ESTIMASI

Topik :

1. Pendahuluan
   point estimator
   sifat-sifat estimator
2. Estimasi Interval
   - Estimasi interval rata-rata
   - Estimasi interval beda rata-rata
   - Estimasi interval proporsi
   - Estimasi interval beda proporsi

PENDAHULUAN

Pada bab sebelumnya sudah dibahas mengenai metode-metode statistik. Metode statistik dibagi menjadi dua bagian, yaitu statistika deskriptif dan statistika inferensia. Statistika deskriptif adalah metode-metode yang berkaitan dengan pengumpulan dan penyajian suatu data sedangkan statistika inferensia adalah semua metode yang berhubungan dengan analisis sebagian data untuk kemudian sampai pada penarikan kesimpulan mengenai keseluruhan data.

Pada bab ini, dibahas mengenai estimasi untuk data tunggal dan estimasi interval. Estimator adalah suatu statistik sampel yang digunakan untuk menduga suatu parameter yang tak diketahui. Untuk lebih jelasnya, dibahas pada subbab berikutnya.

POINT ESTIMATOR

Suatu angka tunggal yang digunakan untuk untuk menduga parameter dinamakan point estimator. Contohnya, rata-rata IPK mahasiswa Universitas Esa Unggul adalah 3,25 ($\bar{X} = 3{,}25$ adalah estimator point untuk μ), atau Proporsi mahasiswa yang tidak puas dengan pelayanan kampus adalah 15% ($\hat{p} = 0{,}15$ adalah estimator point untuk P). Dari dua contoh tersebut dapat disimpulkan bahwa $\bar{X}$ dan $\hat{p}$ adalah estimator point untuk μ dan p.

SIFAT-SIFAT ESTIMATOR

Sifat-sifat dari suatu estimator adalah sebagai berikut.

1. Tidak Bias.
   Suatu estimator dikatakan tidak bias apabila estimator tersebut secara tepat dapat menduga nilai parameternya.
2. Konsistensi
   Suatu estimator dikatakan konsisten apabila besarnya sampel semakin bertambah mendekati tidak berhingga maka estimator tersebut akan semakin berkonsentrasi secara sempurna pada parameter yang diduga.
3. Efisiensi
   Suatu estimator akan dikatakan efisien apabila memiliki varians yang kecil.
4. Sufisiensi
   Suatu estimator dikatakan sufisien apabila estimator itu mempunyai informasi yang lengkap dan cukup tentang parameter yang akan diduga. Dengan kata lain tidak ada ukuran statistik lain sebagai estimator yang lebih baik untuk menduga paramater.

ESTIMASI INTERVAL

Estimasi interval adalah suatu pendugaan terhadap parameter berdasarkan suatu interval , di dalam interval mana kita harapkan dengan keyakinan tertentu parameter itu akan terletak. Hasil dari pendugaan interval ini diharapkan akan lebih obyektif.

Pendugaan interval memberikan nilai interval, bukan nilai tunggal.

Untuk membuat estimasi interval, harus ditentukan terlebih dahulu besarnya tingkat keyakinan, yang diberi simbol $1-\alpha$. Umumnya, tingkat keyakinan yang digunakan adalah 90%, 95%, dan 99%.

Berdasarkan Teorema sentral limit, pernyataan probabilitas dapat dituliskan sebagai berikut.

$$P\left(-Z_{\frac{\alpha}{2}} < Z < Z_{\frac{\alpha}{2}}\right) = 1-\alpha$$

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

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

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

Misalkan $a = \bar{X}-Z_{\frac{\alpha}{2}}\frac{\sigma}{\sqrt{n}}$ adalah batas bawah dan $b = \bar{X}+Z_{\frac{\alpha}{2}}\frac{\sigma}{\sqrt{n}}$ adalah batas atas. Persamaan $P(a < \mu < b) = 1-\alpha$, yang menyatakan bahwa interval $a-b$ memuat rata-rata μ.

Misalnya, batas bawahnya diberikan 2,5 dan batas atasnya 3,5 dengan $1-\alpha = 90\%$. maka $P(2{,}5 < \mu < 3{,}5) = 90\%$, menunjukkan adanya probabilitas sebesar 90% bahwa pada interval 2,5 dan 3,5 akan memuat nilai rata-rata populasi yang sebenarnya yaitu μ. Kesalahan yang mungkin terjadi, probabilitasnya adalah 10%. Artinya, kemungkinan interval tersebut tidak memuat μ, artinya μ bisa lebih kecil dari 2,5 atau lebih besar dari 3,5.

ESTIMASI INTERVAL UNTUK RATA-RATA

Ada tiga rumus estimasi interval rata-rata μ.

Pertama,

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

Rumus ini berlaku jika sampel besar ($n \geq 30$) dari populasi yang tak terbatas atau dari populasi terbatas, tetapi penarikan sampel dilakukan dengan pengembalian.

Kedua,

$$\bar{X}-Z_{\frac{\alpha}{2}}\frac{\sigma}{\sqrt{n}}\sqrt{\frac{N-n}{N-1}} < \mu < \bar{X}+Z_{\frac{\alpha}{2}}\frac{\sigma}{\sqrt{n}}\sqrt{\frac{N-n}{N-1}}$$

Rumus ini digunakan kalau populasi terbatas, akan tetapi sampel sebanyak n diambil tanpa pengembalian dari populasi dengan N elemen dan σ diketahui.

Ketiga,

$$\bar{X}-t_{\frac{\alpha}{2}}\frac{s}{\sqrt{n}} < \mu < \bar{X}+t_{\frac{\alpha}{2}}\frac{s}{\sqrt{n}}$$

dengan

$$s = \sqrt{\frac{1}{n-1}\sum(X_i-\bar{X})^2}$$

Rumus ini berlaku bagi sampel kecil ($n < 30$) yang diambil dari populasi (σ tak diketahui) dengan pengembalian. Rumus ini pada dasarnya diperoleh dari rumus pertama dengan mengganti σ dengan s dan mengganti $Z_{\frac{\alpha}{2}}$ dengan $t_{\frac{\alpha}{2}}$. Nilai dari $Z_{\frac{\alpha}{2}}$ diperoleh dari tabel normal dan nilai $t_{\frac{\alpha}{2}}$ diperoleh dari tabel t dengan derajat kebebasan sebesar $(n-1)$.

Keterangan.
$\bar{X}$ = rata-rata sampel
σ = standar deviasi sampel
μ = rata-rata populasi
n = banyaknya sampel
N = banyaknya populasi
s = estimator σ

CONTOH 1.

Seratus orang calon mahasiswa Akademi Ilmu Statistik sebagai sampel acak, yang sudah mengikuti tes IQ, mempunyai rata-rata IQ sebesar 110 dan diketahui mempunyai deviasi standar sebesar 20. Dengan menggunakan tingkat keyakinan sebesar 95%, buatlah estimasi interval dari rata-rata IQ.

Penyelesaian.

Diketahui

$\bar{X} = 110$
$\sigma = 20$
$n = 100$
$1 - \alpha = 95\%$

Pertama-tama, dihitung nilai dari $Z_{\frac{\alpha}{2}}$ dari Tabel 1 yang ada pada Lampiran 1. sehingga diperoleh $Z_{\frac{\alpha}{2}} = 1{,}96$. Selanjutnya dipilih rumus yang mana akan digunakan. Rumus yang digunakan adalah rumus pertama karena σ diketahui. sehingga diperoleh

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

$$110 - 1{,}96 \left(\frac{20}{\sqrt{100}}\right) < \mu < 110 + 1{,}96 \left(\frac{20}{\sqrt{100}}\right)$$

$$106{,}08 < \mu < 113{,}92$$

Jadi, interval antara $106{,}08$ dan $113{,}92$ akan memuat rata-rata IQ sebenarnya dengan tingkat keyakinan 95%.

CONTOH 2.

Lima orang mahasiswa Fasilkom Universitas Esa Unggul, dipilih secara acak untuk kemudian diukur tinggi badannya. diperoleh data-datanya sebagai berikut. 160, 170, 165, 175, 180. Buatlah estimasi interval tentang rata-rata tinggi mahasiswa Fasilkom Universitas Esa Unggul dengan tingkat keyakinan 95%.

Penyelesaian.

Diketahui

$n = 5$
$1 - \alpha = 95\%$

Rumus yang digunakan adalah rumus ketiga karena σ tidak diketahui dan $n < 30$.

Pertama-tama dihitung nilai rata-rata dari sampel. diperoleh

$$\bar{X} = \frac{1}{n}\sum X_i = \frac{1}{5}(160 + 170 + 165 + 175 + 180) = \frac{1}{5}(850) = 170$$

dan dihitung deviasi standar dari sampel, diperoleh

$$\begin{aligned} s &= \sqrt{\frac{1}{n-1}\sum (X_i - \bar{X})^2} \\ &= \sqrt{\frac{1}{4}\left((160-170)^2 + (170-170)^2 + (165-170)^2 + (175-170)^2 + (180-170)^2\right)} \\ &= \sqrt{\frac{1}{4}(100 + 0 + 25 + 25 + 100)} = \sqrt{\frac{1}{4}(250)} = 7{,}906 \end{aligned}$$

Setelah itu, dihitung nilai dari $t_{\frac{\alpha}{2}}$ dengan menggunakan tabel yang ada pada Lampiran 1, dengan nilai $\alpha = 0{,}05$ dan derajat kebebasan $= n - 1 = 5 - 1 = 4$. Cara membaca tabelnya adalah mencari baris yang derajat kebebasannya 4, dan mencari kolom yang $\alpha = 0{,}05$ dan yang dua arah (two tails), sehingga diperoleh

$$t_{\frac{\alpha}{2}} = 2{,}776$$

Selanjutnya, dihitung estimasi intervalnya, dengan menggunakan rumus yang ketiga. sehingga diperoleh

$$\bar{X} - t_{\frac{\alpha}{2}}\frac{s}{\sqrt{n}} < \mu < \bar{X} + t_{\frac{\alpha}{2}}\frac{s}{\sqrt{n}}$$

$$170 - 2{,}776\left(\frac{7{,}906}{\sqrt{5}}\right) < \mu < 170 + 2{,}776\left(\frac{7{,}906}{\sqrt{5}}\right)$$

$$170 - 9{,}815 < \mu < 170 + 9{,}815$$

$$160{,}185 < \mu < 179{,}815$$

Jadi, interval antara $160{,}185$ dan $179{,}815$ akan memuat rata-rata tinggi mahasiswa Fasilkom Universitas Esa Unggul dengan tingkat keyakinan 95%.

ESTIMASI INTERVAL UNTUK BEDA RATA-RATA

Misalkan $\bar{X}_1$ dan $\bar{X}_2$ adalah beda rata-rata sampel dari dua populasi dengan distribusi sampling masing-masing mendekati distribusi normal, maka estimasi interval selisih rata-rata dua populasi yang merupakan sumber pengambilan sampel, dirumuskan sebagai berikut.

$$(\bar{X}_1 - \bar{X}_2) - Z_{\frac{\alpha}{2}}\,\sigma_{(\bar{X}_1 - \bar{X}_2)} < \mu_1 - \mu_2 < (\bar{X}_1 - \bar{X}_2) + Z_{\frac{\alpha}{2}}\,\sigma_{(\bar{X}_1 - \bar{X}_2)}$$

dengan

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

jika $n \geq 30$, serta σ_1 dan σ_2 diketahui.

Jika $n < 30$, serta σ_1 dan σ_2 tidak diketahui, maka rumus yang digunakan adalah sebagai berikut.

$$(\bar{X}_1 - \bar{X}_2) - t_{\frac{\alpha}{2}} s_{(\bar{X}_1 - \bar{X}_2)} < \mu_1 - \mu_2 < (\bar{X}_1 - \bar{X}_2) + t_{\frac{\alpha}{2}} s_{(\bar{X}_1 - \bar{X}_2)}$$

dengan

$$s_{(\bar{X}_1 - \bar{X}_2)} = \sqrt{\frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2}} \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}$$

$$s_1^2 = \frac{1}{n_1 - 1} \sum_{i=1}^{n_1} (X_{i1} - \bar{X}_1)^2$$

$$s_2^2 = \frac{1}{n_2 - 1} \sum_{i=1}^{n_2} (X_{i2} - \bar{X}_2)^2$$

Keterangan.

μ_1 = rata-rata populasi 1
μ_2 = rata-rata populasi 2
n_1 = banyaknya sampel 1
n_2 = banyaknya sampel 2
$\bar{X}_1$ = rata-rata sampel 1
$\bar{X}_2$ = rata-rata sampel 2
σ_1 = standar deviasi sampel 1
σ_2 = standar deviasi sampel 2
s_1^2 = varians sampel 1
s_2^2 = varians sampel 2
X_{i1} = data ke i dari sampel 1
X_{i2} = data ke i dari sampel 2

CONTOH 3.

Lampu pijar merek ampuh memiliki rata-rata daya tahan 4500 jam dengan deviasi standar 500 jam, sedangkan lampu pijar merek baik memiliki rata-rata daya tahan 4000 jam dengan deviasi standar 400 jam. Jika diambil sampel masing-masing 100 buah lampu pijar dan diteliti, buatlah estimasi interval untuk selisih rata-rata daya tahan dari kedua lampu pijar tersebut dengan tingkat keyakinan 90%.

Penyelesaian.

Misalkan

1 = lampu pijar merk ampuh

2 = lampu pijar mek baik

maka diketahui

$\bar{X}_1 = 4500$, $\sigma_1 = 500$

$\bar{X}_2 = 4000$, $\sigma_2 = 400$

$n_1 = n_2 = 100$

$1 - \alpha = 90\%$

Karena nilai σ_1 dan σ_2 diketahui, serta sampel berukuran besar maka rumus yang digunakan untuk menduga selisih rata-rata daya tahan dari kedua lampu pijar adalah sebagai berikut.

$$(\bar{X}_1 - \bar{X}_2) - Z_{\frac{\alpha}{2}}\, \sigma_{(\bar{X}_1 - \bar{X}_2)} < \mu_1 - \mu_2 < (\bar{X}_1 - \bar{X}_2) + Z_{\frac{\alpha}{2}}\, \sigma_{(\bar{X}_1 - \bar{X}_2)}$$

Pertama-tama dihitung terlebih dahulu nilai dari $Z_{\frac{\alpha}{2}}$ dengan menggunakan tabel pada Lampiran 1. sehingga diperoleh

$$Z_{\frac{\alpha}{2}} = 1{,}645$$

Selanjutnya, dihitung nilai dari $\sigma_{(\bar{X}_1 - \bar{X}_2)}$, sehingga diperoleh

$$\sigma_{(\bar{X}_1 - \bar{X}_2)} = \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}} = \sqrt{\frac{500^2}{100} + \frac{400^2}{100}} = \sqrt{4100} = 64{,}031$$

Selanjutnya, dihitung estimasi intervalnya.

$$(\bar{X}_1 - \bar{X}_2) - Z_{\frac{\alpha}{2}}\, \sigma_{(\bar{X}_1 - \bar{X}_2)} < \mu_1 - \mu_2 < (\bar{X}_1 - \bar{X}_2) + Z_{\frac{\alpha}{2}}\, \sigma_{(\bar{X}_1 - \bar{X}_2)}$$

$$(4500 - 4000) - 1{,}645\ (64{,}031) < \mu_1 - \mu_2 < (4500 - 4000) + 1{,}645\ (64{,}031)$$

$$500 - 105{,}331 < \mu_1 - \mu_2 < 500 + 105{,}331$$

$$394{,}669 < \mu_1 - \mu_2 < 605{,}331$$

Jadi, interval antara $349{,}669$ dan $605{,}331$ akan memuat selisih rata-rata daya tahan dari kedua lampu pijar merk baik dan ampuh dengan tingkat keyakinan 90%.

CONTOH 4.

Untuk mengetahui apakah ada perbedaan rata-rata gaji bulanan bagi karyawan dari dua perusahaan A dan B, maka dilakukan terhadap 9 karyawan yang dipilih secara acak sebagai sampel dari masing-masing perusahaan. hasil wawancara adalah sebagai berikut.

Karyawan	Gaji per bulan (puluhan ribu rupiah)	
	Perusahaan A	Perusahaan B
1	40	30
2	46	24
3	50	16
4	36	25
5	38	35
6	34	40
7	42	46
8	44	38
9	30	34

Buatlah estimasi inteval dari selisih rata-rata gaji karyawan dari kedua perusahaan tersebut, dengan tingkat keyakinan sebesar 95%.

Penyelesaian.

Misalkan

1 = perusahaan A

2 = perusahaan B

Karena pada soal tidak diketahui rata-rata dan deviasi standar dari populasi, maka akan diduga nilai rata-rata dan deviasi standarnya dari sampel yang ada dengan menggunakan rumus yang sudah dijelaskan diatas. pertama dicari terlebih dahulu nilai rata-rata dari perusahaan A dan B. sehingga diperoleh

$$\bar{X}_1 = \frac{1}{n}\sum x_i = \frac{1}{9}(40 + 46 + 50 + 36 + 38 + 34 + 42 + 44 + 30) = \frac{1}{9}(360) = 40$$

dan

$$\bar{X}_2 = \frac{1}{n}\sum x_i = \frac{1}{9}(30 + 24 + 16 + 25 + 35 + 40 + 46 + 38 + 34) = \frac{1}{9}(288) = 32$$

Selanjutnya dihitung nilai varians dari masing-masing sampel. sehingga diperoleh

$$\begin{aligned} s_1^2 &= \frac{1}{n_1 - 1}\sum_{i=1}^{n_1}(X_{i1} - \bar{X}_1)^2 \\ &= \frac{1}{9-1}\begin{pmatrix} (40-40)^2 + (46-40)^2 + (50-40)^2 + (36-40)^2 + (38-40)^2 \\ +(34-40)^2 + (42-40)^2 + (44-40)^2 + (30-40)^2 \end{pmatrix} \\ &= \frac{1}{8}(0 + 36 + 100 + 16 + 4 + 36 + 4 + 16 + 100) = \frac{1}{8}(312) = 39 \end{aligned}$$

dan

$$\begin{aligned} s_2^2 &= \frac{1}{n_2 - 1}\sum_{i=1}^{n_2}(X_{i2} - \bar{X}_2)^2 \\ &= \frac{1}{9-1}\begin{pmatrix} (30-32)^2 + (24-32)^2 + (16-32)^2 + (25-32)^2 + (35-32)^2 \\ +(40-32)^2 + (46-32)^2 + (38-32)^2 + (34-32)^2 \end{pmatrix} \\ &= \frac{1}{8}(4 + 64 + 256 + 49 + 9 + 64 + 196 + 36 + 4) = \frac{1}{8}(682) = 85{,}25 \end{aligned}$$

Kemudian dihitung nilai dari deviasi standar selisih dari perusahaan A dan B dengan menggunkaan rumus sebagai berikut.

$$\begin{aligned} s_{(\bar{X}_1 - \bar{X}_2)} &= \sqrt{\frac{(n_1-1)s_1^2 + (n_2-1)s_2^2}{n_1 + n_2 - 2}}\sqrt{\frac{1}{n_1} + \frac{1}{n_2}} = \sqrt{\frac{(9-1)\,39 + (9-1)85{,}25}{9+9-2}}\sqrt{\frac{1}{9} + \frac{1}{9}} \\ &= \sqrt{\frac{994}{16}}\sqrt{0{,}222} = \sqrt{62{,}125}\,\sqrt{0{,}222} = 3{,}714 \end{aligned}$$

Dari soal diketahui bahwa tingkat keyakinannya adalah 95%, artinya nilai dari $\alpha = 5\% = 0{,}05$. Nilai $t_{\frac{\alpha}{2}}$ dengan derajat kebebasan $= n_1 + n_2 - 2 = 9 + 9 - 2 = 16$ adalah 2,12.

Selanjutnya, dihitung estimasi interval dari selisih rata-rata gaji bulanan karyawan perusahaan A dan B dengan menggunakan rumus sebagai berikut.

$$(\bar{X}_1 - \bar{X}_2) - t_{\frac{\alpha}{2}}\, s_{(\bar{X}_1 - \bar{X}_2)} < \mu_1 - \mu_2 < (\bar{X}_1 - \bar{X}_2) + t_{\frac{\alpha}{2}}\, s_{(\bar{X}_1 - \bar{X}_2)}$$

$$(40 - 32) - 2{,}12\ (3{,}714) < \mu_1 - \mu_2 < (40 - 32) + 2{,}12\ (3{,}714)$$

$$8 - 7{,}874 < \mu_1 - \mu_2 < 8 + 7{,}874$$

$$0{,}126 < \mu_1 - \mu_2 < 15{,}874$$

Jadi, interval antara $0{,}126$ dan $15{,}874$ akan memuat selisih rata-rata gaji bulanan karyawan dari perusahaan A dan perusahaan B dengan tingkat keyakinan 95%.

ESTIMASI INTERVAL UNTUK PROPORSI

Proporsi sampel adalah estimator tak bias terhadap proporsi populasi. jika ukuran sampel cukup besar, yaitu nP maupun $n(1 - P)$ lebih besar $5,$ dimana P adalah proporsi populasi, maka distribusi sampling proporsi akan mendekati distribusi normal dengan rata-rata P dan standar deviasi proporsi

$$\sigma = \sqrt{\frac{P(1 - P)}{n}}$$

dalam menghitung σ dibutuhkan pengetahuan tentang P. Umumnya, P tidak diketahui. sehingga nilai dari σ diestimasi dengan s_p dan P diestimasi dengan p. sehingga

$$s_p = \sqrt{\frac{p(1 - p)}{n}}\sqrt{\frac{N - n}{N - 1}}$$

untuk populasi berhingga, dan

$$s_p = \sqrt{\frac{p(1 - p)}{n}}$$

untuk populasi tak berhingga.

p adalah estimator untuk P, maka ukuran sampel dikatakan cukup besar jika np dan $n(1 - p)$ lebih besar dari 15. Rumus untuk estimasi proporsi populasi adalah sebagai berikut.

$$p - Z_{\frac{\alpha}{2}}\, s_p < P < p + Z_{\frac{\alpha}{2}}\, s_p$$

dengan

$p = \frac{X}{n} =$ proporsi untuk sampel
$X =$ banyaknya elemen dengan karakteristik tertentu
$n =$ banyaknya sampel
$s_p =$ standar deviasi untuk sampel
$P =$ proporsi untuk populasi

CONTOH 5.

Tiga puluh lima dari sampel random sebanyak 500 angkatan kerja dijumpai sedang menganggur. Buatlah interval keyakinan proporsi penganggur di daerah itu dengan menggunakan tingkat keyakinan sebesar 90%.

Penyelesaian.

Misalkan

$X =$ banyaknya penganggur

Diketahui.

$X = 35$

$n = 500$

$p = \frac{X}{n} = \frac{35}{500} = 0{,}07$

$\alpha = 10\% = 0{,}05$

$Z_{\frac{\alpha}{2}} = 1{,}645$

Pertama-tama, dihitung deviasi standar dengan menggunakan rumus

$$s_p = \sqrt{\frac{p(1-p)}{n}} = \sqrt{\frac{0{,}07(1-0{,}07)}{500}} = \sqrt{\frac{0{,}0651}{500}} = \sqrt{0{,}0001302} = 0{,}011$$

Selanjutnya, dihitung estimasi interval proporsi dengan rumus

$$p - Z_{\frac{\alpha}{2}}\, s_p < P < p + Z_{\frac{\alpha}{2}}\, s_p$$

$$0{,}07 - 1{,}645\ (0{,}011) < P < 0{,}07 + 1{,}645\ (0{,}011)$$

$$0{,}07 - 0{,}018 < P < 0{,}07 + 0{,}018$$

$$0{,}052 < P < 0{,}088$$

Jadi, interval antara $0{,}052$ dan $0{,}088$ akan memuat proporsi penganggur di daerah tersebut dengan tingkat keyakinan sebesar 90%.

ESTIMASI INTERVAL UNTUK BEDA PROPORSI

Misalkan P_1 dan P_2 adalah beda proporsi sampel dari dua populasi dengan distribusi sampling masing-masing mendekati distribusi normal, maka estimasi interval selisih proporsi dua populasi yang merupakan sumber pengambilan sampel, dirumuskan sebagai berikut.

$$(p_1 - p_2) - Z_{\frac{\alpha}{2}}\, s_{(p_1-p_2)} < P_1 - P_2 < (p_1 - p_2) + Z_{\frac{\alpha}{2}}\, s_{(p_1-p_2)}$$

dengan

$$p_1 = \frac{X_1}{n_1}$$

dan

$$p_2 = \frac{X_2}{n_2}$$

serta

$$s_{(p_1-p_2)} = \sqrt{\frac{p_1(1-p_1)}{n_1} + \frac{p_2(1-p_2)}{n_2}}$$

Keterangan:

P_1 = proporsi populasi 1 dengan karakteristik tertentu
P_1 = proporsi populasi 2 dengan karakteristik tertentu
p_1 = proporsi sampel 1 dengan karakteristik tertentu
p_2 = proporsi elemen 2 dengan karakteristik tertentu
X_1 = banyaknya elemen 1 dengan karakteristik tertentu
X_2 = banyaknya elemen 2 dengan karakteristik tertentu
n_1 = banyaknya sampel 1
n_2 = banyaknya sampel 2
$s_{(p_1-p_2)}$ = deviasi standar selisih proporsi 1 dan 2.

CONTOH 6.

BKKBN melakukan penelitian di dua daerah (D_1 dan D_2) untuk mengetahui apakah ada perbedaan antara persentase penduduk yang setuju KB di daerah tersebut. kemudian akan dibuat estimasi interval mengenai besarnya selisih / perbedaan persentase tersebut. Di daerah D_1 dan D_2 masing-masing dilakukan wawancara terhadap 120 orang, antara lain menanyakan apakah mereka setuju KB atau tidak.

Dari D_1 ada 90 orang dan dari D_2 ada 78 orang yang setuju KB. Buatlah estimasi interval dari perbedaan persentase tentang pendapat penduduk yang setuju dengan KB di kedua daerah tersebut dengan tingkat keyakinan sebesar 90%.

Penyelesaian.

Misalkan

X_1 = penduduk daerah D_1 yang setuju KB

X_2 = penduduk daerah D_2 yang setuju KB

Diketahui

$X_1 = 90$
$X_2 = 78$
$n_1 = 120$
$n_2 = 120$
$\alpha = 10\%$
$Z_{\frac{\alpha}{2}} = 1{,}645$ (Lihat tabel pada Lampiran 1).
Pertama-tama, dihitung nilai dari p_1 dan p_2.

$$p_1 = \frac{X_1}{n_1} = \frac{90}{120} = 0{,}75$$

$$p_2 = \frac{X_2}{n_2} = \frac{78}{120} = 0{,}65$$

Selanjutnya, dihitung nilai standar deviasi dari selisih proporsi kedua daerah tersebut, dengan menggunakan rumus sebagai berikut.

$$s_{(p_1-p_2)} = \sqrt{\frac{p_1(1-p_1)}{n_1} + \frac{p_2(1-p_2)}{n_2}}$$

$$= \sqrt{\frac{0{,}75(1-0{,}75)}{120} + \frac{0{,}65(1-0{,}65)}{120}}$$

$$= \sqrt{\frac{0{,}1875}{120} + \frac{0{,}2275}{120}}$$

$$= \sqrt{0{,}00156 + 0{,}00189}$$

$$= \sqrt{0{,}00345}$$

$$= 0{,}059$$

Selanjutnya dihitung estimasi interval untuk selisih proporsi dari kedua daerah tersebut dengan menggunakan rumus berikut.

$$(p_1 - p_2) - Z_{\frac{\alpha}{2}}\, s_{(p_1-p_2)} < P_1 - P_2 < (p_1 - p_2) + Z_{\frac{\alpha}{2}}\, s_{(p_1-p_2)}$$

$$(0{,}75 - 0{,}65) - 1{,}645\ (0{,}059) < P_1 - P_2 < (0{,}75 - 0{,}65) - 1{,}645\ (0{,}059)$$

$$0{,}1 - 0{,}097 < P_1 - P_2 < 0{,}1 + 0{,}097$$

$$0{,}003 < P_1 - P_2 < 0{,}197$$

Jadi, interval antara $0{,}003$ dan $0{,}197$ akan memuat perbedaan persentase tentang pendapat penduduk yang setuju dengan KB di daerah D_1 dan D_2 dengan tingkat keyakinan sebesar 90%.

DAFTAR PUSTAKA

Daniel,W.W, Cross.C.L. 2013. *Biostatistics*. Edisi kesembilan. Amerika: John Wiley and Sons.

Mulyono, sri. 2006. *Statistika untuk Ekonomi dan Bisnis*, Edisi ketiga. Jakarta: Lembaga Penerbit Fakultas Ekonomi Universitas Indonesia.

Sugiono, Prof. DR. 2017. *Statistika untuk Penelitian*. Bandung: Alfabeta.

Supranto, J. 2008. *Statistik Teori dan Aplikasi,* Edisi ketujuh; Jilid 1. Jakarta: Erlangga.

Supranto, J. 2009. *Statistik Teori dan Aplikasi,* Edisi ketujuh; Jilid 2. Jakarta: Erlangga.

Walpole, Ronald E. 2017. *Pengantar Statistika*, Edisi ketiga. Jakarta: Gramedia Pustaka Utama.

Walpole.R.E, Myers.R.H, Myers.S.L, Ye.K. 2012. *Probability & Statistics for Engineers and Scientists*. Ninth ed. Amerika: Pearson Education.

LAMPIRAN

Tabel Z dan t

t Table

cum. prob	$t_{.50}$	$t_{.75}$	$t_{.80}$	$t_{.85}$	$t_{.90}$	$t_{.95}$	$t_{.975}$	$t_{.99}$	$t_{.995}$	$t_{.999}$	$t_{.9995}$
one-tail	0.50	0.25	0.20	0.15	0.10	0.05	0.025	0.01	0.005	0.001	0.0005
two-tails	1.00	0.50	0.40	0.30	0.20	0.10	0.05	0.02	0.01	0.002	0.001
df											
1	0.000	1.000	1.376	1.963	3.078	6.314	12.71	31.82	63.66	318.31	636.62
2	0.000	0.816	1.061	1.386	1.886	2.920	4.303	6.965	9.925	22.327	31.599
3	0.000	0.765	0.978	1.250	1.638	2.353	3.182	4.541	5.841	10.215	12.924
4	0.000	0.741	0.941	1.190	1.533	2.132	2.776	3.747	4.604	7.173	8.610
5	0.000	0.727	0.920	1.156	1.476	2.015	2.571	3.365	4.032	5.893	6.869
6	0.000	0.718	0.906	1.134	1.440	1.943	2.447	3.143	3.707	5.208	5.959
7	0.000	0.711	0.896	1.119	1.415	1.895	2.365	2.998	3.499	4.785	5.408
8	0.000	0.706	0.889	1.108	1.397	1.860	2.306	2.896	3.355	4.501	5.041
9	0.000	0.703	0.883	1.100	1.383	1.833	2.262	2.821	3.250	4.297	4.781
10	0.000	0.700	0.879	1.093	1.372	1.812	2.228	2.764	3.169	4.144	4.587
11	0.000	0.697	0.876	1.088	1.363	1.796	2.201	2.718	3.106	4.025	4.437
12	0.000	0.695	0.873	1.083	1.356	1.782	2.179	2.681	3.055	3.930	4.318
13	0.000	0.694	0.870	1.079	1.350	1.771	2.160	2.650	3.012	3.852	4.221
14	0.000	0.692	0.868	1.076	1.345	1.761	2.145	2.624	2.977	3.787	4.140
15	0.000	0.691	0.866	1.074	1.341	1.753	2.131	2.602	2.947	3.733	4.073
16	0.000	0.690	0.865	1.071	1.337	1.746	2.120	2.583	2.921	3.686	4.015
17	0.000	0.689	0.863	1.069	1.333	1.740	2.110	2.567	2.898	3.646	3.965
18	0.000	0.688	0.862	1.067	1.330	1.734	2.101	2.552	2.878	3.610	3.922
19	0.000	0.688	0.861	1.066	1.328	1.729	2.093	2.539	2.861	3.579	3.883
20	0.000	0.687	0.860	1.064	1.325	1.725	2.086	2.528	2.845	3.552	3.850
21	0.000	0.686	0.859	1.063	1.323	1.721	2.080	2.518	2.831	3.527	3.819
22	0.000	0.686	0.858	1.061	1.321	1.717	2.074	2.508	2.819	3.505	3.792
23	0.000	0.685	0.858	1.060	1.319	1.714	2.069	2.500	2.807	3.485	3.768
24	0.000	0.685	0.857	1.059	1.318	1.711	2.064	2.492	2.797	3.467	3.745
25	0.000	0.684	0.856	1.058	1.316	1.708	2.060	2.485	2.787	3.450	3.725
26	0.000	0.684	0.856	1.058	1.315	1.706	2.056	2.479	2.779	3.435	3.707
27	0.000	0.684	0.855	1.057	1.314	1.703	2.052	2.473	2.771	3.421	3.690
28	0.000	0.683	0.855	1.056	1.313	1.701	2.048	2.467	2.763	3.408	3.674
29	0.000	0.683	0.854	1.055	1.311	1.699	2.045	2.462	2.756	3.396	3.659
30	0.000	0.683	0.854	1.055	1.310	1.697	2.042	2.457	2.750	3.385	3.646
40	0.000	0.681	0.851	1.050	1.303	1.684	2.021	2.423	2.704	3.307	3.551
60	0.000	0.679	0.848	1.045	1.296	1.671	2.000	2.390	2.660	3.232	3.460
80	0.000	0.678	0.846	1.043	1.292	1.664	1.990	2.374	2.639	3.195	3.416
100	0.000	0.677	0.845	1.042	1.290	1.660	1.984	2.364	2.626	3.174	3.390
1000	0.000	0.675	0.842	1.037	1.282	1.646	1.962	2.330	2.581	3.098	3.300
z	0.000	0.674	0.842	1.036	1.282	1.645	1.960	2.326	2.576	3.090	3.291
	0%	50%	60%	70%	80%	90%	95%	98%	99%	99.8%	99.9%
	Confidence Level										