



Modul 4

FPM 226-Methodologi Penelitian Fisioterapi II

Materi 4

Perbandingan 2 Mean: t-Test

Disusun Oleh

Wahyuddin

UNIVERSITAS ESA UNGGUL

2020

Pendahuluan

Perbandingan eksperimental sederhana melibatkan penggunaan dua kelompok independen yang dilakukan secara acak. Desain ini memungkinkan peneliti untuk berasumsi bahwa semua perbedaan individual merata antara kelompok, sehingga setiap kelompok setara di awal percobaan. Secara statistik, kelompok yang dianggap sampel acak dari populasi yang sama, dan oleh karena itu, diamati perbedaan antara mereka harus menjadi hasil dari sampling error atau karena faktor kebetulan. Setelah aplikasi variabel intervensi ke satu grup, peneliti ingin menentukan jika kelompok-kelompok ini masih dari populasi yang sama, atau jika mean diantara mereka dianggap berbeda secara signifikan.

Perbandingan juga dapat dibuat menggunakan desain pengukuran berulang. Peneliti mungkin tertarik untuk melihat perbedaan antara dua kondisi atau performa oleh subjek pada grup yang sama. Dalam kasus ini, subjek mengontrol diri mereka sendiri, dan peneliti ingin menentukan jika terdapat perbedaan kondisi.

Konsep Dasar Perbandingan Mean Grup

Konsep dasar signifikansi secara statistik untuk membandingkan mean berdasarkan hubungan antara karakteristik dua sampel yaitu mean dan varians. Perbedaan antara mean grup berarti menunjukkan tingkat pemisahan antara kelompok (effect size). Pengukuran varians menginformasikan kepada kita bagaimana variabel nilai dalam setiap kelompok. Kedua karakteristik ini mewakili sumber variabilitas yang digunakan untuk menggambarkan tingkat efek treatment. Misalkan kita ingin membandingkan dua kelompok yang secara acak, satu eksperimental dan satu kontrol, untuk menentukan jika treatment didasarkan pada perbedaan kinerja.

Secara teoritis, jika treatment ini efektif, dan faktor-faktor lain yang sama dan konstan, semua subjek dalam kelompok treatment akan menghasilkan skor yang sama, begitu juga pada kelompok kontrol, tetapi akan terdapat perbedaan antara kelompok. Dengan demikian perbedaan tersebut dapat diasumsikan sebagai dampak treatment yang diberikan, tetapi bukan karena tidak adanya varians dalam kelompok.

Error Variance

Jika kita melihat suatu set data untuk mengambil kesimpulan, terdapat kecenderungan bahwa nilai atau skor pada kelompok treatment lebih tinggi dibanding kelompok kontrol. Namun demikian, skor dapat dipengaruhi oleh beberapa faktor seperti karakteristik personal, inkonsistensi pengukuran, perilaku dan lingkungan. Faktor-faktor tersebut menyebabkan varians dalam kelompok yang tidak dapat dijelaskan akibat faktor-faktor yang tidak dapat diketahui. Ini yang disebut error variance.

Konsep kesalahan statistik tidak berarti terdapat kesalahan perhitungan. Ini merujuk kepada semua sumber variabilitas dalam suatu dataset yang tidak dapat dijelaskan melalui variabel bebas. Semua skor merupakan gabungan dari efek treatment dan error variance. Metode randomisasi memberikan asumsi kepada kita bahwa adanya kesalahan merupakan variasi yang tidak sistematis adalah pengaruh dari efek treatment

Faktor-faktor selain variabel treatment yang menyebabkan respon subyek sangat berbeda antara satu dengan yang lain. Di sini kita menemukan banyak tumpang tindih, menunjukkan bahwa banyak subjek dari kedua kelompok memiliki skor yang sama, terlepas dari apakah mereka menerima treatment seperti pada kelompok eksperimen atau kontrol .

Rasio Statistik

Pertimbangan subyektif tentang distribusi terkadang tidak memadai, untuk membuat keputusan penelitian tentang efektivitas treatment. Kita dapat mengetahui perbedaan antar kelompok dapat terjadi karena faktor kebetulan. Jadi bagaimana kita mengetahui secara objektif untuk menentukan jika perbedaan antara kelompok adalah akibat perbedaan populasi atau perbedaan karena faktor kebetulan? Dengan kata lain, bagaimana kita memutuskan jika kita harus menolak null hypothesis? Kita dapat membuat keputusan ini berdasarkan probabilitas yang benar. Ini menjadi dasar rancangan terkait signifikansi secara statistik. Pentingnya perbedaan antara kelompok berarti dinilai dengan rasio berasal sebagai berikut:

rasio = perbedaan antara mean kelompok berarti/variabilitas dalam kelompok

Jika terjadi peningkatan efek treatment nilai absolut rasio tersebut semakin besar.

Jika error variance meningkat, rasio semakin kecil, mendekati 1. Jika kita ingin menunjukkan bahwa dua kelompok berbeda secara signifikan, rasio ini harus sebesar mungkin. Dengan demikian, kita ingin pemisahan mean antar kelompok besar dan variabilitas dalam kelompok kecil. Kita menekankan varians antara dan dalam kelompok sebagai elemen penting makna pengujian yang akan digunakan secara berulang-ulang dan banyak uji statistik didasarkan pada hubungan ini.

Hipotesis Statistik

Desain statistik hipotesis untuk null hypothesis menyatakan bahwa dua populasi sama sebagai berikut:

$$H_0: \mu_1 = \mu_2$$

Sementara untuk alternative dapat dinyatakan dalam nondirectional format seperti berikut ini:

$$H_1: \mu_1 \neq \mu_2$$

atau juga dalam bentuk directional format:

$$H_1: \mu_1 > \mu_2 \text{ or } H_1: \mu_1 < \mu_2$$

Nondirectional hypothesis menggunakan tingkat signifikansi two-tailed test. Meskipun secara aktual kita membandingkan mean sampel, hipotesis dibuat menggunakan parameter populasi.

Equality of Variances

Statistik parametrik memerlukan asumsi equal variances antar grup atau homogenitas varians. Dengan kata lain jika dua kelompok memiliki varians berbeda, maka asumsi homogenitas varians tidak terpenuhi. Banyak prosedur statistik yang membandingkan mean termasuk tes yang akan menentukan apakah perbedaan dalam

komponen varians signifikan. Kita dapat mengharapkan beberapa perbedaan dalam varians secara kebetulan. Dengan metode acak, semakin besar jumlah sampel akan memiliki peluang yang lebih baik dibanding jumlah sampel yang kecil.

T-Test untuk Independent Samples

T-test adalah prosedur statistik yang digunakan untuk membandingkan dua means. Independent/unpaired t-test digunakan ketika dua grup independent subjek dibandingkan. Grup-grup tersebut dipilih berdasarkan metode acak, meskipun dapat menggunakan metode convenience. Grup dipertimbangkan sebagai independent karena setiap komponen subjek tidak berhubungan dari hasil pengukuran.

T Statistic: Equal Variances

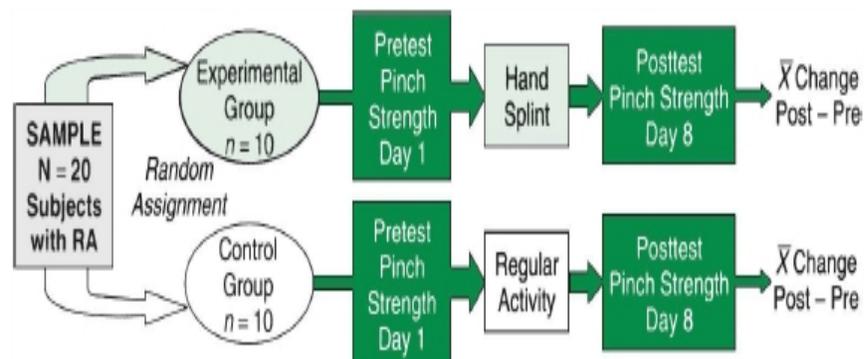
Tes statistik untuk unpaired t-test dihitung dengan menggunakan rumus sebagai berikut:

$$t = \frac{\bar{X}_1 - \bar{X}_2}{s\sqrt{\bar{X}_1 - \bar{X}_2}}$$

Gambar 1. Rumus unpaired t-test

Degrees of Freedom

Jumlah degrees of freedom diasosiasikan dengan independent t-test adalah jumlah untuk kedua grup. Misalnya kita tertarik untuk melakukan tes terkait hipotesis bahwa suatu jenis splint model baru akan meningkatkan fungsi tangan pada pasien dengan rheumatoid arthritis, yang diukur dengan menggunakan pinch strength seperti pada gambar berikut:



Gambar 2. Contoh desain studi eksperimental

Kita menggunakan alternative hypothesis karena lebih menekankan pada peningkatan fungsi. Hasil yang menunjukkan penurunan fungsi atau tidak ada perubahan akan disimpulkan sebagai tidak signifikan. Data hasil pengukuran ditampilkan sebagai berikut:

TABLE 19.1 COMPUTATION OF THE UNPAIRED t-TEST (EQUAL VARIANCES): CHANGE IN PINCH STRENGTH FOLLOWING HAND SPLINTING

A. DATA	Group 1 (splint)		Group 2 (Control)	
	$\bar{X}_1 = 10.11$		$\bar{X}_2 = 5.45$	
	$n_1 = 10$		$n_2 = 10$	
	$s_1^2 = 13.81$		$s_2^2 = 15.58$	

B. COMPUTATIONS	
$s_p^2 = \frac{s_1^2(n_1 - 1) + s_2^2(n_2 - 1)}{n_1 + n_2 - 2} = \frac{13.81(10 - 1) + 15.58(10 - 1)}{10 + 10 - 2} = 14.695$	
$s_{\bar{X}_1 - \bar{X}_2} = \sqrt{\frac{s_p^2}{n_1} + \frac{s_p^2}{n_2}} = \sqrt{\frac{(14.695)}{10} + \frac{14.695}{10}} = 1.714 \quad t = \frac{\bar{X}_1 - \bar{X}_2}{s_{\bar{X}_1 - \bar{X}_2}} = \frac{10.11 - 5.45}{1.7143} = 2.718$	

C. HYPOTHESIS TEST	
$H_0: \mu_1 = \mu_2$	$H_1: \mu_1 > \mu_2$
Reject H_0	$(\alpha_1 = .05)t_{(18)} = 1.734$ (Table A.2)

D. OUTPUT					
Group Statistics					
GROUP	N	Mean	Std. Deviation	Std. Error Mean	
STRENGTH 1	10	10.1100	3.7159	1.1751	
2	10	5.4500	3.9472	1.2482	

Independent Samples Test									
	Levene's Test for Equality of Variances		t-test for Equality of Means						
	F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	95% Confidence Interval of the Mean	
								Lower	Upper
Equal variances assumed	.685	① .419	2.718	18	② .014	③ 4.6600	④ 1.7143	⑤ 1.0584	8.2616
Equal variances not assumed			2.718	17.935	.014	4.6600	1.7143	1.0575	8.2625

① Levene's test compares the variances of the two groups (see s^2 in section A). This difference is not significant ($p = .419$). Therefore, we will use the t-test for equal variances.

② The two-tailed significance is .014. Because this analysis was based on a one-tailed test (directional alternative hypothesis), we use half of the two-tailed value. Therefore, $p = .007$. This is a significant test, and we reject H_0 .

③ The difference between the means of group 1 and 2 (the numerator of the t-test).

④ The standard error of the difference between the means (denominator of the t-test).

Hasil Analisis

Hasil pada tabel menunjukkan unpaired t-test terkait strength dan hand splints. Ada beberapa informasi yang di pertimbangkan. Pertama, output menyediakan ringkasan statistik deskriptif yang terkait dengan masing-masing kelompok. Informasi ini berguna pada tahap awal untuk mengkonfirmasi bahwa jumlah subyek, dan untuk melihat nilai mean dan varians. Begitu pula dengan nilai standar deviasi untuk mendapatkan nilai varians.

Selanjutnya, terdapat dua baris output untuk independent samples test, diberi label menurut asumsi varians yang sama. Program secara otomatis melakukan prosedur t-tes untuk varians yang sama atau tidak sehingga kita perlu menentukan di awal, dan peneliti harus memilih mana yang harus digunakan untuk analisis. Kolom berlabel Levene's Test for Equality of Variances akan menginformasikan apakah varians berbeda secara signifikan. Mengacu pada probabilitas yang terkait dengan Levene's test pada tabel dimana nilai $p=0.419$. Nilai ini lebih besar dari 0.05 , dan kita akan menyimpulkan bahwa varians tidak ada perbedaan signifikan atau sama. Oleh karena itu, kita akan menggunakan baris pertama dari data untuk "Equal variances assumed."

Berdasarkan data pada tabel, kita melakukan one-tailed test. Output komputer hanya melaporkan 2-tailed significance. Oleh karena itu, untuk mendapatkan two-tailed probability kita membagi 0.014 dibagi 2 , atau nilai $p = 0.007$ untuk tes ini. Karena nilai ini kurang dari 0.05 , kita menolak null hypothesis.

Confidence Intervals

Confidence interval adalah interval atau rentang skor mean populasi. Secara umum kita dapat menggunakan 95% CI. Sekarang cek bagaimana kita dapat menggunakan confidence interval untuk menarik kesimpulan. Meskipun dalam studi tersebut menggunakan directional hypothesis, secara definisi confidence interval melihat pada nilai two-tailed test:

$$\begin{aligned} 95\% \text{ CI} &= (10.11 - 5.45) \pm 2.101(1.714) \\ &= 4.66 \pm 3.60 \end{aligned}$$

$$= 1.06, 8.26$$

Dengan perhitungan tersebut, secara konfiden perbedaan mean terletak antara nilai 1.06 dan 8.26.

The Null Value

Nilai null hypothesis menyatakan bahwa perbedaan antara dua cara akan menjadi nol. Jika kita melihat hati-hati 95% confidence interval data ini, kita melihat bahwa nilai null, nol, tidak terkandung di dalamnya. Seperti kita yakin bahwa perbedaan berarti benar terletak di suatu tempat dalam interval ini 95%, tidak mungkin bahwa perbedaan berarti benar adalah nol. Oleh karena itu, kita dapat cukup menolak H₀. Ini menegaskan hasil t-tes untuk ini data yang sama.

TABLE 19.2 COMPUTATION OF THE UNPAIRED t-TEST (UNEQUAL VARIANCES): CHANGE IN PINCH STRENGTH FOLLOWING HAND SPLINTING

A. DATA	Group 1 (splint)	Group 2 (Control)
	$\bar{X}_1 = 10.80$	$\bar{X}_2 = 5.65$
	$n_1 = 15$	$n_2 = 10$
	$s_1^2 = 25.17$	$s_2^2 = 4.89$

B. COMPUTATIONS

$$t = \frac{\bar{X}_1 - \bar{X}_2}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}} = \frac{10.80 - 5.65}{\sqrt{\frac{25.17}{15} + \frac{4.89}{10}}} = \frac{5.15}{1.472} = 3.498 \quad (19.4)$$

C. HYPOTHESIS TEST $H_0: \mu_1 = \mu_2$ $H_1: \mu_1 > \mu_2$ $(\alpha = .05)_{(20,6)} \approx 1.723$ (Table A.2)
 Reject H_0

D. OUTPUT

Group Statistics					
	GROUP	N	Mean	Std. Deviation	Std. Error Mean
STRENGTH	1	15	10.8000	5.0171	1.2954
	2	10	5.6500	2.2117	.6994

Independent Samples Test

	Levene's Test for Equality of Variances		t-test for Equality of Means						
	F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	95% Confidence Interval of the Mean	
								Lower	Upper
Equal variances assumed	4.866	.038	3.039	23	.006	5.150	1.6949	1.6439	8.6561
Equal variances not assumed			3.498	20.625	.002	5.150	1.4722	2.0851	8.2149

❶ Levene's test is significant ($p = .038$), indicating that the variance of group 1 is significantly different from the variance of group 2. Therefore, we will use the t -test for unequal variances.
 ❷ With 25 subjects, the adjusted total degrees of freedom for the test of unequal variances is 20.625.
 ❸ The two-tailed probability for the t -test is .002. This analysis was based on a one-tailed test (directional alternative hypothesis). The one-tailed probability level is half of the two-tailed value. Therefore, $p = .001$. This is significant, and we reject H_0 .
 ❹ Difference between the means of group 1 and 2 (numerator of the t -test).
 ❺ Standard error of the difference (denominator of the t -test).

T-test pada paired samples

Peneliti sering menggunakan langkah-langkah yang berulang-ulang atau disesuaikan desain untuk meningkatkan tingkat kontrol atas asing variabel dalam studi. Dalam desain ini subjek mungkin cocok pada variabel yang relevan, seperti usia, tingkat intelegensia, atau variabel lain yang berpotensi berkorelasi dengan dependent variable. Dalam jenis penelitian ini, data dianggap berpasangan atau berkorelasi, karena setiap pengukuran memiliki nilai yang cocok untuk setiap subjek. Untuk menentukan apakah nilai-nilai ini sangat berbeda dari satu sama lain, dilakukan paired t-tes. Tes ini menganalisis perbedaan nilai dalam setiap pasangan, sehingga skor subjek dibandingkan hanya dengan diri mereka sendiri atau dengan pertandingan mereka.

Secara statistik, ini memiliki efek mengurangi varians total kesalahan dalam data karena sebagian besar faktor-faktor tambahan yang mempengaruhi data akan sama di pada kedua kondidi treatmen. Oleh karena itu, tes signifikansi melibatkan perbandingan berpasangan cenderung lebih kuat dibanding unpaired test. Contoh sebuah studi untuk menguji efek penggunaan lumbar support pillow terhadap posisi angular pelvic pada posisi rileks. Hipotesis yang ditegakkan adalah pelvic tilting akan berubah dengan menggunakan pillow (nondirectional hypothesis).

Analisis data diperoleh sebagai berikut:

TABLE 19.3 COMPUTATION OF THE PAIRED *t*-TEST: ANGLE OF PELVIS WITH AND WITHOUT A LUMBAR SUPPORT PILLOW

A. DATA			
Subject	X_1 (without pillow)	X_2 (with pillow)	d
1	108	112	-4
2	102	96	6
3	98	105	-7
4	112	110	2
5	100	106	-6
6	85	98	-13
7	92	90	2
8	95	102	-7
	$\Sigma X_1 = 819$ $\bar{X}_1 = 102.375$	$\Sigma X_2 = 792$ $\bar{X}_2 = 99.00$	$\Sigma d = -27$ $\bar{d} = -3.375$ $s_d = 6.232$

B. COMPUTATIONS			
$s_{\bar{d}} = \frac{s_d}{\sqrt{n}} = \frac{6.232}{\sqrt{8}} = 2.203$	$t = \frac{\bar{d}}{s_{\bar{d}}} = \frac{-3.375}{2.203} = -1.532$		

C. HYPOTHESIS TEST			
$H_0: \mu_1 = \mu_2$	$H_0: \mu_1 \neq \mu_2$	$(\alpha_2 = .05)^{t(7)} = 2.365$ (Table A.2)	
Do not reject H_0			

D. OUTPUT								
Paired Samples Statistics				Paired Samples Correlations				
	Mean	N	Std. Deviation	Std. Error Mean		N	Corr	Sig.
Pair 1 PILLOW	99.0000	8	8.63548	3.05310		8	.708	.049
Pair 1 NONE	102.3750	8	7.40536	2.61819				

Paired Samples Test								
	Paired Differences				95% Confidence Interval of the Difference			
	Mean	Std. Deviation	Std. Error Mean	Lower				Upper
	Pair 1 NONE-PILLOW	② -3.37500	③ 6.23212	2.20339	⑤ -8.5818	1.83518	-1.532	7

- ① The analysis includes the correlation of pillow and non-pillow scores and significance of that correlation.
- ② Mean of the difference scores, \bar{d} (numerator of the *t*-test).
- ③ Standard error of the difference scores, $s_{\bar{d}}$ (denominator of the *t*-test)
- ④ Because we proposed a non-directional alternative hypothesis, this two-tailed value is used. This test is not significant.
- ⑤ 95% confidence interval of the difference scores contains zero, indicating no significant difference.

Referensi:

Keppel G. Design and Analysis: A Researcher's Handbook (4th ed.). Englewood Cliffs, NJ:Prentice Hall, 2004.

Dallal GE. Multiple comparison procedures. Available at: <<http://www.tufts.edu/>>

-gdallal/mc.htm> Accessed February 20, 2006.

May K. A note on the use of confidence intervals. *Understanding Statistics*
2003;2:133-135.

Leslie G. Portney, Mary P. Watkins. *Foundations of Clinical Research, Applications to
Practice* 2nd ed. F.A Davis Company, Philadelphia 2015