 = 250$$

Selanjutnya nilai dari f_o dan f_e dari setiap cell disajikan dalam satu tabel.

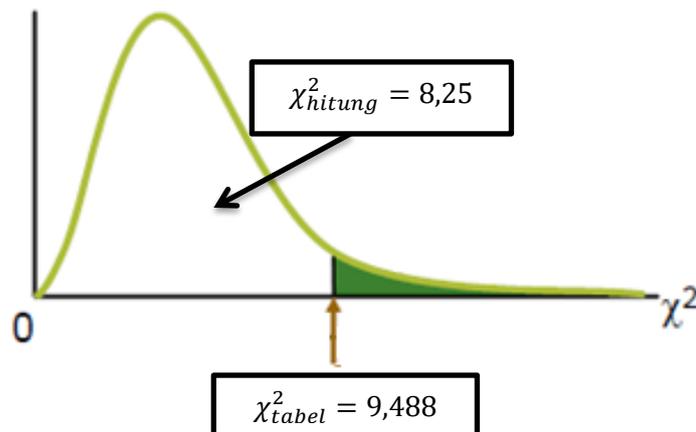
Cell	f_o	f_e
A	210	200
B	310	300
C	170	150
D	85	100
E	225	250

Selanjutnya dihitung nilai dari f_e , $f_o - f_e$, $(f_o - f_e)^2$, dan $\frac{(f_o - f_e)^2}{f_e}$ serta dinyatakan dalam satu tabel sebagai berikut.

	f_o	f_e	$f_o - f_e$	$(f_o - f_e)^2$	$\frac{(f_o - f_e)^2}{f_e}$
A	210	200	10	100	0,5
B	310	300	10	100	0,333
C	170	150	20	400	2,667
D	85	100	-15	225	2,25
E	225	250	-25	625	2,5

$$\chi^2_{hitung} = \sum \frac{(f_o - f_e)^2}{f_e} = 0,5 + 0,333 + 2,667 + 2,25 + 2,5 = 8,25$$

4) Penarikan keputusan



Dari poin 2 diketahui nilai dari $\chi^2_{tabel} = 9,488$ dan dari poin 3 diketahui nilai

$\chi^2_{hitung} = 8,25$. Berdasarkan kriteria keputusan, jika nilai dari $\chi^2_{hitung} < \chi^2_{tabel}$ maka H_0 diterima, artinya distribusi jumlah konsumen yang diamati rokok merk A 20%, rokok merk B 30%, rokok merk C 15%, rokok merk D 10%, dan rokok merk E 25%.

Contoh 5.

Suatu perusahaan membungkus kue dalam kotak yang berisi 20 potong. semua kue dari suatu sampel sebanyak 100 bungkus diteliti dan jumlah kue yang cacat pada masing-masing bungkus dicatat. data sampel adalah sebagai berikut.

Banyak kue cacat dalam satu kotak	Banyak kotak
0	5
1	36
2	28
3	19
4	9
5	3
6	0

Dengan taraf nyata 5%, ujlil apakah jumlah kue yang cacat dalam setiap bungkus mengikuti distribusi binomial?

Penyelesaian:

Diketahui:

banyak kue yang diteliti = $20 \times 100 = 2000$ potong.

banyak kue yang cacat = $(0 \times 5) + (1 \times 36) + (2 \times 28) + (3 \times 19) + (4 \times 9) + (5 \times 3) + (6 \times 0) = 200$.

Jadi proporsi dari kue yang cacat $p = \frac{200}{2000} = 0,1$.

Karena yang diuji adalah kue yang cacat, maka $n = 20$.

Prosedur pengujian hipotesis

- 1) Perumusan Hipotesis

H_0 : Data berdistribusi normal.

H_a : Data tidak berdistribusi normal.

- 2) $\alpha = 5\%$. dengan derajat kebebasannya $r - 1 = 5 - 1 = 4$, sehingga

$$\chi_{tabel}^2 = 9,488$$

- 3) Menghitung nilai χ_{hitung}^2

Dengan menggunakan tabel Binomial yang ada pada lampiran, diperoleh data sebagai berikut

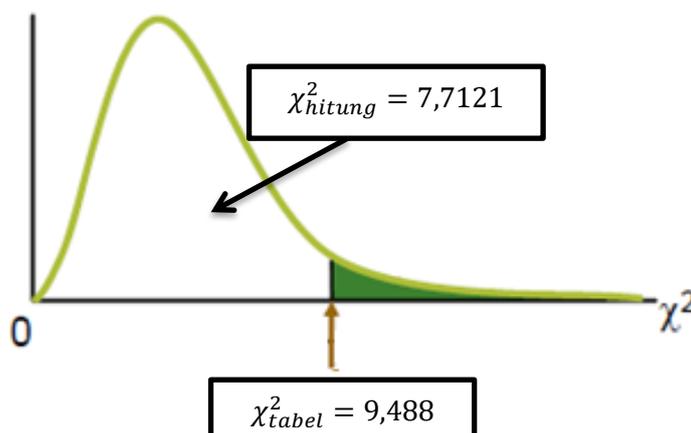
Banyak kue cacat	Probabilitas	f_e
0	0,1216	12,16
1	0,2701	27,01
2	0,2852	28,52
3	0,1901	19,01
4	0,0898	8,98
5	0,0318	3,18
6 atau lebih	0,0113	1,13

Selanjutnya, dihitung nilai dari χ_{hitung}^2 yang disajikan dalam tabel berikut.

Banyak cacat	f_o	f_e	$f_o - f_e$	$(f_o - f_e)^2$	$\frac{(f_o - f_e)^2}{f_e}$
0	5	12,16	-7,16	51,2656	4,215921
1	36	27,01	8,99	80,8201	2,992229
2	28	28,52	-0,52	0,2704	0,009481
3	19	19,01	-0,01	0,0001	5,26E-06
4	9	8,98	0,02	0,0004	4,45E-05
5 atau lebih	3	4,49	-1,49	2,2201	0,494454
					7,712135

Jadi, nilai dari $\chi_{hitung}^2 = 7,7121$

- 4) Penarikan keputusan



Dari poin 2) diketahui nilai dari $\chi^2_{tabel} = 9,488$ dan dari poin 3) diketahui nilai

$\chi^2_{hitung} = 7,7121$. Berdasarkan kriteria keputusan, jika nilai dari $\chi^2_{hitung} < \chi^2_{tabel}$ maka H_0 diterima. Artinya distribusi data kue cacat dalam masing-masing kotak mengikuti distribusi binomial.

Latihan Soal

1. Sebuah perguruan tinggi swasta yakin bahwa untuk mata kuliah statistika, persentasi mahasiswa yang akan mendapat nilai A adalah 10%, nilai B adalah 20%, nilai C adalah 40%, nilai D 20% dan yang mendapat nilai E sebesar 10%. Dari hasil ujian akhir sebanyak 50 mahasiswa didapat hasil sebagai berikut :

Nilai	Jumlah mahasiswa
A	10
B	10
C	10
D	14
E	6
Jumlah	50

- Dengan melihat hasil tersebut, benarkah pernyataan dosen perguruan tinggi swasta tersebut pada taraf nyata 5%?
2. Berikut ini dilakukan penelitian untuk mengetahui apakah terdapat hubungan antara status dengan pengeluaran per bulan. Dari 200 mahasiswa 100 mengaku punya pacar, dan sisanya jomblo. Dari 100 mahasiswa yang punya pacar , 83 mengaku pengeluaran besar, 5 pengeluaran sedang dan 12 pengeluaran tetap rendah. Sedangkan dari mahasiswa yang jomblo, 87 mengaku pengeluaran tinggi, 11 pengeluaran sedang dan 2 pengeluaran rendah. Ujilah pada alpha 10% apakah terdapat hubungan antara status dengan tingkat pengeluaran mahasiswa?
 3. Direktur pemasaran sebuah surat kabar harian ibukota sedang melakukan studi tentang hubungan antara lingkungan tempat tinggal pembaca dengan jenis artikel surat kabar yang dibaca pertama kali oleh pembaca. Data yang diperoleh adalah sebagai berikut:

Asal pembaca	News	Sport	Hiburan	Iklan
Kota	80	65	42	36
Desa	47	52	95	12

- Ujilah pada $\alpha = 0,05$ bahwa tidak terdapat hubungan antara asal pembaca dengan jenis artikel surat kabar.
4. Berikut ini adalah jumlah organisme tertentu yang ditemukan dalam 100 sampel air dari kolam.

Jumlah organisme per sampel	Frekuensi
0	15
1	30
2	25
3	20
4	5
5	4
6	1
7	0

Ujilah hipotesis nol, bahwa data ini diambil dari distribusi Poisson.

5. Seorang peneliti keselamatan lalu lintas melakukan pengamatan pada 500 kendaraan di lampu merah di suatu perkotaan. Berdasarkan hasil pengamatan diperoleh data jenis kendaraan dan perilaku di lampu merah seperti pada tabel di bawah. Dengan taraf nyata 0,05, dapatkah disimpulkan bahwa terdapat hubungan antara perilaku pengendara dan jenis kendaraan?

		Perilaku pengendara pada lampu merah			Total
		berhenti	Berjalan perlahan	Meluncur	
Tipe kendaraan	Car	183	107	60	350
	SUV	54	27	19	100
	Pick up	14	20	16	50
Total		251	154	95	500

6. Lima keping mata uang logam dilempar 100 kali. Pada setiap lemparan, banyaknya sisi kepala yang muncul dicatat seperti tabel berikut.

Banyak sisi kepala	Banyak lemparan
0	2
1	15
2	35
3	30
4	15
5	3

Dengan $\alpha = 5\%$, apakah munculnya sisi kepala mengikuti distribusi binomial?

7. Seorang administrator rumah sakit, ingin menguji hipotesis nol bahwa penerimaan pasien gawat darurat mengikuti distribusi Poisson dengan $\lambda = 3$. Misalkan selama periode 90 hari, jumlah penerimaan pasien Gawat darurat disajikan dalam tabel berikut.

Hari ke	Jumlah pasien						
1	2	24	5	47	4	70	3
2	3	25	3	48	2	71	5
3	4	26	2	49	2	72	4
4	5	27	4	50	3	73	1
5	3	28	4	51	4	74	1

6	2	29	3	52	2	75	6
7	3	30	5	53	3	76	3
8	0	31	1	54	1	77	3
9	1	32	3	55	2	78	5
10	0	33	2	56	3	79	2
11	1	34	4	57	2	80	1
12	0	35	2	58	5	81	7
13	6	36	5	59	2	82	7
14	4	37	0	60	7	83	1
15	4	38	6	61	8	84	5
16	4	39	4	62	3	85	1
17	3	40	4	63	1	86	4
18	4	41	5	64	3	87	4
19	3	42	1	65	1	88	9
20	3	43	3	66	0	89	2
21	3	44	1	67	3	90	3
22	4	45	2	68	2		
23	3	46	3	69	1		

8. Sebuah survei berminat menyelidiki determinasi orang dalam mencegah factor-faktor risiko penyakit jantung koroner. Setiap subjek dari sampel berukuran 200 orang diminta menyatakan sikapnya terhadap sebuah pertanyaan kuesioner sebagai berikut “ apakah anda yakin dapat menghindari makanan berkolesterol tinggi” dengan hasil 70 orang sangat yakin, 50 orang yakin, 45 orang ragu-ragu, dan 35 orang sangat ragu-ragu. Dapatkah kita menarik kesimpulan berdasarkan data tersebut, bahwa keempat sikap yang berbeda menyebar merata di dalam populasi asal sampel?