BAB VI

STATISTIK INFERENSI MEAN SATU POPULASI

A. Tujuan Perkuliahan

Melakukan uji hipotesis terhadap rata-rata satu populasi dengan variansi diketahui dan tidak diketahui dengan software R.

B. Kompetensi Mahasiswa

1. Mahasiswa dapat melakukan pengujian hipotesis dan penarikan kesimpulan terhadap uji hipotesis rata-rata satu populasi dengan variansi diketahui (Uji Z) dengan software R.
2. Mahasiswa dapat melakukan pengujian hipotesis dan penarikan kesimpulan terhadap uji hipotesis rata-rata satu populasi dengan variansi tidak diketahui (Uji t) dengan software R.

C. Hipotesis Statistik

Dalam suatu penelitian, seorang penelti dihadapkan pada suatu masalah yang ingin diselesaikan. Masalah tersebut muncul dari suatu pernyataan yang merupakan dugaan dengan kemungkinan benar atau salah sehingga perlu dilakukan pembuktian lebih lanjut. Pernyataan atau dugaan tersebut dinamakan hipotesis yang digunakan sebagai dasar pengambilan keputusan. Hipotesis statistic merupakan suatu pernyataan atau dugaan yang mungkin benar atau salah tentang parameter dari satu atau lebih populasi yang dapat diuji secara empiris. Contoh hipotesis penelitian yang dapat diuji secara empiris diantaranya :

1. Apakah ada hubungan antara pendidikan pemilih dengan calon presiden yang dipilih?
2. Apakah aturan lalu lintas yang baru diterapkan di suatu ruas jalan dapat mengurai kemacetan yang biasa terjadi di ruas jalan tersebut?
3. Apakah bibit unggul padi varietas baru telah dapat meningkatkan produksi padi per hektar?

4. Dan sebagainya**.**

Jenis hipotesis statistic dibedakan menjadi 2 yaitu :

1. Hipotesis Null (H_0) : Pernyataan yang akan dibuktikan
2. Hipotesis Alternatif (H_1) : Hipotesis yang berlawanan dengan Hipotesis Null.

Beberapa hal yang berkaitan dengan uji hipotesis diantaranya :

1. Tingkat signifikansi (α) : peluang untuk melakukan kesalahan Tipe I
2. Tingkat kepercayaan $(1-\alpha)$: peluang keyakinan dapat menolak H_0
3. Peluang kesalahan tipe II (β)
4. Tingkat kekuatan uji $(1-\beta)$ atau seberapa besar peluang menolak H_0 jika H_0salah.

Skema tipe kesalahan pada uji hipotesis dapat dilihat pada table berikut.

Keputusan	**H_0 Benar**	**H_0 Salah**
Menerima H_0	Keputusan Benar $(1-\alpha)$	Kesalahan Tipe II (β)
Menolak H_0	Kesalahan Tipe I (α)	Keputusan Benar $(1-\beta)$

D. Uji Normalitas Data

Uji Normalitas digunakan untuk mengetahui apakah data yang ada berdistribusi normal atau tidak. Adapun langkah pengujian hipotesis untuk uji kenormalan (normalitas) data sebagai berikut.

1. Hipotesis :

   H_0 : Data berdistribusi normal

   H_1 : Data tidak berdistribusi normal
2. Tingkat signifikansi : α
3. Statistik Penguji : Perhatikan nilai $p-value$
4. Daerah kritis/ Kriteria penolakan H_0 :

   H_0 ditolak jika $p-value<\alpha$
5. Kesimpulan

Uji Normalitas berguna untuk menentukan data yang telah dikumpulkan berdistribusi normal atau diambil dari populasi normal. Metode klasik dalam pengujian normalitas suatu data tidak begitu rumit. Berdasarkan pengalaman empiris beberapa pakar statistik, data yang banyaknya lebih dari 30 angka (n > 30), maka sudah dapat diasumsikan berdistribusi normal. Biasa dikatakan sebagai sampel besar. Namun untuk memberikan kepastian, data yang dimiliki berdistribusi normal atau tidak, sebaiknya digunakan uji normalitas. Karena belum tentu data yang lebih dari 30 bisa dipastikan berdistribusi normal, demikian sebaliknya data yang banyaknya kurang dari 30 belum tentu tidak berdistribusi normal, untuk itu perlu suatu pembuktian.

Contoh :

Diberikan data umur sebanyak 44 Pasien yang berobat ke Puskesmas sbb:

76	18	45	50	60	22
17	26	27	50	38	18
42	35	19	41	7	60
24	50	60	13	37	62
80	52	10	21	60	28
12	30	9	39	38	45
8	9	45	22	33	24
50	25				

Ujilah data tersebut apakah data tersebut berdistribusi normal atau tidak! (α=0,05)

Penyelesaian :

Dengan menggunakan Rstudio, install package normtest, kemudian jalankan syntax sebagai berikut.

```
> Data.umur<- c(76, 18, 45, 50, 60, 22, 17, 26, 27, 50, 38, 18,
42, 35, 19, 41, 7, 60, 24, 50, 60, 13, 37, 62, 80, 52, 10, 21,
60, 28, 12, 30, 9, 39, 38, 45, 8, 9, 45, 22, 33, 24, 50, 25)

> Data.umur
```

```
 [1] 76 18 45 50 60 22 17 26 27 50 38 18 42 35 19 41  7 60 24
50 60 13 37 62 80 52
[27] 10 21 60 28 12 30  9 39 38 45  8  9 45 22 33 24 50 25
> #Uji Normalitas
> shapiro.test(Data.umur)
     Shapiro-Wilk normality test
data:  Data.umur
W = 0.9588, p-value = 0.1174
```

Langkah Uji Hipotesis :

1. Hipotesis : H0 = Data berdistribusi normal vs H1 = Data tidak berdistribusi normal
2. Tingkat signifikansi : $\alpha = 0{,}05$
3. Statistik Penguji : $p - value = 0{,}1174$
4. Kriteria penolakan H0 : karena $p - value = 0{,}1174 > \alpha = 0{,}05$ maka H0 tidak ditolak.
5. Kesimpulan : Karena H0 tidak ditolak berarti data tersebut berdistribusi normal.

E. Pengujian Hipotesis Untuk Mean Satu Populasi Dengan Simpangan Baku Diketahui (Uji Z)

Uji Z dapat digunakan untuk data yang simpangan bakunya diketahui, data berdistribusi normal dan dengan jumlah data (n) cukup besar (n>30).

Teorema 3.1.

Jika $X_1, X_2, \cdots, X_n$ adalah sampel random dari Populasi yang berdistribusi normal dengan mean μ dan variansi σ^2 maka variable random

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

akan mendekati normal standar dengan mean 0 dan simpangan baku 1.

Teorema 3.2.

Jika $X_1, X_2, \cdots, X_n$ adalah sampel random dari Populasi sembarang dengan mean μ dan variansi σ^2 maka untuk sampel berukuran cukup besar $(n > 30)$, berdasarkan Teorema 3.1., variable random

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

akan mendekati normal standar dengan mean 0 dan simpangan baku 1.

Dengan demikian, Berdasarkan Teorema 3.1. dan Teorema 3.2. , untuk melakukan uji rata-rata 1 populasi tidak perlu dilakukan uji normalitas. Akan dilakukan uji hipotesis bahwa mean suatu populasi sama dengan nilai tertentu μ_0, dengan n besar $(n > 30)$. Adapun langkah-langkah pengujian hipotesisnya sebagai berikut :

1. Tentukan Hipotesis
   a. $H_0 \quad : \mu = \mu_0$ (Uji Dua Sisi)
      $H_0 \quad : \mu \neq \mu_0$
   b. $H_0 \quad : \mu \leq \mu_0$ (Uji Satu Sisi)
      $H_0 \quad : \mu > \mu_0$
   c. $H_0 \quad : \mu \geq \mu_0$ (Uji Satu Sisi)
      $H_0 \quad : \mu < \mu_0$
2. Tentukan tingkat signifikansi : α
3. Statistik Penguji

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

4. Daerah Kritis (Kriteria Uji)
   H_0 ditolak jika (sesuaikan dengan hipotesis yang digunakan)
   a. $Z < -Z_{\frac{\alpha}{2}}$ atau $Z > \ -Z_{\frac{\alpha}{2}}$
   b. $Z > Z_\alpha$
   c. $Z < -Z_\alpha$

   Atau

   H_0 ditolak jika p-value < α
5. Kesimpulan

Kesimpulan apakah H_0 ditolak atau tidak ditolak, diambil berdasarkan langkah 4 dan hasil penghitungan pada langkah 3 pada tingkat signifikansi α.

Contoh :

Diberikan data umur sebanyak 44 Pasien yang berobat ke Puskesmas sbb:

76	18	45	50	60	22
17	26	27	50	38	18
42	35	19	41	7	60
24	50	60	13	37	62
80	52	10	21	60	28
12	30	9	39	38	45
8	9	45	22	33	24
50	25				

Misalkan diketahui bahwa simpangan baku dari data tersebut adalah 20, dengan tingkat signifikansi $\alpha = 5\%$, Apakah dapat disimpulkan bahwa pasien yang berobat rata-rata berusia 35 tahun?

Penyelesaian :

Untuk menyelesaikan masalah ini, kita gunakan uji z menggunakan Rstudio dengan terlebih dahulu menginstall package BSDA sbb :

```
> library(BSDA)

Loading required package: lattice

Attaching package: ‘BSDA’

The following object is masked from ‘package:datasets’:

Orange
```

Selanjutnya, ketikkan data tersebut pada Rstudio untuk mendapatkan ringkasan datanya juga.

```
> Data.umur<- c(76, 18, 45, 50, 60, 22, 17, 26, 27, 50, 38,
18, 42, 35, 19, 41, 7, 60, 24, 50, 60, 13, 37, 62, 80, 52,
10, 21, 60, 28, 12, 30, 9, 39, 38, 45, 8, 9, 45, 22, 33, 24,
50, 25)

> Data.umur

 [1] 76 18 45 50 60 22 17 26 27 50 38 18 42 35 19 41  7 60 24
50 60 13 37 62 80 52

[27] 10 21 60 28 12 30  9 39 38 45  8  9 45 22 33 24 50 25

> summary(Data.umur)

   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.

   7.00   20.50   34.00   34.93   50.00   80.00
```

Untuk memunculkan $p - value$ dari uji z satu populasi digunakan syntax berikut :

```
> zsum.test(mean.x=34.93, sigma.x = 20, n.x = 44,

+           alternative = "two.sided", mu = 35,

+           conf.level = 0.95)


     One-sample z-Test


data:  Summarized x

z = -0.023216, p-value = 0.9815

alternative hypothesis: true mean is not equal to 35

95 percent confidence interval:

 29.02049 40.83951

sample estimates:

mean of x
```

```
    34.93
```

Atau

```
> t.test(Data.umur, alternative='two.sided', mu=35)


      One Sample t-test


data:  Data.umur
t = -0.023971, df = 43, p-value = 0.981
alternative hypothesis: true mean is not equal to 35
95 percent confidence interval:
 29.19566 40.66798
sample estimates:
mean of x
 34.93182
```

Langkah pengujian :

1. Hipotesis : $H_0 : \mu = 35$ vs $H_1 : \mu \neq 35$
2. Tingkat Signifikansi : $\alpha = 5\% = 0{,}05$
3. Nilai $p - value = 0{,}9815$
4. Kriteria Penolakan : H_0 ditolak jika $p - value < \alpha$

   Karena $p - value = 0{,}9815 > \alpha = 0{,}05$ maka H_0 tidak ditolak.
5. Kesimpulan :

   Berdasarkan hasil yang diperoleh pada langkah 4, H_0 tidak ditolak berarti benar bahwa rata-rata pasien yang berobat rata-rata berusia 35 tahun.

Latihan :

Gunakan data umur pasien di atas dengan tingkat signifikansi $\alpha = 0{,}1$. Apakah dapat disimpulkan bahwa pasien yang berobat rata-rata berusia lebih dari 30 tahun?

F. Pengujian Hipotesis Untuk Mean Satu Populasi Dengan Variansi Tidak Diketahui (Uji t)

Uji t dapat digunakan untuk data yang simpangan bakunya diketahui, data berdistribusi normal dan dengan jumlah data (n) cukup kecil (n<30). Akan dilakukan uji hipotesis bahwa mean suatu populasi sama dengan nilai tertentu μ_0, dengan n besar $(n > 30)$. Adapun langkah-langkah pengujian hipotesisnya sebagai berikut :

1. Tentukan Hipotesis

   a. $H_0 \quad : \mu = \mu_0$ (Uji Dua Sisi)

      $H_0 \quad : \mu \neq \mu_0$

   b. $H_0 \quad : \mu \leq \mu_0$ (Uji Satu Sisi)

      $H_0 \quad : \mu > \mu_0$

   c. $H_0 \quad : \mu \geq \mu_0$ (Uji Satu Sisi)

      $H_0 \quad : \mu < \mu_0$

2. Tentukan tingkat signifikansi : α
3. Statistik Penguji

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

4. Daerah Kritis

   H_0 ditolak jika (sesuaikan dengan hipotesis yang digunakan)

   a. $t < -t_{\frac{\alpha}{2}}$ atau $t > \; -t_{\frac{\alpha}{2}}$

   b. $t > t_\alpha$

   c. $t < -t_\alpha$

   Atau

   H_0 ditolak jika p-value < α

5. Kesimpulan

   Kesimpulan apakah H_0 ditolak atau tidak ditolak, diambil berdasarkan langkah 4 dan hasil penghitungan pada langkah 3 pada tingkat signifikansi α.

Contoh :

Di bawah ini disajikan data tekanan darah sistolik (mmHg) dari 10 laki-laki dewasa.

183, 152, 178, 157, 194, 163, 144, 114, 178, 152

Apabila diasumsikan tekanan darah sistolik laki-laki dewasa berdistribusi normal dengan tingkat signifikansi 0,05, dapatkah kita simpulkan berdasarkan data di atas bahwa rata-rata tekanan darah sistolik laki-laki dewasa kurang dari 140mmHg?

Penyelesaian :

- Uji Normalitas

```
> Tekanan.darah<- c(183, 152, 178, 157, 194, 163, 144, 114,
178, 152)

> Tekanan.darah

 [1] 183 152 178 157 194 163 144 114 178 152

> #uji normalitas

> shapiro.test(Tekanan.darah)

     Shapiro-Wilk normality test

data:  Tekanan.darah

W = 0.95015, p-value = 0.6702
```

Langkah Uji Hipotesis :

1. Hipotesis : H0 = Data berdistribusi normal vs H1 = Data tidak berdistribusi normal
2. Tingkat signifikansi : $\alpha = 0{,}05$
3. Statistik Penguji : $p - value = 0{,}6702$

4. Kriteria penolakan H0 : karena $p - value = 0{,}6702 > \alpha = 0{,}05$ maka H0 tidak ditolak.
5. Kesimpulan : Karena H0 tidak ditolak berarti data tersebut berdistribusi normal.

- Uji t

```
> Tekanan.darah<- c(183, 152, 178, 157, 194, 163, 144, 114,
178, 152)
> Tekanan.darah
 [1] 183 152 178 157 194 163 144 114 178 152
> summary(Tekanan.darah)
   Min. 1st Qu.  Median    Mean 3rd Qu.    Max. 
  114.0   152.0   160.0   161.5   178.0   194.0 
> t.test(Tekanan.darah, Alternative='less', mu=140)

	One Sample t-test

data:  Tekanan.darah
t = 2.9353, df = 9, p-value = 0.01661
alternative hypothesis: true mean is not equal to 140
95 percent confidence interval:
 144.9306 178.0694
sample estimates:
mean of x 
    161.5 
```

Atau

```
> tsum.test(mean.x=160, s.x = 23.16, n.x = 10,
```

```
+           alternative = "greater", mu = 140,
+           var.equal = TRUE, conf.level = 0.95)


     One-sample t-Test


data:  Summarized x
t = 2.7308, df = 9, p-value = 0.0116
alternative hypothesis: true mean is greater than 140
95 percent confidence interval:
 146.5746       NA
sample estimates:
mean of x
      160
```

Langkah Pengujian Hipotesis :

1. Hipotesis : $H_0 : \mu \geq 35$ vs $H_1 : \mu < 35$
2. Tingkat Signifikansi : $\alpha = 5\% = 0{,}05$
3. Nilai $p - value = 0{,}0116$
4. Kriteria Penolakan : H_0 ditolak jika $p - value < \alpha$
5. Karena $p - value = 0{,}0116 > \alpha = 0{,}05$ maka H_0 ditolak.
6. Kesimpulan :
7. Berdasarkan hasil yang diperoleh pada langkah 4, H_0 ditolak berarti rata-rata tekanan darah sistolik laki-laki dewasa tidak kurang dari 140mmHg.

G. Latihan

Gunakan data tekanan sistolik laki-laki dewasa di atas dengan tingkat signifikansi $\alpha = 0{,}1$. Apakah dapat disimpulkan bahwa rata-rata tekanan darah laki-laki dewasa tidak lebih dari 155 mmHg?

1. Diberikan data jarak yang ditempuh sebuah mobil per bulannya di Jakarta (km) :

2059	2101	2419	2109
1945	1940	2028	1683
1840	2503	1969	2252
2377	1917	1959	2418
2310	2233	2065	2707
2194	1987	2023	1595
2005	1964	2670	2285
1938	2296	2378	2010

   Apabila standar deviasinya adalah 280 km dan tingkat signifikansinya adalah 5%. Selidiki apakah data tersebut berdistribusi normal atau tidak, kemudian apakah dapat disimpulkan bahwa jarak yang ditempuh mobil tersebut kurang dari 2100 km ?

2. Dari suatu sampel acak 100 catatan kematian di USA selama tahun lalu menunjukkan bahwa umur kematian rata-rata adalah 71,8 tahun dan simpangan bakunya 8,9 tahun. Apakah pernyataan ini menunjukkan bahwa harapan umur saat ini adalah lebih dari 70 tahun? (asumsikan tingkat signifikansi yang digunakan adalah 10%)

3. Menggunakan data state.x77 yang ada pada R untuk varibel Income, dengan tingkat signifikansi 5%, lakukan uji normalitas data dan selidiki apakah dapat disimpulkan bahwa rata-rata income di negara-negara tersebut tidak lebih dari $ 15.000?

4. Gunakan data ‘chickwts’ yang tersedia di R, dengan menggunakan tingkat signifikansi 5%, selidiki normalitas data dan selidiki apakah dapat disimpulkan bahwa rata-rata berat (weight) tidak kurang dari 150 ?