TEORI ESTIMASI

Salah satu tugas statistik inferensia adalah menemukan keterangan tentang populasi berdasarkan keterangan yang diperoleh dari sampelnya.

Biasanya nilai parameter populasi itu tidak diketahui dan sering tidak dapat dicari. Oleh karena itu usaha yang dilakukan adalah menghampiri atau mendekati nilai parameter itu.

Pendekatan nilai parameter populasi berdasar sampel disebut **estimasi (pendugaan)**

Teori Pendugaan

Suatu proses menduga parameter populasi dengan menggunakan statistik sampel

Pengujian Hipotesis

Suatu proses untuk memutuskan apakah hasil dugaan tersebut diterima atau tidak

Statistik yang dibentuk dari sampel mempunyai distribusi yang mengandung nilai-nilai parameter seperti yang terdapat dalam populasinya.

Melalui distribusi statistik tersebut ciri-ciri populasinya dapat dikenali.

Ada dua pokok bahasan statistik yang dapat dipelajari yaitu estimasi dan uji hipotesis.

Estimasi / Pendugaan

Suatu pernyataan mengenai parameter populasi yang diketahui berdasarkan informasi sampel.

Penduga atau Estimator

Suatu statistik (harga sampel) yang digunakan untuk menduga parameter.

Teori Pendugaan dikenal dua jenis pendugaan (estimasi) yaitu :

- Pendugaan Titik (Estimasi Titik).
  - Suatu pendugaan titik (point estimator) adalah suatu nilai (suatu titik) yang digunakan untuk menduga suatu parameter populasi.
  - Bila nilai parameter θ dari populasi hanya diduga dengan memakai satu nilai statistik θ dari sampel yang diambil dari populasi tersebut.
- Pendugaan Interval (Estimasi Interval).
  - Menunjukkan pada interval berapa suatu parameter populasi akan berada.
  - Bila nilai parameter θ dari populasi diduga dengan memakai beberapa nilai statistik θ yang berada dalam suatu interval, misalnya $\theta 1 < \theta < \theta 2$.

Estimasi titik

Sampel random banyaknya produksi (Ton) selama 20 hari :

170 165 169 157 172 ... 157

Estimasi titik :

Dalam hal ini, tidak dapat diketahui seberapa besar derajat kepercayaannya.

Penduga yang baik

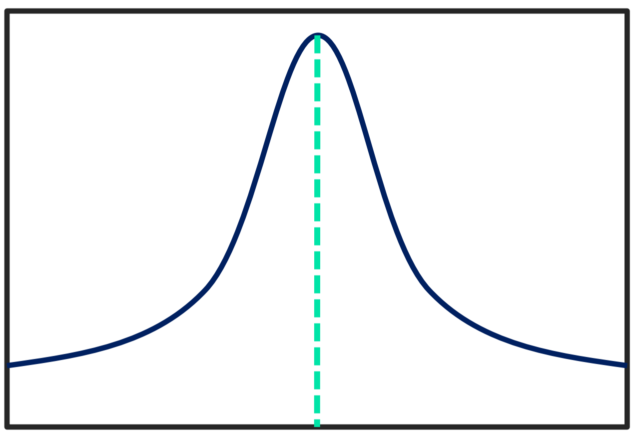
Ciri - ciri

- Tidak Bias
- Efisien
- Konsisten

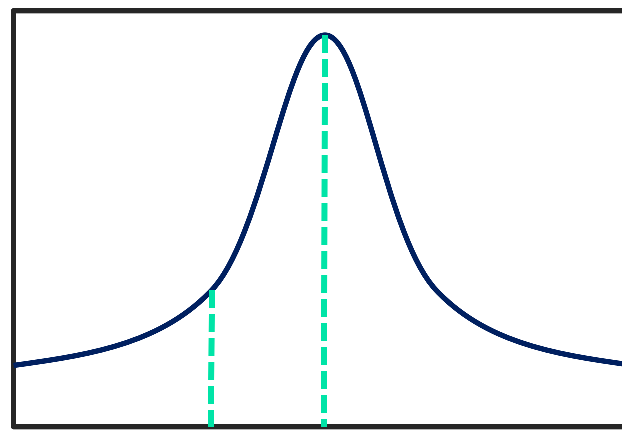
Tidak Bias

Tidak bias jika : E(harga statistik) = harga parameter

Penduga titik dikatakan tidak bias (unbiased estimator) jika di dalam sampel random yang berasal dari populasi, rata-rata atau nilai harapan (expected value) dari statistik sampel sama dengan parameter populasi



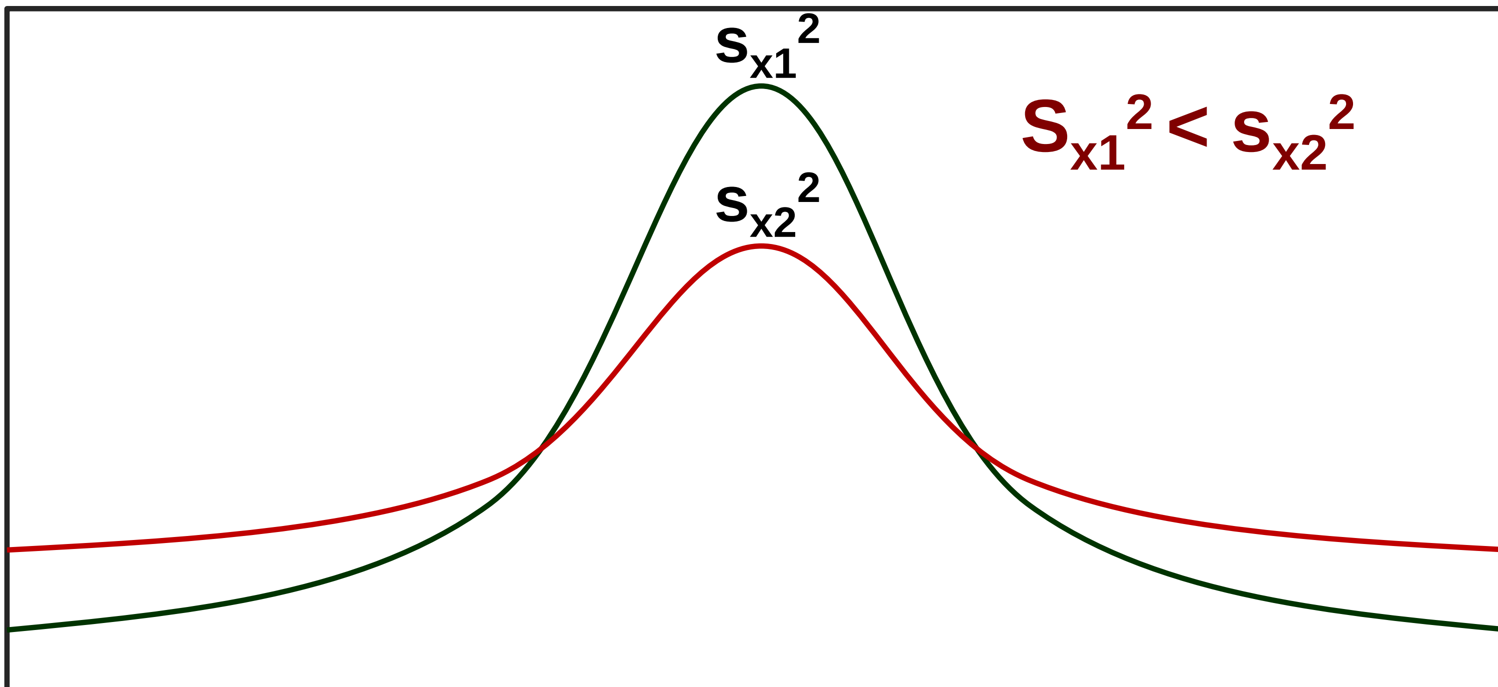
$$E(\overline{x}) = \mu$$



$$\mu \quad E(\overline{X})$$

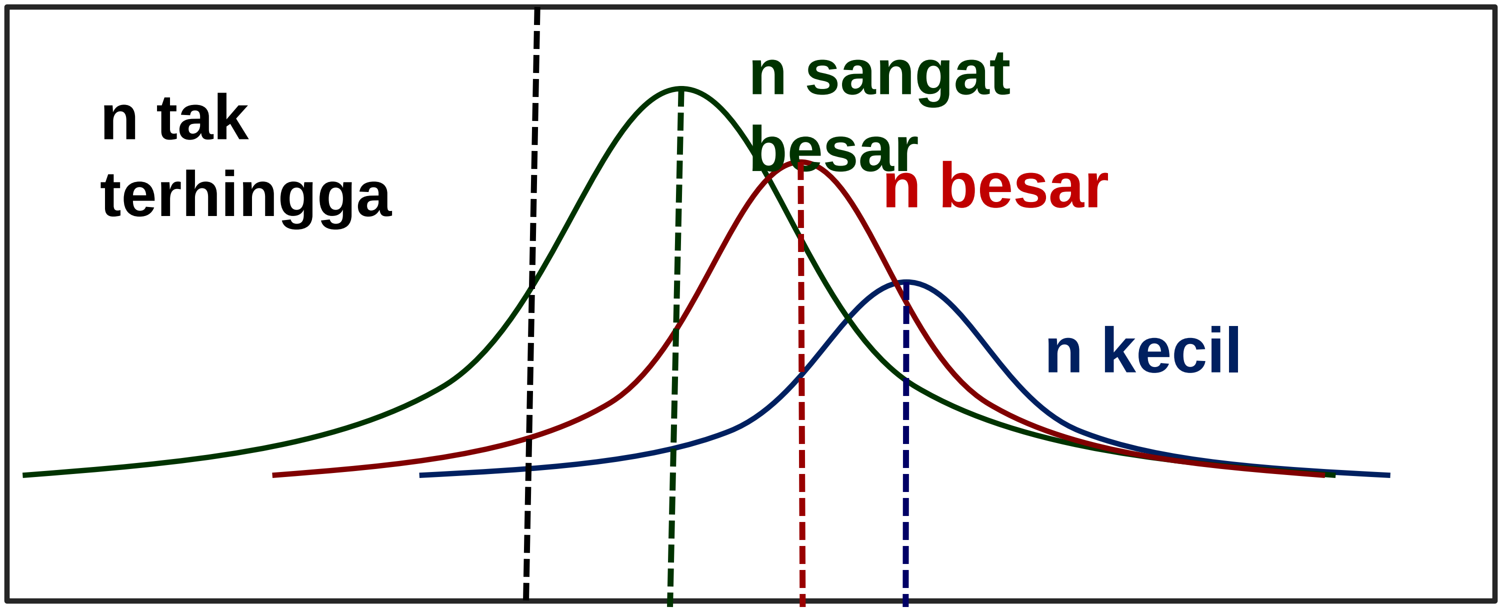
Efisien

Penduga yang efisien (efficient estimator) adalah penduga yang tidak bias dan mempunyai varian terkecil (s_x^2) dari penduga lainnya.



Konsisten

Penduga yang konsisten (consisten estimator) adalah nilai dugaan ($\overline{x}$) yang semakin mendekati nilai yang sebenarnya μ dengan semakin bertambahnya jumlah sample (n).



RUMUS INTERVAL PENDUGAAN

$$(s - Zs_x < P < s + Zs_x) = C$$

Di mana:

S	**: Statistik yang merupakan penduga parameter populasi (P)**
P	**: Parameter populasi yang tidak diketahui**
s_x	**: Standar deviasi distribusi sampel statistik**
Z	**: Suatu nilai yang ditentukan oleh probabilitas yang berhubungan dengan pendugaan interval, nilai Z diperoleh dari tabel luas di bawah kurva normal**
C	**: Probabilitas atau tingkat keyakinan yang dalam praktek sudah ditentukan dahulu.**
$s - Zs_x$	**: Nilai batas bawah keyakinan**
$s + Zs_x$	**: Nilai batas atas keyakinan**

Faktor – faktor yang penting dalam menyusun interval keyakinan rata-rata hitung :

- **Jumlah pengamatan dalam sampel (n).**
- **Besarnya standard deviasi.**
- **Tingkat keyakinan (C) dimana $\alpha = 1 - C$**

STANDAR ERROR

Merupakan besarnya maksimum error (kesalahan) yang dapat terjadi dalam menaksir mean populasi yang diperoleh berdasarkan pengamatan sejumlah n sample yang dipilih secara random.

$$\varepsilon = Z_{\frac{1}{2}\alpha} \frac{S}{\sqrt{n}}$$

SAMPLING ERROR

Sampling (Kesalahan penarikan sampel) merupakan perbedaan nilai statistik sampel dengan nilai parameter dari populasi.

DISTRIBUSI SAMPEL RATA-RATA

Jika diketahui rata-rata hitung (X) dan standard deviasi (S) suatu sample, jika ditransformasikan kedalam persamaan normal baku (Z), maka peluang suatu kejadian dapat diketahui dengan persamaan :

$$\mathbf{Z = (X - u) / S_x}$$

Jenis Populasi

Populasi terbatas (finite) : $n/N > 0{,}05$

Populasi tidak terbatas (infinite): $n/N < 0{,}05$

Untuk populasi yang tidak terbatas n/N < 0,05:

$$s_x = \frac{\sigma}{\sqrt{n}}$$

untuk populasi yang terbatas dan n/N> 0,05:

$$s_x = \frac{\sigma}{\sqrt{n}}\sqrt{\frac{N-n}{N-1}}$$

Di mana:

σ : Standar deviasi populasi

s_x : Standar error / kesalahan standar dari rata-rata hitung sampel

n : Jumlah atau ukuran sampel

N : Jumlah atau ukuran populasi

Estimasi Mean dan Proporsi

- Estimasi Mean n > 30
- Estimasi Mean n < 30
- Estimasi Selisih 2 Mean : data berpasangan
- Estimasi Selisih 2 Mean : n1 dan n2 > 30
- Estiamsi Selisih 2 Mean : n1 dan n2 < 30

- Estimasi Proposi n > 30
- Estimasi Proporsi < 30
- Estimasi Selisih 2 Proporsi : data berpasangan
- Estimasi Selisih 2 Proporsi : n1 dan n2 > 30
- Estiamsi Selisih 2 Proporsi : n1 dan n2 < 30

Estimasi Mean Populasi Besar.

$(X \pm Z_{½α}\ S\sqrt{n}) = (1 - α)$

atau:

$$\mathbf{(X - Z_{½α}\ S\sqrt{n}) \leq μ \leq (X + Z_{½α}\ S\sqrt{n})}$$

$(X – 1{,}96\ S\sqrt{n}) \leq μ \leq (X + 1{,}96\ S\sqrt{n})\ )$

Interval keyakinan untuk beberapa tingkat kepercayaan:

99% : $x \pm 2{,}58\ S\sqrt{n})$

98% : $x \pm 2{,}33\ S\sqrt{n})$

95% : $x \pm 1{,}96\ S\sqrt{n})$

90% : $x \pm 1{,}65\ S\sqrt{n})$

Estimasi Mean Populasi Kecil (n < 30).

$$t = (X - \mu) / S \sqrt{n}$$

Interval estimasi μ :

$$(X - t_{\alpha/2 \text{ df } n-1} \, S \sqrt{n}) \leq \mu \leq (X + t_{\alpha/2 \text{ df } n-1} \, S \sqrt{n})$$

Estimasi Selisih 2 Mean untuk data berpasangan (sample yang sama)

$$t = \frac{\bar{d} - \mu_d}{Sd/\sqrt{n}}$$

$\bar{d} = x1 - x2$

$\mu_d = \mu 1 - \mu 2$

$- t_{\alpha/2,df} \leq t \leq t_{\alpha/2,df}$

substitusi t , sehingga diperoleh rumus berikut:

$$\bar{d} - t_{\alpha/2,df}\, Sd/\sqrt{n} \leq \mu 1 - \mu 2 \leq \bar{d} + t_{\alpha/2,df}\, Sd/\sqrt{n}$$

df = n - 1

Estimasi Selisih 2 Mean untuk data tidak berpasangan (sample berbeda)

Sampel besar (n>30), σ1 dan σ2 diketahui, maka selisih mean akan berdistribusi normal.

$$Z = \frac{(\overline{X1} - \overline{X2}) - (\mu 1 - \mu 2)}{\sqrt{\dfrac{\sigma^2_1}{n1} + \dfrac{\sigma^2_2}{n2}}}$$

maka estimasi (μ1 - μ2) adalah :

$$(\overline{X1} - \overline{X2}) - Z_{\alpha/2}\sqrt{\frac{\sigma^2_1}{n1} + \frac{\sigma^2_2}{n2}} \leq (\mu 1 - \mu 2) \leq \text{dst}$$

Estimasi Selisih 2 Mean untuk data tidak berpasangan (sample berbeda)

Sampel kecil (n<30), σ1 dan σ2 tidak diketahui tetapi bisa diasumsikan nilai kedua ragaman populasi tersebut sama, maka rata-rata akan menyebar mengikuti distribusi t student.

$$t = \frac{(\overline{X1} - \overline{X2}) - (\mu 1 - \mu 2)}{Sp\sqrt{\dfrac{1}{n1} + \dfrac{1}{n2}}}$$

$$Sp = \sqrt{\frac{(n1 - 1)\, S^2_1 + (n2 - 1)\, S^2_2}{n1 + n2 - 2}}$$

maka estimasi (μ1 - μ2) adalah :

$$(\overline{X1} - \overline{X2}) - t_{\alpha/2,df}\ Sp\sqrt{\frac{1}{n1} + \frac{1}{n2}} \le (\mu 1 - \mu 2) \le (X1 - X2) + t_{\alpha/2,df}\ Sp$$

Contoh:

- Saham Bank Danamon di bursa terus mengalami fluktuasi. Harga saham pernah turun mencapai 1200 dan sempat naik mencapai 1600. Selama pengamatan 60 hari terakhir, harga saham Bank Danamon mencapai 1400 dengan standard deviasi 98. Berapa peluang saham Bank Danamon turun di bawah 1350 dan berapa peluang harganya meningkat di atas 1500?

SOAL LATIHAN

1. Taksir rata-rata IQ mahasiswa di Perguruan Tinggi bila dari sebanyak 100 sampel mahasiswa yang dipilih secara random memperlihatkan rata-rata IQ = 112 dan variance = 100.

   a. Pada Derajat Kepercayaan 95%.

   b. Pada Derajat Kepercayaan 99%.

2. Dari hasil sample survey terhadap 900 petani di daerah Labuhan Batu memperlihatkan bahwa rata-rata per tahun pengeluaran keluarga untuk pakaian Rp. 50.000,- dengan standard deviasi sample (S) = Rp. 10.000,-

   Hitunglah:

   a. Interval estimasi rata-rata pengeluaran petani untuk pakaian pada CL=95%.
   b. Dengan tingkat kepercayaan berapakah diperoleh hasil estimasi rata-rata pengeluaran untuk pakaian antara Rp.49.500,- hingga Rp.50.500,-

3. Kecepatan 4 mahasiswa yang dipilih secara random dapat selesai mengerjakan soal ujian dalam waktu 75 menit, 100 menit, 80 menit dan 65 menit. Dengan tingkat kepercayaan 95%, hitunglah interval estimasi rata-rata waktu mengerjakan ujian tersebut.

4. Pabrik mobil Peugeot ingin mengetahui berapa besarnya rata-rata jarak yang ditempuh oleh mobil tersebut untuk setiap liter bahan bakar yang digunakannya. Untuk menguji hal tersebut manager produksi pabrik melakukan pengujian terhadap 25 mobil yang dipilih secara random.

   Setelah dilakukan pengujian, ternyata rata-rata jarak yang ditempuh oleh mobil-mobil tersebut untuk setiap liter bahan bakar yang digunakan mencapai 15 km dengan variance sebesar 6,25 km.

   Dengan tingkat kepercayaan 95%, taksirlah berapa rata-rata jarak yang ditempuh oleh mobil-mobil keluaran pabrik Peugeot tersebut.

5. Sebuah perusahaan meluncurkan sebuah produk baru " Slimtrim " yaitu sebuah produk yang mengeluarkan aroma untuk mengurangi bobot badan. Untuk mengetahui efektivitas produk tersebut, dilakukan penelitian dengan mencoba pada 7 orang wanita sebagai sample.

Catatan bobot badan ke tujuh wanita tersebut sebelum dan sesudah menggunakan slimtrim selama dua minggu adalah sebagai berikut:

Wanita ke-	Bobot sebelum (kg)	Bobot sesudah (kg)
1	56	57
2	60	54
3	61	57
4	69	62
5	64	58
6	62	58
7	55	53

Estimasi penurunan berat badan rata-rata selama dua minggu setelah menggunakan "Slimtrim" pada tingkat kepercayaan 95%.

6. Sebuah peternakan sapi yang cukup besar ingin mengetahui sampai berapa besar perbedaan rata-rata berat sapi yang mereka pelihara bila dalam pemeliharaannya, sapi-sapi tersebut dibagi ke dalam 2 kelompok yang diberi makanan dan vitamin yang berbeda. Dari hasil pengamatan selama setengah tahun, pengusaha tersebut memilih secara random sebanyak 9 sapi dari tiap-tiap kelompok. Setelah diadakan penimbangan diperoleh informasi bahwa sapi pada kelompok 1 memiliki berat rata-rata sebesar 380 kg, sedang kelompok 2 sebesar 350 kg dengan variance sample kelompok 1 = 625 kg dan variance pada kelompok 2 = 400 kg.

   Estimasi berapa besarnya perbedaan berat rata-rata sapi antara kelompok 1 dan kelompok 2 pada tingkat kepercayaan 95%.

7. Investor pada saat ini dapat memilih investasi dalam bentuk deposito dan reksadana. Survei terhadap 18 bank dari 138 bank menunjukkan hasil deposito sebesar 7,71% dengan standard deviasi 0,73%. Sedang hasil reksadana pada 11 perusahaan dari 58 perusahaan adalah 13,17% dan standard deviasi 1,83%.

Dengan menggunakan tingkat keyakinan 95%, tentukan interval keyakinan dari selisih rata-rata hasil investasi tersebut?

8. Diketahui bahwa rata-rata curah hujan selama 15 tahun belakangan di daerah Bogor pada bulan Desember adalah 1,94 inchi sedangkan standard deviasinya = 0,45 inchi. Sebaliknya di daerah Bandung memiliki rata-rata curah hujan 1,04 inchi dan standard deviasinya 0,26 inchi selama 10 tahun ini.

   Hitunglah interval estimasi perbedaan rata-rata curah hujan di kedua daerah tersebut pada tingkat kepercayaan 95%.

9. Sebuah perusahaan taksi ingin mengetahui apakah ada perbedaan daya tahan umur dua merek ban yaitu ban A dan ban B. Untuk itu dilakukan percobaan dengan mengambil sample acak masing-masing 12 unit untuk setiap merek.

   Percobaan yang dilakukan menunjukkan hasil sebagai berikut :

   Rata-rata usia ban A adalah 110,3 km dengan standard deviasi 5 km.

   Rata-rata usia ban B adalah 100,2 km dengan standard deviasi 6 km.

   Estimasi beda rata-rata usia kedua ban mobil tersebut pada tingkat kepercayaan 95%.

10. Seorang pemilik pabrik ingin menduga rata-rata jam kerja yang hilang akibat terjadinya kerusakan mesin. Pencatatan selama satu tahun terakhir menunjukkan terjadinya sembilan kali kerusakan. Dari sample sebanyak sembilan tersebut diketahui rata-rata jam kerja yang hilang adalah 512,5 jam dengan standard deviasi sebesar 150,5 jam

    Buatlah interval (selang) bagi rata-rata jam kerja yang hilang akibat kerusakan mesin dengan tingkat kepercayaan 95%.