

Statistika Non Parametrik

1. Pendahuluan

Kelebihan Uji Non Parametrik:

- Perhitungan sederhana dan cepat
- Data dapat berupa data kualitatif (Nominal atau Ordinal)
- Distribusi data tidak harus Normal

Kelemahan Uji Non Parametrik:

- Tidak memanfaatkan semua informasi dari sampel (Tidak efisien)

Kelemahan diperbaiki dengan menambah ukuran sampel

Beberapa Uji Non Parametrik yang akan dipelajari :

- Uji tanda berpasangan
- Uji Peringkat 2 Sampel Mann-Whitney
- Uji Peringkat 2 Sampel Wilcoxon
- Uji Korelasi Peringkat Spearman
- Uji Konkordansi Kendall
- Uji Run(s)

2. Uji Tanda Berpasangan

Uji dilakukan pada 2 sampel terpisah (independen)

- tanda (+) → data pada sampel 1 > pasangannya sampel 2
- tanda (-) → data pada sampel 1 < pasangannya sampel 2
- tanda Nol (0) → data pada sampel 1 = pasangannya sampel 2

Tanda Nol *tidak digunakan* dalam perhitungan

Notasi yang digunakan :

n = banyak tanda (+) dan tanda (-) dalam sampel

\bar{p} = proporsi SUKSES dalam sampel

$$\bar{q} = 1 - \bar{p}$$

p_0 = proporsi SUKSES dalam H_0

$$q_0 = 1 - p_0$$

$$\text{Standar Error} = \text{Galat Baku} = \sigma_{\bar{p}} = \sqrt{\frac{p_0 \times q_0}{n}}$$

$$\text{Rata-Rata Sampel} = \mu_{\bar{p}} = p_0$$

Statistik Uji $z_{hitung} = \frac{\bar{p} - \mu_{\bar{p}}}{\sigma_{\bar{p}}}$ $z_{hitung} = \frac{\bar{p} - p_0}{\sqrt{\frac{p_0 \times q_0}{n}}}$

SUKSES tergantung dari apa yang ditanyakan (ingin diuji) dalam soal.

Jika yang ingin diuji sampel 1 > sampel 2 maka SUKSES adalah banyak tanda (+)

Jika yang ingin diuji sampel 1 < sampel 2 maka SUKSES adalah banyak tanda (-)

Nilai p_0 disesuaikan dengan nilai pengujian p yang diinginkan dalam soal

atau jika ingin diuji proporsi sampel 1 = proporsi sampel 2 maka $p_0 = q_0 = 0.50$

Penetapan Penetapan H_0 dan H_1 :

Terdapat 3 alternatif H_0 dan H_1 :

(a) $H_0: p = p_0$ dan $H_1: p < p_0$

Uji 1 arah dengan daerah penolakan $H_0: z < -z_{\alpha}$

(b) $H_0: p = p_0$ dan $H_1: p > p_0$

Uji 1 arah dengan daerah penolakan $H_0: z > z_{\alpha}$

(c) $H_0: p = p_0$ dan $H_1: p \neq p_0$

Uji 2 arah dengan daerah penolakan $H_0: z < -z_{\alpha/2}$ dan $z > z_{\alpha/2}$

Contoh 1a:

Berikut adalah nilai preferensi konsumen terhadap 2 Merk Sabun Mandi. Dengan taraf nyata 1%, ujilah apakah proporsi preferensi konsumen pada kedua merk bernilai sama?

No. Responden	LUXE	GIVE	Tanda
1.	4	2	+
2.	2	3	-
3.	3	3	0
4.	2	3	-
5.	3	2	+
6.	1	2	-
7.	2	3	-
8.	3	4	-
9.	3	2	+
10.	2	1	+
11.	4	1	+
12.	1	1	0
13.	4	2	+
14.	3	2	+
15.	4	3	+

Banyak tanda (+) = 8

Banyak tanda (-) = 5

$n = 8 + 5 = 13$

Jika kita asumsikan LUXE lebih disukai dibanding GIVE maka SUKSES dalam sampel adalah \bar{p} = proporsi banyak tanda (+) dalam sampel

$$\bar{p} = \frac{\text{banyak positif}}{n} = \frac{8}{13} = 0.62$$

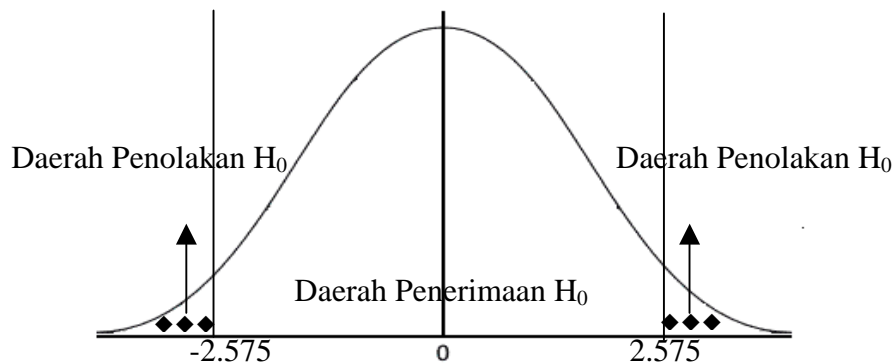
$$\bar{q} = 1 - \bar{p} = 1 - 0.62 = 0.38$$

Karena ingin diuji proporsi yang suka LUXE = GIVE maka $p_0 = q_0 = 0.50$

Langkah Pengujian:

1. $H_0: p = 0.50$ $H_1: p \neq 0.50$
2. Statistik Uji : z
3. Uji: 2 Arah
4. Taraf Nyata Pengujian = $\alpha = 1\% \rightarrow \alpha/2 = 0.5\% = 0.005$
5. Daerah Penolakan H_0

$$z < -z_{0.005} \rightarrow z < -2.575 \quad \text{dan} \quad z > z_{0.005} \rightarrow z > 2.575$$



6. Nilai statistik Uji :

$$z_{hitung} = \frac{\bar{p} - p_0}{\sqrt{\frac{p_0 \times q_0}{n}}} = \frac{0.62 - 0.50}{\sqrt{\frac{0.50 \times 0.50}{13}}} = \frac{0.12}{\sqrt{\frac{0.25}{13}}} = \frac{0.12}{\sqrt{0.0192...}} = \frac{0.12}{0.13867...} = 0.8653...$$

$$\approx 0.87$$

7. Kesimpulan:
z hitung = 0.87 ada di daerah penerimaan H_0 H_0 diterima
Proporsi konsumen yang menyukai LUXE masih sama dengan yang menyukai GIVE.

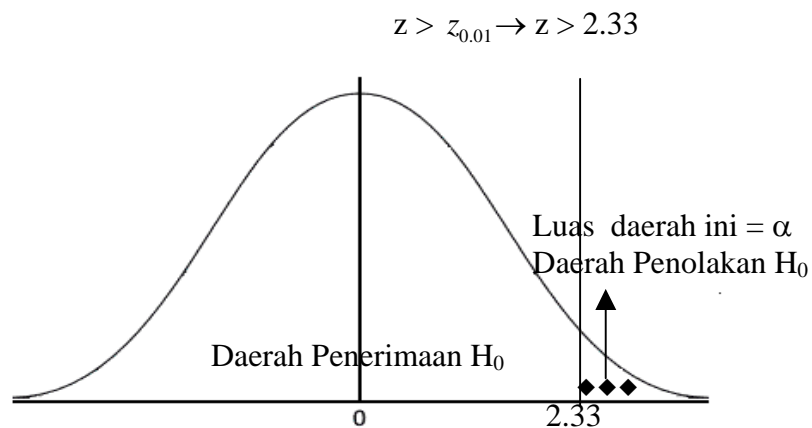
Contoh 1b:

Dengan menggunakan data pada Tabel 1 dan taraf nyata 1% ujilah apakah proporsi preferensi konsumen pada sabun LUXE dibanding sabun GIVE sudah lebih dari 0.30?

$$p_0 = 0.30$$

$$q_0 = 1 - 0.30 = 0.70$$

1. $H_0: p = 0.30$ $H_1: p > 0.30$
2. Statistik Uji : z
3. Uji 1 Arah
4. Taraf Nyata Pengujian = $\alpha = 1\% = 0.01$
5. Daerah Penolakan H_0



6. Nilai statistik Uji :

$$z_{hitung} = \frac{\bar{p} - p_0}{\sqrt{\frac{p_0 \times q_0}{n}}} = \frac{0.62 - 0.30}{\sqrt{\frac{0.30 \times 0.70}{13}}} = \frac{0.32}{\sqrt{\frac{0.21}{13}}} = \frac{0.32}{\sqrt{0.0161...}} = \frac{0.32}{0.1270...} = 2.5177...$$

$$\approx 2.52$$

7. Kesimpulan:
 z hitung = 2.52 ada di daerah penolakan H_0 ,
 H_0 ditolak H_1 diterima
 Proporsi konsumen yang menyukai LUXE sudah lebih dari 0.30

3. Uji Peringkat 2 Sampel Mann-Whitney

Uji ini merupakan alternatif uji beda 2 rata-rata Parametrik dengan menggunakan t (Sampel-sampel berukuran kecil).

Langkah pertama pengujian ini adalah pengurutan nilai mulai dari yang terkecil hingga terbesar. Pengurutan dilakukan tanpa pemisahan kedua sampel.

Selanjutnya lakukan penetapan Rank (Peringkat) dengan aturan berikut:
 Peringkat ke -1 diberikan pada nilai terkecil di urutan pertama
 Peringkat tertinggi diberikan pada nilai terbesar

Jika tidak ada nilai yang sama maka urutan = peringkat
 Jika ada nilai yang sama, maka ranking dihitung dengan rumus

$$\text{Peringkat (R)} = \frac{\sum \text{urutan data yg bernilai sama}}{\text{banyak data yg bernilai sama}}$$

Contoh 2a: Berikan peringkat (ranking) data dalam tabel berikut ini!

Tabel 2. Nilai UAS Statistika 2

Mahasiswa Fak. Ekonomi			Mahasiswa Fak. Ilmu Komputer		
Nilai	Urutan	Rangking	Nilai	Urutan	Ranking
30	2	2	25	1	1
55	4	4	50	3	3
65	5	5	70	6	7
70	8	7	70	7	7
75	10	9.5	75	9	9.5
88	16	15.5	78	11	11
90	17	17	80	12	12
95	18	18	85	13	13.5
98	19	19	85	14	13.5
100	20	20	88	15	15.5
$R_1 =$		117	$R_2 =$		93

$$\text{Ranking untuk Nilai 70} = \frac{6 + 7 + 8}{3} = \frac{21}{3} = 7$$

$$\text{Ranking untuk Nilai 75} = \frac{9 + 10}{2} = \frac{19}{2} = 9.5$$

Notasi yang digunakan

Prinsip pengerjaannya sama dengan Uji Peringkat 2 Sampel Mann-Whitney, hanya fokus kini dialihkan sampel dengan ukuran terkecil.

Notasi yang digunakan :

n_1 = ukuran sampel ke 1

n_2 = ukuran sampel ke 2

$n_1 < n_2$ ukuran sampel ke 1 selalu lebih kecil dari sampel ke 2

W = jumlah peringkat pada sampel berukuran terkecil

$$\text{Nilai Ekspektasi (W)} = E(W) = \frac{n_1(n_1 + n_2 + 1)}{2}$$

$$\text{Standar Error} = SE = \sqrt{\frac{n_1 \times n_2 \times (n_1 + n_2 + 1)}{12}}$$

$$\text{Statistik Uji } z = \frac{W - E(W)}{SE}$$

Penetapan urutan, peringkat dan H_0 dan H_1 sama dengan Uji Mann-Whitney

Contoh 3: Berikut adalah data pendapatan di 2 kelompok pekerja

Tabel 3. Pendapatan Karyawan

Departemen Q			Departemen Z		
Income (ribu USD/tahun)	Urutan	Rangking	Income (ribu USD/tahun)	Urutan	Ranking
6	1	1	12	3	3
10	2	2	13	4	4
15	7	6	15	5	6
32	10	10	15	6	6
	W =	19	20	8	8
			31	9	9
			38	11	11
			40	12	12

5. Uji Korelasi Peringkat Spearman

Dua uji terakhir (Mann-Whitney dan Wilcoxon) ditujukan untuk 2 sampel yang saling bebas (independen), sedangkan Uji Peringkat Spearman ditujukan untuk penetapan peringkat data berpasangan.

Konsep dan interpretasi nilai Korelasi Spearman (R_s) sama dengan konsep Koefisien Korelasi pada Regresi (Linier Sederhana).

Notasi yang digunakan:

n = banyak pasangan data

d_i = selisih peringkat pasangan data ke i

R_s = Korelasi Spearman

$$R_s = 1 - \frac{6 \sum_{i=1}^n d_i^2}{n(n^2 - 1)}$$

$$\text{Statistik Uji } z = R_s \times (\sqrt{n - 1})$$

Penetapan H_0 dan H_1 :

Terdapat 3 alternatif H_0 dan H_1 :

(a) H_0 : $R = 0$ (korelasi bernilai 0, tidak ada hubungan /tidak ada kecocokan)

H_1 : $R < 0$ (korelasi negatif)

Uji 1 arah dengan daerah penolakan H_0 : $z < -z_\alpha$

(b) H_0 : $R = 0$ (korelasi bernilai 0, tidak ada hubungan /tidak ada kecocokan)

H_1 : $R > 0$ (korelasi positif)

Uji 1 arah dengan daerah penolakan H_0 : $z > z_\alpha$

(c) H_0 : $R = 0$ (korelasi bernilai 0, tidak ada hubungan /tidak ada kecocokan)

H_1 : $R \neq 0$ (ada korelasi/ada kecocokan, korelasi tidak sama dengan 0)

Uji 2 arah dengan daerah penolakan H_0 : $z < -z_{\alpha/2}$ dan $z > z_{\alpha/2}$

Peringkat diberikan tergantung kategori penilaian.

Jika ada item yang dinilai ber-peringkat sama, maka penetapan peringkat seperti dalam Mann-Whitney dapat dilakukan (ambil rata-rata peringkatnya!)

Contoh 5:

Dua orang pakar (ahli) diminta memberikan peringkat kinerja pada 10 Bank di Indonesia. Peringkat diberikan mulai dari bank terbaik = peringkat 1 sedang yang terburuk diberi peringkat 10. Hasilnya disajikan dalam Tabel 4.

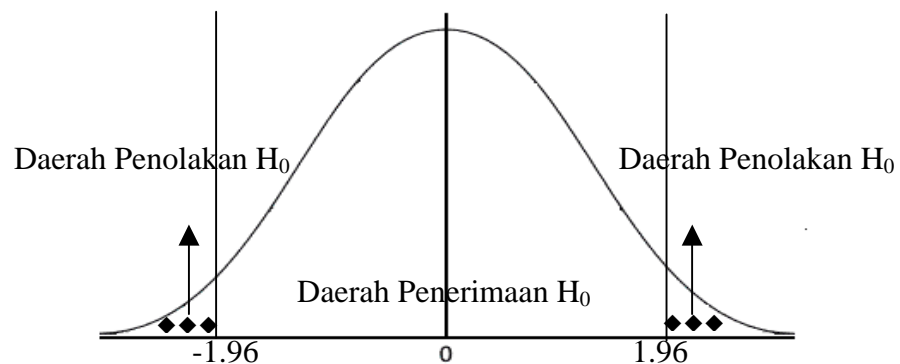
Tabel 4. Hasil peringkat 10 Bank oleh 2 Pakar

Bank	Ranking Pakar I	Rangking Pakar II	d_i	d_i^2
A	4	3	1	1
B	5	1	4	16
C	3	4.5	-1.5	2.25
D	7	6	1	1
E	10	8	2	4
F	1	2	-1	1
G	6	4.5	1.5	2.25
H	2	7	-5	25
I	8.5	10	-1.5	2.25
J	8.5	9	-0.5	0.25
$\Sigma d_i^2 =$				55

Dengan taraf nyata 5% ujilah apakah ada korelasi antara peringkat yang diberikan kedua pakar?

1. $H_0: R = 0$ $H_1: R \neq 0$
2. Statistik Uji : z
3. Uji 2 Arah
4. Taraf Nyata Pengujian = $\alpha = 5\% \rightarrow \alpha/2 = 2.5\% = 0.025$
5. Daerah Penolakan H_0

$$z < -z_{0.025} \rightarrow z < -1.96 \quad \text{dan} \quad z > z_{0.025} \rightarrow z > 1.96$$



6. Nilai statistik Uji :

$$R_s = 1 - \frac{6 \sum_{i=1}^n d_i^2}{n(n^2 - 1)} = 1 - \frac{6 \times 55}{10 \times (10^2 - 1)} = 1 - \frac{330}{990} = 1 - 0.33... = 0.67$$

$$z = R_s \times (\sqrt{n-1}) = 0.67 \times (\sqrt{10-1}) = 0.67 \times \sqrt{9} = 0.67 \times 3 = 2.01$$

7. Kesimpulan:

z hitung = 2.01 ada di daerah penolakan H_0

H_0 ditolak H_1 diterima

Ada korelasi/ada kecocokan pemberian peringkat oleh kedua pakar,

6. Uji Konkordansi Kendall

Pengujian sampel berpasangan ganda (multiple-paired samples).

Orang yang memberi peringkat lebih dari 2.

Statistik Uji yang digunakan : χ^2 (chi kuadrat) dengan derajat bebas (db) = n-1

Notasi yang digunakan

n = banyak pasangan data, $n \geq 8$

R = jumlah peringkat

k = banyak orang yang memberi peringkat ($k > 2$)

$$\text{Statistik Uji } \chi^2 = \frac{12 \sum R^2 - (3n(k(n+1))^2)}{kn(n+1)} **)$$

**) sumber di Diktat Statistika-2, Gunadarma agak rancu...?

Sumber lain belum saya temukan. Yang paling mendekati ada di

http://www.analystsoft.com/en/products/statplus/content/help/src/analysis_nonparametric_statistics_comparing_multiple_dependent_samples_friedman_anova_kendall_concordance.html

Contoh 6:

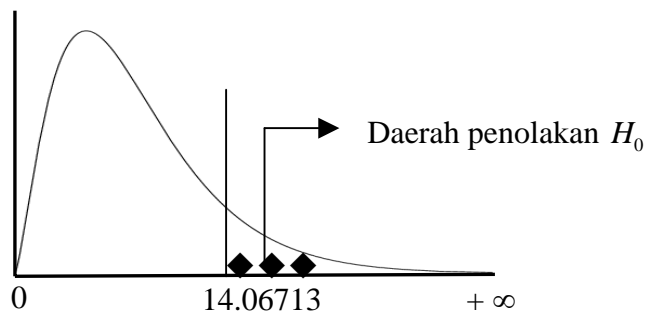
Tiga konsultan Teknologi Informasi (TI) diminta memberi peringkat pada 8 merk laptop.

Dengan taraf nyata 5% ujilah apakah terdapat kecocokan peringkat? (lihat Tabel di bawah)

Merk Laptop	Pakar 1	Pakar 2	Pakar 3	R	R ²
A	3	2	4	9	81
B	2	5	3	10	100
C	1	1	2	4	16
D	5	3	1	9	81
E	8	4	7	19	361
F	6	7	5	18	324
G	7	6	8	21	441
H	4	8	6	18	324
					$\Sigma R^2 = 1728$

Jawab:

1. $H_0: R_{\text{Kendall}} = 0$ (tidak ada korelasi/tidak ada kecocokan)
 $H_1: R_{\text{Kendall}} \neq 0$ (ada korelasi/ada kecocokan)
2. Statistik Uji : χ^2
3. Taraf Nyata Pengujian = $\alpha = 5\% = 0.05$
4. $db = n - 1 = 8 - 1 = 7$ dan $\chi^2_{\text{tabel}(db; \alpha)} = 14.06713$
5. Daerah Penolakan H_0 jika $\chi^2 > \chi^2_{\text{tabel}(db; \alpha)}$ $\chi^2 > 14.06713$



6. Nilai statistik Uji :

$$\chi^2 = \frac{12 \sum R^2 - (3n(k(n+1)^2))}{kn(n+1)} = \frac{(12 \times 1728) - ((3 \times 8) \times (3 \times (8+1)^2))}{(3 \times 3) \times (8+1)} = 15$$

7. Kesimpulan:
 $\chi^2_{\text{hitung}} = 15$ ada di daerah penolakan H_0 maka H_0 ditolak dan H_1 diterima
Ada kecocokan peringkat.
7. Uji Run(s)

Uji Run(s) digunakan untuk menguji keacakan dalam suatu sampel.

Run adalah satu atau lebih lambang-lambang yang identik yang didahului atau diikuti oleh suatu lambang yang berbeda atau tidak ada lambang sama sekali.

Misal: LLL PPP L P L PPPP L P LLLLLL terdapat 9 runs
 Run ke 1 2 3 4 5 6 7 8 9

Statistik Uji yang digunakan = z

Notasi yang digunakan

n_1 = banyak lambang 1 dalam sampel $n_1 > 10$

n_2 = banyak lambang 2 dalam sampel $n_2 > 10$

$n = n_1 + n_2$

n_r = banyak run(s)

Rata-rata Run(s) = $\mu_r = \frac{2n_1n_2}{n} + 1$

Standar Deviasi Run(s) = $\sigma_r = \sqrt{\frac{2n_1n_2(2n_1n_2 - n)}{n^2(n-1)}}$

Statistik Uji z = $z = \frac{n_r - \mu_r}{\sigma_r}$

Penetapan H_0

H_0 : Susunan Acak (Random)

H_1 : Susunan Tidak Acak (Tidak Random)

Uji 2 arah dengan daerah penolakan H_0 : $z < -z_{\alpha/2}$ dan $z > z_{\alpha/2}$

Contoh 7:

Berikut adalah urutan duduk mahasiswa dan mahasiswi dalam suatu kelas:

LL P L PP L P L P L P LL P LLLLLLL PP L P LL PP LLLLLL

L = Laki-laki, P = Perempuan

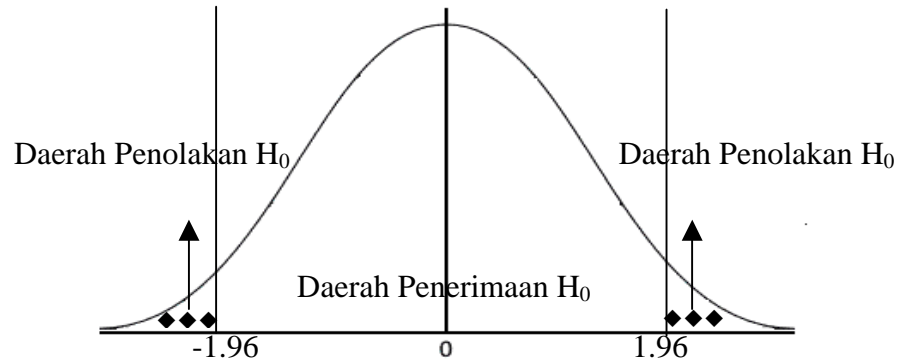
Dengan taraf nyata 5%, ujilah apakah urutan ini sudah random?

n_1 = banyak L = 24 n_2 = banyak P = 12 n_r = banyak runs = 19

1. H_0 : susunan acak H_1 : susunan tidak acak

2. Statistik Uji : z
3. Uji 2 Arah
4. Taraf Nyata Pengujian = $\alpha = 5\% \rightarrow \alpha/2 = 2.5\% = 0.025$
5. Daerah Penolakan H_0

$$z < -z_{0.025} \rightarrow z < -1.96 \quad \text{dan} \quad z > z_{0.025} \rightarrow z > 1.96$$



6. Nilai statistik Uji :

$$\mu_r = \frac{2n_1n_2}{n} + 1 = \frac{2 \times 24 \times 12}{36} + 1 = 17$$

$$\sigma_r = \sqrt{\frac{2n_1n_2(2n_1n_2 - n)}{n^2(n-1)}} = \sqrt{\frac{2 \times 24 \times 12 \times (2 \times 24 \times 12 - 36)}{36^2 \times (36-1)}} = \sqrt{\frac{576 \times 540}{1296 \times 35}} =$$

$$\sqrt{6.857143} = 2.618615 \approx 2.62$$

$$n_r = 19$$

$$z = \frac{n_r - \mu_r}{\sigma_r} = \frac{19 - 17}{2.62} = 0.76$$

7. Kesimpulan:

z hitung = 0.76 ada di daerah penerimaan H_0
 H_0 diterima. Susunan acak.

Catatan akhir:

Terdapat banyak ragam perhitungan Statistika Non-parametrik lainnya, mahasiswa sangat dianjurkan mempelajari sendiri berbagai teknik perhitungan Statistika Non Parametrik tersebut.

Selesai