

Uji Rank Mann-Whitney Dua Tahap

TETI SOFIA YANTI

Dosen Jurusan Statistika FMIPA UNISBA.

ABSTRAK

Uji rank Mann-Whitney adalah salah satu bentuk pengujian dalam analisis statistika non parametrik, di mana pengujian digunakan untuk menguji kesamaan distribusi dua populasi yang saling bebas dengan asumsi distribusi dari kedua populasi adalah kontinu dan skala pengukuran dari data minimal ordinal. Uji rank Mann-Whitney dua tahap merupakan perluasan uji rank Mann-Whitney, di mana pengambilan sampel dilakukan sebanyak dua tahap. Jika pada tahap pertama sudah diperoleh keputusan apakah H_0 ditolak atau diterima maka tahap kedua tidak usah dilakukan, dengan demikian uji rank Mann-Whitney dua tahap dianggap lebih efisien dibandingkan uji rank Mann-Whitney karena dapat mereduksi ukuran sampel sehingga dapat menghemat biaya dan waktu.

Kata kunci: Dua sampel saling bebas, uji rank Mann-Whitney

1. Pendahuluan

Uji rank Mann-Whitney adalah salah satu bentuk pengujian dalam analisis statistika non parametrik, di mana pengujian digunakan untuk menguji kesamaan distribusi dua populasi yang saling bebas dengan asumsi distribusi dari kedua populasi adalah kontinu.

Pengujian dilakukan terhadap sampel-sampel yang diambil dari kedua populasi yang saling bebas dengan teknik sampling tertentu. Salah satu faktor melakukan penyampelan adalah factor biaya dan ketersediaan waktu.

Pada uji rank Mann-Whitney proses pengambilan sampel dari kedua populasi dilakukan dalam satu kali pengambilan. Mann dan Whitney (dalam Suprierr dan Hewett, 1976) memperkenalkan uji rank Mann Whitney dua tahap di mana proses pengambilan sampel terhadap kedua populasi dilakukan dalam dua tahap, hal ini dimaksudkan selain dapat mereduksi ukuran sampel dari data yang akan dianalisis dapat juga menghemat waktu dan biaya. Proses penyampelan tahap kedua dilakukan jika hipotesis pengujian pada tahap pertama ditolak atau diterima, tetapi jika menghasilkan keputusan dilanjutkan maka dilakukan pengambilan sampel tahap kedua.

2. Uji Rank Mann-Whitney

Dua populasi yang akan diuji mempunyai ukuran populasi yaitu M untuk populasi pertama dan N untuk populasi kedua, dari masing-masing populasi diambil sampel sebanyak m dan n . Proses pengambilan sampel dilakukan menggunakan sampling random. Randomisasi dapat dilakukan dengan menggunakan bilangan acak. Skala pengukuran untuk uji Rank Mann-Whitney minimal ordinal.

Misalkan X adalah peubah acak dari populasi 1 dengan fungsi distribusi F_x dan Y adalah peubah acak dari populasi 1 dengan fungsi distribusi G_y , di mana populasi 1 dan populasi 2 saling bebas, maka bentuk hipotesis yang akan diuji adalah:

$$H_0: F_x = G_y$$

$$H_1: F_x < G_y$$

Dengan melakukan teknik sampling random diperoleh sampel pada tahap pertama x_1, x_2, \dots, x_{m1} dari populasi 1 dengan fungsi distribusi F_x dan y_1, y_2, \dots, y_{n1} dari populasi 2 dengan fungsi distribusi G_x . Statistik uji yang digunakan untuk menguji hipotesis di atas diklasifikasikan menjadi dua yaitu jika sampel yang terambil dari kedua populasi masing-masing sebesar $m \leq 20$ dan $9 \leq n \leq 20$ disebut dengan pengujian terhadap sampel kecil, sedangkan jika sampel yang terambil dari kedua populasi masing-masing sebesar $m > 20$ dan $n > 20$ disebut dengan pengujian terhadap sampel besar.

Sampel Kecil

Statistik uji yang digunakan adalah pilih nilai U yang terkecil, di mana statistic uji U adalah:

$$U = mn + \frac{m(m+1)}{2} - R_1 \quad (1)$$

atau

$$U = mn + \frac{n(n+1)}{2} - R_2 \quad (2)$$

Di mana: R_1 = Jumlah peringkat pada kelompok sampel pertama

R_2 = Jumlah peringkat pada kelompok sampel kedua.

Pemberian peringkat diurutkan pada sampel secara keseluruhan, nilai terkecil diberi peringkat 1 dan terbesar diberi peringkat (m+n)

Kriteria ujinya adalah:

Tolak H_0 jika $U \leq U_\alpha$, U_α adalah nilai kritis pada uji rank Mann-Whitney

Sampel Besar

Statistik uji yang digunakan untuk sampel besar melalui pendekatan distribusi normal baku yaitu:

$$Z = \frac{U - E(U)}{\sqrt{\text{var}(U)}} = \frac{U - mn/2}{\sqrt{[mn(m+n+1)/12]}} \quad (3)$$

dengan kriteria uji tolak H_0 jika $Z < Z_\alpha$.

3. Prosedur Uji Rank Mann-Whitney Dua Tahap

Teknik pengambilan Sampel

Dua populasi saling bebas yang akan diuji mempunyai ukuran populasi yaitu M untuk populasi pertama dan N untuk populasi kedua, dari masing-masing populasi diambil sampel sebanyak m dan n. Teknik pengambilan Sampel uji rank Mann Whitney dua tahap dilakukan sebanyak dua tahap yaitu:

Tahap 1

Pengambilan sampel tahap satu masing-masing sampel berukuran m_1 dan n_1 di mana perbandingannya sebesar $\frac{m_1}{m} = \frac{n_1}{n} = \frac{3}{5}$. Proses pengambilan sampel dilakukan menggunakan sampling random. Randomisasi dapat dilakukan dengan menggunakan bilangan acak, jika pada tahap pertama belum diambil keputusan hipotesis diterima atau ditolak maka dilanjutkan pada tahap kedua.

Tahap 2

Pada tahap pertama telah diambil sampel dari kedua populasi masing-masing berukuran m_1 dan n_1 , pada tahap kedua perlu dilakukan pengambilan sampel kembali terhadap kedua populasi masing-masing sebesar $m_2 = (m - m_1)$ dan $n_2 = (n - n_1)$, sehingga ukuran sampel yang digunakan pada tahap dua sebesar m dan n..

Uji Rank Mann-Whitney Dua Tahap

Misalkan X adalah peubah acak dari populasi 1 dengan fungsi distribusi F_x dan Y adalah peubah acak dari populasi 2 dengan fungsi distribusi G_y , di mana populasi 1 dan populasi 2 saling bebas, maka bentuk hipotesis yang akan diuji adalah:

$$H_0: F_x = G_y$$

$$H_1: F_x < G_y$$

Dengan melakukan teknik sampling random diperoleh sampel pada tahap pertama x_1, x_2, \dots, x_{m_1} dari populasi 1 dengan fungsi distribusi F_x dan y_1, y_2, \dots, y_{n_1} dari populasi 2

dengan fungsi distribusi Gx. Statistik uji yang digunakan untuk menguji hipotesis di atas diklasifikasikan menjadi dua yaitu jika sampel yang terambil dari kedua populasi masing-masing sebesar $m \leq 8$ atau $n \leq 8$ disebut dengan pengujian terhadap sampel kecil, sedangkan jika sampel yang terambil dari kedua populasi masing-masing sebesar $m > 8$ dan $n > 8$ disebut dengan pengujian terhadap sampel besar.

Sampel Kecil

Statistik uji yang digunakan pada tahap pertama adalah:

$$U_1 = \sum_{i=1}^{m_1} \sum_{j=1}^{n_1} D_{ij} \tag{4}$$

$$\text{Di mana } D_{ij} = \begin{cases} 1, & \text{untuk } x_i < y_j \\ 0, & \text{untuk lainnya} \end{cases}$$

Kriteria ujinya adalah:

- i. Tolak H_0 , jika $U_1 \leq c_1$
- ii. Terima H_0 , jika $U_1 \geq c_2$
- iii. Lakukan pemilihan sampel tahap dua, jika $c_1 < U_1 < c_2$

Lakukan pemilihan sampel tahap dua, jika $c_1 < U_1 < c_2$

Nilai c_1 dan c_2 terdapat pada lampiran.

Jika yang terjadi adalah kriteria (iii) maka harus dilakukan pengambilan sampel tahap dua, dan dilakukan pengujian terhadap sampel tersebut. Pilih sampel tahap kedua yaitu $x_{m_1+1}, x_{m_1+2}, \dots, x_m$ dari populasi 1 dan $y_{m_1+1}, y_{m_1+2}, \dots, y_m$ dari populasi 2. Statistik uji Mann-Whitney dua tahap dengan sampel kecil adalah:

$$U_p = \sum_{i=1}^{m_1+m_2} \sum_{j=1}^{n_1+n_2} D_{ij} \tag{5}$$

$$\text{Di mana } D_{ij} = \begin{cases} 1, & \text{untuk } x_i < y_j \\ 0, & \text{untuk lainnya} \end{cases}$$

Kriteria ujinya adalah:

- iv. Tolak H_0 , jika $U_p \leq c_3$
- v. Terima H_0 , jika $U_p \geq c_3$ Nilai c_3 terdapat pada lampiran.

Sampel Besar

Uji rank Mann-Whitney dua tahap untuk sampel besar (>8) dapat digunakan pendekatan distribusi normal baku (Spurierr dan Hewet, 1976). Statistik ujinya adalah sebagai berikut:

$$V_1 = \frac{U_1 - E(U_1)}{\sqrt{\text{var}(U_1)}} = \frac{U_1 - m_1 n_1 / 2}{\sqrt{\left[\frac{m_1 n_1 (m_1 + n_1 + 1)}{12} \right]}} \tag{6}$$

Kriteria ujinya adalah:

- i. Tolak H_0 , jika $V_1 \leq c_1$
- ii. Terima H_0 , jika $V_1 \geq c_2$

Lakukan pemilihan sampel tahap dua, jika $c_1 < V_1 < c_2$

Jika yang terjadi adalah kriteria (iii) maka harus dilakukan pengambilan sampel tahap dua, dan dilakukan pengujian terhadap sampel tersebut. Statistik uji Mann-Whitney dua tahap dengan sampel besar adalah:

$$V_1 = \frac{U_1 - E(U_1)}{\sqrt{\text{var}(U_1)}} = \frac{U_1 - m_1 n_1 / 2}{\sqrt{[m_1 n_1 (m_1 + n_1 + 1) / 12]}} \quad (7)$$

Di mana: $m_p = m_1 + m_2$ dan $n_p = n_1 + n_2$

Kriteria ujinya adalah:

- iii. Tolak H_0 , jika $V_p \leq c_3$
- iv. Terima H_0 , jika $V_p \geq c_3$

Nilai c_3 terdapat pada lampiran.

4. Aplikasi

Data yang digunakan adalah jumlah gula darah (mg/dl) penderita stroke infark berdasarkan jenis kelamin, dimana ingin diketahui apakah jumlah gula darah penderita stroke infark laki-laki dan perempuan berbeda atau tidak. Data diperoleh dari hasil penelitian yang dilakukan di Poliklinik dan Rawat Inap di bagian/SMF Ilmu Penyakit Saraf Fakultas Kedokteran UNPAD/RSHS Bandung.

Berdasarkan teknik sampling acak sederhana sampel untuk laki-laki sebanyak 5 dan untuk perempuan sebanyak 10. Data sampel terdapat pada table 1.

Tabel 1 Data Jumlah Gula Darah (mg/dl) Penderita Stroke Infark Berdasarkan Jenis Kelamin

No	Laki-laki	Perempuan
1	67	94
2	94	71
3	105	92
4	102	66
5	72	137
6		95
7		82
8		98
9		68
10		77

Sumber: Poliklinik dan Rawat Inap Bag. Ilmu Penyakit Saraf FK UNPAD

Hipotesis yang akan diuji adalah:

$H_0: F_x = G_y$; jumlah gula darah penderita stroke infark laki-laki dan perempuan tidak berbeda.

$H_1: F_x < G_y$

Uji Rank Mann-Whitney

Statistik uji yang digunakan adalah pilih nilai U yang terkecil, di mana statistik uji U adalah:

$$U = 5 \times 8 + \frac{5(5+1)}{2} - 44 = 11 \quad \text{atau} \quad U = 5 \times 8 + \frac{10(10+1)}{2} - 76 = 19, \text{ sehingga nilai statistik}$$

ujinya sebesar 11. Berdasarkan kriteria uji tolak H_0 jika $U < U_\alpha$, karena nilai kritis (U_α) sebesar 11 maka H_0 diterima.

Uji Rank Mann-Whitney Dua Tahap

Pada tahap pertama diambil sampel dengan perbandingan $m_1/m = n_1/n = 3/5$, sehingga untuk laki-laki sebanyak 3 orang penderita dan untuk perempuan diambil sampel 6 orang, diperoleh data sebagai berikut:

Tabel 3.1 Data Jumlah Gula Darah (mg/dl) Penderita Stroke Infark Berdasarkan Jenis Kelamin Tahap 1

No	Laki-laki	Perempuan
1	105	82
2	94	71
3	67	77
4		66
5		137
6		68

Sumber: Poliklinik dan Rawat Inap Bag. Ilmu Penyakit Saraf FK UNPAD

Statistik uji yang digunakan pada tahap pertama adalah:

$$U_1 = \sum_{i=1}^{3_1} \sum_{j=1}^{6_1} D_{ij}$$

$$\text{Di mana } D_{ij} = \begin{cases} 1, & \text{untuk } x_i < y_j \\ 0, & \text{untuk lainnya} \end{cases}$$

$$\begin{aligned} U_1 &= [(D_{11}+D_{12}+D_{13}+D_{14}+D_{15}+D_{16}) + (D_{21}+D_{22}+D_{23}+D_{24}+D_{25}+D_{26}) \\ &\quad + (D_{31}+D_{32}+D_{33}+D_{34}+D_{35}+D_{36})] \\ &= [(0+0+0+0+1+0) + (0+0+0+0+1+0) + (1+1+1+0+1+1)] = 7 \end{aligned}$$

Berdasarkan tabel nilai kritis yang diberikan oleh Spurrierr dan Hewett dengan perbandingan $m_1/m = n_1/n = 3/5$ diperoleh nilai kritis untuk $c_1=1.886$ dan untuk $c_2=0.710$. Berdasarkan kriteria uji yang ada karena $U_1 > c_2$ maka H_0 diterima yang berarti jumlah gula darah penderita stroke infark laki-laki dan perempuan tidak berbeda.

5. Kesimpulan

- 1) Dengan menggunakan uji Rank Mann-Whitney, di mana ukuran sampel untuk kelompok pertama sebesar 5 dan ukuran sampel kelompok kedua sebesar 10 diperoleh hasil hipotesis nol diterima.
- 2) Dengan menggunakan uji Rank Mann-Whitney dua tahap, dengan pengujian tahap pertama di mana ukuran sampel untuk kelompok pertama sebesar 3 dan ukuran sampel kelompok kedua sebesar 6 diperoleh hasil hipotesis nol diterima.
- 3) Pada uji Rank Mann-Whitney dua tahap, dengan pengujian tahap pertama sudah bisa diambil suatu keputusan yaitu hipotesis nol ditolak. Hal ini menunjukkan bahwa uji Rank Mann-Whitney dua tahap lebih efisien dibandingkan uji Rank Mann-Whitney karena uji Rank Mann-Whitney dua tahap dapat mereduksi ukuran sampel yang berimplikasi waktu yang digunakan dan biaya yang dikeluarkan lebih sedikit juga teknik perhitungan lebih cepat.

6. Daftar Pustaka

Conover, W.J., 1980, Practical Nonparametric Statistic. Second Edition, John Wiley&Son, Newyork Chichester, Brisbane-Toronto-Singapore
 Daniel, Wayne W ,1989, Statistik Nonparametrik Terapan. PT Gramedia, Jakarta
 Spurrier, J.D. and Hewett, J.E., 1975, Double Sampel Test for the Mean of a Normal Population, Journal of the American Statistical association.

Spurrier, J.D. and Hewett, J.E., 1976, Two Stage Wilcoxon Test of Hypothesis, Journal of the American Statistical association.

Lampiran

c_1, c_2, c_3 dan Maximum ESSR for Double Sampel Test in the Limiting Case dengan $\alpha=0.05$ dan 0.01

$\alpha=0.05$						
Design No.	m1/m	n1/n	c1	c2	c3	Max(ESSR)
1	3/5	3/5	1.886	0.710	1.783	0.866
2	2/5	4/5	1.984	0.179	1.782	0.907
3	1/5	1	2.073	-0.482	1.784	0.999
4	55/100	55/100	2.050	0.438	1.716	0.869
5	2/3	2/3	1.781	0.950	1.868	0.882
6	7/10	7/10	1.749	1.045	1.909	0.893

$\alpha=0.01$						
Design No.	m1/m	n1/n	c1	c2	c3	Max(ESSR)
1	3/5	3/5	2.499	1.259	2.493	0.879
2	2/5	4/5	2.558	0.635	2.496	0.931
3	1/5	1	2.600	-0.146	2.502	1.030
4	55/100	55/100	2.635	0.966	2.411	0.878
5	2/3	2/3	2.415	1.520	2.600	0.897
6	7/10	7/10	2.390	-1.828	2.651	0.908

Spurrier, J.D. and Hewett, J.E., 1975, Double Sampel Test for the Mean of a Normal Population, Journal of the American Statistical association