Statistika Inferensia Lanjut
Teori & Praktik

Pertemuan 6,7

Uji Hipotesis komparatif k sampel dependen

Uji Cochran (Q-test)

- Digunakan untuk menguji hipotesis komparatif $k - sampel$ berpasangan jika datanya nominal & bernilai dikotomi.
- Uji ini merupakan pengembangan dari Uji Mac Nemar
- Rumus :

$$Q = \frac{(k-1)\left[k \sum_{j=1}^{k} G_j^2 - \left(\sum_{j=1}^{k} G_j\right)^2\right]}{k \sum_{i=1}^{N} L_i - \sum_{i=1}^{N} L_i^2}$$

Keterangan :

k : banyak sampel

G_j : jumlah jawaban sukses atas perlakuan (kolom) ke-j

L_i : jumlah jawaban sukses atas subyek (baris) ke-i

	Perlakuan						L_i	L_i^2
Subjek	1	2	$\cdots$	j	$\cdots$	k		
1	x_{11}	x_{12}	$\cdots$	x_{1j}	$\cdots$	x_{1k}	L_1	L_1^2
2	x_{21}	x_{22}	$\cdots$	x_{2j}	$\cdots$	x_{2k}	L_2	L_2^2
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
i	x_{i1}	x_{i2}	$\cdots$	x_{ij}	$\cdots$	x_{ik}	L_i	L_i^2
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
N	x_{N1}	x_{N2}	$\cdots$	x_{Nj}	$\cdots$	x_{Nk}	L_N	L_N^2
Jumlah	G_1	G_2	$\cdots$	G_j	$\cdots$	G_k		

- Karena distribusi sampling Q mendekati distribusi χ^2 maka untuk uji signifikansi Q perlu dibandingkan dengan nilai kritis χ^2.
- Kriteria penolakan : H_0 ditolak jika $Q_{hitung} \geq \chi^2_{tabel}$ dengan $df = k - 1$

Contoh

Dilakukan survey aktivitas warga masyarakat dalam rangka menekan penyebaran penyakit demam berdarah. Angka 0 menyatakan tidak melakukan aktifitas dan angka 1 menyatakan melakukan aktifitas. Data hasil survey seperti disajikan dalam table. Selidikilah apakah terdapat perbedaan banyaknya masyarakat yang melakukan setiap aktivitas dalam menekan penularan demam berdarah! Gunakan tingkat signifikansi $\alpha = 5\%$

Abatisasi	Menutup Penampungan Air	Menguras Penampungan Air
0	1	0
0	0	1
0	0	0
1	1	1
0	1	1
1	0	0
1	1	1
1	1	1
0	0	1
0	1	0
1	1	1

Pengujian Hipotesis

- Hipotesis :

H_0 : Tidak ada perbedaan banyaknya masyarakat yang melakukan setiap aktivitas dalam menekan penularan demam berdarah

H_1 : Ada perbedaan banyaknya masyarakat yang melakukan setiap aktivitas dalam menekan penularan demam berdarah

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

- Statistik Penguji :

	Abatisasi	**Menutup**	**Menguras**	L_i	L_i^2
1	0	1	0	1	1
2	0	0	1	1	1
3	0	0	0	0	0
4	1	1	1	3	9
5	0	1	1	2	4
6	1	0	0	1	1
7	1	1	1	3	9
8	1	1	1	3	9
9	0	0	1	1	1
10	0	1	0	1	1
11	1	1	1	3	9
Jumlah	G1=5	G2=7	G3=7	$\sum L_i = 19$	$\sum L_i^2 = 45$

$$Q = \frac{(k-1)\left[k\ \sum_{j=1}^{k} G_j^2 - \left(\sum_{j=1}^{k} G_j\right)^2\right]}{k\sum_{i=1}^{N} L_i - \sum_{i=1}^{N} L_i^2}$$

$$= \frac{2[3(25+49+49)-(5+7+7)^2]}{3\cdot 19-45}$$

$$= \frac{16}{12} = 1{,}33$$

Dengan $\chi^2_{tabel} = 4{,}605$.

- Kriteria Penolakan :

Karena $Q_{\text{hitung}} = 1{,}33 < Q_{\text{tabel}} = 4{,}605$ maka H_0 tidak ditolak

- Kesimpulan :

Tidak ada perbedaan banyaknya masyarakat yang melakukan setiap aktivitas dalam menekan penularan demam berdarah

```
> Abatisasi<- c(0,0,0,1,0,1,1,1,0,0,1)
> Menutup<- c(1,0,0,1,1,0,1,1,0,1,1)
> Menguras<- c(0,1,0,1,1,0,1,1,1,0,1)
> data<- data.frame(Abatisasi, Menutup,Menguras)
> data
   Abatisasi Menutup Menguras
1          0       1        0
2          0       0        1
3          0       0        0
4          1       1        1
5          0       1        1
6          1       0        0
7          1       1        1
8          1       1        1
9          0       0        1
10         0       1        0
11         1       1        1
> library(nonpar)
> cochrans.q(data)

 Cochran's Q Test

 H0: There is no difference in the effectiveness of treatments.
 HA: There is a difference in the effectiveness of treatments.

 Q = 1.33333333333333

 Degrees of Freedom = 2

 Significance Level = 0.05
 The p-value is  0.513417119032592
```

Latihan

Sebuah perusahaan konveksi sedang mempertimbangkan pembelian tiga buah mesin untuk mengerjakan beberapa pesenannya. Manajer perusahaan itu memutuskan untuk mengambil 8 orang pekerja sebagai sampel dan masing-masing dari ketiga mesin tersebut. Setiap pekerja akan memberi nilai pada setiap mesin. 0 jika kinerja mesin dirasakan tidak memuaskan 1 jika kinerja mesin dirasakan memuaskan.

Pekerja	Mesin A	Mesin B	Mesin C
1	0	0	1
2	0	1	0
3	1	1	0
4	1	1	1
5	1	0	0
6	0	0	1
7	0	1	1
8	0	0	0

Uji Friedman (Anava Ranking Two Way)

Fungsi Pengujian :

Menguji perbedaan ranking populasi berdasarkan ranking k (k > 2) sampel berpasangan.

Persyaratan Data :

Data berskala *ordinal*.

Prosedur Pengujian:

1. Masukan data skor hasil penelitian ke dalam Tabel Silang k x n, dimana k adalah kelompok sampel yang berpasangan dijadikan kolom, dan n adalah banyaknya kasus/sampel dijadikan baris.
2. Buat ranking ke arah baris dari skor tersebut, mulai dari ranking 1 untuk skor terendah dan seterusnya sampai ranking k. Jika ada angka kembar buat ranking rata-ratanya.
3. Jumlahkan ranking ke arah kolom, pada masing-masing kolom (Rj)
4. Cari harga χ_r dengan memakai rumus:

$$\chi_r = \frac{12}{nk\,(k+1)} \sum_{j=1}^{k} (R_j)^2 - 3n\,(k+1)$$

5. Jika $2 \leq n \leq 9$ dan $k = 3$ atau $2 \leq n \leq 4$ dan $k = 4$, gunakan Tabel N.
6. Untuk n dan k yang lebih besar dari yang disebut pada nomor 5, gunakan Tabel C.
7. Jika langkah ke-5 dan ke-6 memberikan harga $p \leq \alpha$, maka tolak H_o.

Contoh

Sebuah Perusahaan biscuit ingin meluncurkan empat rasa baru dalam produk biskuitnya. Keempat rasan tersebut terdiri dari rasa coklat, rasa strowberi, rasa keju,dan rasa kelapa. Perusahaan ingin mengetahui bagaimana tanggapan konsumen terhadap keempat rasa tersebut 10 orang diminta untuk mencicipi keempat rasa biskuit tersebut kemudian memberikan nilai untuk setiap rasa yangada.Nilai yang diberikan ditentukan antara 0-100.

No	Coklat	Stroberi	Keju	Kelapa
1	78	80	84	71
2	82	76	85	73
3	81	78	80	70
4	80	77	88	71
5	82	74	86	75
6	83	81	89	70
7	85	78	84	70
8	79	73	85	72
9	82	70	87	73
10	78	71	88	70

Pengujian Hipotesis

- Hipotesis :

H_0 : Keempat rasa biscuit mempunyai penilaian yang sama

H_1 : Ada perbedaan penilaian mengenai keempat rasa biscuit

- Tingkat Signifikansi : $\alpha = 0{,}05$

Statistik Penguji ; Tentukan ranking tiap barisnya dari 1 s.d. k

No	Coklat	Rank	Stroberi	Rank	Keju	Rank	Kelapa	Rank
1	78	2	80	3	84	4	71	1
2	82	3	76	2	85	4	73	1
3	81	4	78	2	80	3	70	1
4	80	3	77	2	88	4	71	1
5	82	3	74	1	86	4	75	2
6	83	3	81	2	89	4	70	1
7	85	4	78	2	84	3	70	1
8	79	3	73	2	85	4	72	1
9	82	3	70	1	87	4	73	2
10	78	3	71	2	88	4	70	1
Jumlah		R1=31		R2=19		R3=38		R4=12

$$F_R = \frac{12}{nk(k+1)} \sum_{j=1}^{k} R_j^2 - 3n(k+1)$$

$$= \frac{12}{10 \cdot 4 \cdot 5} [31^2 + 19^2 + 38^2 + 12^2] - 3 \cdot 10 \cdot 5$$

$$= 174{,}6 - 150 = 24{,}6$$

Kriteria Penolakan :

H_0 ditolak jika $F_R \geq \chi^2_{tabel}$. Karena $\chi^2_{tabel} = 7{,}815 < F_R = 24{,}6$ maka H_0 ditolak

Kesimpulan :

Ada perbedaan penilaian mengenai keempat rasa biscuit

```
> Coklat<- c(78,82,81,80,82,83,85,79,82,78)
> Stroberi<- c(80,76,78,77,74,81,78,73,70,71)
> Keju<- c(84,85,80,88,86,89,84,85,87,88)
> Kelapa<- c(71,73,70,71,75,70,70,72,73,70)
> data<- data.frame(Coklat, Stroberi,Keju,Kelapa)
> data
   Coklat Stroberi Keju Kelapa
1      78       80   84     71
2      82       76   85     73
3      81       78   80     70
4      80       77   88     71
5      82       74   86     75
6      83       81   89     70
7      85       78   84     70
8      79       73   85     72
9      82       70   87     73
10     78       71   88     70
> friedman.test(data.matrix(data))

        Friedman rank sum test

data:  data.matrix(data)
Friedman chi-squared = 24.6, df = 3, p-value = 1.872e-05
```

Latihan

Seorang mahasiswa Fakultas Peternakan, ingin melakukan penelitian mengenai penilaian para peternak terhadap kebijakan sektor peternakan pada tiga masa pemerintahan yang berbeda. Penilaian dilakukan dengan memakai skor antara 1-10.

Peternak yang diteliti hanya 10 orang yang dipilih secara random. Mereka diminta untuk melakukan penilaian terhadap kebijakan pemerintah yang menyangkut sektor peternakan pada masa pemerintahan Presiden H, A, dan M.

Peneliti *menduga*, penilaian peternak terhadap kebijakan pada ketiga masa pemerintahan Presiden H, A, dan M akan berlainan.

Tabel. Skor Terhadap Kebijakan Sektor Peternakan pada Masa Pemerintahan Presiden H, A, dan M

Nomor Responden	Pres. H	Pres. A	Pres. M
1	10	3	7
2	9	4	6
3	7	8	3
4	2	9	2
5	9	7	5
6	6	9	4
7	**7**	**7**	4
8	9	7	5
9	8	6	4
10	6	8	4

Uji Komparasi k Sampel Independen

- Uji Chi-Square k Sampel
- Uji Median Extension
- Uji Kruskal Wallis

Uji Chi-Square k Sampel

Fungsi Pengujian :

Menguji perbedaan proporsi populasi berdasarkan proporsi k sampel tidak berpasangan.

Persyaratan Data :

Data berskala *nominal*.

Prosedur Pengujian:

1. Buat Tabel Silang k x r, k untuk kelompok sampel yang tidak berpasangan dan r untuk kategori dari variabel.
2. Masukan data hasil pengamatan ke dalam sel Tabel Silang sesuai dengan kelompok dan kategori masing-masing.
3. Tentukan frekuensi harapan dari masing-masing sel dengan cara mengalikan total baris dengan total kolom, kemudian dibagi dengan *grand* totalnya.

4. Hitung χ^2 dengan rumus:

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

5. Gunakan Tabel C. Tentukan probabilitas (p) yang dikaitkan dengan harga χ^2 untuk harga db = (k-1) x (r-1). Jika $p \leq \alpha$, maka tolak H_o.

Contoh

BKKBN ingin meneliti apakah alat kontrasepsi yang disukai masyarakat berbeda berdasarkan kelompok umur peserta KB. Untuk itu dilakukan suatu survei pemakaian alat KB pada 3 kelompok umur peseta KB. Data yang diperoleh sebagai berikut:

Kelompok umur	Alat kontrasepsi yang digunakan			
	Pil KB	Suntik	IUD	Kontrasepsi mantap
20 th -30 th	55	60	65	12
30 th -40 th	33	60	81	35
> 40 th	23	36	45	80

Apakah dapat disimpulkan bahwa populasi-populasi peserta KB pada tiga kelompok umur tersebut homogen dalam hal alat kontrasepsi yang dipakai.

Pengujian Hipotesis

- Hipotesis :

H_0 : Jenis alat kontraasepsi yang digunakan oleh ketiga kelompok umur peserta KB bersifat homogen

H_1 : Paling tidak salah satu diantara ketiga kelompok umur peserta KB tersebut menggunakan alat kontrasepsi yang berbeda

- Tingkat Signifikansi : $\alpha = 0{,}05$

- Statistik Penguji :

a. Tabel Kontingensi :

Kelompok Umur	Alat Kontrasepsi Yang Digunakan				Jumlah
	Pil KB	Suntik	IUD	Kontrasepsi Mantap	
20 – 30 th	55 (36,43)	60 (51,2)	65 (62,68)	12 (41,68)	192
30 – 40 th	33 (39,66)	60 (55,73)	81 (68,24)	35 (45,37)	209
> 40 th	23 (34,91)	36 (49,07)	45 (60,08)	80 (39,95)	184
Jumlah	111	156	191	127	585

b. Nilai χ^2

$$\chi^2 = \sum\sum \frac{\left(O_{ij}-E_{ij}\right)^2}{E_{ij}} = \frac{(55-36{,}43)^2}{36{,}43} + \frac{(60-51{,}2)^2}{51{,}2} + \frac{(65-62{,}68)^2}{62{,}68} + \frac{(12-41{,}68)^2}{41{,}68}$$

$$+ \frac{(33-39{,}66)^2}{39{,}66} + \frac{(60-55{,}73)^2}{55{,}73} + \frac{(81-68{,}24)^2}{68{,}24} + \frac{(35-45{,}37)^2}{45{,}37} + \frac{(23-34{,}91)^2}{34{,}91}$$

$$+ \frac{(36-49{,}07)^2}{49{,}07} + \frac{(45-60{,}08)^2}{60{,}08} + \frac{(80-39{,}95)^2}{39{,}95} = 89.91$$

Karena $df = (r-1) \times (c-1) = 6$ menggunakan table χ^2 diperoleh $\chi^2_{tabel} = 12{,}59159$

- Kriteria penolakan : H_0 ditolak jika $\chi^2_{hitung} > \chi^2_{tabel}$

  Karena $\chi^2_{hitung} = 89{,}91 > 12.59 = \chi^2_{tabel}$ maka H_0 ditolak

- Kesimpulan :

Jadi, Paling tidak salah satu diantara ketiga kelompok umur peserta KB tersebut menggunakan alat kontrasepsi yang berbeda

Tabel χ^2

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

Running Rstudio

```
> Kontrasepsi<- matrix(c(55, 33, 23, 60,60,36, 65, 81, 45, 12, 35, 80), nrow=3, dimname
s=list(c("20-30", "30-40", "> 40"), c("Pil KB", "Suntik", "IUD", "Kontrasepsi Manta
p")))
> Kontrasepsi
      Pil KB Suntik IUD
20-30     55     60  65
30-40     33     60  81
> 40      23     36  45
      Kontrasepsi Mantap
20-30                 12
30-40                 35
> 40                  80
> chisq.test(Kontrasepsi)

        Pearson's Chi-squared test

data:  Kontrasepsi
X-squared = 89.894, df = 6, p-value
< 2.2e-16
```

Latihan

Misalkan sebuah mesin pencampur adonan kue menghasilkan perbandingan tepung, susu, telur, dan gula secara berturut-turut adalah 5:2:2:1. Seorang pembuat kue mengambil adonan yang dihasilkan oleh mesin tersebut 500kg, ternyata dalam adonan kue tersebut mengandung 275 kg tepung, 95 kg susu, 70 kg telur, dan 60 kg gula. Pada tingkat signifikansi 1%, ujilah hipotesis apakah proporsi adonan kue yang telah dihasilkan sesuai dengan proporsi adonan yang telah ditetapkan pada mesin pencampur kue.

Latihan

Penelitian dilakukan untuk mengetahui apakah terdapat perbedaan alasan dalam membeli pupuk antara petani lada, karet, dan kelapa sawit. Sebanyak 30 petani lada terpilih sebagai responden, 12 responden membeli pupuk berdasarkan alasan harga, 8 responden memiliki alasan unsur kimia yang dikandung pupuk, dan 10 responden karena bentuk pupuk (cair atau padat) (Tabel 8.1). Dari 25 petani karet yang terpilih sebagai responden terdapat 6 responden yang membeli pupuk berdasarkan harga, sedangkan 10 responden karena unsur kimia, dan 9 responden karena alasan bentuk pupuk. Selanjutnya dari 20 petani kelapa sawit yang menjadi responden terdapat 5 responden yang membeli pupuk karena harganya, 7 responden karena unsur kimianya, dan 8 responden karena bentuknya. Lakukan pengujian hipotesis untuk mengetahui apakah petani lada, karet, dan kelapa sawit memiliki alasan yang sama dalam pembelian pupuk.

Tabel 8.1. Jumlah petani lada, karet, dan kelapa sawit berdasarkan alasan dalam pembelian pupuk.

Alasan	Petani		
	Lada	Karet	Kelapa sawit
Harga	12	6	5
Unsur kimia	8	10	7
Bentuk	10	9	8
Jumlah	30	25	20

Uji Median Extension (Perluasan Uji Median)

Fungsi Pengujian :

Menguji perbedaan median populasi berdasarkan median k sampel tidak berpasangan.

Persyaratan Data :

Data berskala *ordinal*.

Prosedur Pengujian:

1. Tentukan median bersama skor dari seluruh sampel
2. Skor di atas median beri tanda + (plus) dan skor di bawah median tanda – (minus).
3. Masukan frekuensi (+) dan (-) ke dalam Tabel Silang k x 2, k adalah kelompok sampel.

4. Tentukan frekuensi harapan dari masing-masing sel dengan cara mengalikan total baris dengan total kolom, kemudian dibagi dengan *grand* totalnya.
5. Hitung χ^2 dengan rumus:

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

6. Gunakan Tabel C. Tentukan probabilitas (p) yang dikaitkan dengan harga χ^2 untuk harga db = (k-1. Jika $p \leq \alpha$, maka tolak H_o.

Contoh

Seorang peneliti pendidikan ingin mempelajari apakah ada hubungan antara tingkat pendidikan ibu dengan banyaknya kunjungan ke sekolah anaknya. Diambil sampel secara random sebanyak 10% dari 440 anak yang terdaftar di sekolah. Dari sampel tersebut didapat nama dari 44 ibu-ibu yang kemudian dijadikan sampel. Hipotesisnya adalah banyaknya ibu ke sekolah bervariasi menurut tingkat pendidikan yang ditamatkanya. Datanya adalah sebagai berikut: (gunakan $\alpha = 5\%$)

Jumlah kunjungan ke sekolah oleh ibu digolongkan menurut tingkat pendidikan

SD	SMP	SMA	P.T	SD	SMP	SMA	P.T
4	2	2	9	2	0	0	4
3	4	0	4	0	2	5	5
0	1	4	2	3	5	2	2
7	6	3	3	5	1	1	2
1	3	8	2	1	2	7	6
	1	6					
		5					
		1					

Pengujian Hipotesis

Hipotesis dari permasalahan diatas, dapat dirumuskan sebagai berikut :

H_0 : Tidak ada perbedaan banyaknya kunjungan diantara para Ibu dengan variasi tingkat pendidikannya

H_1 : Ada perbedaan banyaknya kunjungan diantara para ibu dengan variasi tingkat pendidikannya.

Tingkat Signifikansi : $\alpha = 0,05$

Statistik Penguji

- Ditentukan median gabungan dari data tersebut :

Jumlah Kunjungan	Frekuensi Kunjungan
0	5
1	7
2	10
3	5
4	4
5	6
6	3
7	2
8	1
9	1

Median Gabungan = $\frac{data\ ke\ 22 + data\ ke\ 23}{2} = 2{,}5$

Diperoleh Tabel Kontingensi sbb :

	Tingkat Pendidikan				Jumlah
	SD	SMP	SMA	PT	
Diatas Median	5 (5)	4 (5,5)	7 (6,5)	6 (5)	22
Dibawah Median	5 (5)	7 (5,5)	6 (6,5)	4 (5)	22
Jumlah	10	11	13	10	44

$$\chi^2 = \sum \frac{\left(O_{ij} - E_{ij}\right)^2}{E_{ij}} = \frac{(5-5)^2}{5} + \frac{(4-5,5)^2}{5,5} + \frac{(7-6,5)^2}{6,5} + \frac{(6-5)^2}{5}$$

$$+ \frac{(5-5)^2}{5} + \frac{(7-5,5)^2}{5,5} + \frac{(6-6,5)^2}{6,5} + \frac{(4-5)^2}{5}$$

$$= 0 + 0,41 + 0,038 + 0,2 + 0 + 0,41 + 0,038 + 0 = 1,296$$

Karena $dk = (r-1) \times (c-1) = 3$ maka diperoleh $\chi^2_{table} = 7,81473.$

- Kriteria Penolakan : H_0 ditolak jika $\chi^2_{hitung} > \chi^2_{tabel}$

  karena $\chi^2_{hitung} = 1,296 < \chi^2_{tabel} = 7,81473$ maka H_0tidak ditolak.

- Kesimpulan :

  Jadi, tidak ada perbedaan banyak kunjungan diantara para ibu dengan tingkat Pendidikan yang berbeda

```
>
> Kunjungan <- c(0,0,1,1,2,3,3,4,5,7,0,1,1,1,2,2,2,3,4,5,6,0,0,1,1,2,2,3,4,5,5,6,7,8,2,
2,2,2,3,4,5,5,6,9)
> Pendidikan<- c(rep("SD", 10), rep("SMP", 11), rep("SMA", 13), rep("PT", 10))
> data<- data.frame(Kunjungan, Pendidikan)
> data
   Kunjungan Pendidikan
1          0         SD
2          0         SD
3          1         SD
4          1         SD
5          2         SD
6          3         SD
7          3         SD
8          4         SD
9          5         SD
10         7         SD
11         0        SMP
12         1        SMP
13         1        SMP
14         1        SMP
15         2        SMP
16         2        SMP
17         2        SMP
18         3        SMP
19         4        SMP
20         5        SMP
21         6        SMP
22         0        SMA
23         0        SMA
24         1        SMA
25         1        SMA
26         2        SMA
27         2        SMA
28         3        SMA
29         4        SMA
30         5        SMA
31         5        SMA
32         6        SMA
33         7        SMA
```

```
33         7        SMA
34         8        SMA
35         2         PT
36         2         PT
37         2         PT
38         2         PT
39         3         PT
40         4         PT
41         5         PT
42         5         PT
43         6         PT
44         9         PT
> library(agricolae)
> Median.test(data$Kunjungan,data$Pendidikan)

The Median Test for data$Kunjungan ~ data$Pendidikan

Chi Square = 1.295105   DF = 3   P.Value 0.7302958
Median = 2.5

    Median  r Min Max Q25  Q75
PT     3.5 10   2   9   2 5.00
SD     2.5 10   0   7   1 3.75
SMA    3.0 13   0   8   1 5.00
SMP    2.0 11   0   6   1 3.50

Post Hoc Analysis

Groups according to probability of treatment differences and alpha level.

Treatments with the same letter are not significantly different.

    data$Kunjungan groups
PT             3.5      a
SMA            3.0      a
SD             2.5      a
SMP            2.0      a
```

Atau Menggunakan Rstudio

```
> Kunjungan<- matrix(c(5,5,4,7,7,6,6,4),nrow=2, dimnames=list(c("Di Atas Rata-Rata", "D
i Bawah Rata-Rata"), c("SD","SMP","SMA","PT")))
> Kunjungan
                   SD SMP SMA PT
Di Atas Rata-Rata   5   4   7  6
Di Bawah Rata-Rata  5   7   6  4
> chisq.test(Kunjungan)

        Pearson's Chi-squared test

data:  Kunjungan
X-squared = 1.2951, df = 3, p-value = 0.7303
```

Latihan

Suatu penelitian dilakukan untuk mengetahui adakah hubungan "golongan gaji pegawai" dengan "jumlah media cetak" yang dibaca. Dalam hal ini golongan gaji dikelompokkan menjadi 4 tingkat yaitu Golongan I,II,III,IV. Dalam penelitian tersebut digunakan sampel pegawai Golongan I 11 orang, Gol II 11 orang, Gol III 12 orang dan Gol IV 12 orang

Jumlah Media Cetak Yang Dibaca			
Golongan I	**Golongan II**	**Golongan III**	**Golongan IV**
0	1	2	5
1	2	3	3
2	2	4	4
1	2	5	6
4	6	3	8
1	1	2	5
1	3	3	6
1	4	3	4
2	2	3	3
2	3	2	3
1	2	1	4
		2	4
$n_1 = 11$	$n_2 = 11$	$n_3 = 12$	$n_4 = 12$

Anova One Way Kruskal-Wallis

Uji Kruskal-Wallis *(Kruskal-Wallis one-way analysis of variance by ranks)* adalah teknik statistika nonparametrik yang digunakan untuk menguji hipotesis awal bahwa beberapa contoh berasal dari populasi yang sama/identik. Jika hanya melibatkan dua contoh, uji Kruskal-Wallis ekuivalen dengan uji Mann-Whitney. Uji Kruskal-Wallis digunakan untuk rancangan acak lengkap.

Tabel : Rancangan untuk uji Kruskal-Wallis

Contoh/Perlakuan			
1	**2**	……	***k***
$X_{1.1}$	$X_{2.1}$	……	$X_{k.1}$
$X_{1.2}$	$X_{2.2}$	……	$X_{k.2}$
$\vdots$	$\vdots$	$\ddots$	$\vdots$
$\vdots$	$\vdots$	$\ddots$	$\vdots$
$X_{1.n1}$	$X_{2.n2}$	……	$X_{k.nk}$
R_1	R_2	……	R_k

Asumsi

a. Data terdiri dari contoh acak X_1, X_2, ..., X_n yang berasal dari populasi 1 dengan median M_x, dan contoh acak Y_1, Y_2, ..., Y_n dari populasi 2 dengan median M_y. Nilai M_x dan M_y tidak diketahui.
b. Kedua contoh saling bebas
c. Peubah acak bersifat kontinu
d. Skala pengukuran minimal ordinal
e. Fungsi sebaran dari kedua populasi hanya dipisahkan oleh lokasi parameter

Hipotesis

H_0 : $M_1 = M_2 = \cdots = M_k$ atau k populasi mempunyai fungsi sebaran yang identik
H_1 : Ada minimal satu $M_i \neq M_j$ dimana $i \neq j$ dan $i, j = 1, 2, ..., k$

Statistik Uji

Statistik uji Kruskal-Wallis dapat ditentukan melalui prosedur berikut :

1. Seperti halnya uji Mann-Whitney, gabungkan seluruh data contoh, sehingga akan ada sebanyak $n_1 + n_2 + \cdots + n_k = N$ pengamatan.
2. Peringkatkan setiap pengamatan dari yang terkecil hingga terbesar. Jika terdapat *ties* (nilai yang sama), beri peringkat tengah *(mid-rank)*.
3. Hitung jumlah peringkat untuk setiap contoh, nyatakan masing-masing sebagai R_i.
4. Statistik uji Kruskal-Wallis dapat diperoleh melalui rumus :

$$H = \frac{12}{N(N+1)} \sum_{i=1}^{k} \frac{1}{n_i} \left(R_i - \frac{n_i (N+1)}{2} \right)^2 \quad \text{atau} \quad H = \frac{12}{N(N+1)} \sum_{i=1}^{k} \frac{R_i^{\,2}}{n_i} - 3(N+1)$$

Dalam hal ini R_i adalah jumlah peringkat untuk contoh ke-i, n_i adalah jumlah pengamatan pada contoh ke-i, dan N adalah total pengamatan.

Jika ada *ties*, statistik uji perlu dikoreksi dengan faktor :

$1 - \frac{\Sigma T}{N^3 - N}$ dalam hal ini $T = t^3 - t$ dan t adalah banyaknya *ties*. Sehingga statistik uji Kruskal-Wallis terkoreksi menjadi :

$$H_C = \frac{H}{1 - \Sigma T / (N^3 - N)}$$

Kaidah Keputusan

a. Jika hanya melibatkan tiga contoh/perlakuan *(k=3)* dan setiap contoh terdiri dari lima atau kurang pengamatan, gunakan tabel Kruskal-Wallis (A.12). Tolak H_0 jika H atau $H_C > H_\alpha$.

b. Jika tabel A.12 tidak dapat digunakan, gunakan tabel Khi-Kuadrat (A.11). Tolak H_0 jika H atau $H_C > \chi^2_{\alpha,k-1}$.

Contoh

Torre *et al.* mencatat adanya perubahan *serotonin* (5-HT) (*platelet*) serebral dan ekstraserebral tikus sesudah pemberian LSD-25 dan *1-methyl-dlysergic acid butanclamide* (UML) secara *intraperitoneal*. Pengukuran yang sama mereka lakukan pada 11 kontrol. Hasil percobaan disajikan pada Tabel di bawah ini. Apakah data ini cukup memberikan bukti untuk menunjukan adanya perbedaan di antara ketiga perlakuan tersebut (α=5%)? Hitung pula nilai *p-value* (Daniel 1990).

Tabel *serotonin* otak (5-HT), nanogram per gram, pada tiga kelompok anak tikus

Kontrol	340	340	356	386	386	402	402	417	433	495	557
LSD 0.5 mg/kg	294	325	325	340	356	371	385	402			
UML 0.5 mg/kg	263	309	340	356	371	371	402	417			

Kontrol	Rangking	LSD 0,5 mg/Kg	Rangking	UML 0,5 mg/Kg	Rangking
340	7,5	294	2	263	1
340	7,5	325	4,5	309	3
356	11	325	4,5	340	7,5
386	17,5	340	7,5	356	11
386	17,5	356	11	371	14
402	20,5	371	14	371	14
402	20,5	385	16	402	20,5
417	23,5	402	20,5	417	23,5
433	25				
495	26				
557	27				
	$R_1 = 203{,}5$		$R_2 = 80$		$R_3 = 94{,}5$

Ties :

Untuk 340 ada sebanyak $t_1 = 4$

Untuk 356 ada sebanyak $t_2 = 3$

Untuk 386 ada sebanyak $t_3 = 2$

Untuk 402 ada sebanyak $t_4 = 4$

Untuk 417 ada sebanyak $t_5 = 2$

Untuk 325 ada sebanyak $t_6 = 2$

Untuk 371 ada sebanyak $t_7 = 3$

Dengan menggunakan rumus $H = \frac{12}{N(N+1)} \sum_{i=1}^{k} \frac{R_i^{\,2}}{n_i} - 3(N+1)$ diperoleh :

$$H = \frac{12}{27(27+1)}\left[\frac{203.5^2}{11} + \frac{80^2}{8} + \frac{94.5^2}{8}\right] - 3(27+1) = 6.18$$

Karena terdapat *ties*, maka dikoreksi dengan rumus $H_C = \frac{H}{1 - \Sigma T/\left(N^3 - N\right)}$ sehingga diperoleh :

$$H_C = \frac{6.18}{1 - 186/\left(27^3 - 27\right)} = 6.24$$

Catatan : *ties* ΣT = $(2^3\text{-}2)+(4^3\text{-}4)+(3^3\text{-}3)+(3^3\text{-}3)+(2^3\text{-}2)+(4^3\text{-}4)+(2^3\text{-}2)=186$

Ukuran contoh lebih dari 5 pengamatan sehingga harus digunakan tabel Khi-Kuadrat. Nilai kritis khi-kuadrat untuk derajat bebas $k-1=3-1=2$ pada taraf nyata 5% adalah 5.991. Sehingga dengan Hc = 6.23 kita dapat menolak H_0 pada taraf nyata 5%, dan simpulkan bahwa ada minimal satu perlakuan yang memberikan pengaruh yang berbeda terhadap *serotonin* otak (5-HT) anak tikus. Pada kasus ini, $0.025 < p\text{-value} < 0.05$.

```
> Skor<- c(340,340,356, 386, 386, 402, 402, 417, 433, 495, 557, 294, 325, 325, 340, 35
6, 371, 385, 402, 263, 309, 340, 356, 371, 371, 402, 417 )
> Perlakuan<- c(rep("Kontrol", 11), rep("LSD", 8), rep("UML", 8))
> Serotonin<- data.frame(Skor, Perlakuan)
> Serotonin
   Skor Perlakuan
1   340   Kontrol
2   340   Kontrol
3   356   Kontrol
4   386   Kontrol
5   386   Kontrol
6   402   Kontrol
7   402   Kontrol
8   417   Kontrol
9   433   Kontrol
10  495   Kontrol
11  557   Kontrol
12  294       LSD
13  325       LSD
14  325       LSD
15  340       LSD
16  356       LSD
17  371       LSD
18  385       LSD
19  402       LSD
20  263       UML
21  309       UML
22  340       UML
23  356       UML
24  371       UML
25  371       UML
26  402       UML
27  417       UML
> kruskal.test(Skor~Perlakuan, data= Serotonin)

        Kruskal-Wallis rank sum test

data:  Skor by Perlakuan
Kruskal-Wallis chi-squared = 6.2341, df = 2, p-value = 0.04429
```

Latihan

Suatu penelitian dilakukan untuk mengetahui adanya perbedaan prestasi kerja pekerja yang rumahnya jauh atau dekat dengan kantor. Misalkan jarak rumah dikategorikan menjadi 3 yaitu: **I** (*untuk jarak s.d 5 km*), **II** (*>5 s.d 10 km*) dan **III** (*>10 km*). Penelitian dilakukan pada tiga kelompok pekerja berdasarkan jarak rumah dari kantornya dan sampel diambil secara acak. Data pengamatan ada pada tabel di bawah ini. Gunakan taraf nyata pengujian 5%.

Prestasi Kerja		
I	**II**	**III**
78	82	69
92	89	79
68	71	65
56	57	60
77	62	72
82	75	74
81	64	83
62	77	56
91	84	59
53	56	90
85	88	
	69	

Latihan

Apakah terdapat perbedaan kemampuan 30 peserta penyuluhan yang mengikuti penyuluhan dengan waktu berbeda, jika nilai kemampuan peserta tercantum pada Tabel 8.8 di bawah ini. Lakukan Uji Kruskal-Wallis dengan $\alpha = 0,05$. Bagaimanakah rumusan hipotesis nol dan hipotesis alternatifnya? Asumsi apakah yang dianggap berlaku untuk penelitian ini?

Tabel 8.8. Nilai kemampuan peserta penyuluhan.

1 minggu	2 minggu	3 minggu
60	70	68
65	60	75
71	73	76
76	77	77
80	82	80
83	80	84
85	82	63
74	85	65
75	65	69
68	67	72