

# M.K. Statistika dan Probabilitas

## Minggu ke 13: **Estimasi**

Haryadi  
haryadi\_ump@yahoo.co.id

Universitas Muhammadiyah Palangka Raya  
May 24, 2026

Data sampel dari suatu populasi dapat digunakan untuk mencari informasi tentang karakter populasi. Misalnya, jika mean populasi  $\mu$  yang tidak diketahui, maka untuk memperoleh informasi tentang parameter  $\mu$  dapat digunakan mean sampel  $\bar{x}$ . Ini berarti statistik  $\bar{x}$  digunakan untuk mengestimasi (meduga) parameter mean populasi  $\mu$ . Dalam hal ini  $\bar{x}$  dinamakan **estimator** (penduga) untuk  $\mu$ . Secara umum, estimator untuk parameter populasi  $\theta$ , ditulis  $\hat{\theta}$ . Misalnya  $\hat{\mu} = \bar{x}$ .

**Estimator titik** suatu parameter adalah estimator yang berupa sebuah nilai tunggal. Sebagai contoh, dalam pernyataan "rata-rata hasil pengukuran kecepatan cahaya adalah 301.000 km/detik", nilai tersebut adalah suatu estimator titik.

**Definisi 1.** Suatu statistik dikatakan estimator **unbiased** parameter  $\theta$  jika nilai harapannya sama dengan  $\theta$ . Jika tidak demikian maka statistik tersebut dikatakan *bias*.

Pada kuliah sebelumnya telah disampaikan bahwa statistik

$$\bar{X} = \frac{\sum_{i=1}^n X_i}{n}$$

memiliki nilai harapan sama dengan mean populasi  $\mu$ . Dengan demikian  $\bar{X}$  merupakan estimator tak bias untuk parameter  $\mu$ . Demikian pula varian

sampel

$$S^2 = \frac{\sum_{i=1}^n (X_i - \bar{X})^2}{n - 1}$$

juga merupakan estimator tak bias parameter varian  $\sigma^2$ .

**Teorema 1.** Diberikan data sampel  $X_1, X_2, \dots, X_n$ .

- (i) Statistik  $\bar{X}$  merupakan estimator tak bias parameter mean populasi  $\mu$
- (ii) Statistik  $S^2$  merupakan estimator tak bias parameter varian populasi  $\sigma^2$

Selain kriteria tak bias, salah satu kriteria yang sering digunakan dalam estimasi adalah efisiensi.

**Definisi 2.** Suatu estimator  $\theta_1$  dikatakan lebih efisien dari pada estimator  $\theta_2$  jika varian  $\theta_1$  lebih kecil dibanding varian  $\theta_2$ .

Suatu estimator yang berupa interval dimana parameter diduga berada dinamakan **estimator interval**. Sebagai contoh, pernyataan "kecepatan cahaya berkisar antara 299.000 km/detik sampai dengan 305.000 km/detik", berarti nilai kecepatan cahaya yang sebenarnya dipercaya berada di antara kedua batas interval.

**Interval kepercayaan** untuk suatu parameter adalah suatu interval dalam mana parameter dipercaya berada. Misalkan  $\theta$  adalah parameter yang tidak diketahui nilainya. Untuk membentuk interval kepercayaan  $\theta$ , perlu dicari statistik  $U$  dan  $L$  sehingga peluang

$$P(L \leq \theta \leq U) = 1 - \alpha$$

adalah benar. Interval

$$L \leq \theta \leq U \tag{1}$$

dinamakan **interval kepercayaan** 100(1- $\alpha$ ) persen untuk parameter  $\theta$ .

Interval kepercayaan dapat diinterpretasikan sebagai berikut: jika kita mengambil sampel random berulang-ulang, maka  $100(1 - \alpha)$  persen dari semua nilai data akan memuat nilai  $\theta$  yang sebenarnya.

Di dalam persamaan 1,  $L$  dan  $U$  berturut-turut dinamakan batas bawah dan batas atas interval. Misalnya untuk  $\alpha = 0.05$ , maka persamaan 1 dinamakan interval kepercayaan 95 persen untuk  $\theta$ . Pada bagian berikut, kita akan belajar membentuk interval kepercayaan untuk paramter mean populasi  $\mu$ , varian populasi  $\sigma^2$  dan selisih dua mean populasi.

## 1 Interval kepercayaan $\mu$ dengan $\sigma$ diketahui

Diketahui  $X_1, X_2, \dots, X_n$  adalah sampel random dari populasi berdistribusi normal dengan mean  $\mu$  dan varian  $\sigma^2$ . Telah disampaikan bahwa  $\bar{X}$  merupakan estimator  $\mu$ . Namun demikian tidak dapat dipastikan bahwa  $\bar{X} = \mu$ , melainkan hanya dapat dinyatakan bahwa  $\mu$  berada di dalam interval tertentu.

Seperti telah diketahui di dalam materi sebelumnya bahwa  $\bar{X}$  berdistribusi normal dengan mean  $\mu$  dan varian  $\sigma^2/n$ . Oleh karena itu

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

berdistribusi normal standar. Dengan demikian

$$P\left(-B < \frac{(\bar{X} - \mu)}{\sigma/\sqrt{n}} < B\right) = 0.95$$

jika  $B = 1.96$ , yaitu

$$P\left(-1.96 < \frac{(\bar{X} - \mu)}{\sigma/\sqrt{n}} < 1.96\right) = 0.95. \quad (2)$$

Jika ketiga ruas pada ketaksamaan

$$-1.96 < \frac{(\bar{X} - \mu)}{\sigma/\sqrt{n}} < 1.96$$

dikalikan  $\sigma/\sqrt{n}$  kemudian ditambahkan  $\bar{X}$ , maka diperoleh

$$\bar{X} - 1.96 \frac{\sigma}{\sqrt{n}} < \mu < \bar{X} + 1.96 \frac{\sigma}{\sqrt{n}}$$

Dengan demikian (2) ekuivalen dengan

$$P\left(\bar{X} - 1.96\frac{\sigma}{\sqrt{n}} < \mu < \bar{X} + 1.96\frac{\sigma}{\sqrt{n}}\right) = 0.95$$

Interval

$$\left(\bar{X} - 1.96\frac{\sigma}{\sqrt{n}}, \bar{X} + 1.96\frac{\sigma}{\sqrt{n}}\right)$$

dinamakan *interval kepercayaan 95 persen* untuk  $\mu$ .

**Teorema 2.** *Jika  $x_1, x_2, \dots, x_n$  adalah data sampel dari distribusi dengan  $\sigma$  diketahui, dan  $\bar{x}$  adalah rata-rata sampel, maka interval kepercayaan 95 persen untuk  $\mu$  adalah*

$$\left(\bar{x} - 1.96\frac{\sigma}{\sqrt{n}}, \bar{x} + 1.96\frac{\sigma}{\sqrt{n}}\right).$$

**Contoh 1.** Dari data sampel random berat badan 100 orang dewasa di Palangka Raya diperoleh  $\bar{x} = 67.45$  kg dan penelitian sebelumnya menyatakan bahwa  $\sigma^2 = 8.6136$  kg. Carilah interval kepercayaan 95 persen rata-rata berat badan orang dewasa di Palangkaraya.

**Penyelesaian.** Berdasarkan yang diketahui, interval kepercayaan 95 persen untuk mean populasi adalah

$$\begin{aligned}\left(\bar{x} - 1.96\frac{\sigma}{\sqrt{n}}, \bar{x} + 1.96\frac{\sigma}{\sqrt{n}}\right) &= \left(67.45 - 1.96\frac{2.93}{\sqrt{100}}, 67.45 + 1.96\frac{2.93}{\sqrt{100}}\right) \\ &= (66.8748, 68.0252).\end{aligned}$$

Ini berarti 95 persen dapat dipercaya bahwa berat badan rata-rata orang dewasa di Palangka Raya berada di antara 66.8748 kg sampai dengan 68.0252 kg.

Dengan cara yang serupa, interval kepercayaan 99 persen untuk mean populasi  $\mu$  adalah

$$\left(\bar{X} - 2.58\frac{\sigma}{\sqrt{n}}, \bar{X} + 2.58\frac{\sigma}{\sqrt{n}}\right).$$

Secara umum, interval kepercayaan  $(1 - \alpha)$  persen untuk mean populasi  $\mu$  jika  $\sigma$  diketahui adalah

$$\left( \bar{X} - z_{\alpha/2} \frac{\sigma}{\sqrt{n}}, \bar{X} + z_{\alpha/2} \frac{\sigma}{\sqrt{n}} \right).$$

Nilai  $z_{\alpha/2}$  untuk beberapa tingkat kepercayaan yang sering digunakan adalah

Table 1: Tabel  $z_{\alpha/2}$

Tingkat kepercayaan	99.73 %	99 %	98 %	96 %	95.45 %	95 %	90 %
$z_{\alpha/2}$	3.00	2.58	2.33	2.05	2.00	1.96	1.645

## 2 Interval kepercayaan $\mu$ dengan $\sigma$ tidak diketahui

Misalkan  $X_1, X_2, \dots, X_n$  sampel random dari populasi berdistribusi normal dengan  $\mu$  dan  $\sigma^2$  keduanya tidak diketahui. Dapat dibuktikan bahwa

$$\frac{(\bar{X} - \mu)\sqrt{n}}{S}$$

berdistribusi  $t$  dengan derajat bebas  $n - 1$ .

Interval kepercayaan  $(1 - \alpha)$  persen untuk  $\mu$  dapat dibentuk sebagai berikut.

$$P\left(-t_{\alpha/2, n-1} < \frac{(\bar{X} - \mu)\sqrt{n}}{S} < t_{\alpha/2, n-1}\right) = 1 - \alpha$$

atau ekuivalen dengan

$$P\left(\bar{X} - t_{\alpha/2, n-1} \frac{S}{\sqrt{n}} < \mu < \bar{X} + t_{\alpha/2, n-1} \frac{S}{\sqrt{n}}\right) = 1 - \alpha.$$

**Teorema 3.** Diketahui  $x_1, x_2, \dots, x_n$  adalah data sampel dari distribusi dengan  $\sigma$  tidak diketahui. Jika  $\bar{X} = \bar{x}$  dan  $S = s$ , maka interval kepercayaan  $100(1 - \alpha)$  persen parameter  $\mu$  adalah

$$\left( \bar{x} - t_{\alpha/2, n-1} \frac{S}{\sqrt{n}}, \bar{x} + t_{\alpha/2, n-1} \frac{S}{\sqrt{n}} \right).$$

**Contoh 2.** Hasil pengamatan terhadap kinerja suatu jaringan memberikan hasil *packet loss* (dalam byte) sebagai berikut:  
189, 184, 183, 182, 181, 181, 187, 181, 185, 183, 184.  
Carilah interval kepercayaan 95 persen untuk rata-rata *packet loss* pada jaringan tersebut.

**Penyelesaian.** Mean sampel dan standar deviasi sampel data tersebut adalah  $\bar{x} = 183.67$  dan  $s = 2.58$ . Berdasarkan tabel  $t$  dengan  $\alpha = 0.05$  dan  $n = 11$  diperoleh  $t_{0.025,10} = 2.228$ . Oleh karena itu

$$t_{\alpha/2, n-1} \cdot \frac{s}{\sqrt{n}} = 2.228 \cdot \frac{2.58}{\sqrt{11}} = 1.73.$$

Dengan demikian berdasarkan Teorema 3, interval kepercayaan 95 persen *packet loss* tersebut adalah

$$(183.67 - 1.73, 183.67 + 1.73) = (181.94, 185.40).$$

Ini bermakna bahwa *packet loss* yang sebenarnya berada antara 181.94 byte sampai dengan 184.40 byte dengan keyakinan 95 persen.

**Contoh 3.** Berdasarkan data sampel random pengukuran 10 diameter pipa menghasilkan mean  $\bar{x} = 2.38$  cm dan deviasi standar  $s = 0.06$  cm. Carilah interval kepercayaan 95 persen untuk rata-rata diameter pipa yang sebenarnya.

**Penyelesaian.** Berdasarkan tabel  $t$  dengan  $\alpha = 0.05$  dan  $n = 10$  diperoleh  $t_{0.025,9} = 2.262$ . Interval kepercayaan 95 persen untuk diameter pipa yang sebenarnya adalah

$$\begin{aligned} \left( 2.38 - 2.262 \frac{0.06}{\sqrt{10}}, 2.38 + 2.262 \frac{0.06}{\sqrt{10}} \right) &= (2.38 - 0.04292, 2.38 + 0.04292) \\ &= (2.3371, 2.4229) \end{aligned}$$

yang berarti bahwa 95 persen dapat dipercaya diameter pipa yang sebenarnya antara 2.3371 dan 2.4229.

### 3 Interval kepercayaan untuk $\sigma^2$

Diketahui  $X_1, X_2, \dots, X_n$  sampel random dari populasi berdistribusi normal dengan mean  $\mu$  dan varian  $\sigma^2$  yang tidak diketahui. Dalam pembahasan sebelumnya tentang distribusi sampling, telah dibuktikan bahwa

$$(n-1) \frac{S^2}{\sigma^2}$$

berdistribusi Chi-square dengan derajat bebas  $n - 1$ .  
 Interval kepercayaan  $(1 - \alpha)$  persen untuk  $\sigma^2$  dapat dibentuk sebagai berikut.

$$P\left(\chi_{1-\alpha/2, n-1}^2 < (n-1)\frac{S^2}{\sigma^2} < \chi_{\alpha/2, n-1}^2\right) = 1 - \alpha$$

atau ekuivalen dengan

$$P\left(\frac{(n-1)S^2}{\chi_{\alpha/2, n-1}^2} < \sigma^2 < \frac{(n-1)S^2}{\chi_{1-\alpha/2, n-1}^2}\right) = 1 - \alpha.$$

**Teorema 4.** *Jadi jika  $S^2 = s^2$ , interval kepercayaan  $100(1 - \alpha)$  persen untuk parameter varian populasi  $\sigma^2$  adalah*

$$\left(\frac{(n-1)S^2}{\chi_{\alpha/2, n-1}^2}, \frac{(n-1)S^2}{\chi_{1-\alpha/2, n-1}^2}\right).$$

**Contoh 4.** Kapasitas 10 batere diukur dan hasilnya sebagai berikut (dalam ampere - jam):

140, 136, 150, 144, 148, 152, 138, 141, 143, 151.

(a) Carilah estimasi untuk varian populasi  $\sigma^2$ , dan (b) hitunglah interval kepercayaan 99 persen untuk  $\sigma^2$ .

**Penyelesaian.** (a) Dari data tersebut diperoleh  $\bar{x} = 144.3$ . Estimasi untuk varian populasi

$$S^2 = \sum_{i=1}^{10} \frac{(X_i - 144.4)^2}{10 - 1} = 32.23.$$

(b) Karena  $1 - \alpha = 0.99$  maka  $\alpha/2 = 0.01/2 = 0.005$ . Berdasarkan tabel Chi-square diperoleh  $\chi_{0.005, 9}^2 = 23.589$  dan  $\chi_{1-0.005, 9}^2 = \chi_{0.995, 9}^2 = 1.735$ . Jadi interval kepercayaan 99 persen untuk  $\sigma^2$  adalah

$$\left(\frac{(n-1)S^2}{\chi_{\alpha/2, n-1}^2}, \frac{(n-1)S^2}{\chi_{1-\alpha/2, n-1}^2}\right) = \left(\frac{9 \times 32.23}{23.589}, \frac{9 \times 32.23}{1.735}\right) = (12.30, 167.19),$$

yang berarti bahwa dapat dipercaya 99 persen nilai varian populasi berada pada interval (12.30, 167.19).

## 4 Interval kepercayaan selisih dua mean

Dalam suatu penelitian mungkin peneliti ingin membentuk interval kepercayaan selisih dua mean populasi. Misalkan  $X_1, X_2, \dots, X_n$  adalah sampel random dari populasi berdistribusi normal dengan mean  $\mu_1$  dan varian  $\sigma_1^2$ , dan  $Y_1, Y_2, \dots, Y_m$  adalah sampel random dari populasi berdistribusi normal dengan mean  $\mu_2$  dan varian  $\sigma_2^2$ . Akan dibentuk interval kepercayaan selisih kedua mean populasi, yaitu  $\mu_1 - \mu_2$ .

Dari pembahasan sebelumnya  $\bar{X}$  berdistribusi normal dengan mean  $\mu_1$  dan varian  $\sigma_1^2/n$ ; dan  $\bar{Y}$  berdistribusi normal dengan mean  $\mu_2$  dan varian  $\sigma_2^2/m$ . Oleh karena itu  $\bar{X} - \bar{Y}$  berdistribusi normal dengan mean  $\mu_1 - \mu_2$  dan varian  $\sigma_1^2/n + \sigma_2^2/m$ . Selanjutnya pembahasan akan dibagi menjadi dua, yaitu jika  $\sigma_1^2$  dan  $\sigma_2^2$  diketahui dan tidak diketahui.

Jika  $\sigma_1^2$  dan  $\sigma_2^2$  diketahui, maka

$$\frac{\bar{X} - \bar{Y} - (\mu_1 - \mu_2)}{\sqrt{\frac{\sigma_1^2}{n} + \frac{\sigma_2^2}{m}}}$$

berdistribusi normal standar. Dengan demikian diperoleh hasil berikut.

**Teorema 5.** *Interval kepercayaan 100(1- $\alpha$ ) persen untuk  $\mu_1 - \mu_2$  adalah*

$$\left( \bar{x} - \bar{y} - z_{\alpha/2} \sqrt{\frac{\sigma_1^2}{n} + \frac{\sigma_2^2}{m}}, \quad \bar{x} - \bar{y} + z_{\alpha/2} \sqrt{\frac{\sigma_1^2}{n} + \frac{\sigma_2^2}{m}} \right).$$

**Contoh 5.** Hasil pengamatan IQ yang diambil dari dua populasi anak yang tinggal di perkotaan (A) dan di pedesaan (B) adalah sebagai berikut:

Sampel A : 100 104 108 99 115 120 103 112 124 120 110

Sampel B : 121 101 120 98 99 104 104 107 111 123

Jika diketahui varians IQ anak-anak di perkotaan dan pedesaan berturut-turut adalah 7,02 dan 6,8, carilah interval kepercayaan 90 persen selisih rata-rata kedua kelompok anak.

**Penyelesaian.** Diketahui  $n_1 = 11$ ,  $n_2 = 10$ ,  $\sigma_1^2 = 7.02$  dan  $\sigma_2^2 = 6.8$ . Dapat dihitung bahwa

$$\bar{x}_1 = 110.45 \quad \bar{x}_2 = 108.8$$

Karena  $z_{\alpha/2} = z_{0.05} = 1.645$ , maka interval kepercayaan 90 persen untuk  $\mu_1 - \mu_2$  adalah

$$\left( 110.45 - 108.8 - (1.645)\sqrt{\frac{7.02}{11} + \frac{6.8}{10}}, 110.45 - 108.8 + (1.645)\sqrt{\frac{7.02}{11} + \frac{6.8}{10}} \right) \\ = (-0.24, 3.54)$$

yang berarti bahwa 90 persen dapat dipercaya selisih IQ anak di perkotaan dan di pedesaan berada pada interval  $-0.24$  sampai dengan  $3.54$ .

Jika varian kedua populasi tidak diketahui, namun nilainya sama, maka dapat digunakan varian sampel,

$$s_x^2 = \frac{1}{n-1} \sum_i (x_i - \bar{x})^2 \quad s_y^2 = \frac{1}{m-1} \sum_i (y_i - \bar{y})^2 \\ s_p^2 = \frac{(n-1)s_x^2 + (m-1)s_y^2}{n+m-2}. \quad (3)$$

**Teorema 6.** *Jika varian kedua populasi tidak diketahui, maka interval kepercayaan  $100(1 - \alpha)$  persen untuk  $\mu_1 - \mu_2$  adalah*

$$\left( \bar{x} - \bar{y} - t_{\alpha/2, n+m-2} s_p \sqrt{1/n + 1/m}, \bar{x} - \bar{y} + t_{\alpha/2, n+m-2} s_p \sqrt{1/n + 1/m} \right)$$

**Contoh 6.** Jika varian kedua populasi pada contoh 5 tidak diketahui tetapi dianggap sama, carilah interval kepercayaan 90 persen untuk beda mean kedua populasi.

**Penyelesaian.** Misalkan  $X$  dan  $Y$  berturut-turut menyatakan variabel random IQ anak perkotaan dan anak pedesaan. Berdasarkan data diatas,  $n = 11$  dan  $m = 10$ . Selanjutnya dapat dihitung bahwa

$$\bar{x} = 110.45 \quad \bar{y} = 108.8 \\ s_x^2 = 73.2727 \quad s_y^2 = 89.2889 \\ s_p^2 = \frac{(10)(73.2727 + (9)(89.2889))}{19} = 80.8593, \quad s_p = 8.9923$$

Berdasarkan tabel, nilai  $t_{\alpha/2, n+m-2} = t_{0.05, 19} = 1.729$ .

$$\begin{aligned} t_{\alpha/2, n+m-2} s_p \sqrt{1/n + 1/m} &= (1.729)(8.9923) \sqrt{\frac{1}{10} + \frac{1}{9}} \\ &= 7.1437, \end{aligned}$$

dan

$$\bar{x} - \bar{y} = 110.45 - 108.8 = 1.65.$$

Oleh karena itu interval 90 persen untuk  $\mu_1 - \mu_2$  adalah

$$(1.65 - 7.1437, 1.65 + 7.1437) = (-5.4937, 8.7937),$$

yang berarti dapat dipercaya 90 persen bahwa selisih mean IQ antara anak perkotaan dan anak pedesaan berada di antara -5.4937 dan 8.7937.

## Tugas

Suatu studi ingin mengetahui mean (rata-rata) tinggi badan mahasiswa UM Palangkaraya. Hasil pengukuran tinggi badan 10 mahasiswa UM Palangkaraya memberikan hasil (dalam cm)

150, 160, 165, 168, 155, 164, 168, 167, 170, 163.

- Hitunglah estimator tak bias untuk mean populasi  $\mu$
- Hitunglah estimator tak bias untuk varian populasi  $\sigma^2$
- Carilah interval kepercayaan 95 persen untuk mean populasi.