Uji Hipotesa Dua Sampel

Tjipto Juwono, Ph.D.

November 2017

SURYA UNIVERSITY

Membandingkan Dua Populasi

Contoh

1. Apakah ada perbedaan jumlah rata-rata penjualan rumah oleh agen pria dan agen wanita?
2. Apakah ada perbedaan jumlah rata-rata produk cacat antara shift pagi dan shift malam?
3. Apakah ada perbedaan jumlah rata-rata absen pada pekerja muda (di bawah 21 tahun) dan pekerja senior (di atas 40 tahun)?
4. Apakah ada perbedaan proporsi mahasiswa yang lulus dari UI dan yang lulus dari ITB dalam 10 tahun terakhir?

Membandingkan dua populasi

Dua populasi, masing-masing di ambil satu sample, sehingga kita mempunyai dua sample.

- Ukuran sample: n_1, n_2
- Standard deviasi: σ_1, σ_2
- Hubungan antara kedua populasi: *Independent* atau *Dependent*.
- Mean dari masing-masing sample: $\bar{X}_1$, $\bar{X}_2$.

Membandingkan dua populasi

Standard Deviasi Diketahui

Digunakan jika $n_1,\ n_2 > 30$ **atau** σ_1 dan σ_2 diketahui

$$Z = \frac{\bar{X}_1 - \bar{X}_2}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}} \tag{1}$$

Membandingkan dua populasi

Standard Deviasi Tidak Diketahui

Digunakan jika $n_1,\ n_2 > 30$ **dan** σ_1 dan σ_2 tidak diketahui

$$Z = \frac{\bar{X}_1 - \bar{X}_2}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}} \tag{2}$$

Contoh 1

Contoh

Peralatan U-Scan baru saja dipasang di sebuah supermarket. Manager ingin mengetahui apakah waktu checkout rata-rata memang lebih panjang jika menggunakan metode checkout standard. Gunakan level of significance 0.01.

Type	Sample Mean	Pop. St. Dev	Sample Size
Standard	5.5 minutes	0.4 minutes	50
U-Scan	5.3 minutes	0.3 minutes	100

Tabel 1: Perbandingan metode standard dan U-Scan

Contoh 1

1. Nyatakan Null Hypothesis dan Alternate Hypothesis:

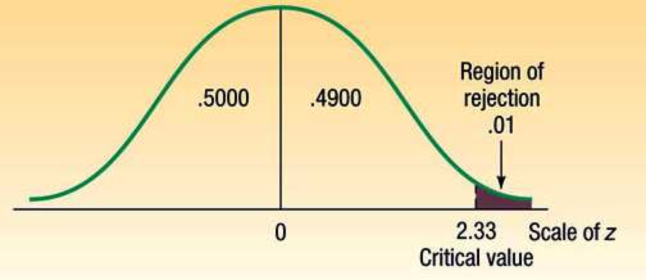
$$\begin{aligned} H_0 : \mu_{\mathrm{S}} &\leq \mu_{\mathrm{U}} \\ H_1 : \mu_{\mathrm{S}} &> \mu_{\mathrm{U}} \end{aligned} \tag{3}$$

2. Tentukan level of significance.
   Dari soal: $\alpha = 0.01$
3. Tentukan Statistik. Karena $\sigma_1, \ \sigma_2$ diketahui, maka kita gunakan distribusi-Z.

Contoh 1

4. Tentukan aturan pengambilan keputusan. H_0 ditolak, jika:

$$\begin{aligned} Z &> Z_\alpha \\ Z &> 2.33 \end{aligned} \qquad (4)$$



Contoh 1

5. Hitung Z dan ambil keputusan.

$$\begin{aligned} Z &= \frac{\bar{X}_s - \bar{X}_u}{\sqrt{\frac{\sigma_s^2}{n_s} + \frac{\sigma_u^2}{n_u}}} \\ &= \frac{5.5 - 5.3}{\sqrt{\frac{0.4^2}{50} + \frac{0.3^2}{100}}} \\ &= \frac{0.2}{0.064} \\ &= 3.13 \end{aligned} \tag{5}$$

Kesimpulan: $Z > Z_\alpha \rightarrow 3.13 > 2.33$
Keputusan: H_0 ditolak. Metode U-Scan lebih cepat.

Uji Dua Sampel: Proporsi

Uji Proporsi

Kita menyelidiki apakah dua sampel berasal dari dua populasi yang mempunyai proporsi sukses yang sama.

Uji Dua Sampel: Proporsi

Uji Proporsi

Proporsi dari kedua sampel disatukan menjadi:

$$p_c = \frac{x_1 + x_2}{n_1 + n_2} \tag{6}$$

Dengan: X_1 adalah jumlah yang mempunyai sifat yang diselidiki pada sampel pertama
X_2 adalah jumlah yang mempunyai sifat yang diselidiki pada sampel kedua
n_1 adalah jumlah observasi pada sampel pertama
n_2 adalah jumlah observasi pada sampel kedua

Uji Proporsi

$$\begin{aligned} Z &= \frac{p_1 - p_2}{\sqrt{\frac{p_c(1-p_c)}{n_1} + \frac{p_c(1-p_c)}{n_2}}} \\ p_c &= \frac{x_1 + x_2}{n_1 + n_2} \\ p_1 &= \frac{x_1}{n_1} \\ p_2 &= \frac{x_2}{n_2} \end{aligned} \tag{7}$$

Contoh 2: Uji Proporsi

Uji Proporsi

Sebuah perusahaan parfum ingin memasarkan sebuah produk baru. Namun sebelumnya mereka ingin mengetahui apakah produk itu akan lebih disukai oleh wanita yang berumur ataukah yang berusia lebih muda. Sebuah sampel yang terdiri dari 100 orang wanita muda dan sampel lain yang terdiri dari 200 wanita berumur diminta mencium bau parfum itu dan menyampaikan apakah mereka menyukainya atau tidak. Diperoleh hasil: 19 orang wanita muda menyukai produk itu, dan 62 orang wanita berumur menyukainya. Gunakan level of significance $\alpha = 0.05$.

Contoh 2

1. Nyatakan Null Hypothesis dan Alternate Hypothesis:

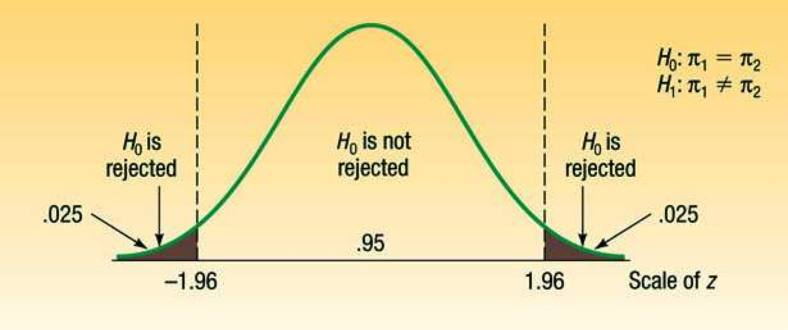
$$\begin{aligned} H_0 : \pi_1 &= \pi_2 \\ H_1 : \pi_1 &\neq \pi_2 \end{aligned} \tag{8}$$

2. Tentukan level of significance.
   Dari soal: $\alpha = 0.05$
3. Tentukan Statistik. Kita gunakan distribusi-Z.

Contoh 2

4. Tentukan aturan pengambilan keputusan. H_0 ditolak, jika:

$$\begin{aligned} Z &> Z_{\alpha/2} \text{ atau } Z < -Z_{\alpha/2} \\ Z &> Z_{0.025} \text{ atau } Z < -Z_{0.025} \\ Z &> 1.96 \text{ atau } Z < -1.96 \qquad (9) \end{aligned}$$



Contoh 2

5. Hitung Z, dan ambil keputusan.

$$
\begin{aligned}
p_1 &= \frac{x_1}{n_1} = \frac{19}{100} = 0.19 \\
p_2 &= \frac{x_2}{n_2} = \frac{62}{200} = 0.31 \qquad (10) \\
p_c &= \frac{x_1 + x_2}{n_1 + n_2} = \frac{19 + 62}{100 + 200} = \frac{81}{300} = 0.27 \qquad (11) \\
Z &= \frac{p_1 - p_2}{\sqrt{\frac{p_c(1-p_c)}{n_1} + \frac{p_c(1-p_c)}{n_2}}} \\
&= \frac{0.19 - 0.31}{\sqrt{\frac{0.27(1-0.27)}{100} + \frac{0.27(1-0.27)}{200}}} \\
&= -2.21 \qquad (12)
\end{aligned}
$$

Contoh 2

5. Hitung Z, dan ambil keputusan.
   Hasil: $Z = -2.21$ yang berada pada daerah penolakan H_0.
   Keputusan: H_0 ditolak.
   Artinya: Kita menolak hipotesa bahwa proporsi wanita muda yang menyukai produk parfum tersebut adalah sama dengan proporsi wanita berumur yang menyukai produk itu.

Membandingkan dua populasi. Standard deviasi sama, tetapi tidak diketahui. Sampel kecil.

Distribusi-t digunakan jika satu atau lebih sampel berukuran $n < 30$. Asumsi-asumsi yang disyaratkan adalah:

1. Kedua populasi harus mempunyai distribusi normal
2. Kedua populasi harus mempunyai standard deviasi yang sama
3. Sampel diambil dari populasi-populasi yang independen satu sama lain

Membandingkan dua populasi. Standard deviasi sama, tetapi tidak diketahui. Sampel kecil.

Menghitung s_p^2 dan t

$$s_p^2 = \frac{(n_1-1)s_1^2 + (n_2-1)s_2^2}{n_1+n_2-2} \tag{13}$$

$$t = \frac{\bar{X}_1 - \bar{X}_2}{\sqrt{s_p^2(\frac{1}{n_1}+\frac{1}{n_2})}} \tag{14}$$

Contoh 3

Sebuah mesin pemangkas rumput dapat dirakit dengan dua macam cara, yaitu "prosedur Welles", dan "prosedur Atkins". Satu sampel terdiri dari 5 karyawan diminta merakit dengan metode Welles dan diukur waktunya. Sampel lain terdiri dari 6 karyawan merakit dengan prosedur Atkins dan diukur waktunya. Hasilnya ditunjukkan pada Tabel (2). Dengan level of significance 0.1, tentukan apakah ada perbedaan signifikan antara rata-rata lamanya merakit menurut kedua metode tersebut.

Welles (minutes)	Atkins (minutes)
2	3
4	7
9	5
3	8
2	4
	3

Tabel 2: Dua sampel, untuk dua metode

Contoh 3

1. Nyatakan Null Hypothesis dan Alternate Hypothesis:

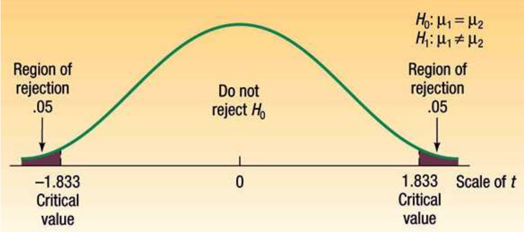
$$\begin{aligned} H_0: \mu_1 &= \mu_2 \\ H_1: \mu_1 &\neq \mu_2 \end{aligned} \tag{15}$$

2. Tentukan level of significance.
   Dari soal: $\alpha = 0.1$
3. Tentukan Statistik. Kita gunakan distribusi-t (standard deviasi kedua populasi tidak diketahui, tetapi diasumsikan sama).

Contoh 3

4. Tentukan aturan pengambilan keputusan. H_0 ditolak, jika:

$$\begin{aligned} t &> t_{\alpha/2,n_1+n_2-2} \text{ atau } t < -t_{\alpha/2,n_1+n_2-2} \\ t &> t_{0.05,9} \text{ atau } t < -t_{0.05,9} \\ t &> 1.833 \text{ atau } t < -1.833 \end{aligned} \tag{16}$$



Lihat tabel distribusi-t dengan $df = n_1 + n_2 - 2$

Contoh 3

5. Hitung t dan buat keputusan

Welles Method		Atkins Method	
X_1	$(X_1 - \bar{X}_1)^2$	X_2	$(X_2 - \bar{X}_2)^2$
2	$(2-4)^2 = 4$	3	$(3-5)^2 = 4$
4	$(4-4)^2 = 0$	7	$(7-5)^2 = 4$
9	$(9-4)^2 = 25$	5	$(5-5)^2 = 0$
3	$(3-4)^2 = 1$	8	$(8-5)^2 = 9$
2	$(2-4)^2 = 4$	4	$(4-5)^2 = 1$
20	34	3	$(3-5)^2 = 4$
		30	22

$$\begin{aligned} \bar{X}_1 &= \frac{\sum X_1}{n_1} = \frac{20}{5} = 4 \\ \bar{X}_2 &= \frac{\sum X_2}{n_2} = \frac{30}{6} = 5 \end{aligned} \tag{17}$$

Contoh 3

5. Hitung t dan buat keputusan

Welles Method		Atkins Method	
X_1	$(X_1 - \bar{X}_1)^2$	X_2	$(X_2 - \bar{X}_2)^2$
2	$(2-4)^2 = 4$	3	$(3-5)^2 = 4$
4	$(4-4)^2 = 0$	7	$(7-5)^2 = 4$
9	$(9-4)^2 = 25$	5	$(5-5)^2 = 0$
3	$(3-4)^2 = 1$	8	$(8-5)^2 = 9$
2	$(2-4)^2 = 4$	4	$(4-5)^2 = 1$
20	34	3	$(3-5)^2 = 4$
		30	22

$$s_1 = \sqrt{\frac{\sum(X_1 - \bar{X}_1)^2}{n_1 - 1}} = \sqrt{\frac{34}{5-1}} = 2.9155$$

$$s_2 = \sqrt{\frac{\sum(X_2 - \bar{X}_2)^2}{n_2 - 1}} = \sqrt{\frac{22}{6-1}} = 2.0976 \tag{18}$$

Contoh 3

5. Hitung t dan buat keputusan

Welles Method		Atkins Method	
X_1	$(X_1 - \bar{X}_1)^2$	X_2	$(X_2 - \bar{X}_2)^2$
2	$(2-4)^2 = 4$	3	$(3-5)^2 = 4$
4	$(4-4)^2 = 0$	7	$(7-5)^2 = 4$
9	$(9-4)^2 = 25$	5	$(5-5)^2 = 0$
3	$(3-4)^2 = 1$	8	$(8-5)^2 = 9$
2	$(2-4)^2 = 4$	4	$(4-5)^2 = 1$
20	34	3	$(3-5)^2 = 4$
		30	22

$$
\begin{aligned}
s_p^2 &= \frac{(n_1-1)s_1^2 + (n_2-1)s_2^2}{n_1+n_2-2} \\
&= \frac{(5-1)(2.9155)^2 + (6-1)(2.0976)^2}{5+6-2} \qquad (19) \\
&= 6.2222
\end{aligned}
$$

Contoh 3

5. Hitung t dan buat keputusan

$$
\begin{aligned}
t &= \frac{\bar{X}_1 - \bar{X}_2}{\sqrt{s_p^2(\frac{1}{n_1} + \frac{1}{n_2})}} \\
&= \frac{4.00 - 5.00}{\sqrt{6.2222(\frac{1}{5} + \frac{1}{6})}} \qquad (20) \\
&= -0.662
\end{aligned}
$$

Kita peroleh:

$$-1.833 < -0.662 < 1.833 \qquad (21)$$

Kesimpulan: H_0 tidak ditolak. Tidak ada perbedaan signifikan dalam waktu rata-rata perakitan mesin dari dua metode yang digunakan.

Kuliah berikutnya

Standard Deviasi Tidak Diketahui, saling berbeda, $n < 30$

$$t = \frac{\bar{X}_1 - \bar{X}_2}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}} \tag{22}$$

Kuliah berikutnya

Dua sample saling dependent

$$t = \frac{\bar{d}}{s_d/\sqrt{n}} \tag{23}$$

Dengan: $\bar{d}$ mean dari selisih
s_d standard deviasi dari selisih
n jumlah pasangan