

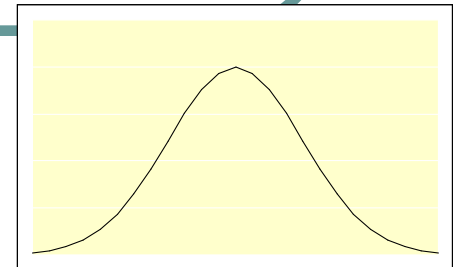
Statistika Nonparametrik

(Nonparametric Statistics)

Beberapa uji median dan ukuran asosiasi
(Some median tests and measures of association)

Prof. Ir. Sigit Nugroho, M.Sc., Ph.D.

Universitas Bengkulu



Skala Pengukuran Data

- **Nominal** : angka hanya menunjukkan kategori
 - Jenis Kelamin, Status Pernikahan, Kepegawaian
- **Ordinal** : angka selain menunjukkan kategori, tetapi juga mengandung peringkat atau urutan
 - Urutan juara, data diperingkatkan
- **Interval / Selang** : selain sifat yang dimiliki Ordinal, bisa diukur beda / jarak nya
 - Skala pengukur temperatur (Celcius, Reamur, Fahrenheit) dan pengukur gempa (Richter)
- **Rasio / Nisbah** : semua sifat interval plus bisa dibandingkan (rasio)
 - Waktu datangnya nasabah, rata-rata tabungan, luas bangunan, total produksi hasil pertanian

Metode Statistika Nonparametrik

Suatu Metode Statistika dikatakan nonparametrik jika memenuhi sedikitnya salah satu dari kriteria di bawah ini

- Metode menggunakan skala pengukuran **nominal**
- Metode menggunakan skala pengukuran **ordinal**
- Metode menggunakan skala pengukuran **interval** atau **rasio**, dimana fungsi distribusi variabel acak yang menghasilkan data tidak diketahui (*distribution-free*) atau diketahui kecuali parameternya yang tidak diketahui (*nonparametric*)

Peta Uji Nonparametrik

Uji Median

Tipe Sampel yg diambil	Nominal	Ordinal	Interval
Satu populasi	Binomial Test	Quantil Test	Wilcoxon Test
Pengamatan Berpasangan	McNemar Test	Sign Test	Wilcoxon Test
RAKL	Cochran Test	Friedman Test	
Dua Populasi	Chi-Sq Test	Mann-Whitney	
Lebih dari dua populasi	Chi-Sq Test	Kruskal-Wallis	

Teladan Binomial Test

Pengaruh didirikannya PLTN terhadap kesehatan pekerja dan penduduk yang tinggal di sekitarnya, akhir-akhir ini menjadi bahan perdebatan. Salah satu bahayanya adalah kemungkinan paparan radiasi akan meningkatkan kematian karena kanker. Problem yang biasa dijumpai ketika kita melakukan studi kasus seperti ini adalah sedikitnya jumlah kematian karena kanker (dari semua jenis) maupun suatu jenis kanker tertentu, sehingga kebermaknaan statistik akan sulit dicapai, kecuali jika studi dilakukan dalam jangka waktu lama. Salah satu alternatif adalah melakukan studi mortalitas proporsional, yang didalamnya proporsi kematian karena suatu penyebab pada kelompok yang terpapar dibandingkan dengan proporsi pada populasi umum. Andaikan hasil penelitian menemukan bahwa 4 dari 13 kematian pada para pekerja berusia 55-64 tahun di PLTN disebabkan karena kanker. Jika berdasarkan laporan statistik disebutkan bahwa 20% dari semua kematian disebabkan kanker, dapatkah kita simpulkan bahwa hasil penelitian tersebut berbeda secara bermakna terhadap statistik populasi umum ? Taraf nyata pengujian 5%.

Teladan Binomial Test

$$H_0: p = p_0 = 0,20 \quad \text{vs} \quad H_1: p \neq p_0 = 0,20$$

Dimana p adalah proporsi kematian karena kanker di antara para pekerja di PLTN dan p_0 adalah proporsi kematian karena kanker pada populasi umum

Dengan $\alpha = 0,05$, dan karena $p = 4/13 = 0,31 > 0,20$, maka H_0 kita tolak jika peluang untuk memperoleh sebanyak 4 kematian atau lebih diantara 13 kematian disebabkan kanker, adalah kurang dari 0,05. Sebaliknya kita terima apabila nilainya sama dengan atau lebih besar dari 0,05

$$p = 2 \sum_{k=4}^{13} C_k^{13} (0,20)^k (0,80)^{13-k} = 2 \left(1 - \sum_{k=0}^3 C_k^{13} (0,20)^k (0,80)^{13-k} \right)$$

Dari Tabel Binomial $n=13$ $x-1=3$ dan $p_0 = 0,20$, maka nilai peluangnya = $2(1-0,7473) = 0,5054$ Hipotesis nol **tidak** ditolak.

Teladan McNemar Test

Suatu perusahaan ingin mengetahui pengaruh sponsor yang diberikan dalam suatu pertandingan olahraga terhadap volume penjualan barangnya. Dalam penelitian ini digunakan sampel yang dipilih secara acak sebanyak 200. Sebelum sponsor diberikan, terdapat 50 orang yang membeli barang tersebut, dan sisanya tidak membeli. Setelah sponsor diberikan, ternyata dari 200 orang tersebut terdapat 125 orang yang membeli dan 75 orang yang tidak membeli. Dari 125 orang tersebut terdiri atas pembeli tetap sebanyak 40 orang, dan yang berubah dari tidak membeli menjadi membeli 85. Selanjutnya dari 75 orang yang tidak membeli terdiri dari yang berubah dari membeli menjadi tidak 10 orang, dan yang tetap tidak membeli ada 65 orang. Adakah terdapat perbedaan volume penjualan sebelum dan sesudah ada sponsor ?

Teladan McNemar Test

Sebelum Ada Sponsor	Setelah Ada Sponsor
Tidak Membeli 150	75 = 65 tetap + 10 berubah
Membeli 50	125 = 40 tetap + 85 berubah

H_0 : Tidak terdapat perbedaan volume penjualan sebelum dan sesudah ada sponsor

H_1 : Terdapat perbedaan volume penjualan sebelum dan sesudah ada sponsor

Tolak hipotesis nol jika χ^2 hitung $> \chi^2_{1;0.05} = 3,84$

Teladan McNemar Test

Perilaku Konsumen	Membeli	Tidak Membeli
Tidak Membeli	84 (A)	65 (B)
Membeli	40 (C)	10 (D)

$$\chi_{hit}^2 = \frac{(|A - D| - 1)^2}{A + D} = \frac{(|84 - 10| - 1)^2}{84 + 10} = 57,642$$

Tolak hipotesis nol. Terdapat cukup bukti dengan taraf nyata 5%, bahwa volume penjualan sebelum dan sesudah ada sponsor berbeda.

Teladan Cochran Test

Penelitian dilakukan untuk mengetahui efektivitas tiga metode kerja baru yang diadopsi dari konsultan. Untuk itu, tiga kelompok karyawan dipilih secara acak, yang masing-masing terdiri dari 15 karyawan. Efektivitas metode akan diukur dari gagal-tidaknya pegawai tersebut menyelesaikan pekerjaan dalam waktu 1 jam. Hasil eksperimen data seperti pada tabel.

No	Kel 1	Kel 2	Kel 3
1	0	1	1
2	1	0	1
3	1	1	1
4	0	0	1
5	0	1	1
6	0	0	0
7	1	1	1
8	0	1	1
9	1	0	1
10	0	1	1
11	1	0	0
12	0	1	1
13	0	0	1
14	0	0	0
15	1	0	1

Teladan Cochran Test

H_0 : tidak terdapat perbedaan efektivitas ke tiga metode kerja baru

H_1 : terdapat perbedaan efektivitas ke tiga metode kerja baru

Tolak hipotesis nol jika nilai Q lebih besar dari χ^2 -tabel = $\chi^2_{-2;0,05} = 5,99$

$$Q = \frac{(k-1) \left[k \sum_{j=1}^k G_j^2 - \left(\sum_{j=1}^k G_j \right)^2 \right]}{k \sum_{i=1}^N L_i - \sum_{i=1}^N L_i^2} = \frac{(3-1) [3(6^2 + 7^2 + 12^2) - 25^2]}{(3)(25) - 53} = 5,64$$

Terima hipotesis nol. Tidak terdapat cukup bukti untuk menyatakan bahwa ke tiga metode kerja baru berbeda, pada taraf uji 5%.

Teladan Chi-Squared Test-2

Penelitian dilakukan untuk mengetahui adakah hubungan antara tingkat pendidikan masyarakat dengan jenis Bank yang dipilih untuk menyimpan uangnya. Pendidikan masyarakat dikelompokkan menjadi 2, yaitu lulusan SLTA dan Perguruan Tinggi. Sampel pertama sebanyak 80 orang lulusan SLTA, dan sampel kedua sebanyak 70 orang lulusan Perguruan Tinggi. Berdasarkan angket yang diberikan kepada sampel lulusan SLTA, 60 orang memilih bank pemerintah dan 20 bank swasta. Selanjutnya dari kelompok sampel Perguruan Tinggi, sebanyak 30 orang memilih bank pemerintah dan 40 bank swasta. [Seluruh data diatas adalah fiktif dan hanya untuk keperluan pembelajaran Uji Chi-Squared 2 sampel].

Teladan Chi-Squared Test-2

H_0 : tidak terdapat perbedaan tingkat pendidikan masyarakat dalam memilih dua jenis bank (atau tidak ada hubungan antara jenjang pendidikan dengan jenis bank yang dipilih).

H_1 : terdapat perbedaan tingkat pendidikan masyarakat dalam memilih dua jenis bank (atau tidak ada hubungan antara jenjang pendidikan dengan jenis bank yang dipilih).

Tolak hipotesis nol jika $\chi^2_{\text{hit}} > \chi^2_{\text{tabel}} = \chi^2_{1;\alpha}$

Sampel	Bank Pemerintah	Bank Swasta	Jumlah Sampel
Lulusan PT	a =60	b =20	80
Lulusan SLTA	c =30	d =40	70
Jumlah	90	60	150

Teladan Chi-Squared Test-2

$$\chi^2 = \frac{n \left(|ad - bc| - \frac{n}{2} \right)^2}{(a+b)(a+c)(b+d)(c+d)} = 14,76$$

Pada taraf nyata 1%, nilai χ^2 -tabel = 6,635. Dengan demikian hipotesis nol ditolak.

Artinya, pada taraf nyata pengujian 1% terdapat cukup bukti untuk mengatakan bahwa terdapat perbedaan tingkat pendidikan dalam memilih jenis bank.

Teladan Chi-Squared Test-k

Suatu contoh acak siswa dipilih dari sekolah menengah umum negeri dan contoh acak yang lain dipilih dari sekolah menengah umum swasta. Mereka diberikan uji potensi akademik dan hasilnya adalah seperti pada tabel di bawah ini.

	0-275	276-350	351-425	426-500
Swasta	6	14	17	9
Negeri	30	32	17	3

Apakah distribusi nilai ujian di kedua kelompok sekolah (negeri dan swasta) sama ? Gunakan taraf nyata pengujian 5%.

Teladan Chi-Squared Test-k

Observasi	0-275	276-350	351-425	426-500
Swasta	6	14	17	9
Negeri	30	32	17	3

Ekspektasi	0-275	276-350	351-425	426-500
Swasta	6	14	17	9
Negeri	30	32	17	3

Teladan Chi-Squared Test-k

$$T = \sum_{i=1}^2 \sum_{j=1}^4 \frac{(O_{ij} - E_{ij})^2}{E_{ij}} = \frac{(6 - 12,9)^2}{12,9} + \dots + \frac{(3 - 7,7)^2}{7,7} = 17,3$$

$H_0: p_{1j} = p_{2j}$ untuk $j = 1, 2, 3, 4$

$$\chi_{0,05;3}^2 = 7,815$$

$H_1: p_{1j} \neq p_{2j}$ untuk $j = 1, 2, 3, 4$

Kriteria Penolakan hipotesis nol : Tolak hipotesis nol, jika nilai $T >$ Kai-kuadrat tabel dengan derajat bebas $(2-1)(4-1)$ dan $\alpha = 0,05$.

Kesimpulan: Karena $T = 17,3 > 7,815$ maka hipotesis nol diolak.

Teladan Sign Test

Suatu perusahaan ingin mengetahui pengaruh adanya kenaikan uang insentif terhadap kesejahteraan karyawan. Dalam penelitian itu, dipilih 20 pegawai beserta istrinya secara acak, sehingga terdapat 20 pasangan data suami-istri. Perubahan peringkat kesejahteraan keluarga menurut suami-istri disajikan pada tabel berikut. Apakah terdapat perbedaan pengaruh insentif yang signifikan terhadap kesejahteraan keluarga menurut suami-istri? Gunakan taraf nyata pengujian sebesar 5%.

Jumlah tanda + ada sebanyak 7. Akan menghasilkan nilai peluang 0,132. Karena nilai ini lebih besar dari 0,05 maka hipotesis nol diterima

Nomor	Isteri	Suami	Arah	Tanda
1	4	1	>	-
2	5	4	>	-
3	4	5	<	+
4	4	5	<	+
5	5	4	>	-
6	4	3	>	-
7	4	3	>	-
8	2	1	>	-
9	1	3	<	+
10	4	5	<	+
11	3	2	>	-
12	2	3	<	+
13	4	1	>	-
14	4	2	>	-
15	3	2	>	-
16	2	1	>	-
17	4	5	<	+
18	1	2	<	+
19	5	4	>	-
20	5	4	>	-

Teladan Kruskal-Wallis Test

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 sebelah.

Prestasi Kerja

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

Teladan Kruskal-Wallis Test

Peringkat	II	III
21.0	24.5	13.5
33.0	30.0	22.0
12.0	15.0	11.0
3.0	5.0	7.0
19.5	8.5	16.0
24.5	18.0	17.0
23.0	10.0	26.0
8.5	19.5	3.0
32.0	27.0	6.0
1.0	3.0	31.0
28.0	29.0	
	13.5	
205.5	203.0	152.5

H_0 : Tidak terdapat perbedaan prestasi kerja pekerja berdasarkan jarak rumah dengan kantor

H_1 : Terdapat perbedaan prestasi kerja pekerja berdasarkan jarak rumah dengan kantor

Tolak hipotesis nol jika nilai H lebih besar dari kai-kuadrat tabel

$$H = \frac{12}{N(N+1)} \sum_{j=1}^k \frac{R_j^2}{n_j} - 3(N+1)$$

Teladan Kruskal-Wallis Test

$$H = \frac{12}{N(N+1)} \sum_{j=1}^k \frac{R_j^2}{n_j} - 3(N+1) = \frac{12}{(33)(34)} \left[\frac{(205.5)^2}{11} + \frac{(203)^2}{12} + \frac{(152.5)^2}{10} \right] - 3(34) = 0.66$$

$$\chi_{0.05;2}^2 = 5.99$$

Kesimpulan: Karena $H = 0.66 < 5.99$ maka hipotesis nol diterima. Tidak terdapat prestasi kerja berdasarkan jarak rumah. Pekerja yang rumahnya jauh dari kantor memiliki prestasi kerja yang sama dengan pekerja yang rumahnya dekat dengan kantor. Atau tidak ada hubungan antara jarak rumah –kantor dengan prestasi kerja pekerja.

Teladan Friedman Test

Empat subyek penelitian mengikuti sebuah eksperimen untuk meneliti perbedaan efektifitas tiga metode terapi stres. Masing-masing subyek mengalami beban stres yang sama pada tiga kesempatan. Pada tiap kali kesempatan, subyek daiberi sebuah metode terapi stres. Variabel respon yang diukur ialah jumlah penurunan tingkat stres sebelum dan sesudah diberi terapi. Hasilnya terlihat pada tabel disamping

Subyek	Metode		
	A	B	C
1	16	26	22
2	16	20	23
3	17	21	22
4	28	19	36

Tolak H_0 : semua metode memiliki pengaruh yang sama jika nilai-p < 0,05

Teladan Friedman Test

$$F_R = \frac{12}{bk(k+1)} \sum_{j=1}^k \left[R_j - \frac{b(k+1)}{2} \right]^2$$

Dengan menggunakan Tabel Friedman dan nilai $F_R = 6,5$ maka nilai-p = 0,042 maka hipotesis nol ditolak

	A	B	C
1	1	3	2
2	1	2	3
3	1	2	3
4	1	2	3
	4	9	11

b = banyaknya blok atau kategori baris, dalam contoh ini b = 4

k = banyaknya perlakuan atau kategori kolom, k = 3

Teladan Mann-Whitney U Test

Sebuah penelitian dilakukan untuk mengetahui perbedaan kualitas manajemen 2 kelompok bank (kelas A dan kelas B). Sampel sebanyak 12 bank kelas A dan 15 bank kelas B diambil secara acak dari sejumlah bank-bank kelas A dan B. Kualitas manajemen kedua kelompok bank diukur dengan menggunakan sebuah instrumen. Skor terendah 0 dan tertinggi 40. Hasil penelitian dapat dilihat pada tabel.

B	A
16	19
18	19
10	21
12	25
16	26
14	27
15	23
10	27
12	19
15	19
16	25
11	27
	23
	19
	29

Teladan Mann-Whitney U Test

$$U_1 = n_1 n_2 + \frac{n_1(n_1 + 1)}{2} - R_1$$

$$U_2 = n_1 n_2 + \frac{n_2(n_2 + 1)}{2} - R_2$$

Tolak hipotesis nol jika $\min(U_1, U_2) > U\text{-tabel}$

Karena $\min(78, 100) = 78 > U\text{ tabel}$ yang nilainya 42 (untuk alpha 0,05 dan $n_1 = 12$ dan $n_2 = 15$) maka H_0 ditolak

B	A
10	15
12	15
1.5	18
4.5	21.5
10	23
6	25
7.5	19.5
1.5	25
4.5	15
7.5	15
10	21.5
3	25
	19.5
	15
	27
U1 =78	U2 =300

Teladan Wilcoxon Test

Di sebuah kantor dilakukan penelitian untuk mengetahui pengaruh ruangan ber-AC terhadap produktivitas kerja. Pengumpulan data terhadap produktivitas kerja pegawai dilakukan pada waktu AC sebelum dan sesudah dipasang. Data produktivitas kerja pegawai sebelum dan sesudah dipasang AC dapat dilihat pada tabel.

Tolak H_0 jika $\min(W^-, W^+) < W$ tabel, dimana utk alpha 0,05, $n = 10$ diperoleh W tabel = 8

Sebelum	Sesudah
100	105
98	94
76	78
90	98
87	90
89	85
77	86
92	87
78	80
82	83

Teladan Wilcoxon Test

Sebelum	Sesudah	Beda	Abs	Peringkat	+	-
100	105	5	5	7.5	7.5	
98	94	-4	4	5.5		5.5
76	78	2	2	2.5	2.5	
90	98	8	8	9	9	
87	90	3	3	4	4	
89	85	-4	4	5.5		5.5
77	86	9	9	10	10	
92	87	-5	5	7.5		7.5
78	80	2	2	2.5	2.5	
82	83	1	1	1	1	
					36.5	18.5

$$W_- = 18,5 \quad W_+ = 36,5$$

Kesimpulan : Terima Hipotesis Nol

Ukuran Asosiasi

skala nominal

Phi Coefficient

$$\Phi = \sqrt{\frac{\chi^2}{N}}$$

$$\Phi = \frac{ad - bc}{\sqrt{(a+b)(c+d)(a+c)(b+d)}}$$

Pearson's Coefficient of Contingency

$$C = \sqrt{\frac{\chi^2}{N + \chi^2}}$$

Tschuprow's Coefficient

$$C_{Ts} = \sqrt{\frac{\chi^2}{N\sqrt{(r-1)(c-1)}}$$

Ukuran Asosiasi

skala nominal

Kramer V Coefficient

$$V = \sqrt{\frac{\chi^2}{N(k-1)}}$$

k = minimum baris atau kolom

Mean Square Contingency

$$C_{MS} = \frac{\chi^2}{N}$$

Ukuran Asosiasi

skala ordinal

Spearman's ρ

$$\rho = 1 - \frac{6 \sum_{i=1}^n [R(X_i) - R(Y_i)]^2}{n(n^2 - 1)}$$

$$\rho = 1 - \frac{(6)(75)}{(12)(143)} = 0,7378$$

X	Y	R(X)	R(Y)	(R(X)-R(Y)) ²
86	88	8	10	4.00
71	77	3.5	7	12.25
77	76	6.5	6	0.25
68	64	1	1	0.00
91	96	11.5	12	0.25
72	72	5	4.5	0.25
77	65	6.5	2.5	16.00
91	90	11.5	11	0.25
70	65	2	2.5	0.25
71	80	3.5	8	20.25
88	81	10	9	1.00
87	72	9	4.5	20.25
				75.00

Ukuran Asosiasi

skala ordinal

X	Y	Conc	Disc
68	64	11	0
70	65	9	0
71	77	4	4
71	80	4	4
72	72	5	1
77	76	5	0
77	65	4	1
86	88	2	2
87	72	3	0
88	81	2	0
91	96	0	0
91	90	0	0
		49	12

$$\tau = \frac{2(N_c - N_d)}{n(n-1)}$$

$$\tau = \frac{2(49 - 12)}{(12)(11)} = 0,5606$$