Two Samples Test (Uji Beda Dua Kelompok)

Ali Muhson, M.Pd.

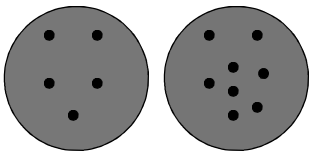
Jenisnya

- Uji Beda Rata-rata
  - Uji z
  - Uji t
- Uji Beda Proporsi
  - Uji z

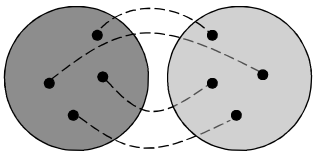
Jenis Uji Beda Rata-rata dua kelompok

- Dua Kelompok Saling Bebas (Independent Samples):
  - Uji z untuk uji populasi
  - Uji t untuk sampel kecil:
    - Pooled t test
    - Separate t test
- Dua kelompok berpasangan (Paired samples)
  - Paired samples t test

Illustrasi dua kelompok

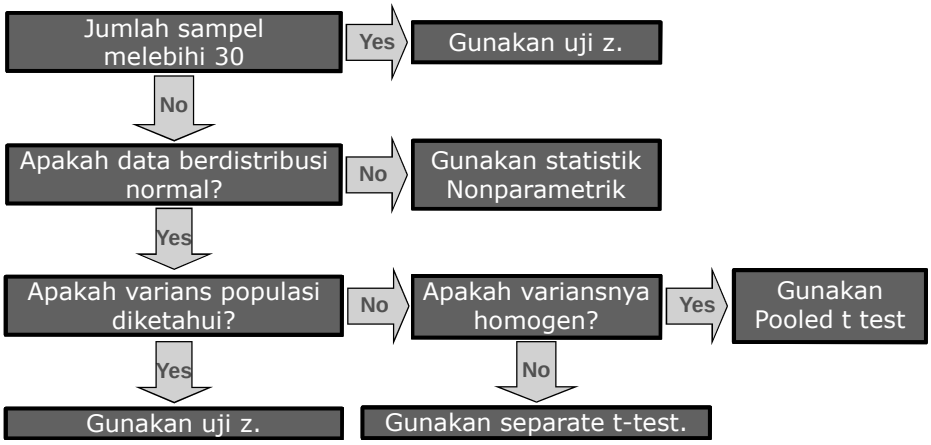


Dua Kelompok saling bebas



Dua Kelompok Berpasangan

Jenis uji dua kelompok saling bebas



Uji z Dua Kelompok Saling Bebas

- Tujuan → Menguji perbedaan rerata dua kelompok populasi yang saling bebas.
- Syarat:
  - Sampel harus diambil secara random
  - Data berskala interval
  - Data pada masing-masing kelompok berdistribusi normal
  - Dua kelompok tersebut tidak saling berhubungan
  - Varians populasi diketahui, atau sampel minimal 30

Contoh Masalah

- Apakah ada perbedaan tinggi badan antara mahasiswa pria dan wanita?
- Benarkah bahwa hasil ujian siswa kelas A lebih baik daripada kelas B?
- Apakah benar bahwa mobil dengan sistem injeksi lebih hemat BBM daripada yang tidak menggunakan sistem injeksi?

$$\begin{cases} H_0: \mu_1 = \mu_2 \\ H_a: \mu_1 \neq \mu_2 \end{cases}$$

$$\begin{cases} H_0: \mu_1 \leq \mu_2 \\ H_a: \mu_1 > \mu_2 \end{cases}$$

$$\begin{cases} H_0: \mu_1 \geq \mu_2 \\ H_a: \mu_1 < \mu_2 \end{cases}$$

Uji z dua kelompok

Rumus

$$z = \frac{\overline{X}_1 - \overline{X}_2}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}$$

Standar Error Perbedaan Rerata

$(\sigma_{\overline{X}_1 - \overline{X}_2})$

Nilai Kritis (z tabel)

- Z (1-α)

Estimasi Parameter

- Dengan tingkat keyakinan (1-α) tertentu dapat estimasi nilai parameter.

$$\left(\overline{X}_1 - \overline{X}_2\right) - z_{1-\alpha} SE \le \left(\mu_1 - \mu_2\right) \le \left(\overline{X}_1 - \overline{X}_2\right) + z_{1-\alpha} SE$$

Margin Error (ME)

Pooled t test

- Tujuan → Menguji perbedaan rerata dua kelompok populasi yang saling bebas.
- Syarat:
  - Sampel harus diambil secara random
  - Data berskala interval
  - Data pada masing-masing kelompok berdistribusi normal
  - Dua kelompok tersebut tidak saling berhubungan
  - Varians populasi tidak diketahui, atau sampel kurang dari 30
  - Varians kedua kelompok bersifat homogen

Pooled t test

$$t = \frac{\overline{X}_1 - \overline{X}_2}{S_p\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}$$

Standar Error Perbedaan Rerata $(S_{\overline{X}_1 - \overline{X}_2})$

$$S_p = \sqrt{\frac{(n_1 - 1)SD_1^2 + (n_2 - 1)SD_2^2}{n_1 + n_2 - 2}}$$

- Nilai kritis t(α; n_1+n_2-2)

Estimasi Parameter

- Dengan tingkat keyakinan (1-α) tertentu dapat estimasi nilai parameter.

$$\left(\overline{X}_1 - \overline{X}_2\right) - (t)(SE) \le (\mu_1 - \mu_2) \le \left(\overline{X}_1 - \overline{X}_2\right) + (t)(SE)$$

Margin Error (ME)

Separate t test

- Tujuan → Menguji perbedaan rerata dua kelompok populasi yang saling bebas.
- Syarat:
  - Sampel harus diambil secara random
  - Data berskala interval
  - Data pada masing-masing kelompok berdistribusi normal
  - Dua kelompok tersebut tidak saling berhubungan
  - Varians populasi tidak diketahui, atau sampel kurang dari 30
  - Varians kedua kelompok bersifat tidak homogen

Separate t test

$$t = \frac{\overline{X}_1 - \overline{X}_2}{\sqrt{\frac{SD_1^2}{n_1} + \frac{SD_2^2}{n_2}}}$$

$$t\ tabel = \frac{w_1 t_1 + w_2 t_2}{w_1 + w_2}$$

$$w_1 = \frac{SD_1^2}{n_1}$$

$$w_2 = \frac{SD_2^2}{n_2}$$

- T_1 :
  - Alpha (α)
  - db = $n_1 - 1$
- T_2 :
  - Alpha (α)
  - db = $n_2 - 1$

Alternatif menghitung t tabel

- Dengan menghitung rumus db sebagai berikut:

$$df = \frac{\left(\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}\right)^2}{\frac{1}{n_1 - 1}\left(\frac{s_1^2}{n_1}\right)^2 + \frac{1}{n_2 - 1}\left(\frac{s_2^2}{n_2}\right)^2}$$

Estimasi Parameter

- Dengan tingkat keyakinan (1-α) tertentu dapat estimasi nilai parameter.

$$\left(\overline{X}_1 - \overline{X}_2\right) - (t)(SE) \le \left(\mu_1 - \mu_2\right) \le \left(\overline{X}_1 - \overline{X}_2\right) + (t)(SE)$$

Margin Error (ME)

Uji Homogenitas Varians

$$F = \frac{SD_B^2}{SD_K^2}$$

- F tabel:
  - Alpha (α)
  - $db_1 = n_b - 1$
  - $db_2 = n_k - 1$
  - Biasa ditulis F(α; n_b–1; n_k–1)

Contoh Soal

- Seorang guru Ekonomi di SMA menyatakan bahwa siswa kelas akselerasi memiliki nilai yang lebih tinggi dibandingkan dengan kelas reguler. Guna membuktikan pernyataan tersebut dilakukan penelitian dengan mengambil sampel secara acak di kedua kelas tersebut. Dari 20 sampel siswa kelas akselerasi diperoleh nilai rerata 82,4 dengan standar deviasi 5,3 sementara 25 sampel siswa di kelas reguler diperoleh nilai rerata 76,8 dengan standar deviasi 8,1. Dengan α = 0,05, benarkah pernyataan guru tersebut?

Contoh Hasil Analisis

	Kelas	N	Mean	Std. Deviation	Std. Error Mean
IPK	A	7	3.0871	.28459	.10756
	B	10	2.7760	.29463	.09317

Contoh Hasil Analisis

		Levene's Test for Equality of Variances		t-test for Equality of Means						
									95% Confidence Interval of the Difference	
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	Lower	Upper
IPK	Equal variances assumed	.305	.589	2.172	15	.046	.3111	.14324	.00584	.61644
	Equal variances not assumed			2.186	13.365	.047	.3111	.14230	.00456	.61772

Paired t test

- Tujuan → Menguji perbedaan rerata dua kelompok populasi yang berpasangan.
- Syarat:
  - Sampel harus diambil secara random
  - Data berskala interval
  - Data pada masing-masing kelompok berdistribusi normal
  - Dua kelompok tersebut saling berhubungan atau berpasangan

Paired t test

$$t = \frac{\bar{d}}{\frac{SD_d}{\sqrt{n}}}$$

$$\bar{d} = \bar{X}_1 - \bar{X}_2$$

$$\bar{d} = \frac{\sum d}{n}$$

$$d = X_1 - X_2$$

- Nilai Kritis t:
  - Alpha (α)
  - db = n - 1

Estimasi Parameter

- Dengan tingkat keyakinan (1-α) tertentu dapat estimasi nilai parameter.

$$\left(\overline{X}_1 - \overline{X}_2\right) - (t)(SE) \le (\mu_1 - \mu_2) \le \left(\overline{X}_1 - \overline{X}_2\right) + (t)(SE)$$

Margin Error (ME)

Paired t test

Example:
A reading center claims that students will perform better on a standardized reading test after going through the reading course offered by their center. The table shows the reading scores of 6 students before and after the course. At $\alpha = 0.05$, is there enough evidence to conclude that the students' scores after the course are better than the scores before the course?

Student	1	2	3	4	5	6
Score (before)	85	96	70	76	81	78
Score (after)	88	85	89	86	92	89

H_0: $\mu_d \le 0$

H_a: $\mu_d > 0$ (Claim)

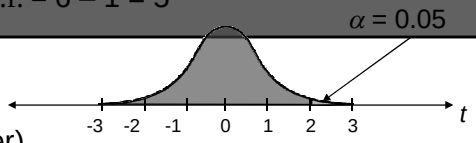
Continued.

t-Test for the Difference Between Means

d.f. = 6 – 1 = 5

H_0: $\mu_d \leq 0$

H_a: $\mu_d > 0$ (Claim)



$t_0 = 2.015$

d = (score before) – (score after)

Student	1	2	3	4	5	6
Score (before)	85	96	70	76	81	78
Score (after)	88	85	89	86	92	89
d	–3	11	–19	–10	–11	–11
d^2	9	121	361	100	121	121

$\sum d = -43$

$\sum d^2 = 833$

$$\bar{d} = \frac{\sum d}{n} = \frac{-43}{6} \approx -7.167$$

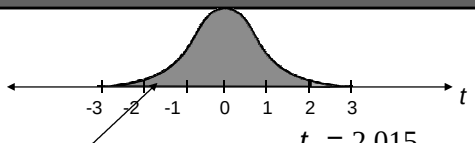
$$s_d = \sqrt{\frac{n(\sum d^2) - (\sum d)^2}{n(n-1)}} = \sqrt{\frac{6(833) - 1849}{6(5)}} \approx \sqrt{104.967} \approx 10.245$$

Continued.

t-Test for the Difference Between Means

H_0: $\mu_d \leq 0$

H_a: $\mu_d > 0$ (Claim)



$t_0 = 2.015$

The standardized test statistic is

$$t = \frac{\bar{d} - \mu_d}{s_d / \sqrt{n}} = \frac{-7.167 - 0}{10.245 / \sqrt{6}} \approx -1.714.$$

Fail to reject H_0.

There is not enough evidence at the 5% level to support the claim that the students' scores after the course are better than the scores before the course.

Hasil Analisis:

		Mean	N	Std. Deviation	Std. Error Mean
Pair 1	Nilai Pre Test	54.4000	10	7.33636	2.31996
	Nilai Post Test	68.7000	10	8.75658	2.76908

		Paired Differences					t	df	Sig. (2-tailed)
		Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference				
					Lower	Upper			
Pair 1	Nilai Pre Test - Nilai Post Test	-14.300	8.94489	2.82862	-20.699	-7.9012	-5.055	9	.001

Uji Beda Proporsi

- Tujuan → menguji perbedaan proporsi dua kelompok
- Syarat:
  - Sampel harus diambil secara random
  - Dua kelompok tersebut saling bebas
  - Jumlah sampel harus cukup besar untuk dapat menggunakan distribusi normal, yaitu:

$$n_1p_1 \geq 5,\ n_1q_1 \geq 5,$$

$$n_2p_2 \geq 5,\ \text{dan}\ n_2q_2 \geq 5.$$

Uji Beda Proporsi

$$z = \frac{p_1 - p_2}{\sqrt{pq\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}}$$

$$p_1 = \frac{X_1}{n_1}$$

$$p_2 = \frac{X_2}{n_2}$$

$$p = \frac{X_1 + X_2}{n_1 + n_2}$$

$$q = 1 - p$$

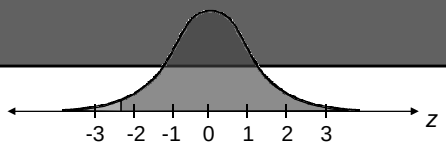
Example

- A recent survey stated that male college students smoke less than female college students. In a survey of 1245 male students, 361 said they smoke at least one pack of cigarettes a day. In a survey of 1065 female students, 341 said they smoke at least one pack a day. At α = 0,01, can you support the claim that the proportion of male college students who smoke at least one pack of cigarettes a day is lower then the proportion of female college students who smoke at least one pack a day?

Two Sample z-Test for Proportions

H_0: $p_1 \geq p_2$

H_a: $p_1 < p_2$ (Claim)



$-z_0 = -2.33$

$$\bar{p} = \frac{x_1 + x_2}{n_1 + n_2} = \frac{n_1\hat{p}_1 + n_2\hat{p}_2}{n_1 + n_2} = \frac{361 + 341}{1245 + 1065} = \frac{702}{2310} \approx 0.304$$

$$\bar{q} = 1 - \bar{p} = 1 - 0.304 = 0.696$$

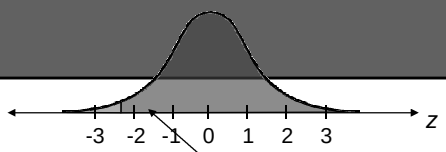
Because 1245(0.304), 1245(0.696), 1065(0.304), and 1065(0.696) are all at least 5, we can use a two-sample *z*-test.

Continued.

Two Sample z-Test for Proportions

H_0: $p_1 \geq p_2$

H_a: $p_1 < p_2$ (Claim)



$-z_0 = -2.33$

$$z = \frac{p_1 - p_2}{\sqrt{pq\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}} = \frac{(0.29 - 0.32) - 0}{\sqrt{(0.304)(0.696)\left(\frac{1}{1245} + \frac{1}{1065}\right)}} \approx -1.56$$

Fail to reject H_0.

There is not enough evidence at the 1% level to support the claim that the proportion of male college students who smoke is lower then the proportion of female college students who smoke.

Home Work

- Halaman 446 soal nomor 16
- Halaman 458 soal nomor 42