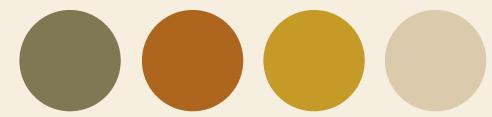


STA1343-PENGANTAR ANALISIS DATA KATEGORIK

RESPONSI STA1343-PADK

PERTEMUAN 2 - 31 AGUSTUS 2023



POKOK BAHASAN

1

Uji Mann-Whitney

2

Uji Tanda Data
Berpasangan

3

Uji Wilcoxon Data
Berpasangan

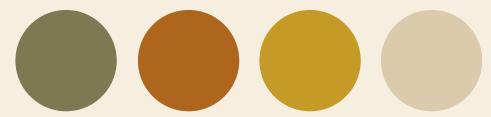
4

Korelasi Spearman



1

Uji Mann Whitney



Uji Mann-Whitney

Asumsi

- Kedua contoh **Saling Bebas**
- Peubah acak bersifat kontinu
- Skala pengukuran minimal ordinal.

Statistik Uji

$$T = S - \frac{n_1(n_1 + 1)}{2}$$

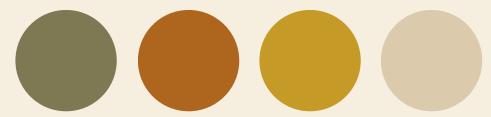
dengan S adalah jumlah peringkat amatan yang berasal dari populasi 1

Hipotesis

- $H_0 : M_x = M_y$ **vs** $H_1 : M_x \neq M_y$ (dua arah)
- $H_0 : M_x \geq M_y$ **vs** $H_1 : M_x < M_y$ (satu arah)
- $H_0 : M_x \leq M_y$ **vs** $H_1 : M_x > M_y$ (satu arah)

Kaidah penolakan H_0

- $T < w_{\frac{\alpha}{2}}$ atau $T > w_{1-\frac{\alpha}{2}}$
di mana $w_{1-\frac{\alpha}{2}} = n_1 n_2 - w_{\frac{\alpha}{2}}$
- $T < w_{\alpha}$
- $T > w_{1-\alpha}$
dimana $w_{1-\alpha} = n_1 n_2 - w_{\alpha}$



Uji Mann-Whitney

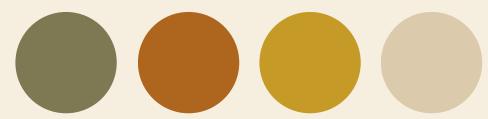
Pendekatan Sebaran Normal

Jika ukuran contoh $n_1, n_2 > 20$ maka dapat didekati dengan sebaran normal dengan statistik uji sebagai berikut:

$$Z = \frac{T - \frac{n_1 n_2}{2}}{\sqrt{\frac{n_1 n_2 (n_1 + n_2 + 1)}{12} - \frac{n_1 n_2 (\sum t^3 - \sum t)}{12(n_1 + n_2)(n_1 + n_2 - 1)}}} \rightarrow \text{Ada ties}$$

$$Z = \frac{T - \frac{n_1 n_2}{2}}{\sqrt{\frac{n_1 n_2 (n_1 + n_2 + 1)}{12}}} \rightarrow \text{Tidak ada ties}$$

Tolak H_0 Jika $Z_{hit} > Z_\alpha$



Contoh Soal

Menteri pertanian mengambil 24 sampel padi di dataran rendah & 24 padi di dataran tinggi. Menteri tersebut mengatakan bahwa produktivitas padi di **dataran rendah lebih besar** dibandingkan dengan dataran tinggi. Ujilah dengan uji Mann-Whitney pernyataan menteri tersebut pada taraf nyata 5%

Rendah	Tinggi
1200	1100
1350	980
1115	995
1400	1200
1000	900
990	1000
1000	980
970	1100
1200	1115
1100	990
1500	1000
1450	1100

Rendah	Tinggi
1350	1200
1200	1000
950	980
1100	970
1000	990
1500	1200
1350	1000
980	1100
1400	1000
1200	1100
1000	980
1000	970

Contoh Soal

Grup	Nilai	Rank
Tinggi	900	1
Rendah	950	2
Rendah	970	4
Tinggi	970	4
Tinggi	970	4
Rendah	980	8
Tinggi	980	8
Rendah	990	12
Tinggi	990	12

Grup	Nilai	Rank
Tinggi	990	12
Tinggi	995	14
Rendah	1000	19.5
Tinggi	1000	19.5

Grup	Nilai	Rank
Rendah	1100	28
Rendah	1100	28
Tinggi	1100	28
Rendah	1115	32.5
Tinggi	1115	32.5
Rendah	1200	37
Rendah	1200	37
Rendah	1200	37

Grup	Nilai	Rank
Rendah	1200	37
Tinggi	1200	37
Tinggi	1200	37
Tinggi	1200	37
Rendah	1350	42
Rendah	1350	42
Rendah	1350	42
Rendah	1400	44.5
Rendah	1400	44.5
Rendah	1450	46
Rendah	1500	47.5
Rendah	1500	47.5

$$n_1 = 24 \quad n_2 = 24$$

- diurutkan dr tertinggi ke \rightarrow rendah

Hipotesis :

$$H_0 = M_x \leq M_y \quad H_1 = M_x > M_y$$

$s = 716$ $s = \text{jumlah peringkat populasi}_1$ (rendah)

$$T = 716 - \frac{24(24+1)}{2} = 416$$

$$Z = \frac{T - \frac{n_1 n_2}{2}}{\sqrt{\frac{n_1 n_2 (n_1 + n_2 + 1)}{12} - \frac{n_1 n_2 (\sum t^3 - \sum t)}{12(n_1 + n_2)(n_1 + n_2 - 1)}}}$$

stat uji =

$$\sum t = 3 + 5 + 3 + \dots + 2 = 44$$

$$\sum t^3 = 3^3 + 5^3 + 3^3 + \dots + 2^3 = 1916$$

$$Z = \frac{T - \frac{n_1 n_2}{2}}{\sqrt{\frac{n_1 n_2 (n_1 + n_2 + 1)}{12} - \frac{n_1 n_2 (\sum t^3 - \sum t)}{12 (n_1 + n_2) (n_1 + n_2 - 1)}}}$$

$$\begin{aligned} T &= \frac{416 - \frac{24 \cdot 24}{2}}{\sqrt{\frac{24 \cdot 24 (24+24+1)}{12} - \frac{24 \cdot 24 (1916 - 44)}{12(24+24)(24+24-1)}}} \\ &= 2166195 \end{aligned}$$

Kaidah penolakan

$$z_{hit} = 2,66$$

$$z_{ta} = 1,64$$

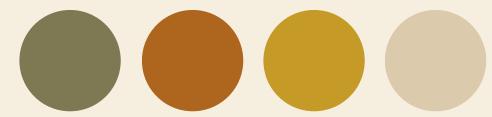
$$z_{hit} > z_{ta} \rightarrow tolak H_0$$

Cukup bukti untuk menyatakan bahwa produktivitas padi di dataran rendah lebih besar dibandingkan dgn dataran tinggi pada taraf nyata 5%



2

**Uji Tanda Data
Berpasangan**



Uji Tanda Data Berpasangan

Asumsi

- Data berpasangan
- Peubah yang diamati adalah selisih data $D = X - Y$
- Antar pasangan saling bebas
- Skala data minimal ordinal

Hipotesis

- a. $H_0 : M = M_0$ vs $H_1 : M \neq M_0$ (dua arah)
- b. $H_0 : M \leq M_0$ vs $H_1 : M > M_0$ (satu arah)
- c. $H_0 : M \geq M_0$ vs $H_1 : M < M_0$ (satu arah)

Prosedur

1. Hitung selisih data berpasangan $D = X - Y$
2. Hitung nilai D yang minus (S^-) dan plus (S^+)
3. Nilai $D=0$ dihilangkan



Uji Tanda Data Berpasangan

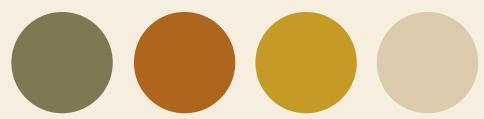
Kaidah Keputusan

Tolak H_0 jika....

- a. $P(X \leq S' | b(n, 0.5)) \leq \alpha/2$
- b. $P(X \leq S^- | b(n, 0.5)) \leq \alpha$
- c. $P(X \leq S^+ | b(n, 0.5)) \leq \alpha$

Statistik Uji

- a. $S = S' = \min(S^-, S^+)$
- b. $S = S^-$
- c. $S = S^+$



Contoh Soal

Uji coba pengaturan suhu ruangan rumah sakit diharapkan dapat menurunkan suhu badan bagi penderita demam. Sebelum pengaturan suhu ruangan dilakukan pengukuran suhu badan awal dan setelah pengaturan suhu ruangan dilakukan pengukuran suhu badan akhir, didapatkan data berikut. Ujilah dengan $\alpha = 10\%$, apakah model pengaturan suhu ruangan mampu menurunkan suhu badan bagi penderita demam

Hipotesis

$$\begin{aligned} H_0 &: Ma \leq Mb \text{ atau } Md \leq 0 \\ H_1 &: Ma > Mb \text{ atau } Md > 0 \end{aligned}$$

No	Awal (a)	Akhir (b)
1	39	38
2	38.5	38.5
3	38.5	37
4	37	39
5	37	37
6	38	38
7	37	38

No	Awal (a)	Akhir (b)
8	38.5	38
9	38	37.5
10	38	37
11	37	37
12	38	38.5
13	39	38
14	37.5	37.5



Contoh Soal

No	a	b	$D = a - b$
1	39	38	1
2	38.5	38.5	0
3	38.5	37	1.5
4	37	39	-2
5	37	37	0
6	38	38	0
7	37	38	-1

No	a	b	$D = a - b$
8	38.5	38	0.5
9	38	37.5	0.5
10	38	37	1
11	37	37	0
12	38	38.5	-0.5
13	39	38	1
14	37.5	37.5	0

Hipotesis
 $H_0: Ma \leq Mb$ atau $Mo \leq 0$
 $H_1: Ma > Mb$ atau $Mo > 0$

- $n = 9$
- $S_+ = 6$
- $S_- = 3$

P-value
 $P(S \leq 3 | b(9, 0.5)) = 0.2539$

P-value
 $0.2539 > 0.1$

p-value > α

Tak tolak H_0

RUMUS EXCEL :

=BINOM.DIST(3, 9, 0.5, TRUE)

= 0.2539

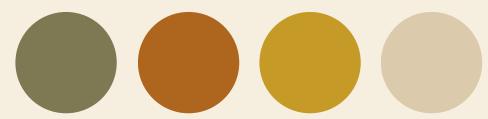
Kaidah Keputusan

TAK MULAK NO



3

Uji Wilcoxon Data Berpasangan



Uji Wilcoxon Data Berpasangan

Asumsi

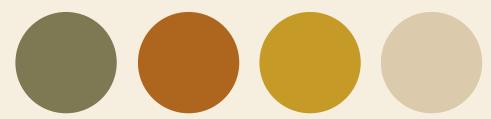
- Data berpasangan dengan selisih $D = Y - X$
- Sebaran populasi dari selisih adalah simetris dengan nilai tengah M_D
- Selisih saling bebas
- Selisih minimal berskala interval

Prosedur

1. Hitung selisih data berpasangan $D = Y - X$
2. Beri peringkat $|D|$
3. Hitung jumlah peringkat bertanda plus (T_+) dan jumlah peringkat bertanda minus (T_-)

Bila n lebih besar dari 30, kita tidak dapat menggunakan tabel A.3 untuk mendapatkan nilai kritis uji statistik. Untuk sampel lebih besar dari 30 dapat dihitung dengan :

$$z = \frac{T_+ - [n(n+1)/4]}{\sqrt{\frac{n(n+1)(2n+1)}{24}}}$$



Uji Wilcoxon Data Berpasangan

1. Hipotesis

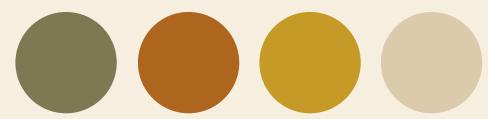
- a. $H_0 : M_x = M_y$ vs $H_1 : M_x \neq M_y$
- b. $H_0 : M_x \leq M_y$ vs $H_1 : M_x > M_y$
- c. $H_0 : M_x \geq M_y$ vs $H_1 : M_x < M_y$

3. Kaidah Penolakan H_0

- a. $T \leq T_{n(\alpha/2)}$
- b. $T \leq T_{n(\alpha)}$
- c. $T \leq T_{n(\alpha)}$

2. Statistik Uji

- a. $T = T' = \min(T^-, T^+)$
- b. $T = T^-$
- c. $T = T^+$



Contoh Soal

BPS Provinsi Jawa Barat telah mengukur IPM pada 10 kota/kabupaten di provinsi Jawa Barat. BPS menyatakan bahwa IPM kota/kabupaten di provinsi Jawa Barat lebih tinggi setelah masa pemerintahan Gubernur X. Ujilah pernyataan tersebut pada taraf nyata 5%.

Sebelum	Sesudah
68.84	77.41
71.16	76.68
68.08	77.39
71.87	76.10
69.99	69.73
72.20	68.88
70.44	67.10
67.52	72.62
66.61	69.46
68.84	70.86

1. Hipotesis

$$H_0 : M_{sesudah} \leq M_{sebelum} \text{ vs } H_1 : M_{sesudah} > M_{sebelum}$$
$$H_0 : M_{selisih} \leq 0 \text{ vs } H_1 : M_{selisih} > 0$$

Contoh Soal

Sebelum	Sesudah	Di	Rank
68.84	77.41	8.57	9
71.16	76.68	5.52	8
68.08	77.39	9.31	10
71.87	76.10	4.23	6
69.99	69.73	-0.26	-1
72.20	68.88	-3.32	-4
70.44	67.10	-3.34	-5
67.52	72.62	5.1	7
66.61	69.46	2.85	3
68.84	70.86	2.02	2

1. Hipotesis

$$H_0 : M_{selisih} = 0 \text{ vs } H_1 : M_{selisih} > 0$$

2. Statistik Uji

$$T = T^- = 10$$

3. Kaidah Penolakan

$$T_{10(0.05)} = 10$$

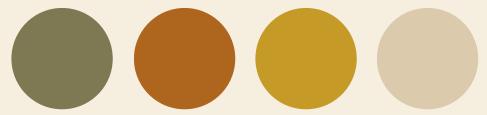
$$T^- \leq T_{n(\alpha)}$$

Cukup bukti untuk menyatakan bahwa IPM kota/kabupaten di provinsi Jawa Barat lebih tinggi setelah masa pemerintahan Gubernur X pada taraf nyata 5%



4

Korelasi
Spearman



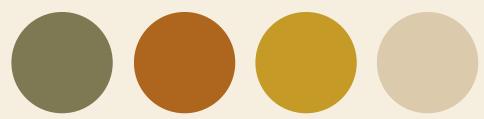
Korelasi Spearman

Asumsi

- 2 peubah (X dan Y) yang diukur dari objek yang sama
- Skala data minimal ordinal
- Data terdiri dari contoh acak

Prosedur

- 1.Urutkan nilai amatan X dari kecil ke besar. Nilai amatan terkecil diberi peringkat 1, dst.
- 2.Urutkan nilai amatan Y.
- 3.Jika ada nilai yang sama, diberi peringkat tengah.
- 4.Hitung selisih peringkat $d=R(X) - R(Y)$.
- 5.Hitung nilai korelasi



Korelasi Spearman

Hipotesis

- a. $H_0: \rho=0$ vs $H_1: \rho \neq 0$ (dua arah)
- b. $H_0: \rho=0$ vs $H_1: \rho > 0$ (satu arah)
- c. $H_0: \rho=0$ vs $H_1: \rho < 0$ (satu arah)

Kaidah Keputusan

Tolak H_0 jika.....

- a. $|r_s| > \text{tabel}(n, \alpha/2)$
- b. $r_s > \text{tabel}(n, \alpha)$
- c. $r_s < \text{tabel}(n, \alpha)$

Statistik Uji

- Jika tidak ada amatan yang sama

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

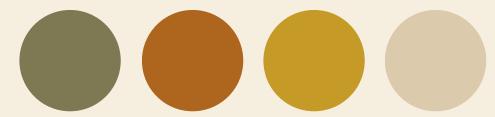
- Jika ada amatan yang sama

$$T_x = \frac{t_x^3 - t_x}{12} \quad T_y = \frac{t_y^3 - t_y}{12}$$

$$\sum x^2 = \frac{n^3 - n}{12} - \sum T_x$$

$$\sum y^2 = \frac{n^3 - n}{12} - \sum T_y$$

$$r_s = \frac{\sum x^2 + \sum y^2 - \sum_{i=1}^n (d_i)^2}{2\sqrt{\sum x^2 \sum y^2}}$$

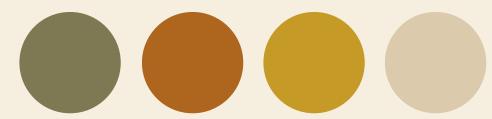


Korelasi Spearman

Pendekatan Sebaran Normal

Jika jumlah data besar ($n > 30$), maka korelasi spearman dapat diaproksimasi menggunakan sebaran normal dengan rumus sebagai berikut

$$z = r_s \sqrt{n - 1} \sim N(0,1)$$



Contoh Soal

Berikut ini adalah data jumlah kehadiran dalam kuliah dan nilai ujian akhir. Hitung korelasi peringkat Spearman antara jumlah kehadiran dalam kuliah dan nilai ujian akhir. Uji apakah kedua peubah tersebut saling bebas?

Kehadiran (X)	13	12	15	15	10	13	15	13	16	16
Nilai Ujian (Y)	53	42	70	69	32	76	73	45	58	45

Hipotesis

$$H_0: \rho = 0$$

$$H_1: \rho \neq 0$$

R(X)	4	2	7	7	1	4	7	4	9.5	9.5
R(Y)	5	2	8	7	1	10	9	3.5	6	3.5
d _i	-1	0	-1	0	0	-6	-2	0.5	3.5	6
d _i ²	1	0	-1	0	0	36	4	0.25	12.25	36



Contoh Soal

R(X)	4	2	7	7	1	4	7	4	9.5	9.5
R(Y)	5	2	8	7	1	10	9	3.5	6	3.5
d _i	-1	0	-1	0	0	-6	-2	0.5	3.5	6
d _i ²	1	0	1	0	0	36	4	0.25	12.25	36

Statistik Uji

$$\sum T_x = \sum \frac{t_x^3 - t_x}{12} = \frac{3^3 - 3}{12} + \frac{3^3 - 3}{12} + \frac{2^3 - 2}{12} = \frac{54}{12}$$

$$\sum T_y = \sum \frac{t_y^3 - t_y}{12} = \frac{2^3 - 2}{12} = \frac{6}{12}$$

$$\sum x^2 = \frac{n^3 - n}{12} - \sum T_x = \frac{10^3 - 10}{12} - \frac{54}{12} \\ = \frac{936}{12} = 78$$

$$\sum y^2 = \frac{n^3 - n}{12} - \sum T_y = \frac{10^3 - 10}{12} - \frac{6}{12} \\ = \frac{984}{12} = 82$$

$$r_s = \frac{\sum x^2 + \sum y^2 - \sum_{i=1}^n d_i^2}{2\sqrt{\sum x^2 \sum y^2}} \\ = \frac{78 + 82 - 90.5}{2\sqrt{(78)(82)}} = 0.4399$$

Keputusan

nilai tabel $(10, 0.025) = 0,6485$

$0,4399 < 0,6485$

stat uji $<$ nilai tabel \rightarrow TAK tolak H_0

Interpretasi :

Tidak cukup bukti untuk menyatakan bahwa ada hubungan antara jumlah kehadiran kuliah dan nilai ujian akhir pada taraf nyata 5%.



Latihan Soal

1

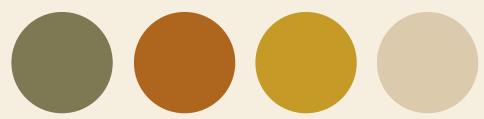
Suatu percobaan dilakukan untuk mengetahui tingkat pengetahuan umum mahasiswa yang sering membaca dan jarang membaca. Tabel berikut menunjukkan skor kuis pengetahuan umum mahasiswa. Dengan data tersebut, apakah benar bahwa skor mahasiswa yang sering membaca lebih tinggi daripada mahasiswa yang jarang membaca?

Sering	95	87	77	73	44	64	68	70	55	59	67	88	89	90	52
Jarang	67	55	51	40	25	18	34	44	52	59	54	53			

2

H. Shani et al. mempelajari pengaruh fenobarbital pada fungsi hati pada pasien dengan Sindrom Dubin-Johnson. Tabel berikut menunjukkan total bilirubin dalam serum pasien sebelum dan setelah perlakuan dengan fenobarbital. Dapatkah kita menyimpulkan atas dasar data ini bahwa fenobarbital mengurangi total tingkat bilirubin?

Pasien	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
Sebelum	4.0	3.2	3.8	1.8	3.0	5.3	5.7	3.0	2.7	2.9	2.8	1.8	2.6	2.8	5.0
Sesudah	3.1	3.0	3.5	1.0	1.8	3.9	2.2	2.1	1.4	2.9	2.6	1.4	2.5	2.3	3.0



Latihan Soal

3

3. Sebuah RS ingin mengetahui efektifitas pelatihan pada staff promosi kesehatan dalam rangka mensukseskan program PKRS. Data skor tingkat pengetahuan dikumpulkan dari 15 staff promosi secara acak untuk mengetahui apakah terdapat perbedaan pengetahuan sebelum mengikuti pelatihan dan sesudah mengikuti pelatihan. Adapun datanya adalah sebagai berikut.

Staff	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O
Sebelum	140	110	100	120	130	120	200	110	140	100	180	240	170	200	220
Sesudah	150	120	100	110	150	140	210	120	160	100	200	250	200	210	220

4

Seorang manager produksi ingin mengetahui apakah ada hubungan antara nilai tes bakat (aptitude test) pada waktu penerimaan kerja dengan performa kerja setelah satu semester. Nilai aptitude test berkisar antara 0 sampai 100. Sedangkan performa kerja mempunyai skala sebagai berikut.

- 1 = pekerja berpenampilan sangat dibawah rata-rata
- 2 = pekerja berpenampilan dibawah rata-rata
- 3 = pekerja berpenampilan sedang (rata-rata)
- 4 = pekerja berpenampilan diatas rata-rata
- 5 = pekerja berpenampilan sangat diatas rata-rata

Nilai tes	59	47	58	66	77	57	62	68	69	36
Peforma kerja	3	2	4	3	2	4	3	3	5	1

Uji apakah kedua peubah tersebut saling berhubungan pada taraf nyata 10%.



**TERIMA
KASIH**