



Modul 8

FPM 226-Metodologi Penelitian Fisioterapi II

Materi 8

Non-Parametric Test

Disusun Oleh

Wahyuddin

UNIVERSITAS ESA UNGGUL

2020

Pendahuluan

Dalam pembahasan-pembahasan sebelumnya telah disajikan beberapa uji statistik yang didasarkan pada asumsi-asumsi tertentu tentang parameter populasi terkait pengambilan sampel. Tes parametrik ini didasarkan pada asumsi terpenuhinya normalitas dan keseragaman untuk analisis validitas. Pada materi ini, akan disajikan serangkaian prosedur statistik yang diklasifikasikan sebagai nonparametric, yang menguji hipotesis perbandingan kelompok tanpa asumsi normalitas atau homogenitas varians. Secara umum, non-parametric sama dengan parametric dalam hal penggunaan rasio/tes statistik dan penentuan nilai alpha untuk level signifikansi

Beberapa prosedur nonparametric yang paling sering digunakan ditampilkan pada tabel berikut ini:

Comparison	Parametric Test	Nonparametric Test
Two independent groups	Unpaired <i>t</i> -test	Mann-Whitney <i>U</i> test
Two related scores	Paired <i>t</i> -test	Sign test Wilcoxon signed-ranks test (<i>T</i>)
Three or more independent groups	One-way analysis of variance (<i>F</i>)	Kruskal-Wallis analysis of variance by ranks (<i>H</i> or χ^2)
Three or more related scores	One-way repeated measures analysis of variance (<i>F</i>)	Friedman two-way analysis of variance by ranks (χ^2)

Tabel 1: Beberapa Jenis Tes Parametrik

Kriteria Pemilihan Non-Parametric Test

Dua kriteria utama umumnya diadopsi untuk memilih tes non-parametric test. Yang pertama adalah bahwa asumsi normalitas populasi dan keseragaman varians Penyelidikan klinis telah dipelajari untuk mendukung asumsi ini. Kriteria kedua untuk memilih non parametric test adalah bahwa data diukur pada skala nominal atau ordinal. Banyak alat-alat penilaian telah dikembangkan berdasarkan skala ini. Non parametric test menyediakan mekanisme yang objektif untuk mendukung hipotesis statistik berdasarkan penggunaan level pengukuran tersebut.

Power pada Non-Parametric Test

Banyak peneliti lebih memilih untuk menggunakan parametric test karena lebih powerfull. Non-parametric test kurang sensitif dibandingkan parametric test karena umumnya meliputi nilai peringkat dibandingkan dengan membandingkan perubahan metrik yang tepat. Metode nonparametric dan parametric telah dibandingkan berdasarkan efisiensi power, yang merupakan sebuah tes kemampuan relatif untuk mengidentifikasi perbedaan yang signifikan terkait ukuran sampel.

Umumnya, peningkatan ukuran sampel yang dibutuhkan untuk membuat nonparametric test sekuat parametric test. Sebagai contoh, nonparametric test mungkin memerlukan ukuran sampel 50 untuk mencapai power yang sama dengan parametric test dengan jumlah sampel 30. Hubungan ini dapat dinyatakan sebagai persentase yang menunjukkan efisiensi power relatif nonparametric test. Misalnya, jika power 60%, maka dengan ukuran sampel yang sama, power nonparametric test adalah 60% terhadap parametric test. Dengan kata lain, untuk mencapai power yang sama dengan nonparametric test, dibutuhkan 10 subjek berbanding 6 subjek pada prosedur parametric.

Sign Test and Wilcoxon Signed Rank Test

Dua prosedur yang umum digunakan untuk menguji perbedaan antara sampel yang berkorelasi yaitu sign test and Wilcoxon signed rank test. Tes ini digunakan dengan with two level repeated measures designs. Tes-tes tersebut analog dengan parametric t-test untuk sampel yang berkorelasi atau berpasangan.

Sign Test

Sign test adalah nonparametric test paling sederhana karena tidak memerlukan perhitungan matematis. Tes ini digunakan pada data binomial, dan tidak memerlukan pengukuran kuantitatif. Seperti namanya, data dianalisis menggunakan tanda plus dan minus. Oleh karena itu, tes ini menyediakan sebuah mekanisme untuk pengujian relatif seperti lebih-kurang, lebih tinggi-lebih rendah atau lebih besar-lebih kecil.

Misalnya studi efek sudut knee terhadap kekuatan ekstensor knee. Menggunakan manual muscle test (MMT), dilakukan studi pada 10 pasien, enam bulan

setelah total knee replacement. Grade MMT dicatat dari 0 (tidak ada aktivitas otot) sampai 12 (kekuatan normal). Hipotesis peneliti bahwa kekuatan ekstensor knee akan berbeda pada fleksi 90° dan 15°. Data hasil pengukuran tercantum pada gambar 3.

A. DATA						
Subject	Angle		Sign	d	Rank of d	Ranks with less frequent sign
	90°	15°				
1	8	8	0	0		
2	10	11	-	-1	-1	-1
3	7	7	0	0		
4	9	7	+	+2	+3	
5	10	8	+	+2	+3	
6	11	7	+	+4	+7	
7	10	8	+	+2	+3	
8	10	7	+	+3	+5.5	
9	8	8	0	0		
10	10	7	+	+3	+5.5	
						T = -1

B. HYPOTHESIS TEST/OUTPUT FOR THE SIGN TEST			
6 plus signs, 1 minus sign, x = 1 (number of fewer signs)		Frequencies	
For x = 1, n = 7, p = .124 (Table A.9)		90-15	N
$z = \frac{ D - 1}{\sqrt{n}} = \frac{ 5 - 1}{\sqrt{7}} = 1.51$		Negative Differences (a)	1
Do not reject H_0 .		Positive Differences (b)	6
		Ties (c)	3
		Total	10
		(a) 90 < 15 (b) 90 > 15 (c) 90 = 15	
		Exact Sig. (2-tailed)	.125

C. HYPOTHESIS TEST/OUTPUT FOR THE SIGNED-RANKS TEST			
Sum ranks with less frequent sign = -1		For n = 7 at $\alpha_2 = .05$, T = 2 (Table A. 10) Reject H_0	
$z = \frac{T - \frac{n(n+1)}{4}}{\sqrt{\frac{n(n+1)(2n+1)}{24}}} = \frac{1 - \frac{7(7+1)}{4}}{\sqrt{\frac{7(7+1)(2(7)+1)}{24}}} = \frac{-13}{\sqrt{35}} = -2.20$			
Ranks		Test Statistics	
90-15	N	Mean Rank	Sum of Ranks
Negative Ranks	1 (a)	1.00	1.00
Positive Ranks	6 (b)	4.50	27.00
Ties	3 (c)		
Total	10		
Z Asymp. Sig. (2-tailed)	-2.217 .027		

Gambar 1. Data dan Output Sign Test

Wilcoxon Signed-Ranks Test

Sign test mengevaluasi perbedaan dalam nilai berpasangan berdasarkan apakah suatu skor lebih besar atau lebih kecil daripada yang lain. Hal ini sering menjadi pendekatan yang terbaik dengan variabel klinik yang subjektif. Namun, untuk dapat memberikan informasi tentang perbedaan relatif, Wilcoxon signed-ranks test dapat digunakan. Tes ini mengkaji dua aspek yaitu arah perbedaan dan jumlah relatif perbedaan.

Misalnya berdasarkan data pada gambar 1, kita dapat menolak null hypothesis dan membuat konklusi bahwa terdapat perbedaan kekuatan ekstensor knee terkait dengan perbedaan sudut dimana sudut 90° lebih besar disbanding sudut 15°

Mann-Whitney U-Test

Mann-Whitney U-test adalah salah satu prosedur nonparametric yang powerful, dirancang untuk menguji null hypothesis bahwa dua sampel independen berasal dari populasi yang sama. Tes ini analog dengan parametric t-test untuk sampel yang independen. Seperti unpaired t-test, tes ini tidak memerlukan ukuran sampel yang sama pada setiap kelompok. Karena itu, tes ini menjadi alternatif yang sangat baik untuk t-test ketika asumsi parametrik tidak terpenuhi.

Misalnya seorang peneliti tertarik dengan efek posisi tubuh terhadap kemampuan rileks seseorang, yang diukur dengan EMG biofeedback pada otot frontalis. 11 subjek direkrut secara acak dan dibagi dalam dua kelompok dengan desain pretest-posttest, dimana satu kelompok diposisikan telentang dan kelompok lain dalam posisi duduk. Hasil rekaman perubahan dalam aktivitas microvolt. Hipotesis peneliti bahwa perbedaan posisi akan memfasilitasi berbagai tingkat relaksasi. Hasil pengukuran dan analisis data dapat dilihat pada gambar 2 berikut ini:

TABLE 22.3 MANN-WHITNEY U TEST: CHANGE IN MICROVOLT ACTIVITY FOLLOWING RELAXATION IN TWO POSITIONS

A. DATA

Supine ($n_1 = 5$)	Rank	Sitting ($n_2 = 6$)	Rank
20	5	10	3
30	7	5	2
50	11	35	8
45	10	25	6
40	9	0	1
		15	4
	$R_1 = 42$		$R_2 = 24$

B. COMPUTATIONS

$$U_1 = R_1 - \frac{n_1(n_1 + 1)}{2} = 42 - \frac{5(5 + 1)}{2} = 27$$

$$U_2 = R_2 - \frac{n_2(n_2 + 1)}{2} = 24 - \frac{6(6 + 1)}{2} = 3 \quad U = 3, \text{ the smaller of the values of } U_1 \text{ or } U_2$$

C. TEST FOR LARGE SAMPLES

$$z = \frac{U - \frac{n_1 n_2}{2}}{\sqrt{\frac{n_1 n_2 (n_1 + n_2 + 1)}{12}}} = \frac{3 - \frac{(5)(6)}{2}}{\sqrt{\frac{(5)(6)(5 + 6 + 1)}{12}}} = \frac{-12}{\sqrt{30}} = -2.19$$

D. HYPOTHESIS TEST

Critical Value of U: For $n_1 = 5, n_2 = 6$, at $\alpha_2 = .05$, $U = 3$ (Table A.8) **Reject H_0 .**

z test: $z = -2.19 > 1.96$ (two-tailed $p = .0286$) (Table A.1) **Reject H_0 .**

E. OUTPUT

Ranks				Test Statistics		
GROUP	N	Mean Rank	Sum of Ranks		ACTIVITY	
ACTIVITY	Supine	5	8.40	42.00	Mann-Whitney U	3.000
	Sitting	6	4.00	24.00	Z	-2.191
	Total	11			Asymp. Sig. (2-tailed)	.028

Gambar 2. Data dan Output Pengukuran dengan Mann-Whitney U-Test

Kruskal-Wallis One-Way Analysis of Variance (ANOVA)

Ketika terdapat tiga atau lebih kelompok dibandingkan ($k \geq 3$), analisis varians nonparametric analysis of variance (ANOVA) merupakan tes yang tepat. Kruskal-Wallis one-way analysis of variance (ANOVA) adalah nonparametric test dari one-way analysis of variance (ANOVA). Ini adalah alternatif yang kuat ketika asumsi variance dan normalitas untuk parametric test tidak terpenuhi. Ini juga cara yang paling tepat untuk menangani data pada level ordinal ketika membandingkan data lebih dari dua kelompok. Dengan $k=2$, tes ini setara Mann-Whitney U-test.

Contoh kita ingin mempelajari pengaruh tiga modalitas untuk menghilangkan nyeri punggung bawah kronis. Secara acak 17 subjek ($N = 17$) menerima intervensi ice ($n = 6$), hot pack ($n = 6$), atau ultrasound ($n = 5$). Nyeri diukur menggunakan visual analogue scale dari 0 mm (tanpa nyeri) sampai 100 mm (nyeri tidak tertahankan). Skor dicatat sebagai perubahan level nyeri dari pre treatment ke post treatment.

Prosedur dari data sampai output dapat dilihat pada gambar 3 berikut ini:

TABLE 22.4 KRUSKAL-WALLIS ONE-WAY ANALYSIS OF VARIANCE BY RANKS: CHANGES IN LEVEL OF PAIN ($N = 14$)

A. DATA

	Group 1: Ice		Group 2: Hot Packs		Group 3: Ultrasound	
	Change score	Rank	Change score	Rank	Change score	Rank
	40	8	35	6	80	14
	60	11	25	2.5	50	10
	10	1	30	4.5	75	13
	25	2.5	40	8	70	12
	30	4.5	40	8		
R		27		29		49
n		5		5		4
\bar{R}		5.4		5.8		12.25

B. COMPUTATIONS

$$H = \frac{12}{N(N+1)} \sum \frac{R^2}{n} - 3(N+1)$$

$$= \frac{12}{14(14+1)} \left[\frac{(27)^2}{5} + \frac{(29)^2}{5} + \frac{(49)^2}{4} \right] - 3(14+1)$$

$$= \frac{12}{210} [914.25] - 45 = 7.243$$

C. HYPOTHESIS TEST

For $df = 2$, $\chi^2 = 5.99$ (Table A.5) Reject H_0

D. OUTPUT

Ranks				Test Statistics	
	GROUP	N	Mean Rank		PAIN
PAIN	Ice	5	5.40	Chi-Square	7.340
	HP	5	5.80	df	2
	US	4	12.25	Asymp. Sig.	.025
	Total	14			

Gambar 3. Data dan Output Pengukuran dengan Kruskal-Wallis One-Way Analysis of Variance (ANOVA)

Beberapa peneliti mungkin akan menghentikan pada tahap ini, berdasarkan pada perbandingan mean rank setiap grup. Pada contoh di atas terlihat skor. Namun demikian, bila dianggap belum cukup, prosedur multiple comparison dilakukan untuk menentukan perbedaan.

Multiple Comparison Kruskal-Wallis One-Way Analysis of Variance (ANOVA)

Untuk menentukan grup mana yang signifikan terhadap grup lain, dapat dilakukan multiple comparison, misalnya dengan Bonferroni.

Friedman Two-way Analysis of Variance (ANOVA)

Tes ini bertujuan untuk menganalisis data dari single-factor repeated measures design dengan tiga atau lebih kondisi eksperimental. Friedman two-way analysis of variance by ranks adalah alternatif yang kuat untuk parametric repeated measures ANOVA ketika menggunakan data ordinal atau ketika asumsi parametric tidak terpenuhi. Tes ini diberi nama "two-way" didasarkan pada interpretasi bahwa "subjek" diperlakukan sebagai variabel independen dengan $n=1$. Hal ini diasumsikan bahwa jumlah pengukuran di masing-masing kondisi eksperimental adalah sama.

Contoh suatu penelitian akan mengukur efek dari perubahan posisi tubuh pada tekanan darah pada enam pasien dengan penyakit paru kronik. Setiap pasien akan ditempatkan di tiga level posisi, kepala ke bawah dan kepala ditinggikan-dalam urutan acak. Tekanan darah akan diukur dalam jarak 1 menit dari setiap posisi. Kita dapat memilih untuk menggunakan analisis nonparametric untuk penelitian ini karena jumlah sampel kecil, dan karena kita tidak punya cukup alasan untuk mengasumsikan bahwa tekanan darah untuk populasi pasien dengan penyakit ini akan berada pada distribusi normal. Selain itu, meskipun pengukuran tekanan darah dapat dianggap data level rasio, kita dapat merasionalisasi bahwa rendahnya reliabilitas data menjadi alasan untuk menggunakan nonparametric test. Data hasil pengukuran sampai output digambarkan pada gambar 4 berikut:

A. DATA						
Subject	(1) Level		(2) Elevated		(3) Down	
	BP	Rank	BP	Rank	BP	Rank
1	110	1	150	2	175	3
2	100	1.5	100	1.5	110	3
3	120	1	140	3	135	2
4	110	1	130	2	155	3
5	120	1	130	2	145	3
6	130	1	155	2	170	3
	<i>R</i>	6.5		12.5		17
	<i>R</i>	1.08		2.08		2.83
	<i>R</i> ²	4225		156.25		289.00

B. COMPUTATIONS	
$\chi_r^2 = \frac{12}{nk(k+1)} \sum R^2 - 3n(k+1)$	
$= \frac{12}{(6)(3)(3+1)} [4225 + 156.25 + 289] - 3(6)(3+1)$	
$= \frac{12}{72} [487.5] - 72 = 9.25$	

C. HYPOTHESIS TEST/OUTPUT			
For $k = 3$ ($df = 2$)		Ranks	
$\chi^2 = 5.99$ at $\alpha = .05$ (Table A.5)		Test Statistics	
Reject H_0 .		Mean Rank	
DOWN	2.83	N	6
ELEV	2.08	Chi-Square	9.652
LEVEL	1.08	df	2
		Asymp. Sig.	.008

Gambar 4: Data Friedman Two-way Analysis of Variance (ANOVA)

Kesimpulan

Prosedur nonparametric prosedur menawarkan kepada peneliti klinis suatu mekanisme statistik yang kuat dan mudah dipahami untuk menganalisis perubahan yang diukur dengan menggunakan alat-alat yang bersifat subjektif. Karena banyaknya penilaian klinis, kemampuan untuk menganalisis data ordinal menjadi sangat penting. Masih ada beberapa perdebatan di kalangan ahli statistik dan peneliti tentang aplikasi statistik yang sesuai dari parametric versus nonparametric pada data yang berbentuk ordinal.

Perspektif sebelumnya mengatakan bahwa prosedur nonparametric harus digunakan dengan pada pengukuran dengan data ordinal, namun, banyak peneliti menggunakan parametric test untuk data ordinal. Hal ini mungkin karena parametric test memiliki power statistik yang lebih besar. Hal ini diperkuat oleh asumsi bahwa interval ordinal bersifat konsisten, meskipun sensitifitas pengukurannya relatif rendah.

Oleh karena itu, analisis akan secara konseptual tidak sesuai dengan asumsi parametric test.

Meskipun beberapa skala penilaian dapat dikonstruksi sedemikian rupa untuk menentukan interval setepat mungkin, hal ini mungkin masuk akal untuk mengasumsikan bahwa konstruksi seperti fungsi, resistensi secara manual, sensasi, dan sebagainya, biasanya diukur sebagai peringkat, juga dapat diukur dengan reliabilitas yang cukup bahwa interval dapat dianggap sama. Hal ini juga mungkin bahwa banyak dari skala nonlinier, sehingga interval pada skala ekstrem akan berbeda. Mereka yang menggunakan pendekatan ini atau yang menafsirkan temuan lain harus mempertimbangkan potensi terkait validitas hasil statistik dengan menggunakan data ordinal sebagai data interval.

Metode nonparametric juga sesuai untuk digunakan pada data interval atau rasio ketika distribusi miring (kanan atau kiri) atau ketika ukuran sampel terlalu kecil untuk merepresentasikan distribusi normal. Oleh karena itu, ketika kriteria untuk menggunakan tes nonparametric didasarkan pada normalitas saja, mungkin berguna untuk mengubah data menggunakan transformasi logaritma untuk menghasilkan data yang distribusi normal dan menerapkan parametric test.

Tes-tes yang telah dibahas hanya sebuah contoh prosedur nonparametric yang tersedia. Ahli statistik terus mengembangkan dan memperbaiki tes ini dan memperluas kemampuan metode nonparametric ke bidang regresi dan faktorial desain. Banyak tes telah dikembangkan dengan tujuan yang sangat spesifik, seperti membandingkan beberapa kelompok pengobatan dengan kontrol tunggal atau melihat perbedaan dalam variabel yang melekat. Statistik nonparametric juga dapat digunakan untuk prosedur korelasi dan data pengujian skala nominal.

Referensi

Conover WJ. *Practical Nonparametric Statistics* (3rd ed.). New York: John Wiley & Sons, 1999.

O'Connell AA. *Logistic Regression Models for Ordinal Response Variables*. Thousand Oaks, CA: Sage Publications, 2005.

Winer BJ, Michels KM, Brown DR. Statistical Principles in Experimental Design (3rd ed.). New York: McGraw-Hill, 1991.

Neave HR, Granger WJ. A Monte Carlo study comparing various two-sample tests for differences in means. *Technometrics* 1968;10:509-522.

Siegel S, Castellan NJ. Nonparametric Statistics for the Behavioral Sciences (2nd ed.). New York: McGraw-Hill, 1988.

Daniel WW. Applied Nonparametric Statistics (2nd ed.). Belmont, CA: Wadsworth Publishing, 1989.

Wang ST, Yu ML, Wang CJ, Huang CC. Bridging the gap between the pros and cons in treating ordinal scales as interval scales from an analysis point of view. *Nurs Res* 1999;48:226-229.

Gaito J. Measurement scales and statistics: Resurgence of an old misconception. *Psycho Bull* 1980;87:564-567.

Royeen CB, Seaver WL. Promise in nonparametrics. *Am J Occup Ther* 1986;40:191-193.