BAB II ESTIMASI STATISTIK

2.1 Pengertian Estimasi

a. Estimasi merupakan suatu metode dimana kita dapat memperkirakan nilai Populasi dengan memakai nilai sampel.

b. Estimasi merupakan kegiatan penarikan kesimpulan statistik yang berawal dari hal-hal yang bersifat umum ke hal – hal yang bersifat khusus, agar penarikan kesimpulan dapat dibenarkan dan mampu mendekati kebenaran maka dibutuhkan suatu alat untuk memproses data secara benar, jika kegiatan estimasi dapat dilakukan secara benar maka semua keputusan yang berkaitan dengan estimasi dapat dilakukan juga dengan benar dan dapat untuk mengatasi segala persoalan statistik.

c. Estimasi adalah menaksir ciri-ciri tertentu dari populasi atau memperkirakan nilai populasi (parameter) dengan memakai nilai sampel (statistik).

2.2 Estimator yang Baik

2.2.1 Pengertian Estimator

a. Estimator adalah nilai penduga.

b. Estimator adalah setiap statistik (mean sampel, presentase sampel, varians sampel, dan lain lain) yang digunakan untuk mengestimasi sebuah parameter.

2.2.2 Kriteria Estimator yang Baik

Estimator yang Baik

Estimator yang baik mempunyai tiga ciri berikut :

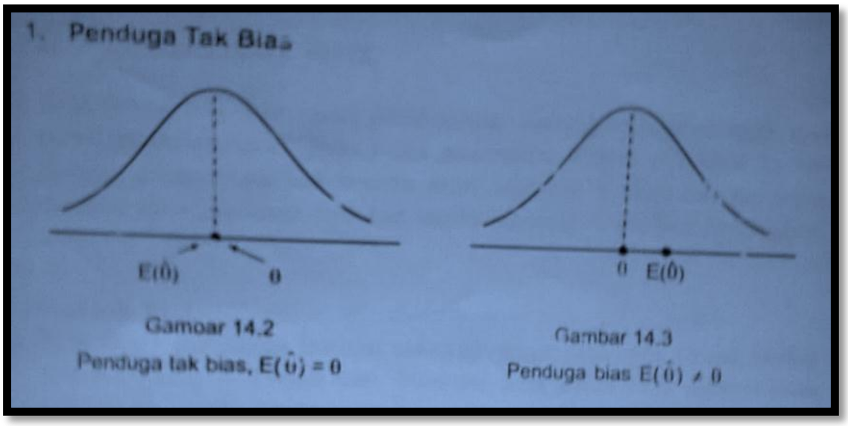
1. $\hat{\theta}$ (statistik dari sampel) merupakan penduga tak bias dari θ, yaitu E ($\hat{\theta}$) = , artinya harapan penduga $\hat{\theta}$ sama dengan θ;

2. $\hat{\theta}$ merupakan penduga yang efisien; artinya bila ada lebih dari satu penduga, maka penduga yang efisien adalah penduga yang mempunyai variansi paling kecil; dan

3. $\hat{\theta}$ merupakan penduga yang konsisten; artinya bila sampel yang diambil makin besar, maka nilai $\hat{\theta}$ akan semakin mendekati θ.

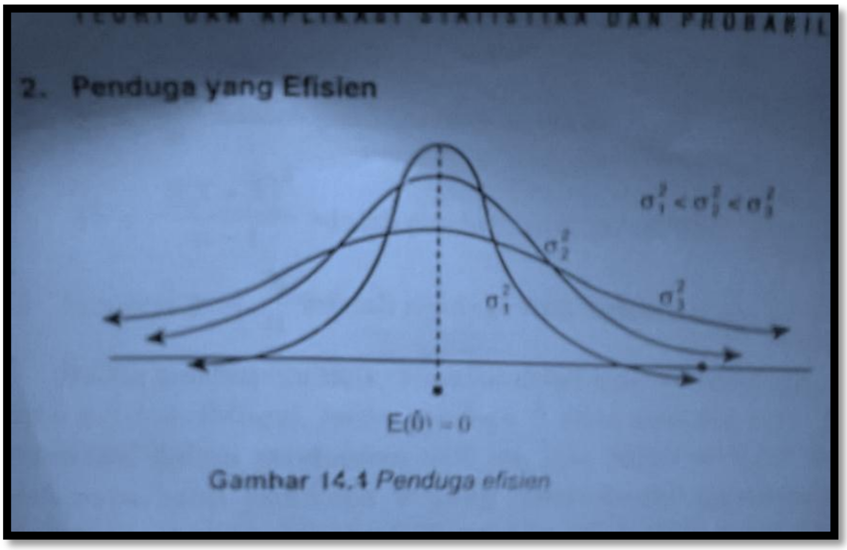
Gambar berikut ini yang berkaitan dengan tiga ciri estimator yang baik tersebut.

1. Penduga Tak Bias



Penduga tak bias artinya penduga yang dengan tepat mengenai sasaran, seperti ditunjukkan oleh gambar 14.2. sedangkan penduga bias artinya penduga yang tidak tepat mengenai sasaran atau disebut meleset , seperti ditunjukkan oleh gambar 14.3

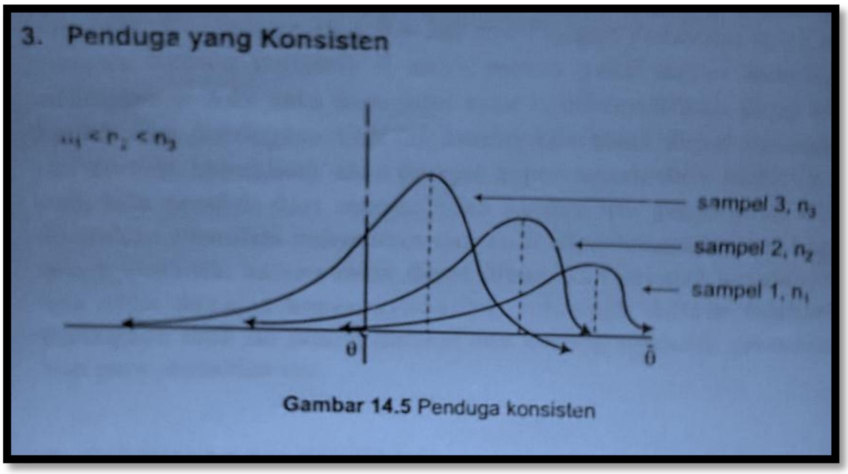
2. Penduga yang Efisien



Gambar 14.4 menunjukkan ada tiga penduga, yaitu $\hat{\theta}_1$, $\hat{\theta}_2$, $\hat{\theta}_3$ yang diperoleh dari tiga sampel, dimana distribusi sampel 1 mempunyai variansi σ_1^2, sampel 2 mempunyai variansi σ_2^2 , dan sampel 3

mempunyai variansi σ_3^2. oleh karena sampel 1 mempunyai variansi paling kecil, maka dikatakan $\hat{\theta}_1$ merupakan penduga yang efisien

3. Penduga yang Konsisten



Gambar 14.5 Penduga konsisten

Pada Gambar 14.5, ditunjukkan bahwa ukuran sampel 1, yaitu n_1, lebih kecil daripada ukuran sampel 2, yaitu n_2 dan lebih kecil dari ukuran sampel 3, yaitu n_3. Terlihat bahwa semakin besar ukuran sampel, maka statistik penduga $\hat{\theta}$ makin mendekati parameter θ dari populasi, dimana distribusi sampel *konsisten bergerak kekiri*.

2.3 Jenis - Jenis Estimasi

Statistik inferensial adalah statistika yang dengan segala informasi dari sample digunakan untuk menarik kesimpulan mengenai karakteristik populasi dari mana sample itu di ambil. Untuk menarik kesimpulan tersebut dapat dilakukan dengan duacar, yaitu: penaksiran parameter dan pengujian hipotesis. Parameter adalah karakteristik dari suatu populasi contohnya: m(mean atau rata-rata), s (varian atau keseragaman) , md (mean differensial atau perbedaan rata-rata), p (Proporsi), pd (Proporsi rata-rata) , sedangkan statistik ialah karakteristik dari data sample, contoh: x (mean atau rata-rata), S (varian atau keseragaman), xd (mean differensial atau perbedaan rata-rata), P(Proporsi), Pd (Proporsi rata-rata). Adanya statistic itu untuk menduga parameter, seperti untuk menduga σ dapat diduga dengan S . Estimasi adalah keseluruhan proses yang menggunakan sebuah estimator(fungsi sample) untuk menghasilkan sebuah estimate(nilai terealisasi dari estimator, yaitu bilangan yang didapat bila sample benar-benar diambil) dari suatu parameter.

Atau dengan kata lain suatu estimator titik adalah sebarang fungsi T (X , X , X , …, X) dari sampel. Ini berarti sebarang titik adalah estimator titik.

Jenis-jenis Estimasi :

a. Estimasi Titik

Titik estimasi merupakan salah satu cara untuk mengadakan estimasi terhadap parameter populasi yang tidak diketahui (Nar: 2011). Titik estimasi ialah nilai tunggal yang digunakan untuk mengadakan estimasi terhadap parameter populasi. Titik estimasi yang dapat digunakan untuk mengadakan estimasi parameter populasi ialah rata-rata sampel terhadap rata-rata populasi, proporsi sampel terhadap proporsi populasi, jumlah variabel tertentu yang terdapat dalam sampel untuk menaksir jumlah variabel tersebut dalam populasi, dan varians atau simpangan baku sampel untuk menaksir simpangan baku populasi.

E (μ) = ; E (σ2) = S2 ; E (p) =

Sebuah estimasi titik dari sebuah parameter q adalah sesuatu angka tunggal yang dapat dianggap sebagai nilai yang masuk akal dari q.

Adapun Hakikat Estimasi secara umum, yaitu:

1) Estimasi adalah taksiran, dan yang diestimasi adalah parameter populasi

2) Data yang digunakan untuk melakukan estimasi parameter populasi adalah statistik sampel sebagai estimator

3) Terdapat prosedur tertentu untuk melaksanakan estimasi

Salah satu Contoh dalam kehidupan sehari-hari dari estimasi titik itu ialah: rata-rata bayiyang baru lahir beratnya 3 kg. Dari pernyataan ini dapat disimpulkan bahwa pernyataan tersebut dapat menggunakan distribusi normal, karena tidak banyak bayi yang baru lahir itu beratnya kurang atau lebih berat dari 3kg. bayi yang beratnya 3kg dapat disebut sebagai estimate, sedangkan yang disebut estimatornya ialah bayi yang baru lahir.

Dalam menentukan estimasi titik(menaksir) itu harus mengetahui distribusinya terlebih dahulu, apabila distribusi tersebut tidak diketahui maka kita tidak dapat menentukan estimasi titiknya.

1.Metode penaksir titik

Metode penaksir titik sebuah parameter dapat ditempuh ddengan menggunakan bebebrapa metode, yaitu: metode momen, metode likelihood, Metode Kuadrat Terkecil, Metode Chikuadrat Minimum, estimator bayes dan lainnya.

Akan tetapi tidak semua metode sering digunakan, yang paling banyak digunakan dalam penentuan parameter titik hanya dua metode saja, yaitu Metode Momen dan Metode Kemungkinan Maksimum atau metode likelihood.

Berikut sedikit tentang kedua metode tersebut dan estimator bayes untuk perbandingan dari kedua metode tersebut kenapa kedua metode itu sering digunakan.

1. Metode momen

Metode moment diciptakan oleh Karl Pearson pada tahun 1800. Metode ini merupakan metode tertua dalam menentukan estimator titik.

Misalkan X adalah peubah acak kontinu(atau diskrit) dengan fungsi kepadatan peluang berbentuk f(x;), dengan adalah k buah para meter yang tidal diketehui, misalkan X , X , X , …, X merupakan sebuah sampel acak berukuran n dan didefinisikan k buah momen sekitar pusat sample pertama sebagai:

; t= 1,2, 3, …, k

Kemudian kita tentukan k buah momen sekitar pusat populasi pertama dengan rumussebagai berikut:

Untuk diskrit : ; t=1, 2, 3,…., k

Untuk diskrit: ; t=1, 2, 3,…., k

Secara umum, momen populasi merupakan fungsi dair k buah parameteryang tidak diketahui. Penyamaan momen sampel dan momen populasi akan menghasilkan k buah persamaan dalam k buah parameter yang tidak diketahui , yaitu:

; t = 1, 2, 3, …, k

Solusi dari persamaan diatas, dinotasikan dengan menghasilkan penaksir momen untuk

Contoh 1:

Misalkan peubah acak X berdistribusi B(1;) dengan tidak diketahui. Tentukan penaksir titik untuk dengan menggunakan metode momen.

Penyelesaian:

Fungsi kepaadatan peluang dari X adalah:

f(1;) = ; x = 0, 1

= 0 ; lainnya

Contoh 2:

Misalkan x , x , x , ….., x adalah sample random dari populasi yang berdistribusi X ~ N (m, s). Dengan menggunakan metode momen, tentukan estimator titik untuk m dan s

Penyelesaian:

X ~ N (m, s) berarti E(X) = m dan Var(Z) = s

Var (X) = E (X)-(E(X))

s = E (X) – m

E (X) = m + s

Sehingga memperoleh persamaan seperti berikut:

E(X) = sehingga m =

E (X) = maka m + s =

s = -

= -`X

=

1.Metode kemungkinan maksimum(likelihood)

Metode yang terbaik untuk menentukan penaksir titik sebuah parameter adalah metode kemungkinan maksimum(likelihood). Metode kemungkinan maksimum merupakan metode untuk memperoleh estimator titik dengan cara memaksimumkan fungsi kemungkinanya.

Misalkan X adalah peubah acak kontinu / diskrit dengan fungsi kepadatan peluang f(x;θ), dengan θ adalah suatu parameter yang tidak diketahui, dan X , X ,…,X sampel acak berukuran n,maka fungsi kemungkinan maksimum(likelihood function) θ adalah:

L( ) = f(x ;). f(x ;). f(x ;)……..f(x ;)

L( ) =

Untuk setiap tiitk sampel x, misalkan (x) adalah harga perameter dimana L( ) sebagai fungsi dengan menganggap x konstan mencapai maksimumnya. Estimator maksimum likelihood (MLE) dari parameter berdasarkan sampel X adalah (X).

Perhatikan bahwa dari konstruksinya, jelajah dari MLE berimpit dengan jelajah dari parameter secara intuitif, MLE adalah estimator yang masuk akal, karena MLE adalah titik parameter sampel terobservasi yang paling mungkin terjadi. Meskipun demikian, perhatikan juga sensitivitas perubahan data.

Langkah-langkah untuk menentukan estimasi titik dengan menggunakan metode kemungkinan maksimum(likelihood):

1.Tentukan L( ) (i)

2.Kedua ruas beri ln (ln L( )) (ii)

3.Turunkan (ii) atau ln L( ) terhadap (iii)

4. Selesaikanlah (iii)

Contoh:

Misalkan X X , … , X adalah sampel acak berukuran n dari distribusi B (1;), dengan tidak diketahui. Tentukan penaksir titik untuk dengan menggunakan metode kemungkinan Maksimum.

Penyelesaian:

Fungsi kepadatan peluangdari X adalah:

F(x;) = ; x = 0, 1

= 0 ; lainnya

Fungsi kemungkinan dari sample acak berukuran n adalah:

L( ) =

L( )=[

L( ) =

Kemudian kedua ruas beri ln, sehingga diperoleh:

Catatan: Tujuan di beri tambahan ln ialah untuk menghilangkan bentuk pangkat.

Selanjutnya kita turunkan ln L( ) terhadap , yaitu:

. – n . = 0

Jadi, penaksir kemungkinan maksimum untuk adalah , yang merupakan rerata sempel.

1. Estimator bayes

Pendekatan Bayesian dalam statistic secra fundamental berbeda dengan pendekatan klasik momen dan likelihood. Meskipun demikian, beberapa aspek dari pendekatan tersebut dapat berguna pada beberapa pendekatan statistic yang lain.

Dalam pendekatan klasik, parameter adalah besaran tetep yang tidak diketaui. Sampel random X , X , X , …, X diambil dari populasi berideks dan berdasarkan harga-harga terobservasi dalam sampel didapat pengetahuan tentang . Dalam pendekatan Bayesian, dipandang sebagai besaran yang variasinya digambarkan dengan distribusi probabilitas(disebut distribusi posterior). Ini adalah distribusi subjektif. Berdasarkan pada keyakinan seseorang dan dirumuskan sebelum data diambil. Kemudian, smapel diambil dari populasi berindeks dan distribusi prior disesuaikan dengan informasi sampel ini. Prior yang telah disesuaikan disebut distribusi posterior. Penyesuaian ini dilakukan dengan menggunakan aturan bayes. Itulah kenapa dinamai statistic Bayesian.

Misalkan X , X , X , …, X merupakan sebuah sampel acak berukuran n dari distribusi yang mempunyai fungsi kepadatan peluang berbentuk f(x;), dalam hal ini, kita akan menentukan taksiran bayes untuk parameter

Langkah-langkah untuk menentukan teksiran bayes bagi adalah:

1. Tentukan fungsi kepadatan peluang gabungan dari X , X , X , …, X (dinotasikan) dengan g(X , X , X , …, X) yang didefinisikan sebagai berikut:

   g(X , X , X , …, X)=f(x ;) . f(x ;) . f(x ;) …..f(x ;)

1. Tentukan fungsi densitas (theta besar), yang besarnya di ambil atau dipilih dan disesuaikan dengan g(X , X , X , …, X). Distribusi yang mempunyai densitas dari , dan dinotasikan dengan , dinamakan distribusi prior.

   1.Penaksir bayes untuk ditentukan oleh:

   1) Jika dari peubah acak berbentuk diskrit, maka:

   2) Jika dari peubah acak berbentuk kontinu, maka:

   δ(, , , …,)=

   Kemudian hasil akhir dari δ(X , X , X , …, X) dilakukan perubahan dari x menjadi X sedangkan untuk menentukan distribusi posteriornya digunakan rumus sebagai berikut:

   Prior tak sejati. Salah satu sifat menarik dari pendekatan Bayes pada statistic adalah kesederhanaannya. Begitu distribusi prior sudah ditentukan, prhitungan aturan bayes dapat langsung dikerjakan. Kesederhanaan ini membawa pada usaha untuk menggunakan pendekatan Bayesian meskipun informasi prior tebatas bahkan tidak ada sama sekali. Dalam situasi demikian, yang diperlukan adalah prior takinformatif, yaitu prior yang tidak memuat informasi tentang .

Contoh :

Misalkan x x ,…,x merupakan sebuah sampel acak dari disrtibusi B(1 ; θ), θ Ω = (0,1). Tentukan penaksir Bayes untuk θ

Penyelesaian:

Fungsi kepadatan peluang dari X adalah:

$f(x;\theta) = \theta\ (1\text{-}\theta)$; x = 0,1

= 0 = lainya.

Fungsi densitas gabungan dari X ,X2,…,X adalah:

= f(x ; θ) . F(X θ). … f(X θ)

B. Estimasi Interval

Dari penelitian dan perhitungan-perhitungan harga statistik suatu sampel, bisa dihitung suatu interval dimana dengan peluang tertentu harga parameter yang hendak ditaksir terletak dalam interval tersebut. Estimasi interval merupakan sekumpulan nilai statistik sampel dam interval tertentu yang digunakan untuk mengadakan estimasi terhadap parameter populasi dengan harapan bahwa nilai parameter populasi terletak dalam interval tersebut. Estimasi Rata – rata : dalam statistik di asumsikan suatu ukuran sampel dikatakan besar apabila $n \geq 30$, sampel dikatakan kecil apabila $n \leq 30$. Estimasi rata-rata untuk sampel kecil $n < 30$, maka interval konfidensi untuk m adalah :

1. – t (n-1 ; a/2). S ≤ μ ≤ + t (n-1 ; α/2) .

S √n √n

Contoh: Winda, Budi, Roni melakukan pengamatan mengenai lama usia pakai baterey merk Alkalin yang digunakan pada alfalinknya masing-masing, menurut mereka dari 4 baterey merk Alfalink tersebut rata-rata bisa dipakai selama 1200 jam dengan simpangan baku 200 jam, dengan interval konfidensi 98% temukan berapa rata-rata usia pakai sebenarnya dari baterey merk Alkalin tersebut ?

Estimasinya:

Xbar = 1200 jam.

N = 4

S = 1200 jam

1-α = 98%

α = 0,02 = 0,01

tn-1;) = t(3;0,01) = 4,451 (dari tabel t).

Xbar – t (n-1;a/2). S ≤ μ ≤ Xbar + t (n-1;α/2) .
S √n √n

1200-[4,451. 200] ≤ μ ≤ 1200 + [4,451 . 200]
√ 4 √ 4

754,9 jam ≤ μ ≤ 1546,1 jam

ternyata setelah di uji dengan interval konfidensi 98%, usia pakai baterey merk Alkalin berkisar (sebenarnya) antara 754,9 jam minimum dan 1645,1 jam maksimum.

2.4 Penghitungan Estimasi

ESTIMASI RATA-RATA (μ)

Kasus Sampel Besar (n ≥ 30) dan atau σ diketahui

⊙ Untuk Infinite Population

$$P\left[\bar{X} - Z_{0.5\alpha}\ \sigma/\sqrt{n} < \mu < \bar{X} + Z_{0.5\alpha}\ \sigma/\sqrt{n}\right] = 1-\alpha$$ jika σ diketahui

$$P\left[\bar{X} - Z_{0.5\alpha}\ s/\sqrt{n} < \mu < \bar{X} + Z_{0.5\alpha}\ s/\sqrt{n}\right] = 1-\alpha$$ jika n ≥ 30

⊙ Untuk Finite Population

$$P\left[\bar{X} - Z_{0.5\alpha}\ \sigma/\sqrt{n}\sqrt{\frac{N-n}{N-1}} < \mu < \bar{X} + Z_{0.5\alpha}\ \sigma/\sqrt{n}\sqrt{\frac{N-n}{N-1}}\right] = 1-\alpha$$ jika σ diketahui

$$P\left[\bar{X} - Z_{0.5\alpha}\ s/\sqrt{n}\sqrt{\frac{N-n}{N-1}} < \mu < \bar{X} + Z_{0.5\alpha}\ s/\sqrt{n}\sqrt{\frac{N-n}{N-1}}\right] = 1-\alpha$$ jika n ≥ 30

Kasus Sampel Kecil (n < 30) dan atau σ tidak diketahui

⊙ Untuk Infinite Population

$$P\left[\bar{X} - t_{0.5\alpha;df}\ {}^{s}\!/\!_{\sqrt{n}} < \mu < \bar{X} + t_{0.5\alpha;df}\ {}^{s}\!/\!_{\sqrt{n}}\right] = 1 - \alpha$$

⊙ Untuk Finite Population

$$P\left[\bar{X} - t_{0.5\alpha;df}\ {}^{s}\!/\!_{\sqrt{n}} \sqrt{\frac{N-n}{N-1}} < \mu < \bar{X} + t_{0.5\alpha;df}\ {}^{s}\!/\!_{\sqrt{n}} \sqrt{\frac{N-n}{N-1}}\right] = 1 - \alpha$$

dengan df = n − 1

SOAL-SOAL ESTIMASI RATA-RATA

1. Telah diambil secara acak sampel yang terdiri dari 100 orang mahasiswa sebuah universitas di Jakarta. Melalui test IQ terhadap 100 mahasiswa tersebut diperoleh rata-rata IQ sebesar 112 dan varians 100. Dengan menggunakan tingkat keyakinan (*confidence level*) sebesar 95%, tentukan interval konfidens untuk nilai rata-rata IQ seluruh mahasiswa universitas tersebut.

Diketahui : n = 100 $\bar{X} = 112$ $s^2 = 100 \rightarrow s = 10$

$$1 - \alpha = 0.95 \rightarrow 0.5\alpha = 0.025 \rightarrow Z_{0.025} = 1.96$$

Ditanyakan : P(. . . < μ < . . .) = 0.95

Jawab : $P\left[\bar{X} - Z_{0.5\alpha}\ {}^{s}\!/\!_{\sqrt{n}} < \mu < \bar{X} + Z_{0.5\alpha}\ {}^{s}\!/\!_{\sqrt{n}}\right] = 1 - \alpha$

$$P\left[112 - 1.96 \frac{10}{\sqrt{100}} < \mu < 112 + 1.96 \frac{10}{\sqrt{100}}\right] = 0.95$$

$$P\left[112-1.96<\mu<112+1.96\right]=0.95$$ $\rightarrow$

$$P\left[110.04<\mu<113.96\right]=0.95$$

Kita merasa yakin sebesar 95% bahwa rata-rata IQ seluruh mahasiswa universitas tersebut antara 110.04 dan 113.96

ESTIMASI PROPORSI (p)

Kasus Sampel Besar (n ≥ 30)

⦿ Untuk Infinite Population

$$P\left[\hat{p}-Z_{0.5\alpha}\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}<p<\hat{p}+Z_{0.5\alpha}\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}\right]=1-\alpha$$

⦿ Untuk Finite Population

$$P\left[\hat{p}-Z_{0.5\alpha}\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}\sqrt{\frac{N-n}{N-1}}<p<\hat{p}+Z_{0.5\alpha}\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}\sqrt{\frac{N-n}{N-1}}\right]=1-\alpha$$

Kasus Sampel Kecil (n < 30)

⦿ Untuk Infinite Population

$$P\left[\hat{p}-t_{0.5\alpha;df}\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}<p<\hat{p}+t_{0.5\alpha;df}\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}\right]=1-\alpha$$

⦿ Untuk Finite Population

$$P\left[\hat{p}-t_{0.5\alpha;df}\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}\sqrt{\frac{N-n}{N-1}}<p<\hat{p}+t_{0.5\alpha;df}\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}\sqrt{\frac{N-n}{N-1}}\right]=1-\alpha$$

dengan df = n − 1

SOAL-SOAL ESTIMASI PROPORSI

1. Dari hasil survey yang dilakukan suatu *research agency* mengenai kebiasaan ibu rumah tangga menyaksikan tayangan iklan di TV Swasta. Ternyata diperoleh hasil bahwa 76 orang dari 180 orang ibu rumah tangga yang dipilih secara acak, biasa menyaksikan tayangan iklan paling sedikit 2 jam per minggu. Jika peneliti tersebut menggunakan taraf konfidens sebesar 90%, maka tentukan interval estimasi seluruh ibu rumah tangga yang biasa menyaksikan tayangan iklan paling sedikit 2 jam per minggu.

Diketahui : Misalkan X adalah ibu rumah tangga yang biasa menyaksikan tayangan iklan paling sedikit 2 jam per hari. n = 180 dan X = 76 sehingga $\hat{p} = 76/180 = 0.42$

$1 - \alpha = 0.90 \rightarrow 0.5\alpha = 0.05 \rightarrow Z_{0.05} = 1.645$

Ditanyakan : P(. . . < p < . . .) = 0.90

Jawab : $$P\left[\hat{p} - Z_{0.5\alpha}\sqrt{\frac{\hat{p}(1-\hat{p})}{n}} < p < \hat{p} + Z_{0.5\alpha}\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}\right] = 1 - \alpha$$

$$P\left[0.42 - 1.645\sqrt{\frac{0.42(1-0.42)}{180}} < p < 0.42 + 1.645\sqrt{\frac{0.42(1-0.42)}{180}}\right] = 0.90$$

$$P[0.42 - 0.060515732 < p < 0.42 + 0.060515732] = 0.90$$

$$P[0.359484268 < p < 0.480515732] = 0.90$$

$$P[0.359 < p < 0.481] = 0.90$$

Kita merasa yakin sebesar 90% bahwa proporsi ibu-ibu yang biasa menyaksikan tayangan iklan paling sedikit 2 jam per hari antara 35.9% dan 48.1%

ESTIMASI BEDA DUA RATA-RATA

A. Kasus $\sigma_1 = \sigma_1 = \sigma$ diketahui :

⊙ **Untuk Infinite Population**

$$P\left[\left(\bar{X}_1 - \bar{X}_2\right) - Z_{0.5\alpha}\sigma\sqrt{\frac{1}{n_1} + \frac{1}{n_2}} < \mu_1 - \mu_2 < \left(\bar{X}_1 - \bar{X}_2\right) + Z_{0.5\alpha}\sigma\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}\right] = 1 - \alpha$$

⊙ **Untuk Finite Population**

$$P\left[\left(\bar{X}_1 - \bar{X}_2\right) - d < \mu_1 - \mu_2 < \left(\bar{X}_1 - \bar{X}_2\right) + d\right] = 1 - \alpha$$

dimana $d = Z_{0.5\alpha}\sigma\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}\sqrt{\frac{(N_1 + N_2) - (n_1 + n_2)}{N_1 + N_2 - 1}}$

B. Kasus $\sigma_1 \neq \sigma_1$ diketahui :

⊙ **Untuk Infinite Population**

$$P\left[\left(\bar{X}_1 - \bar{X}_2\right) - Z_{0.5\alpha}\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}} < \mu_1 - \mu_2 < \left(\bar{X}_1 - \bar{X}_2\right) + Z_{0.5\alpha}\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}\right] = 1 - \alpha$$

⊙ **Untuk Finite Population**

$$P\left[\left(\bar{X}_1 - \bar{X}_2\right) - d < \mu_1 - \mu_2 < \left(\bar{X}_1 - \bar{X}_2\right) + d\right] = 1 - \alpha$$

dimana $d = Z_{0.5\alpha}\sqrt{\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}}\sqrt{\frac{(N_1+N_2)-(n_1+n_2)}{N_1+N_2-1}}$

C. Kasus $\sigma_1 = \sigma_1$ tidak diketahui **:**

⊙ **Untuk Infinite Population**

$$P\left[\left(\bar{X}_1-\bar{X}_2\right)-t_{0.5\alpha;df}\ s_p\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}<\mu_1-\mu_2<\left(\bar{X}_1-\bar{X}_2\right)+t_{0.5\alpha;df}\ s_p\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}\right]=1-\alpha$$

⊙ **Untuk Finite Population**

$$P\left[\left(\bar{X}_1-\bar{X}_2\right)-d<\mu_1-\mu_2<\left(\bar{X}_1-\bar{X}_2\right)+d\right]=1-\alpha$$

dimana $d = t_{0.5\alpha;df}\ s_p\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}\sqrt{\frac{(N_1+N_2)-(n_1+n_2)}{N_1+N_2-1}}$

$s_p = \sqrt{\frac{(n_1-1)s_1^2+(n_2-1)s_2^2}{n_1+n_2-2}}$ $\quad$ $df = n_1 + n_2 - 2$

D. Kasus $\sigma_1 \neq \sigma_1$ tidak diketahui :

⊙ **Untuk Infinite Population**

$$P\left[\left(\bar{X}_1-\bar{X}_2\right)-t'\sqrt{\frac{s_1^2}{n_1}+\frac{s_2^2}{n_2}}<\mu_1-\mu_2<\left(\bar{X}_1-\bar{X}_2\right)+t'\sqrt{\frac{s_1^2}{n_1}+\frac{s_2^2}{n_2}}\right]=1-\alpha$$

⊙ **Untuk Finite Population**

$$P\left[\left(\bar{X}_1 - \bar{X}_2\right) - d < \mu_1 - \mu_2 < \left(\bar{X}_1 - \bar{X}_2\right) + d\right] = 1 - \alpha$$

dimana $d = t'\sqrt{\dfrac{s_1^2}{n_1} + \dfrac{s_2^2}{n_2}}\sqrt{\dfrac{(N_1 + N_2) - (n_1 + n_2)}{N_1 + N_2 - 1}}$ $\quad t' = \dfrac{t_1 w_1 + t_2 w_2}{w_1 + w_2}$ $\quad w_1 = \dfrac{s_1^2}{n_1}$

$w_2 = \dfrac{s_2^2}{n_2}$

$t_1 = t_{0.5\alpha;df=n_1-1} \quad t_2 = t_{0.5\alpha;df=n_2-1}$

SOAL-SOAL ESTIMASI BEDA DUA RATA-RATA

1. Sampel acak yang terdiri dari 22 orang buruh perusahaan A telah diperiksa ternyata rata-rata waktu menyelesaikan pekerjaannya per unit barang adalah 12 menit dengan standar deviasi 2 menit. Sedangkan dari perusahaan B yang sejenis diambil sampel acak berukuran 20, setelah diperiksa ternyata rata-rata menyelesaikan pekerjaan yang sama adalah 11 menit dengan standar deviasi 3 menit. Tentukanlah interval keyakinan sebesar 95% untuk mengestimasi beda rata-rata waktu penyelesaian pekerjaan semua buruh di perusahaan A dan perusahaan B. Asumsi $\sigma_1 = \sigma_2$

Diketahui : $n_1 = 22$ $\quad \bar{X}_1 = 12$ $\quad s = 2$ $\quad n_1 = 20$ $\quad \bar{X}_1 = 11$ $\quad s = 3$

Karena $\sigma_1 = \sigma_2$ tidak diketahui, maka digunakan rumus interval konfidens untuk kasus C. Sehingga $1 - \alpha = 0.95 \rightarrow \alpha = 0.05 \rightarrow$ dengan $0.5\alpha = 0.025$ dan df = 40 dari tabel t diperoleh $t_{0.025;df=40} = 2.021$

Ditanyakan : $P(\ldots < \mu_1 - \mu_2 < \ldots) = 0.95$

Jawab :

$$P\left[\left(\bar{x}_1-\bar{x}_2\right)-t_{0.5\alpha;df}\,s_p\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}<\mu_1-\mu_2<\left(\bar{x}_1-\bar{x}_2\right)+t_{0.5\alpha;df}\,s_p\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}\right]=1-\alpha$$

$$s_p=\sqrt{\frac{(n_1-1)s_1^2+(n_2-1)s_2^2}{n_1+n_2-2}} \rightarrow s_p=\sqrt{\frac{(22-1)2^2+(20-1)3^2}{22+20-2}} \rightarrow s_p = 2.524876235$$

$$P\left[(12-11)-2.021(2.524876235)\sqrt{\frac{1}{22}+\frac{1}{20}}<\mu_1-\mu_2<(12-11)+2.021(2.524876235)\sqrt{\frac{1}{22}+\frac{1}{20}}\right]=$$

$$P\left[1-1.576538987<\mu_1-\mu_2<1+1.576538987\right]=0.95$$

$$P\left[-0.576538987<\mu_1-\mu_2<2.576538987\right]=0.95 \rightarrow$$

$$P\left[-0.58<\mu_1-\mu_2<2.58\right]=0.95$$

Kita merasa yakin sebesar 95% bahwa beda rata-rata waktu penyelesaian pekerjaan semua buruh di perusahaan A dan perusahaan B antara – 0.58 dan 2.58 menit.

ESTIMASI BEDA DUA PROPORSI

Kasus Sampel Besar

⦿ Untuk Infinite Population

$$P\left[(\hat{p}_1-\hat{p}_2)-d<p_1-p_2<(\hat{p}_1-\hat{p}_2)+d\right]=1-\alpha$$

$$d = Z_{0.5\alpha}\sqrt{\frac{\hat{p}_1(1-\hat{p}_1)}{n_1}+\frac{\hat{p}_2(1-\hat{p}_2)}{n_2}}$$

⊙ Untuk Finite Population

$$P\left[(\hat{p}_1-\hat{p}_2)-d < p_1-p_2 < (\hat{p}_1-\hat{p}_2)+\right] = 1-\alpha$$

$$d = Z_{0.5\alpha}\sqrt{\frac{\hat{p}_1(1-\hat{p}_1)}{n_1}+\frac{\hat{p}_2(1-\hat{p}_2)}{n_2}}\sqrt{\frac{(N_1+N_2)-(n_1+n_2)}{N_1+N_2-1}}$$

Kasus Sampel Kecil

⊙ Untuk Infinite Population

$$P\left[(\hat{p}_1-\hat{p}_2)-d < p_1-p_2 < (\hat{p}_1-\hat{p}_2)+d\right] = 1-\alpha$$

$$d = t_{0.5\alpha;df=n_1+n_2-2}\sqrt{\frac{\hat{p}_1(1-\hat{p}_1)}{n_1}+\frac{\hat{p}_2(1-\hat{p}_2)}{n_2}}$$

⊙ Untuk Finite Population

$$P\left[(\hat{p}_1-\hat{p}_2)-d < p_1-p_2 < (\hat{p}_1-\hat{p}_2)+\right] = 1-\alpha$$

$$d = t_{0.5\alpha;df=n_1+n_2-2}\sqrt{\frac{\hat{p}_1(1-\hat{p}_1)}{n_1}+\frac{\hat{p}_2(1-\hat{p}_2)}{n_2}}\sqrt{\frac{(N_1+N_2)-(n_1+n_2)}{N_1+N_2-1}}$$

SOAL-SOAL ESTIMASI DUA PROPORSI

1. Dua sampel acak masing-masing terdiri 700 mahasiswa dan 500 mahasiswi yang mengunjungi suatu bazar buku murah. Ternyata setelah kedua sampel tersebut diperiksa, terdapat 392 mahasiswa dan 325 mahasiswi yang merasa puas dengan adanya bazar tersebut. Tentukan interval konfidens sebesar 98% untuk mengestimasi perbedaan proporsi mahasiswa dan mahasiswi yang merasa puas terhadap bazar buku murah tersebut.

Diketahui : $n_1 = 700$ $x_1 = 392$ $\rightarrow \hat{p}_1 = \frac{392}{700} = 0.56$ $n_2 = 500$ $x_2 = 325 \rightarrow \hat{p}_2 = \frac{325}{500} = 0.65$

Karena sampelnya besar, maka $1 - \alpha = 0.98 \rightarrow 0.5\alpha = 0.01 \rightarrow Z_{0.01} = 2.32$

Ditanyakan : $P(\ldots < p_1 - p_2 < \ldots) = 0.98$

Jawab :

$$d = Z_{0.5\alpha}\sqrt{\frac{\hat{p}_1(1-\hat{p}_1)}{n_1} + \frac{\hat{p}_2(1-\hat{p}_2)}{n_2}} \rightarrow$$

$$d = 2.32\sqrt{\frac{0.56(1-0.56)}{700} + \frac{0.65(1-0.65)}{500}} = 0.065905969$$

$$P[(\hat{p}_1 - \hat{p}_2) - d < p_1 - p_2 < (\hat{p}_1 - \hat{p}_2) + d] = 1 - \alpha$$

$$P[(0.56 - 0.65) - 0.066 < p_1 - p_2 < (0.56 - 0.65) + 0.66] = 0.98$$

$$P[-0.09 - 0.066 < p_1 - p_2 < -0.09 + 0.066] = 0.98 \rightarrow$$

$$P[-0.156 < p_1 - p_2 < -0.024] = 0.98$$

Kita merasa yakin sebesar 98% proporsi mahasiswi yang merasa puas terhadap bazar buku lebih besar daripada mahasiswa antara 2.4% dan 15.6%.

MENENTUKAN UKURAN SAMPEL DALAM ESTIMASI PARAMETER SECARA STATISTIS

- Dalam estimasi parameter secara statistis, ukuran sampel (n) minimal dapat dihitung jika diketahui :

1. Besarnya estimation error (d) maksimal yang dikehendaki
2. Tingkat keyakinan (level of confidence) yang dikehendaki
3. Parameter yang akan diestimasi
4. Standar error (standar deviasi statistik) berdasarkan hasil survey terdahulu atau asumsi

- Jika yang akan diestimasi adalah rata-rata (μ) :

$$Z_{0.5\alpha}\frac{\sigma}{\sqrt{n}} \le d \rightarrow \frac{Z_{0.5\alpha}\sigma}{d} \le \sqrt{n} \rightarrow \sqrt{n} \ge \frac{Z_{0.5\alpha}\sigma}{d} \rightarrow n \ge \left(\frac{Z_{0.5\alpha}\sigma}{d}\right)^2$$

- Jika yang akan diestimasi adalah proporsi (p) :

$$Z_{0.5\alpha}\sqrt{\frac{p(1-p)}{n}} \le d \rightarrow Z_{0.5\alpha}\frac{\sqrt{p(1-p)}}{\sqrt{n}} \le d \rightarrow \frac{Z_{0.5\alpha}\sqrt{p(1-p}}{d} \le \sqrt{n} \rightarrow$$

$$\sqrt{n} \ge \left(\frac{Z_{0.5\alpha}\sqrt{p(1-p)}}{d}\right) \rightarrow$$

$$n \ge \left(\frac{Z_{0.5\alpha}\sqrt{p(1-p)}}{d}\right)^2$$