Statistik Inferensi 2 (Lanjutan)

Inferensi Statistik

Tahap-tahap Uji Hipotesis Secara umum

1. Menentukan model probabilitas yang cocok dari data
2. Menentukan Hipotesis H_0 dan H_1
3. Menentukan Statistik Penguji, yang harus merupakan fungsi dari data dan tidak memuat parameter yang tidak diketahui
4. Menentukan tingkat signifikansi
5. Menentukan daerah kritik berdasarkan tingkat signifikansi
6. Menghitung Statistik Penguji, apakah masuk daerah kritik atau tidak
7. Alternatif: Hitung p-*value* berdasarkan statistik penguji
8. Mengambil kesimpulan berdasarkan 6 atau 7

Inferensi Statistik

- Satu Populasi
  - Populasi sembarang
    - μ
    - p
  - Populasi Normal
    - μ
    - σ^2
- Dua Populasi
  - Populasi sembarang
    - $\mu_1^2,\ \mu_2^2$
    - $p_1^2,\ p_2^2$
  - Populasi Normal
    - $\mu_1^2,\ \mu_2^2$
    - $\sigma_1^2,\ \sigma_2^2$
- $k > 2$ Populasi
  - Analisis Variansi (ANOVA)

Inferensi Statistik Satu Populasi Sembarang

<u>Estimasi interval mean (μ) suatu populasi</u>

Teorema Limit Pusat

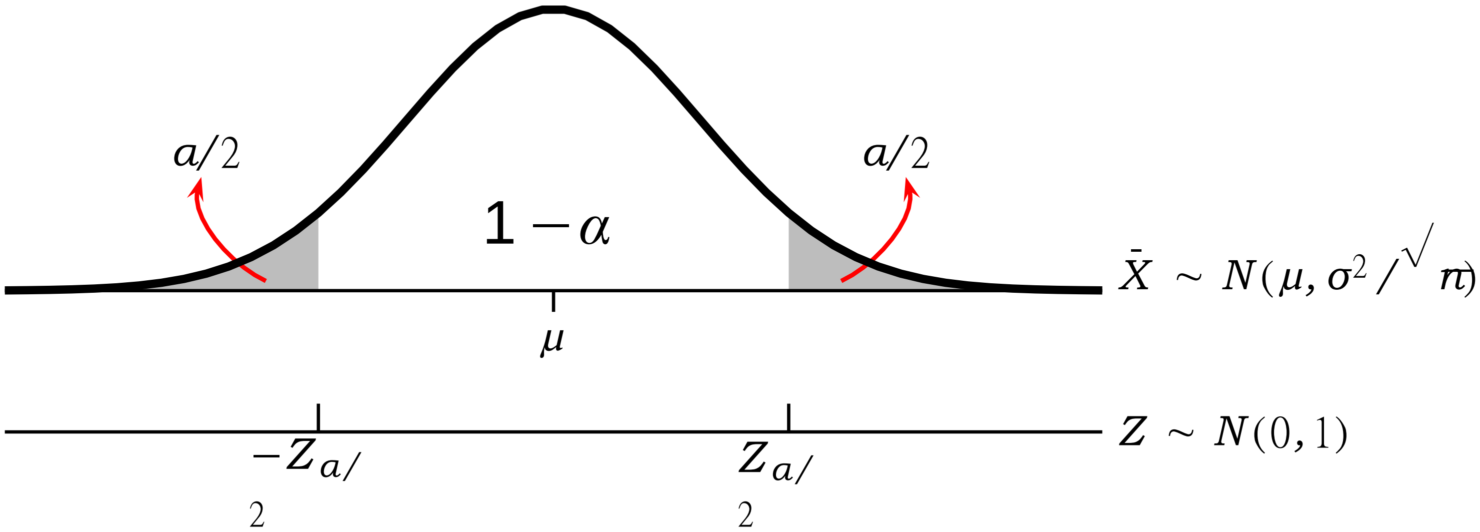
Apabila sampel-sampel random diambil dari suatu populasi yang berdistribusi sembarang, yang mempunyai mean μ dan variansi σ^2, maka untuk n besar, distribusi sampling untuk mean dapat dianggap mendekati Normal dengan $\mu_{\bar{X}} = \mu$ dan variansi $\sigma_{\bar{X}}^2 = \sigma^2/n$, sehingga

$$Z = \frac{\bar{X} - \mu}{\sigma/\sqrt{n}}$$

mendekati Normal Standar.

Inferensi Statistik Satu Populasi Sembarang

Estimasi interval mean (μ) suatu populasi



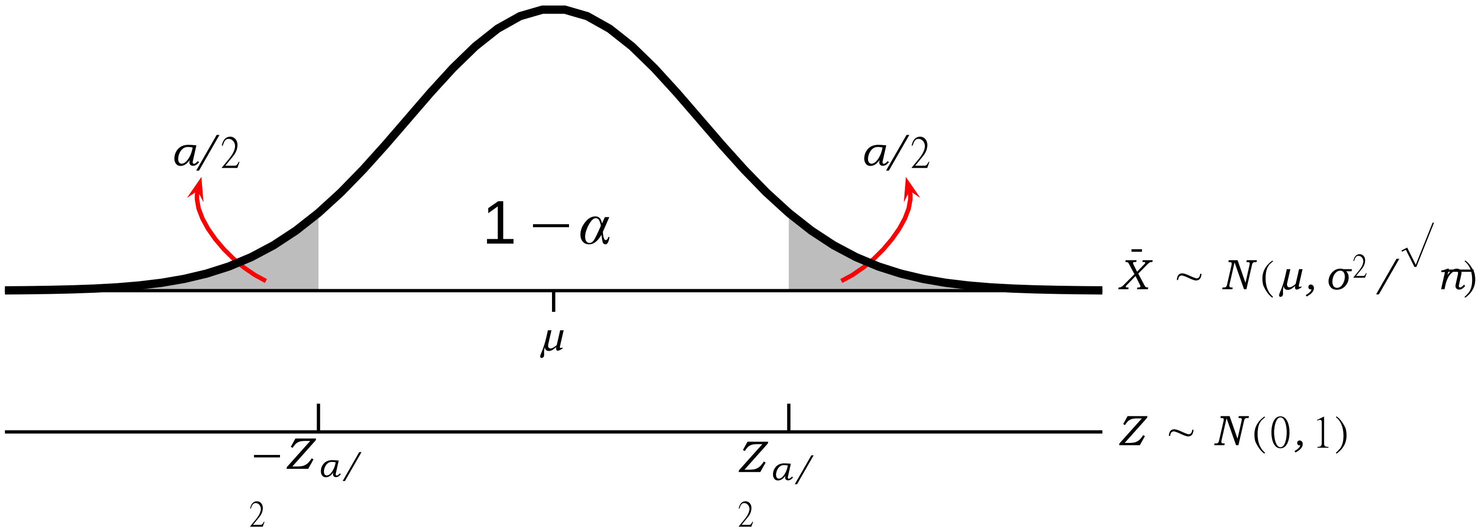
$$P(-Z_{\alpha/2} \le Z \le Z_{\alpha/2}) \approx 1-\alpha$$

$$P\left(-Z_{\alpha/2} \le \frac{\bar{X}-\mu}{\sigma/\sqrt{n}} \le Z_{\alpha/2}\right) \approx 1-\alpha$$

$$P\left(\bar{X} - Z_{\alpha/2}\frac{\sigma}{\sqrt{n}} \le \mu \le \bar{X} + Z_{\alpha/2}\frac{\sigma}{\sqrt{n}}\right) \approx 1-\alpha$$

Inferensi Statistik Satu Populasi Sembarang

<u>Estimasi interval mean (μ) suatu populasi</u>



Interval Konfidensi $(1-\alpha)$100% untuk mean

$\mu\ B \le \mu \le A$

$$B = \bar{X} - Z_{\alpha/2}\frac{\sigma}{\sqrt{n}}$$

$$A = \bar{X} + Z_{\alpha/2}\frac{\sigma}{\sqrt{n}}$$

Inferensi Statistik Satu Populasi Sembarang

Contoh:

Suatu sampel random dengan 150 keluarga petani di suatu desa menunjukkan penghasilan bulanan dari sektor pertanian rata-rata Rp 325 000,00 dengan deviasi standar Rp 25 000,00. Hitung interval konfidensi 95% untuk rata-rata penghasilan pertanian bulanan seluruh keluarga di desa tersebut.

Jawab:

X : penghasilan pertanian bulanan di desa tersebut

$\bar{X} = 325.000$; $s = 25.000$; $n = 150$.

Interval konfidensi 95% untuk rata-rata penghasilan bulanan (μ):

$B = \bar{X} - Z_{\alpha/2} \frac{\sigma}{\sqrt{n}} = 325.000 - 1{,}96 \frac{25}{\sqrt{150}} = 324.996$

$A = \bar{X} + Z_{\alpha/2} \frac{\sigma}{\sqrt{n}} = 325.000 + 1{,}96 \frac{25}{\sqrt{150}} = 325.004$

Interval konfidensi 95%: $324.996 \leq \mu \leq 325.004$

σ dapat diganti s

Inferensi Statistik Satu Populasi Sembarang

Uji Hipotesis Mean (μ) Populasi

1. Hipotesis
   A. $H_0 : \mu = \mu_0$ vs. $H_1 : \mu \neq \mu_0$ B. $H_0 : \mu \leq \mu_0$ vs. $H_1 : \mu > \mu_0$ C. $H_0 : \mu \geq \mu_0$ vs. $H_1 : \mu < \mu_0$
2. Tingkat signifikansi α
3. Statistik Penguji

$$Z = \frac{\bar{X} - \mu_0}{\sigma / \sqrt{n}}$$

atau

$$Z = \frac{\bar{X} - \mu_0}{s / \sqrt{n}}$$

jika σ tidak diketahui diganti s. Distribusi dari Z adalah Normal Standar.

Inferensi Statistik Satu Populasi Sembarang

Uji Hipotesis Mean (μ) Populasi

4. Daerah penolakan (berdasarkan α dan Hipotesis)

   $A.$ H_0 ditolak apabila $Z > Z_{\alpha/2}$ atau $Z < -Z_{\alpha/2}$

   $B.$ H_0 ditolak apabila $Z > Z_{\alpha}$

   $C.$ H_0 ditolak apabila $Z < -Z_{\alpha}$

Inferensi Statistik Satu Populasi Sembarang

Uji Hipotesis proporsi (p) Populasi

1. Hipotesis
   $A.\ H_0 : p = p_0$ vs. $H_1 : p \neq p_0$
   $B.\ H_0 : p \leq p_0$ vs. $H_1 : p > p_0$
   $C.\ H_0 : p \geq p_0$ vs. $H_1 : p < p_0$
2. Tingkat signifikansi α
3. Statistik Penguji

$$Z = \frac{\hat{p} - p_0}{\sqrt{\frac{p_0(1-p_0)}{n}}}$$

Distribusi dari Z adalah Normal Standar.

Inferensi Statistik Satu Populasi Sembarang

<u>Uji Hipotesis proporsi (p) Populasi</u>

4. Daerah penolakan (berdasarkan α dan Hipotesis)

$A.$ H_0 ditolak apabila $Z > Z_{\alpha/2}$ atau $Z < -Z_{\alpha/2}$

$B.$ H_0 ditolak apabila $Z > Z_{\alpha}$

$C.$ H_0 ditolak apabila $Z < -Z_{\alpha}$

Inferensi Statistik Satu Populasi Sembarang

<u>Hubungan antara Interval Konfidensi dan Uji</u>

Hipotesis Interval Konfidensi $(1-\alpha)100\%$ untuk mean μ

$$\bar{X} - Z_{\alpha/2}\frac{\sigma}{\sqrt{n}} \leq \mu \leq \bar{X} + Z_{\alpha/2}\frac{\sigma}{\sqrt{n}}$$

Daerah penolakan dengan tingkat signifikansi α untuk uji hipotesis $H_0 : \mu = \mu_0$ vs. $H_1 : \mu \neq \mu_0$

$$Z > Z_{\alpha/2} \text{ atau } Z < -Z_{\alpha/2}$$

Daerah penerimaan

$$-Z_{\alpha/2} \leq Z \leq Z_{\alpha/2}$$

Inferensi Statistik Satu Populasi Sembarang

Ringkasan

Parameter	Statistik	Interval Konfidensi (1- α)100%	Hipotesis alternatif	Daerah Kritik
μ mean	$Z = \dfrac{\bar{X}-\mu_0}{\sigma/\sqrt{n}}$ $Z \sim N(0,1)$	$B \le \mu \le A$ $B = \bar{X} - Z_{\alpha/2}\dfrac{\sigma}{\sqrt{n}}$ $A = \bar{X} + Z_{\alpha/2}\dfrac{\sigma}{\sqrt{n}}$	$H_1 : \mu \neq \mu_0$ $H_1 : \mu > \mu_0$ $H_1 : \mu < \mu_0$	$Z > Z_{\alpha/2}$ atau $Z < -Z_{\alpha/2}$ $Z > Z_{\alpha}$ $Z < -Z_{\alpha}$
p proporsi	$Z = \dfrac{\hat{p}-p_0}{\sqrt{\dfrac{p_0(1-p_0)}{n}}}$ $Z \sim N(0,1)$	$B \le p \le A$ $B = \hat{p} - Z_{\alpha/2}\sqrt{\dfrac{\hat{p}(1-\hat{p})}{n}}$ $A = \hat{p} + Z_{\alpha/2}\sqrt{\dfrac{\hat{p}(1-\hat{p})}{n}}$	$H_1 : p \neq p_0$ $H_1 : p > p_0$ $H_1 : p < p_0$	$Z > Z_{\alpha/2}$ atau $Z < -Z_{\alpha/2}$ $Z > Z_{\alpha}$ $Z < -Z_{\alpha}$

Inferensi Statistik Satu Populasi Normal

- Data dianggap berdistribusi
- Normal Ukuran sampel tidak harus
- besar Jenis parameter:
  - mean μ
  - variansi σ^2
- Distribusi
  - Sampling
  - Normal
  - t

  Chi-kuadrat (*Chi-square*)

Inferensi Statistik Satu Populasi Normal

<u>Normal Standar</u>

Jika $X_1, \ldots, X_n$ adalah sampel random berasal dari populasi Normal dengan mean μ dan variansi σ^2 maka variabel random

$$Z = \frac{\bar{X} - \mu}{\sigma/\sqrt{n}}$$

berdistribusi Normal Standar N (0, 1)

Inferensi Statistik Satu Populasi Normal

Distribusi t

Jika $X_1, \ldots, X_n$ adalah sampel random berasal dari populasi Normal dengan mean μ dan variansi σ^2 maka variabel random

$$t = \frac{\bar{X} - \mu}{s/\sqrt{n}}$$

berdistribusi t dengan derajad bebas $n - 1$.

Untuk n yang semakin besar, distribusi t akan mendekati distribusi Normal.

Inferensi Statistik Satu Populasi Normal

<u>Distribusi Chi-kuadrat 2k</u>

Diketahui $X_1, \ldots, X_k$ adalah variabel random yang berdistribusi Normal yang independen satu dengan yang lain. Distribusi variabel random

$$\chi^2 = X_1^2 + \ldots + X_k^2$$

berdistribusi Chi-kuadrat berderajad bebas k dengan mean E(χ^2) = k dan variansi Var(χ^2) = 2k

Inferensi Statistik Satu Populasi Normal

<u>Distribusi Chi-kuadrat $n-1$</u>

Diketahui $X_1, \ldots, X_n$ adalah variabel random yang berdistribusi Normal dengan mean μ dan variansi σ^2 maka variabel random

$$\chi^2 = \frac{(n-1)s^2}{\sigma^2}$$

berdistribusi Chi-kuadrat dengan derajad bebas $n-1$

Inferensi Statistik Satu Populasi Normal

<u>Distribusi Normal Standar</u>

Apabila sampel random berukuran n diambil dari suatu populasi
yang berdistribusi Normal dengan mean μ dan variansi σ^2, maka variabel random

$$Z = \frac{s^2 - \sigma^2}{\sigma^2 \sqrt{\frac{2}{n-1}}}$$

berdistribusi $N(0, 1)$ untuk n besar.

Inferensi Statistik Satu Populasi Normal

Parameter	**Statistik**	**Interval Konfidensi (1-α)100%**	**Hipotesis alternatif**	**Daerah Kritik**
μ mean	Bila σ^2 diketahui $$Z = \frac{\bar{X} - \mu_0}{\sigma/\sqrt{n}}$$ $Z \sim N(0,1)$	$B \le \mu \le A$ $B = \bar{X} - Z_{\alpha/2}\frac{\sigma}{\sqrt{n}}$ $A = \bar{X} + Z_{\alpha/2}\frac{\sigma}{\sqrt{n}}$	$H_1 : \mu \neq \mu_0$ $H_1 : \mu > \mu_0$ $H_1 : \mu < \mu_0$	$Z > Z_{\alpha/2}$ atau $Z < -Z_{\alpha/2}$ $Z > Z_\alpha$ $Z < -Z_\alpha$
	Bila σ^2 tidak diketahui $$t = \frac{\bar{X} - \mu_0}{s/\sqrt{n}}$$ $t \sim$ distribusi t dgn. derajad bebas $n-1$	$B \le \mu \le A$ $B = \bar{X} - t_{(n-1,\alpha/2)}\frac{s}{\sqrt{n}}$ $A = \bar{X} + t_{(n-1,\alpha/2)}\frac{s}{\sqrt{n}}$	$H_1 : \mu \neq \mu_0$ $H_1 : \mu > \mu_0$ $H_1 : \mu < \mu_0$	$t > t_{(n-1,\alpha/2)}$ atau $t < -t_{(n-1,\alpha/2)}$ $t > t_{(n-1,\alpha)}$ $t < -t_{(n-1,\alpha)}$

Inferensi Statistik Satu Populasi Normal

Parameter	Statistik	Interval Konfidensi (1-α)100%	Hipotesis alternatif	Daerah Kritik
σ^2 variansi	$\chi^2 = \frac{(n-1)s^2}{\sigma^2}$ $\chi^2 \sim$ chi-square dgn. derajad bebas $k = n - 1$	$B \leq \sigma^2 \leq A$ $B = \frac{(n-1)s^2}{\chi^2_{(n-1,\alpha/2)}}$ $A = \frac{(n-1)s^2}{\chi^2_{(n-1,1-\alpha/2)}}$	$H_1 : \sigma^2 \neq \sigma_0^2$ $H_1 : \sigma^2 > \sigma_0^2$ $H_1 : \sigma^2 < \sigma_0^2$	$\chi^2 > \chi^2_{(k,\alpha/2)}$ atau $\chi^2 < \chi^2_{(k,1-\alpha/2)}$ $\chi^2 > \chi^2_{(k,\alpha)}$ $\chi^2 < \chi^2_{(k,1-\alpha)}$
	Untuk n besar, $Z = \frac{s^2 - \sigma^2}{\sigma^2 \sqrt{\frac{2}{n-1}}}$ $Z \sim N(0, 1)$	$B \leq \sigma^2 \leq A$ $B = \frac{s^2}{1+Z_{\alpha/2}\sqrt{\frac{2}{n-1}}}$ $A = \frac{s^2}{1-Z_{\alpha/2}\sqrt{\frac{2}{n-1}}}$	$H_1 : \sigma^2 \neq \sigma_0^2$ $H_1 : \sigma^2 > \sigma_0^2$ $H_1 : \sigma^2 < \sigma_0^2$	$Z > Z_{\alpha/2}$ atau $Z < -Z_{\alpha/2}$ $Z > Z_\alpha$ $Z < -Z_\alpha$

Inferensi Statistik Dua Populasi Sembarang

Distribusi sampling selisih dua mean

Misalkan $X_{11}, X_{12}, \ldots, X_{1n_1}$ dan $X_{21}, X_{22}, \ldots, X_{2n_2}$ adalah dua sampel random independen satu sama lain yang diambil dari populasi yang mempunyai mean μ_1 dan μ_2 serta variansi σ_1^2 dan σ_2^2, maka untuk n_1 dan n_2 besar, variabel random

$$Z = \frac{(\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}$$

berdistribusi Normal Standar, dengan

$$\bar{X}_1 = \sum_{i=1}^{n_1} \frac{X_{1i}}{n_1} \qquad \bar{X}_2 = \sum_{i=1}^{n_2} \frac{X_{2i}}{n_2}$$

Inferensi Statistik Dua Populasi Sembarang

Distribusi sampling selisih dua mean

Jika σ_1^2 dan σ_2^2 tidak diketahui, dan diasumsikan $\sigma_1^2 \neq \sigma_2^2$

$$Z = \frac{(\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}$$

berdistribusi Normal Standar dengan s_1^2 dan s_2^2 adalah variansi sampel

Inferensi Statistik Dua Populasi Sembarang

<u>Distribusi sampling selisih dua mean</u>

Jika σ_1^2 dan σ_2^2 tidak diketahui, dan diasumsikan $\sigma_1^2 = \sigma_2^2$

$$Z = \frac{(\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2)}{\sqrt{s_p^2\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}}$$

berdistribusi Normal Standar dengan

s_1^2 dan s_2^2 adalah variansi sampel

$$s_p^2 = \frac{(n_1 - 1)S_1^2 + (n_2 - 1)S_2^2}{n_1 + n_2 - 2}$$

yang disebut sebagai *pooled variance*

Inferensi Statistik Dua Populasi Sembarang

- Distribusi sampling selisih dua proporsi
- Misalkan $X_{11}, X_{12}, \ldots, X_{1n_1}$ dan $X_{21}, X_{22}, \ldots, X_{2n_2}$ adalah dua sampel random independen satu sama lain yang diambil dari populasi yang berdistribusi binomial. Untuk n_1 dan n_2 besar, variabel random

$$Z = \frac{\left(\frac{X_1}{n_1} - \frac{X_2}{n_2}\right) - (p_1 - p_2)}{\sqrt{\frac{\frac{X_1}{n_1}\left(1 - \frac{X_1}{n_1}\right)}{n_1} + \frac{\frac{X_2}{n_2}\left(1 - \frac{X_2}{n_2}\right)}{n_2}}}$$

berdistribusi Normal Standar.

Inferensi Statistik Dua Populasi Sembarang

Parameter	Statistik	Interval Konfidensi (1-α)100%	Hipotesis alternatif	Daerah Kritik
$\mu_1 - \mu_2$ selisih dua mean	σ_1^2 dan σ_2^2 diketahui $Z = \frac{(\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}$ $Z \sim N(0,1)$	$B \le \mu_1 - \mu_2 \le A$ $B = (\bar{X}_1 - \bar{X}_2) - Z_{\frac{\alpha}{2}} \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}$ $A = (\bar{X}_1 - \bar{X}_2) + Z_{\frac{\alpha}{2}} \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}$	$H_1 : \mu_1 - \mu_2 \neq \mu_0$ $H_1 : \mu_1 - \mu_2 > \mu_0$ $H_1 : \mu_1 - \mu_2 < \mu_0$	$Z > Z_{\alpha/2}$ atau $Z < -Z_{\alpha/2}$ $Z > Z_\alpha$ $Z < -Z_\alpha$
	σ_1^2 dan σ_2^2 tdk diketahui, $\sigma_1^2 \neq \sigma_2^2$ $Z = \frac{(\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}$ $Z \sim N(0,1)$	$B \le \mu_1 - \mu_2 \le A$ $B = (\bar{X}_1 - \bar{X}_2) - Z_{\frac{\alpha}{2}} \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}$ $A = (\bar{X}_1 - \bar{X}_2) + Z_{\frac{\alpha}{2}} \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}$	$H_1 : \mu_1 - \mu_2 \neq \mu_0$ $H_1 : \mu_1 - \mu_2 > \mu_0$ $H_1 : \mu_1 - \mu_2 < \mu_0$	$Z > Z_{\alpha/2}$ atau $Z < -Z_{\alpha/2}$ $Z > Z_\alpha$ $Z < -Z_\alpha$

Inferensi Statistik Dua Populasi Sembarang

Parameter	Statistik	Interval Konfidensi (1-α)100%	Hipotesis alternatif	Daerah Kritik
	σ_1^2 dan σ_2^2 tdk diketahui, $\sigma_1^2 = \sigma_2^2$ $Z = \frac{(\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2)}{\sqrt{s_p^2\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}}$ $Z \sim N(0,1)$ $s_p^2 = \frac{(n_1-1)S_1^2 + (n_2-1)S_2^2}{n_1+n_2-2}$	$B \leq \mu_1 - \mu_2 \leq A$ $B = (\bar{X}_1 - \bar{X}_2) - Z_{\frac{\alpha}{2}}\sqrt{S_p^2\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}$ $A = (\bar{X}_1 - \bar{X}_2) + Z_{\frac{\alpha}{2}}\sqrt{S_p^2\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}$	$H_1: \mu_1 - \mu_2 \neq \mu_0$ $H_1: \mu_1 - \mu_2 > \mu_0$ $H_1: \mu_1 - \mu_2 < \mu_0$	$Z > Z_{\frac{\alpha}{2}}$ atau $Z < -Z_{\frac{\alpha}{2}}$ $Z > Z_\alpha$ $Z < -Z_\alpha$
$p_1 - p_2$ Selisih dua proporsi	$Z = \frac{(\hat{p}_1 - \hat{p}_2) - (p_1 - p_2)}{\sqrt{\frac{\hat{p}_1(1-\hat{p}_1)}{n_1} + \frac{\hat{p}_2(1-\hat{p}_2)}{n_2}}}$ $Z \sim N(0,1)$ $\hat{p}_1 = \frac{X_1}{n_1}$; $\hat{p}_2 = \frac{X_2}{n_2}$	$B \leq p_1 - p_2 \leq A$ $B = (\hat{p}_1 - \hat{p}_2) - Z_{\frac{\alpha}{2}}\sqrt{\frac{\hat{p}_1(1-\hat{p}_1)}{n_1} + \frac{\hat{p}_2(1-\hat{p}_2)}{n_2}}$ $A = (\hat{p}_1 - \hat{p}_2) + Z_{\frac{\alpha}{2}}\sqrt{\frac{\hat{p}_1(1-\hat{p}_1)}{n_1} + \frac{\hat{p}_2(1-\hat{p}_2)}{n_2}}$	$H_1: p_1 - p_2 \neq p_0$ $H_1: p_1 - p_2 > p_0$ $H_1: p_1 - p_2 < p_0$	$Z > Z_{\frac{\alpha}{2}}$ atau $Z < -Z_{\frac{\alpha}{2}}$ $Z > Z_\alpha$ $Z < -Z_\alpha$

Inferensi Statistik Dua Populasi Normal

<u>Distribusi sampling selisih dua mean</u>

Misalkan $X_{11}, X_{12}, \ldots, X_{1n_1}$ dan $X_{21}, X_{22}, \ldots, X_{2n_2}$ adalah dua sampel random independen satu sama lain yang berdistribusi Normal dengan mean μ_1 dan μ_2 serta variansi σ_1^2 dan σ_2^2, maka variabel random

$$Z = \frac{(\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}$$

berdistribusi Normal Standar, dengan

$$\bar{X}_1 = \sum_{i=1}^{n_1} \frac{X_{1i}}{n_1} \qquad \bar{X}_2 = \sum_{i=1}^{n_2} \frac{X_{2i}}{n_2}$$

Inferensi Statistik Dua Populasi Normal

<u>Distribusi sampling selisih dua mean</u>

Jika σ_1^2 dan σ_2^2 tidak diketahui, dan diasumsikan $\sigma_1^2 \neq \sigma_2^2$

$$t = \frac{(\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}$$

berdistribusi t dengan derajad bebas

$$k = \frac{(s_1^2/n_1 + s_2^2/n_2)^2}{\frac{(s_1^2/n_1)^2}{n_1+1} + \frac{(s_2^2/n_2)^2}{n_2+1}} - 2, \quad \text{atau} \quad k = \frac{(s_1^2/n_1 + s_2^2/n_2)^2}{\frac{(s_1^2/n_1)^2}{n_1-1} + \frac{(s_2^2/n_2)^2}{n_2-1}}$$

dengan s_1^2 dan s_2^2 adalah variansi sampel

Inferensi Statistik Dua Populasi Normal

Distribusi sampling selisih dua mean

Jika σ_1^2 dan σ_2^2 tidak diketahui, dan diasumsikan $\sigma_1^2 = \sigma_2^2$

$$t = \frac{(\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2)}{\sqrt{s_p^2\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}}$$

berdistribusi t dengan derajad bebas $n_1 + n_2 - 2$

dan

$$s_p^2 = \frac{(n_1 - 1)S_1^2 + (n_2 - 1)S_2^2}{n_1 + n_2 - 2}$$

yang disebut sebagai *pooled variance*

Inferensi Statistik Dua Populasi Normal

Distribusi sampling Perbandingan dua variansi

Misalkan $X_{11}, X_{12}, \ldots, X_{1n_1}$ dan $X_{21}, X_{22}, \ldots, X_{2n_2}$ adalah dua sampel random independen satu sama lain yang berdistribusi Normal dengan mean μ_1 dan μ_2 serta variansi σ_1^2 dan σ_2^2, maka variabel random

$$F = \frac{s_1^2/\sigma_1^2}{s_2^2/\sigma_2^2}$$

berdistribusi F dengan derajad bebas pembilang $n_1 - 1$, derajad bebas penyebut $n_2 - 1$

Inferensi Statistik Dua Populasi Normal

Parameter	Statistik	Interval Konfidensi (1-α)100%	Hipotesis alternatif	Daerah Kritik
$\mu_1 - \mu_2$ Selisih dua mean	σ_1^2 dan σ_2^2 diketahui $Z=\frac{(\bar{X}_1-\bar{X}_2)-(\mu_1-\mu_2)}{\sqrt{\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}}}$ $Z \sim N(0,1)$	$B \le \mu_1 - \mu_2 \le A$ $B=(\bar{X}_1-\bar{X}_2)-Z_{\frac{\alpha}{2}}\sqrt{\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}}$ $A=(\bar{X}_1-\bar{X}_2)+Z_{\frac{\alpha}{2}}\sqrt{\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}}$	$H_1: \mu_1-\mu_2 \neq \mu_0$ $H_1: \mu_1-\mu_2 > \mu_0$ $H_1: \mu_1-\mu_2 < \mu_0$	$Z>Z_{\frac{\alpha}{2}}$ atau $Z<-Z_{\frac{\alpha}{2}}$ $Z>Z_{\alpha}$ $Z<-Z_{\alpha}$
	σ_1^2 dan σ_2^2 tdk diketahui dan $\sigma_1^2 \neq \sigma_2^2$ $t=\frac{(\bar{X}_1-\bar{X}_2)-(\mu_1-\mu_2)}{\sqrt{\frac{s_1^2}{n_1}+\frac{s_2^2}{n_2}}}$ $t \sim t_k$ dgn $\frac{(s^2/n_1+s^2/n_2)^2}{\frac{\quad}{n_1+1}+\frac{\quad}{n_2+1}}$ atau $(s^2/n_1+s^2/n_2)^2$	$B \le \mu_1 - \mu_2 \le A$ $B=(\bar{X}_1-\bar{X}_2)-t_{\frac{\alpha}{2},k}\sqrt{\frac{s_1^2}{n_1}+\frac{s_2^2}{n_2}}$ $A=(\bar{X}_1-\bar{X}_2)+t_{\frac{\alpha}{2},k}\sqrt{\frac{s_1^2}{n_1}+\frac{s_2^2}{n_2}}$	$H_1: \mu_1-\mu_2 \neq \mu_0$ $H_1: \mu_1-\mu_2 > \mu_0$ $H_1: \mu_1-\mu_2 < \mu_0$	$t>t_{\frac{\alpha}{2},k}$ atau $t<-t_{\frac{\alpha}{2},k}$ $t>t_{\alpha,k}$ $t<-t_{\alpha,k}$

Inferensi Statistik Dua Populasi Normal

Parameter	Statistik	Interval Konfidensi (1-α)100%	Hipotesis alternatif	Daerah Kritik
	σ_1^2 dan σ_2^2 tdk diketahui dan $\sigma_1^2 = \sigma_2^2$ $t = \frac{(\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2)}{\sqrt{S_p^2\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}}$ $t \sim t_k$ dgn. $k = n_1 + n_2 - 2$ $S_p^2 = \frac{(n_1-1)S_1^2 + (n_2-1)S_2^2}{n_1 + n_2 - 2}$	$B \le \mu_1 - \mu_2 \le A$ $B = (\bar{X}_1 - \bar{X}_2) - t_{\frac{\alpha}{2},k}\sqrt{S_p^2\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}$ $A = (\bar{X}_1 - \bar{X}_2) + t_{\frac{\alpha}{2},k}\sqrt{S_p^2\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}$	$H_1: \mu_1 - \mu_2 \ne \mu_0$ $H_1: \mu_1 - \mu_2 > \mu_0$ $H_1: \mu_1 - \mu_2 < \mu_0$	$t > t_{\frac{\alpha}{2},k}$ atau $t < -t_{\frac{\alpha}{2},k}$ $t > t_{\alpha,k}$ $t < -t_{\alpha,k}$
σ_1^2 / σ_2^2 Perban-dingan dua variansi	$F = s_1^2 / s_2^2$ dengan $F \sim F_{\frac{\alpha}{2},k_1,k_2}$ $k_1 = n_1 - 1,\ k_2 = n_2 - 1$	$B \le \sigma_1^2/\sigma_2^2 \le A$ $B = \frac{s_1^2/s_2^2}{F_{(k_1,k_2,\frac{\alpha}{2})}}$ $A = \frac{s_1^2}{s_2^2} F_{(k_1,k_2,\frac{\alpha}{2})}$ catatan: $F(1-\frac{\alpha}{2},k_1,k_2) = 1/F(\frac{\alpha}{2},k_2,k_1)$	$H_1: \sigma_1 \ne \sigma_2$ $H_1: \sigma_1 > \sigma_2$ $H_1: \sigma_1 < \sigma_2$	$F > F_{\frac{\alpha}{2},k_1,k_2}$ atau $F < 1/F_{\frac{\alpha}{2},k_2,k_1}$ $F > F_{\alpha,k_1,k_2}$ $F < 1/F_{\alpha,k_2,k_1}$

Inferensi Statistik Dua Populasi Normal

Parameter	**Statistik**	**Interval Konfidensi (1-α)100%**	**Hipotesis alternatif**	**Daerah Kritik**
μ_d mean selisih data berpasangan	$t = \frac{\bar{D} - \mu_D}{s_D/\sqrt{n}}$ dengan $t \sim$ distribusi t dgn derajad bebas $k = n - 1$	$B \leq \mu \leq A$ $B = \bar{X} - t_{(n-1,\alpha/2)} \frac{s}{\sqrt{n}}$ $A = \bar{X} + t_{(n-1,\alpha/2)} \frac{s}{\sqrt{n}}$	$H_1 : \mu_D \neq \mu_0$ $H_1 : \mu_D > \mu_0$ $H_1 : \mu_D < \mu_0$	$t > t_{(n-1,\alpha/2)}$ atau $t < -t_{(n-1,\alpha/2)}$ $t > t_{(n-1,\alpha)}$ $t < -t_{(n-1,\alpha)}$

Analisis Variansi Satu Arah

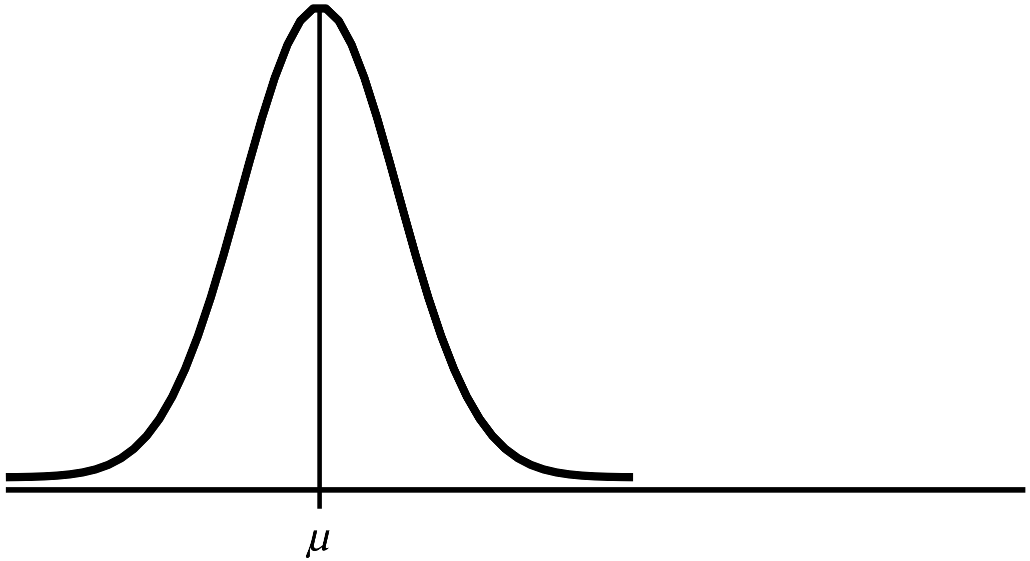
- Perluasan dari uji mean dua populasi Normal (data berasal dari populasi Normal)
- Ada k mean populasi yang
- dibandingkan Berdasarkan pada

  pemecahan variansi

Analisis Variansi Satu Arah

Inferensi mean populasi

Normal



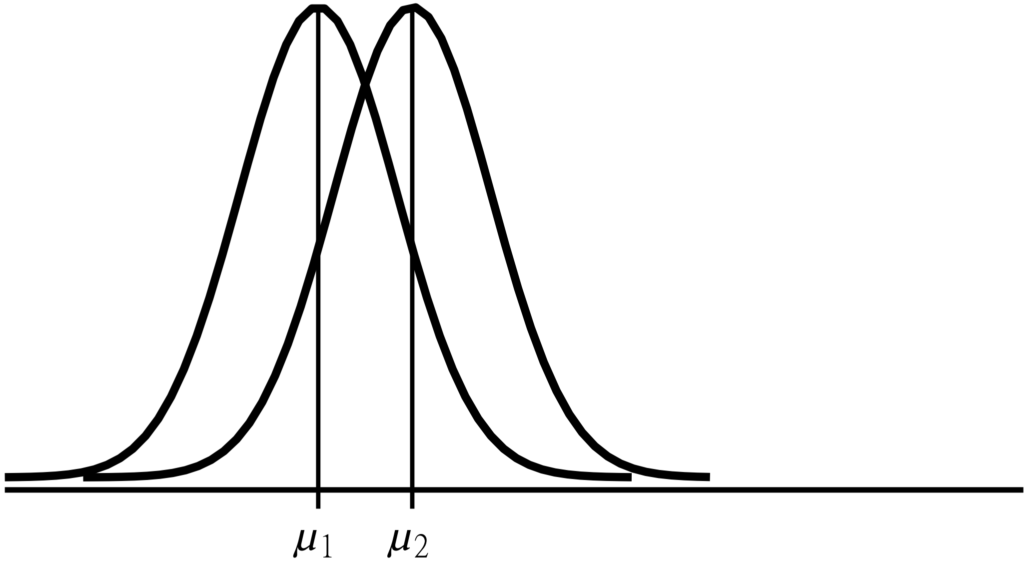
Uji mean satu populasi

$H_0 : \mu = \mu_0$

Analisis Variansi Satu Arah

Inferensi mean populasi

Normal



Uji mean satu populasi

H$_0$: $\mu = \mu_0$

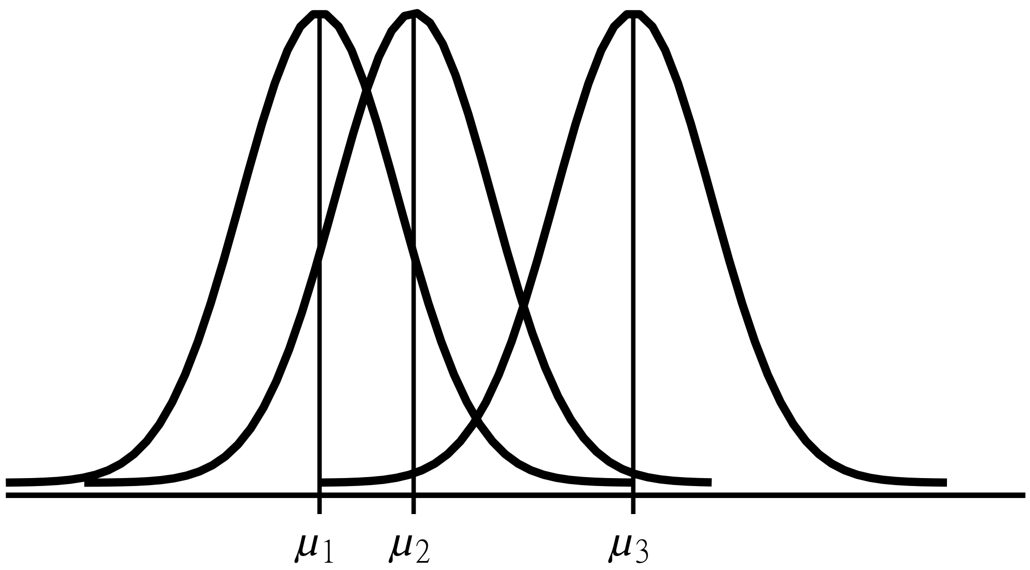
Uji mean dua populasi

H$_0$: $\mu_1 = \mu_2$

Analisis Variansi Satu Arah

Inferensi mean populasi

Normal



Uji mean satu populasi
$H_0 : \mu = \mu_0$

Uji mean dua populasi
$H_0 : \mu_1 = \mu_2$

Uji mean k populasi
$H_0 : \mu_1 = \mu_2 = \mu_3$

Analisis Variansi Satu Arah

Uji Hipotesis

H_0 : $\mu_1 = \mu_2 = \ldots = \mu_k$

H_1 : minimal ada dua mean yang tidak sama

Statistik Penguji

$$F = \frac{\text{MST}}{\text{MSE}}$$

dimana $F \sim F_{(k-1,n-k)}$

MST: *mean square treatment* (kuadrat rata-rata perlakuan) MSE: *mean square error* (kuadrat rata-rata sesatan)
yang diperoleh dari Tabel **Anova** (Analisis Variansi)

Daerah Kritis

H_0 ditolak jika $F > F_{(k-1,n-k)}$

Analisis Variansi Satu Arah

<u>Tabel Anova</u>

Sumber Variansi	derajad bebas	Jumlah Kuadrat (SS)	Rata-rata Jumlah Kuadrat (MS)	rasio F
Perlakuan	$k-1$	SST = $\sum_{i=1}^{k} n_i(\bar{X}_i - \bar{X})^2$	MST = $\frac{SST}{k-1}$	$F = \frac{MST}{MSE}$
Sesatan	$N-k$	SSE = $\sum_{i=1}^{k} (n_i - 1)S_i^2$	MSE = $\frac{SSE}{N-k}$	

$$\bar{X}_i = \frac{1}{n_i}\sum_{j=1}^{n_i} x_{ij}$$

$$S_i^2 = \frac{1}{n_i-1}\sum_{j=1}^{n_i}(X_{ij} - \bar{X}_i)^2$$

$$\bar{X} = \frac{1}{N}\sum_{i=1}^{k}\sum_{j=1}^{n_i} x_{ij}, \quad N = \sum n_i$$

Analisis Variansi Satu Arah

Contoh:

Dipunyai empat varitas padi yang akan kita uji produktivitasnya. Dua puluh empat petak tanah yang kira-kira mempunyai kesuburan yang sama dipilih. Kemudian 24 petak itu dibagi secara random menjadi empat kelompok, masing-masing 6 petak yang selanjutnya tiap kelompok ditanami satu varitas padi. Apakah rata-rata produktivitas 4 varitas padi tersebut sama?

varitas			
A	B	C	D
24	13	21	27
13	21	13	30
18	11	26	24
24	23	23	29
16	28	16	26
23	18	12	34

Analisis Variansi Satu Arah

	varitas			
	A	B	C	D
	24	13	21	27
	13	21	13	30
	18	11	26	24
	24	23	23	29
	16	28	16	26
	23	18	12	34
n_i	6	6	6	6
$\bar{X}_i$	19,67	19,00	18,50	28,33
S_i^2	21,87	40,40	32,30	12,27

$$\bar{X} = \frac{1}{N}\sum_{i=1}^{k}\sum_{j=1}^{n_i} x_{ij} = \frac{513}{24} = 21,38$$

$$SST = \sum_{i=1}^{k} n_i(\bar{X}_i - \bar{X})^2$$

$$= 391,46$$

$$SSE = \sum_{i=1}^{k}(n_i - 1)S_i^2$$

$$= 534,17$$

Analisis Variansi Satu Arah

Tabel Anova Hasil 4 Varitas Padi

Sumber Variansi	derajad bebas	Jumlah Kuadrat (SS)	Rata-rata Jumlah Kuadrat (MS)	rasio F
Varitas	3	$\mathrm{SST} = 391,46$	$\mathrm{MST} = 130,49$	$4,8856$
Sesatan	20	$\mathrm{SSE} = 534,17$	$\mathrm{MSE} = 26,71$	

Statistik Penguji

$F = 4,8856$

Daerah Kritik ($\alpha = 0,05$)

H_0 ditolak jika $F > F_{(3,20)} = 3,10$

Kesimpulan

$F = 4,8856 > 3,10 \Rightarrow H_0$ ditolak, paling tidak ada dua mean yang tidak sama

Analisis Variansi Satu Arah

Pembandingan Ganda (*Multiple Comparisons*)

Merupakan analisis lanjutan bila H_0 ditolak dalam Anova. Metode:

- Tukey
- Scheffé
- Bonferron
- i

  Newman - Keuls

Analisis Variansi Satu Arah

Contoh

Pembandingan ganda untuk varitas padi contoh di muka. Diketahui (dari hitungan di muka):

	varitas			
	A	B	C	D
n_i	6	6	6	6
$\bar{X}_i$	19,67	19,00	18,50	28,33

MSE = $26,71$

Pembandingan ganda Scheffé:

Pembandingan	$\lvert \bar{X}_i - \bar{X}_j \rvert$	$\sqrt{248,403 \left(\frac{1}{n_i} + \frac{1}{n_j} \right)}$	Kesimpulan
μ_A vs. μ_B	0,67	9,1	H_0 diterima
μ_A vs. μ_C	1,17	9,1	H_0 diterima
μ_A vs. μ_D	8,66	9,1	H_0 diterima
μ_B vs. μ_C	0,50	9,1	H_0 diterima
μ_B vs. μ_D	9,33	9,1	H_0 ditolak
μ_C vs. μ_D	9,83	9,1	H_0 ditolak