

Statistika Inferensia Lanjut

Teori & Praktik

Pertemuan 4,5

Uji Hipotesis Komparatif Dua Sampel Independen

Uji Hipotesis Komparatif 2 Sampel

- Pengujian hipotesis komparatif dua sampel independen dapat dilakukan dengan beberapa cara.
- Jika data berbentuk nominal maka dapat menggunakan Fisher Exact Probability Test dan Chi Square Two Samples.
- Jika data berbentuk ordinal dapat digunakan Median Test, Mann-Whitney U Test, Uji Dua-contoh Kolmogorov-Smirnov atau Wald-Wolfowitz.

Uji Fisher Exact Probability

- **Fungsi Pengujian :**

Untuk menguji perbedaan proporsi dua populasi yang hanya memiliki dua kategori berdasarkan proporsi dua sampel tidak berpasangan. Jumlah n untuk tiap kelompok sampel tidak harus sama.

- **Persyaratan Data :**

Dapat digunakan untuk data berskala nominal dengan dua kategori.

Prosedur Pengujian

1. Buat Tabel Silang , Baris adalah kelompok sampel, dan kolom - dan + untuk menunjukkan kategori yang bersifat *mutually exclusive*.

Sampel	Frekuensi pada		Jumlah sampel
	Obyek I	Objek II	
Sampel A	A	B	$A + B$
Sampel B	C	D	$C + D$
Jumlah	$A + C$	$B + D$	N

2. Masukkan frekuensi-frekuensi hasil pengamatan ke dalam baris dan kolom yang tepat.
3. Hitung jumlah frekuensi ke arah baris dan kolom, N adalah jumlah keseluruhan frekuensi pengamatan.

4. Dapat digunakan untuk $n \leq 40$ dan $E = \frac{1}{2}(A + D) < 5$
5. Untuk $n < 20$ dalam kondisi apapun uji Fisher dapat digunakan (baik sel dengan $E < 5$ ataupun tidak)
6. Jika digunakan uji 2 sisi maka nilai $P_{value} = 2 \times P_{hitung}$
7. Rumus Uji Fisher Exact Probability:

$$P = \frac{(A+B)!(C+D)!(A+C)!(B+D)!}{N!A!B!C!D!}$$

Untuk Tabel Yang Mempunyai Nilai Sel Nol

- Signifikansi Pada Uji Fisher Exact dibandingkan langsung dengan α
- Jika $p_{value} < \alpha$ maka H_0 ditolak
- Jika menggunakan uji 2 sisi dapat dipakai $p_{value} = 2 \times p_{tabel}$

Untuk Tabel Yang Tidak Mempunyai Nilai Sel Nol

- Perlu dibuat kemungkinan deviasi nilai ekstrimnya

$$\text{Nilai } P_{value} = \text{Nilai } P_{kasus} + \text{Nilai } P_{deviasi\ ekstrim\ terkecil}$$

- Jika $p_{value} < \alpha$ maka H_0 ditolak
- Jika menggunakan uji 2 sisi dapat dipakai $p_{value} = 2 \times p_{tabel}$

Tabel 6.2. Harga faktorial.

N	$N!$
0	1
1	1
2	2
3	6
4	24
5	120
6	720
7	5.040
8	40.320
9	362.880
10	36.288.800
11	39.916.800
12	479.001.600
13	6.227.020.800
14	87.178.291.200
15	1.307.674.368.000
16	20.922.789.888.000
17	355.687.428.096.000
18	6.402.373.705.728.000
19	121.645.100.408.832.000
20	2.432.902.008.176.640.000

Contoh 1

Sebuah penelitian dilakukan untuk mengetahui tingkat serangan hama keong mas dan tikus terhadap tanaman padi di Desa A dan B. Rata-rata jumlah keong mas yang menyerang tanaman padi di Desa A sebanyak 4 ekor per m^2 dan di Desa B sebanyak 3 ekor per m^2 . Rata-rata jumlah tikus yang menyerang tanaman padi di Desa A sebanyak 2 ekor per m^2 dan di Desa B sebanyak 5 ekor per m^2 .

Lakukan pengujian hipotesis untuk membuktikan apakah terdapat perbedaan nyata/signifikan tingkat serangan hama keong mas dan tikus di Desa A dan B. Taraf signifikansi (α) ditetapkan sebesar 5% (0,05).

Penyelesaian

□ Hipotesis

H_0 : Tingkat serangan hama keong mas dan tikus di Desa A tidak berbeda dengan desa B ($d_m = 0$)

H_1 : Tingkat serangan hama keong mas dan tikus di Desa A berbeda dengan Desa B ($d_m < 0$)

□ Tingkat Signifikansi: $\alpha = 0,05$

❑ Uji Statistik:

	Hama Tikus	Hama Keong Mas	Jumlah
Desa A	2	4	6
Desa B	5	3	8
Jumlah	7	7	

Karena dalam sel tersebut tidak ada yang bernilai 0 maka perlu dilakukan deviasi nilai ekstrim

	Hama Tikus	Hama Keong Mas	Jumlah
Desa A	2	4	6
Desa B	5	3	8
Jumlah	7	7	



$$P_0 = \frac{(A+B)!(C+D)!(A+C)!(B+D)!}{N!A!B!C!D!}$$

$$= \frac{6!8!7!7!}{14!4!2!3!5!} = 0,2448$$

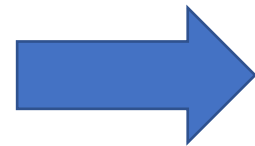
	Hama Tikus	Hama Keong Mas	Jumlah
Desa A	0	6	6
Desa B	7	1	8
Jumlah	7	7	



$$P_1 = \frac{(A+B)!(C+D)!(A+C)!(B+D)!}{N!A!B!C!D!}$$

$$= \frac{6!8!7!7!}{14!0!6!7!1!} = \frac{1}{429} = 0,0023$$

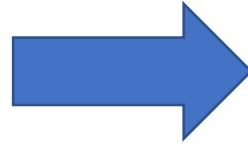
	Hama Tikus	Hama Keong Mas	Jumlah
Desa A	1	5	6
Desa B	6	2	8
Jumlah	7	7	



$$P_2 = \frac{(A+B)!(C+D)!(A+C)!(B+D)!}{N!A!B!C!D!}$$

$$= \frac{6!8!7!7!}{14!1!5!6!2!} = 0,0489$$

	Hama Tikus	Hama Keong Mas	Jumlah
Desa A	6	0	6
Desa B	1	7	8
Jumlah	7	7	



$$P_3 = \frac{(A+B)!(C+D)!(A+C)!(B+D)!}{N!A!B!C!D!}$$

$$= \frac{6!8!7!7!}{14!6!0!1!7!} = \frac{1}{429} = 0,0023$$

Dengan demikian, nilai $P_{value} = P_0 + \min\{P_1, P_2, P_3\} = 0,2448 + 0,0023 = 0,2471$.

Karena menggunakan uji dua sisi maka $p - value = 2 \times P_{value} = 2 \times 0,2471 = 0,4942$

❑ **Kriteria Penolakan** : H_0 ditolak jika $p - value < \alpha$

Karena $p - value = 0,4942 > \alpha = 0,05$ sehingga H_0 tidak ditolak. .

❑ **Kesimpulan** :

Jadi, tidak ada perbedaan tingkat serangan hama keong mas dan tikus di desa A dan B.

```

> Hama<- matrix(c(2, 5, 4, 3), nrow=2, dimnames=list(c("Desa A","Desa B"),c("Hama Tikus", "Hama Keong Mas")))
> Hama
      Hama Tikus Hama Keong Mas
Desa A          2             4
Desa B          5             3
> fisher.test(Hama, alternative="two.sided")

      Fisher's Exact Test for Count Data

data:  Hama
p-value = 0.5921
alternative hypothesis: true odds ratio is not equal to 1
95 percent confidence interval:
 0.01809307 4.13427934
sample estimates:
odds ratio
 0.3285541

```


Contoh 2

Sebuah studi kasus kontrol ingin melihat pengaruh merokok malam dengan kejadian kanker paru, hasil yang diperoleh tersaji pada tabel silang berikut :

Merokok Malam	Kanker Paru		Jumlah
	Ya	Tidak	
Ya	3	0	3
Tidak	1	3	4
Jumlah	4	3	7

Apakah ada Perbedaan antara kejadian kanker paru pada kebiasaan merokok malam dengan dengan kejadian kanker paru pada perokok pada $\alpha = 5\%$?

Penyelesaian

□ Hipotesis

H_0 : Tidak ada perbedaan antara kebiasaan merokok malam dengan kejadian kanker paru pada perokok ($d_m = 0$)

H_1 : ada perbedaan antara kebiasaan merokok malam dengan kejadian kanker paru pada perokok ($d_m > 0$)

□ Tingkat Signifikansi: $\alpha = 0,05$

❑ Uji Statistik:

	Kanker Paru-Paru	Tidak Kanker Paru2	Jumlah
Merokok Malam	3	0	3
Tidak Merokok Malam	1	3	4
Jumlah	4	3	7

$$P = \frac{(A+B)!(C+D)!(A+C)!(B+D)!}{N!A!B!C!D!} = \frac{3!4!4!3!}{7!3!0!1!3!} = 0,1143$$

Karena menggunakan uji 2 sisi maka nilai $P_{value} = 2 \times 0,1143 = 0,2286$

❑ Kriteria Penolakan H_0 :

Karena $P_{value} = 0,2286 > 0,05 = \alpha$ maka H_0 tidak ditolak.

❑ Kesimpulan:

Jadi, Tidak ada perbedaan antara kebiasaan merokok malam dengan kejadian kanker paru pada perokok

```

> kanker<- matrix(c(3, 1, 0, 3), nrow=2, dimnames=list(c("Merokok","Tidak Merokok"),c("Kanker Paru", "Ti
dak Kanker Paru")))
> kanker
      Kanker Paru Tidak Kanker Paru
Merokok           3              0
Tidak Merokok     1              3
> fisher.test(kanker, alternative="greater")

      Fisher's Exact Test for Count Data

data:  kanker
p-value = 0.1143
alternative hypothesis: true odds ratio is greater than 1
95 percent confidence interval:
 0.560439      Inf
sample estimates:
odds ratio
      Inf

```

Latihan

Seorang mahasiswa Fakultas Pertanian ingin meneliti perbedaan latar belakang tingkat pendidikan (sarjana dan bukan sarjana) Kepala BUMN Pertanian dan Kepala Perusahaan Pertanian Swasta. *Dugaan* peneliti, BUMN lebih banyak dipimpin oleh sarjana pertanian dibandingkan dengan Perusahaan Swasta.

Berdasarkan sampel yang dipilih secara random diperoleh 7 BUMN. Dari 7 BUMN tersebut ada 6 buah yang dipimpin sarjana dan ada 1 buah yang dipimpin oleh bukan sarjana. Sedangkan Perusahaan Swasta yang terpilih secara random hanya ada 5 perusahaan, 1 dipimpin oleh sarjana dan 4 lagi dipimpin oleh bukan. Hasilnya diperlihatkan dalam Tabel 5.2.

Tabel 5.2 Frekuensi Tingkat Pendidikan Menurut Jabatan yang Didudukinya Saat Ini

<i>Tingkat Pendidikan</i>	<i>Bukan Sarjana</i>	<i>Sarjana</i>	Total
<i>Ka. BUMN</i>	1	6	7
<i>Ka. Per.Swasta</i>	4	1	5
Total	5	7	12

Uji χ^2 Dua Sampel Independen

- **Fungsi Pengujian :**

- Hampir sama dengan Uji Fisher, yaitu untuk menguji perbedaan proporsi dua buah populasi berdasarkan proporsi dua sampel yang tidak berpasangan.
- Kelebihan Uji χ^2 bisa dipakai untuk dua atau lebih kategori.
- Uji χ^2 sebaiknya digunakan jika $n > 40$.
- Untuk $20 < n < 40$ dengan frekuensi kategori-kategorinya ($O_{ij} \geq 5$) bisa digunakan Uji χ^2 , namun jika ada salah satu frekuensi < 5 Uji χ^2 tidak boleh digunakan.
- Untuk $n < 20$ pilihlah Uji Fisher.

- **Persyaratan Data :**

Dapat digunakan untuk data berskala nominal dengan dua atau lebih dari dua kategori.

Prosedur Pengujian

1. Buat Tabel Silang $(k \times r)$, k adalah kolom = 2 dan r adalah baris ≥ 2 . Kolom dipakai untuk dua pasangan sampel yang tidak berpasangan, sedangkan baris disediakan untuk berbagai kategori.
2. Masukkan frekuensi-frekuensi hasil pengamatan (O_{ij}) ke dalam Tabel.
3. Hitung dan masukan ke dalam Tabel, frekuensi-frekuensi yang diharapkan (E_{ij}) yang dihitung dengan cara mengalikan jumlah baris dan jumlah kolom pada posisi E_{ij} kemudian membaginya dengan total frekuensi (N) .

4. Hitung harga χ^2 memakai rumus

$$\chi^2 = \sum_{i=1}^r \sum_{j=1}^k \frac{(O_{ij} - E_{ij})^2}{E_{ij}}$$

5. Untuk $k = 2$ dan $r = 2$, hitung dengan rumus

$$\chi^2 = \frac{N \left(|AD - BC| - \frac{1}{2}N \right)^2}{(A + B)(C + D)(A + C)(B + D)}$$

6. Gunakan Tabel Chi-square. Tentukan probabilitas (p) yang dikaitkan dengan terjadinya suatu harga sebesar χ^2 pada $db = (r - 1)(k - 1)$. Harga-harga p tersebut dipakai untuk pengujian dua sisi, sedangkan untuk melakukan pengujian satu sisi harga $p = \frac{1}{2} \times p_{Tabel}$.

Contoh

Misalkan suatu penelitian untuk melihat adakah hubungan antara tingkat kelulusan / kegagalan dalam sebuah kelas matematika tertentu dengan jenis kelamin. sampel acak dari 100 siswa dan mengukur kedua jenis kelamin (laki-laki/wanita) dan status kelulusan (lulus/gagal) sebagai variabel kategorik.

Tabel 1. Data tingkat kelulusan kelas matematika berikut

Siswa	Laki-laki	Perempuan	TOTAL
Lulus	30	36	66
Tidak lulus	14	20	34
TOTAL	44	56	100

Penyelesaian

□ Hipotesis :

H_0 : Tidak terdapat perbedaan antara tingkat kelulusan / kegagalan dalam sebuah kelas matematika tertentu dengan jenis kelamin.

H_1 : Terdapat perbedaan antara tingkat kelulusan / kegagalan dalam sebuah kelas matematika tertentu dengan jenis kelamin.

□ Tingkat Signifikansi : $\alpha = 0,05$

❑ **Statistik Penguji** : Data yang ada dimasukkan dalam table kontingensi berikut :

	Laki-Laki	Perempuan	
Lulus	30	36	66
Tidak Lulus	14	20	34
	44	56	100

Karena table berukuran 2×2 maka

$$\chi^2 = \frac{N \left(|AD - BC| - \frac{1}{2}N \right)^2}{(A+B)(C+D)(A+C)(B+D)} = \frac{100 \left(|30 \cdot 20 - 36 \cdot 14| - \frac{1}{2} \cdot 100 \right)^2}{(30+36)(14+20)(30+14)(36+20)} = \frac{100(96-50)^2}{66 \cdot 34 \cdot 44 \cdot 56} = 0,03827$$

Dengan $db = (r - 1) \times (k - 1) = 1 \times 1 = 1$ sehingga berdasarkan table Chi-square diperoleh

$$\chi_{tabel}^2 = 3,841$$

❑ Kriteria Penolakan H_0 :

H_0 ditolak jika $\chi^2_{hitung} < \chi^2_{tabel}$. Karena $\chi^2_{hitung} = 0,03827 < 3,841 = \chi^2_{tabel}$ maka H_0 ditolak.

❑ Kesimpulan:

Jadi, Terdapat perbedaan antara tingkat kelulusan / kegagalan dalam sebuah kelas matematika tertentu dengan jenis kelamin

Contoh

Sekelompok mahasiswa dari Jurusan Sosek dan Produksi Fakultas Peternakan melakukan penelitian bersama untuk mengetahui sektor pekerjaan alumni yang berasal dari kedua jurusan tersebut. Diduga, alumni kedua jurusan yang bekerja di sektor pertanian, industri, dan jasa proporsinya berlainan. Sampel diambil secara random, dengan jumlah sampel alumni Jurusan Sosek 20 orang dan Jurusan Produksi 30 orang. Dari hasil penelitian didapatkan data sebagai berikut:

1. Dari jumlah 30 orang alumni Sosek, sebanyak 15 orang bekerja di sektor Pertanian, 10 orang di sektor industri, dan 5 orang di sektor Jasa.
2. Dari jumlah 60 orang alumni Produksi, sebanyak 15 orang bekerja di sektor Pertanian, 20 orang di sektor industri, dan 25 orang di sektor Jasa.

Penyelesaian

□ Hipotesis :

$$H_0 : p_1 = p_2 = p_3$$

$$H_1 : p_1 \neq p_2 \neq p_3$$

□ Tingkat Signifikansi : $\alpha = 0,05$

❑ **Statistik Penguji** : Data yang ada dimasukkan dalam table kontingensi berikut:

	Sosek	Produksi	
Pertanian	(10) 15	(20) 15	30
Industri	(10) 10	(20) 20	30
Jasa	(10) 5	(20) 25	30
	30	60	90

Nilai E_{ij} diletakkan dalam tanda kurung di dalam sel.

$$E_{11} = \frac{30 \times 30}{90} = 10; E_{12} = \frac{30 \times 60}{90} = 20; E_{21} = \frac{30 \times 30}{90} = 10;$$

$$E_{22} = \frac{30 \times 60}{90} = 20; E_{31} = \frac{30 \times 30}{90} = 10; E_{32} = \frac{30 \times 60}{90} = 20$$

Sehingga diperoleh,

$$\begin{aligned}\chi^2 &= \sum_{i=1}^3 \sum_{j=1}^2 \frac{(O_{ij} - E_{ij})^2}{E_{ij}} \\&= \frac{(O_{11} - E_{11})^2}{E_{11}} + \frac{(O_{12} - E_{12})^2}{E_{12}} + \frac{(O_{21} - E_{21})^2}{E_{21}} + \frac{(O_{22} - E_{22})^2}{E_{22}} + \frac{(O_{31} - E_{31})^2}{E_{31}} + \frac{(O_{32} - E_{32})^2}{E_{32}} \\&= \frac{(15 - 10)^2}{10} + \frac{(15 - 20)^2}{20} + \frac{(10 - 10)^2}{10} + \frac{(20 - 20)^2}{20} + \frac{(5 - 10)^2}{10} + \frac{(25 - 20)^2}{20} \\&= 2,5 + 1,25 + 2,5 + 1,25 = 7,5\end{aligned}$$

df	Pr	0.25	0.10	0.05	0.010	0.005	0.001
1		1.32330	2.70554	3.84146	6.63490	7.87944	10.82757
2		2.77259	4.60517	5.99146	9.21034	10.59663	13.81551
3		4.10834	6.25139	7.81473	11.34487	12.83816	16.26624
4		5.38527	7.77944	9.48773	13.27670	14.86026	18.46683
5		6.62568	9.23636	11.07050	15.08627	16.74960	20.51501
6		7.84080	10.64464	12.59159	16.81189	18.54758	22.45774
7		9.03715	12.01704	14.06714	18.47531	20.27774	24.32189
8		10.21885	13.36157	15.50731	20.09024	21.95495	26.12448
9		11.38875	14.68366	16.91898	21.66599	23.58935	27.87716
10		12.54886	15.98718	18.30704	23.20925	25.18818	29.58830
11		13.70069	17.27501	19.67514	24.72497	26.75685	31.26413
12		14.84540	18.54935	21.02607	26.21697	28.29952	32.90949
13		15.98391	19.81193	22.36203	27.68825	29.81947	34.52818
14		17.11693	21.06414	23.68479	29.14124	31.31935	36.12327
15		18.24509	22.30713	24.99579	30.57791	32.80132	37.69730
16		19.36886	23.54183	26.29623	31.99993	34.26719	39.25235

❑ Kriteria Penolakan :

H_0 ditolak jika $p - value < \alpha$ atau $\chi_{hit}^2 > \chi_{tabel}^2$

Besarnya derajat kebebasan

$$db = (r - 1) \times (k - 1) = 2 \times 1 = 2$$

Karena $\alpha = 0,05$ dan $db = v = 2$ maka menggunakan table di samping didapat $\chi_{tabel}^2 = 5,99$

Terlihat bahwa $\chi_{hit}^2 = 7,50 > \chi_{tabel}^2 = 5,99$

Jadi, H_0 ditolak

❑ Kesimpulan :

Jadi, terdapat perbedaan proporsi alumni yang bekerja di berbagai sektor dari ketiga jurusan tersebut.

```
> Alumni<- matrix(c(15, 10, 5, 15, 20, 25), nrow=3, dimnames=list(c("Pertanian","Industri", "Jasa"),c("Sosek","Produksi")))
```

```
> Alumni
```

	Sosek	Produksi
Pertanian	15	15
Industri	10	20
Jasa	5	25

```
> chisq.test(Alumni)
```

Pearson's Chi-squared test

data: Alumni

X-squared = 7.5, df = 2, p-value = 0.02352

Latihan

Suatu penelitian dilakukan untuk melihat hubungan antara tipe sekolah dengan jenis kelamin apakah ada kecenderungan sekolah swasta lebih banyak murid perempuan jika dibandingkan dengan sekolah negeri. Hasil penelitian dapat dilihat pada tabel berikut :

		Jenis Kelamin		Total
		Laki-laki	Perempuan	
Sekolah	Negeri	77	91	168
	Swasta	14	18	32
Total		91	109	200