BAB III

INTERVAL ESTIMASI DAN UJI HIPOTESIS

1.1. GAMBARAN INTERVAL KONFIDENSI SECARA UMUM

Karena dalam prakteknya titik estimasi yang terdiri dari satu angka saja tidak memberikan gambaran mengenai berapa jarak/selisih dari nilai estimasi tersebutterhadap nilai suatu parameter yang sebenarnya, terkecu ali jika diberikan besarnya kesalahan dalam estimasi yang mungkin terjadi.

Sebagai contoh kalau sampel suatu perusahaan diselidiki dan memberikan nilai estimasi untuk modal ($=\bar{X}$) sebesar Rp. 100 juta, berapa sebetulnya nilai rata - rata modal sebenarnya (=M). Dan kita tahu bahwa $|\bar{X} - M| = E$ dimana E = kesalahan yang terjadi dalam suatu estimasi. Maka itulah sebabnya sering dipergunakan interval estimasi berapa interval yang dibatasi dua nilai, yang tersebut nilai batas bawah dan nilai batas atas.

Sekarang pandanglah suatu sampel random yang terdiri dari 4 observasi yaitu 1,2,3,4;6;5.6 yang diambil dari suatu populasi normal dengan mean μ (tak diketahui nilainya) dan standard deviasi 3.
Estimasi likelihood maximum dari μ adalah rata-rata dari sampel yang diobservasi yaitu $\bar{x} = 2,7$. Jadi yang diinginkan adalah suatuinterval yang memuat nilai parameteryang sebenarnya. Untuk sampel berukuran 4 tersebut maka $Z = \frac{\bar{x} - \mu}{3/2}$ akan didistribusikan secara normal dengan mean o dari varian 1, dimana $\bar{X}$ adalah mean sampel dan 3/2 adalah σ/n jadi Z mempunyai density

$$f_Z(z) = \Phi(z) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2} z^2}$$

yang tidak tergantung dari nilai parameter yang sebenarnya, maka probabilitas bahwa Z akan terletak didalam setiap sebarang 2 bilangan yang dipilih dapat dihitung sebagaicontoh :

$$P\left[-1,96 < Z < 1,96\right] = \int_{-1,96}^{1,96} \Phi(z)\, dz = 0,95 \quad \ldots \quad (1)$$

dimana. $Z = \dfrac{\bar{x} - \mu}{3/2}$

Jadi $\dfrac{\bar{x} - \mu}{3/2} > -1.96 \Rightarrow \mu = \bar{x} + 3/2\,(1.96) = \bar{x} + 2.94$

Persamaan (1) menjadi

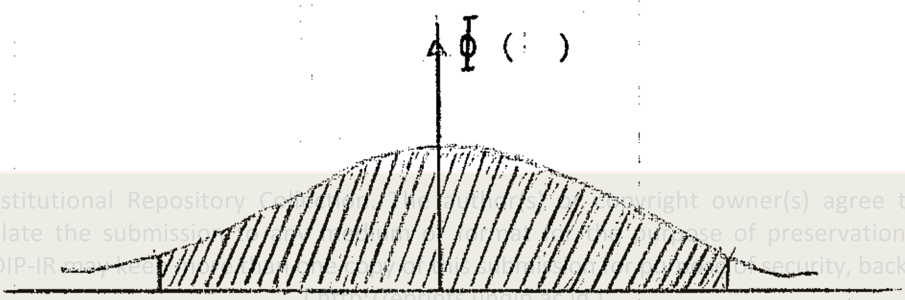
$$P\left[\bar{x} - 2.94 < \mu < \bar{x} + 2.94\right] = 0.95 \quad \ldots\ldots\ldots \quad (2)$$

Jika $\bar{x} = 2.7$ maka interval yang diperoleh adalah (- 0.24, 5.64).

Interval random ($\bar{x}$- 2.94 , $\bar{x}$ + 2.94) dan (-0.24 , 5.64) kedua-duanya dinamakan interval konfidensi atau lebih tepatnya interval konfidensi 95 %, dimana 0.95 dinamakankoefisien konfidensi.

Sebarang 2 bilangan yaitu a dan b dimana 95 % dari luasan dibawah Φ (z) terletak diantara a dan b akan menentukan suatu interval konfidensi 95 %.

Dan jarak b - a akan minimum untuk suatu luasan tertentu jika Φ (a) = Φ (b). (lihat gambar 1).



GAMBAR 1.

Karena ϕ (z) simetris terhadap z = 0, maka b - a adalah minimum untuk suatu luasan tertentu jika b = - a. Jadi untuk $\bar{x}$ = 2.7 maka interval (- 0.24, 5.64) memberikan - suatu interval konfidensi 95 % yang paling pendek.

Jadi contoh tersebut diatas adalah suatu cara untuk mendapatkan interval konfidensi yang umum selanjutnya akan didefinisikan suatu definisi mengenai interval konfidensi.

DEFINISI 1.

Jika $X_1, \ldots, X_n$ adalah n sampel random darisuatu density f(. ; θ) dan jika $T_1 = t_1(X_1, \ldots, X_n)$dan $T_2 = t_2(x_1, \ldots, x_n)$ adalah 2 statistik yang memenuhi $T_1 \leq T_2$ dengan $P_o\left[T_1 < \tau(\theta) < T_2\right] = \gamma$,dimana γ tak tergantung pada θ, maka interval random (T_1, T_2) dinamakan interval konfidensi 100 γ persen untuk $\tau(\theta)$, dengan γ adalah koefisienkonfidensi dan T_1 dan T_2 masing-masing dinamakan limit konfidensi bawah dan limit konfidensi atas untuk $\tau(\theta)$. Suatu nilai (t_1, t_2) dariinterval random (T_1, T_2) dinamakan juga interval konfidensi- 100 γ persen untuk $\tau(\theta)$.

Jika salah satu tetapi tidak semuanya dari dua statistik $t_1(X_1, \ldots, X_n)$ dan $t_2(X_1, \ldots, X_n)$ merupakan suatu konstanta maka akan didefinisikan suatu definisi dibawah ini.

DEFINISI 2 INTERVAL KONFIDENSI SATU RUAS.

Jika $X_1, \ldots, X_n$ adalah n sampel random dari suatu density f(. ;θ) dan jika $T_1 = t_1(X_1, \ldots, X_n)$adalah

adalah suatu statistik dengan $P_o \left[T_1 < \tau(\theta) \right] \doteq Y$ maka T_1 dinamakan limit konfidensi bawah satu ruas untuk $\tau(\theta)$. Dengan cara sama, jika $T_2 = t_2(X_1, \ldots, X_n)$ adalah suatu statistik dengan $P_\theta \left[\tau(\theta) < T_2 \right] \doteq Y$ maka T_2 dinamakan limit konfidensi atas satu ruas untuk $\tau(\theta)$, dimana Y tak tergantung pada θ.

C o n t o h :

Misalkan $X_1, \ldots, X_n$ adalah n sampel random dari suatu density $f(x;\theta) = \phi_{\theta,9}(x)$. Statistik

Statistik $T_1 = t_1(X_1, \ldots, X_n) = \bar{X} - 6/\sqrt{n}$ dan $T_2 = t_2(X_1, \ldots, X_n) = \bar{X} + 6/\sqrt{n}$, maka (T_1, T_2) menyusun suatu interval random dan merupakan suatu interval konfidensiuntuk $\tau(\theta) = \theta$, dengan koefisien konfidensi.

$$Y = P_o \left[\bar{X} - 6/\sqrt{n} < \theta < \bar{X} + 6/\sqrt{n} \right] = P_\theta \left[-2 < (\bar{X} - \theta)/(3/\sqrt{n}) < 2 \right]$$

$$= \Phi(2) - \Phi(-2)$$

$$= 0{,}9772 - 0{,}0228$$

$$= 0{,}9544$$

Juga, jika suatu sampel random dari 25 observasi mempunyai rata-rata yaitu 17,5, maka interval $(17{,}5 - 6/\sqrt{25}, 17{,}5 + 6/\sqrt{25})$ dinamakan interval konfidensi 95.44 persen untuk θ.

A K I B A T.

Jika suatu interval komfidensi untuk θ telah ditentukan maka suatu keluarga dari seluruh interval konfidensi dapat diperoleh khususnya, jika suatu estimator interval konfidensi 100 Y persen untuk θ telah diberikan.

Maka suatu estimator interval konfidensi 100 Y persen untuk $\tau(\theta)$ dapat diperoleh, dimana $\tau(.)$ adalah sebarang fungsi monoton yang sesuai misal jika $\tau(\ .)$ adalah suatu fungsi naik monoton dan $(T_1 = t_1(x_1, .., x_n), T_2 = t_2(x_1, .., x_n))$ adalah suatu interval konfidensi 100 Y persen untuk θ, maka $(\tau(T_1), \tau(T_2))$ adalah interval konfidensi 100 Y persen untuk $\tau(\theta)$.

Jadi $P_\theta\left[\tau(T_1) < \tau(\theta) < \tau(T_2)\right] = P_\theta\left[T_1 < \theta < T_2\right]$

$= Y$

1.2. METODE JUMLAH PIVOTAL.

Jika $X_1, .., X_n$ adalah n sampel random dari suatu density $f(.\ ;\theta)$ dimana θ adalah parameter, maka yang diinginkan adalah mengestimasi $\tau(\theta)$ dengan menggunakan i interval konfidensi.

DEFINISI 3 JUMLAH PIVOTAL.

Misalkan $X_1, .., X_n$ adalah n sampel random dari suatu density $f(.;\theta)$ jika $Q = q(X_1, .., X_n; \theta)$, dimana merupakan fungsi dari $X_1, .., X_n$ dan θ. Dan jika Q mempunyai distribusi yang tidak tergantung pada θ, maka Q adalah sebagai suatu jumlah pivotal.

C o n t o h:

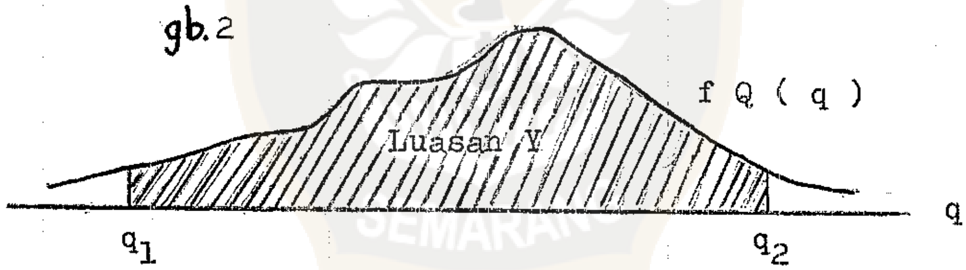
Misalkan $X_1, .., X_n$ adalah suatu sampel random dari density $f(x\ ;\theta) = \phi_{\theta,9}(x)$, maka $(\bar{X} - \theta)$ adalah suatu suatu jumlah pivotal, sebab.

$E(\bar{X} - \theta) = E(\theta - \theta) = 0$

$$\begin{aligned} \mathrm{Var}\ (\bar{X} - \theta) &= E\left[(\bar{X} - \theta)^2 \right] - \left[E\ (\bar{X} - \theta) \right]^2 \\ &= E\ (\bar{X} - \theta)^2 \\ &= \mathrm{Var}\ (\bar{X}) \\ &= 9/n \end{aligned}$$

Jika $Q = q\ (X_1, \ldots, X_n)$ adalah suatu jumlah pivotal dan mempunyai suatu fungsi density probabilitas maka untuk setiap $0 < Y < 1$ yang tertentu akan terdapatlah q_1 dan q_2 yang tergantung pada Y, sedemikian hingga $P\left[q_1 < Q < q_2\right] = Y$ (lihat gambar 2).

gb.2



Sekarang, jika untuk setiap nilai sampel ($x_1, \ldots, x_n$), yang mungkin, $q_1 < q_1\ (x_1, \ldots, x_n, o) < q_2$ jika dan hanya jika $t_1(x_1, \ldots, x_n) < \tau(\theta) < t_2\ (x_1, \ldots, x_n)$ dimana $t_1\ (x_1, \ldots, x_n)$ dan $t_2(x_1, \ldots, x_n)$ adalah fungsi yang tak tergantung pada θ, maka (T_1, T_2) adalah interval konfidensi 100 Y persen untuk $\tau(\theta)$, dimana $T_i = t_i\ (X_1, \ldots, X_n)$, untuk $i = 1,2$ Dan untuk setiap Y tertentu maka terdapatlah demikian banyak pasangan dari bilangan - bilangan q_1 dan q_2 yang dapat dipilih sedemikian hingga - $P\ (q_1 < Q < q_2) = Y$ untuk pasangan q_1 dan q_2 yang berbeda akan menghasilkan t_1 dan t_2 yang berbeda pula.

Dan jika $t_2 (X_1, .. ,X_n) - t_1 (X_1, .. ,X_n)$ adalah panjang interval konfidensi maka q_1 dan q_2 dapat dipilih sehingga panjang interval konfidensi tersebut paling kecil

C o n t o h :

Misal $X_1, .. ,X_n$ adalah n sampel random dari N (e,9) dengan e>o berikan suatu contoh jumlah pivotal dan gunakan jumlah pivotal tersebut untuk mendapatkan suatu estimator interval konfidensi untuk e.

Penyelesaian :

Jika $X_1, .. ,X_n$ adalah n sampel random dari N (θ,9) maka $X_1, .. ,X_n$ didistribusikan secara normal dengan mean θ dan varian 9.

Pandang $(\bar{X} - e)/(\sqrt{9}/\sqrt{n})$. Ini merupakan fungsi dari $X_1, ..., X_n$ dan e.

$$E\left[(\bar{X} - e)/(\sqrt{9}/\sqrt{n})\right] = \frac{\sqrt{n}}{3} E\left[\bar{X} - e\right] = \frac{\sqrt{n}}{3} . o = o$$

$$\begin{aligned}Var\left[(\bar{X}-o)/(\sqrt{9}/\sqrt{n})\right] &= E\left[(\bar{X}-e)/(\sqrt{9}/\sqrt{n})\right]^2 - \left[E(\bar{X}-e)/(\sqrt{9}/\sqrt{n})\right]^2 \\ &= E\left[(\bar{X} - e)/(\sqrt{9}/\sqrt{n})\right]^2 \\ &= \frac{n}{9} E (\bar{X}-e)^2 \\ &= \frac{n}{9} Var (\bar{X}) \\ &= \frac{n}{9} \cdot \frac{9}{n} = 1\end{aligned}$$

Jadi $(\bar{X} - e)/(\sqrt{9}/\sqrt{n})$ mempunyai distribusi yang tidak tergantung pada e.

Jadi $(\bar{X} - e)/(\sqrt{9}/\sqrt{n})$ merupakan jumlah pivotal untuk Y yang diberikan terdapatlah q_1 dan q_2 sedemikian hingga.

$$P\left[q_1 < (\bar{X} - e)/(\sqrt{9}/\sqrt{n}) < q_2\right] = Y$$

Jadi interval konfidensi yang dicari adalah

$$(\bar{X} - 3 q_1/\sqrt{n}, \bar{X} - 3 q_2/\sqrt{n})$$

2. SAMPLING DARI DISTRIBUSI NORMAL.

Jika $X_1, \ldots, X_n$ adalah n sampel random dari distribusi normal dengan mean μ dan varian σ^2, maka yang akan dibicarakan adalah berikut ini :

1. Interval konfidensi untuk μ
2. Interval konfidensi untuk σ^2
3. Interval konfidensi simultan untuk μ dan σ^2

2.1. INTERVAL KONFIDENSI UNTUK MEAN.

Interval konfidensi untuk mean tersebut tergantung atas 2 hal yaitu σ^2, diketahui atau tidak.
Pandang untuk σ^2 tak diketahui.
Jadi yang diinginkan adalah interval konfidensi untuk μ dimana σ^2 tak diketahui, dengan parameternya yaitu θ = (μ, σ), dimana $\tau(\theta) = \mu$. Andaikan bahwa $(\bar{X} - \mu)/(\sigma/\sqrt{n})$ adalah jumlah pivotal yang mempunyai distribusi normal, tetapi $q_1 < (\bar{X} - \mu)/(\sigma/\sqrt{n}) < q_2$ tidak dapat dikerjakan untuk memberikan $t_1(X_1, \ldots, X_n) < \mu < t_2(X_1, \ldots, X_n)$ untuk setiap statistik t_1 dan t_2, sebab $(\bar{X} - \mu)/(\sigma/\sqrt{n})$ mengandung σ, maka dicari suatu jumlah pivotal yang hanya mengikut sertakan μ saja.

Pandanglah :

$$\frac{(\bar{X} - \mu)/(\sigma/\sqrt{n})}{\sqrt{\sum(X_i - \bar{X})^2/(n-1)\sigma^2}} = \frac{\bar{X} - \mu}{S/\sqrt{n}}$$

Yang mempunyai distribusi t dengan n-1 derajat kebebasan (dimana $S^2 = \sum_{i=1}^{n}(X_i - \bar{X})^2/(n-1)$).

Jadi $\frac{\bar{X} - \mu}{S/\sqrt{n}}$ mempunyai suatu density yang tidak tergantung dari μ dan σ^2.

Jadi $\left\{ q_1 < (\bar{X} - \mu)/(S/\sqrt{n}) < q_2 \right\}$ jika dan hanya jika

$\bar{X} - q_2 (S/\sqrt{n}) < \mu < \bar{X} - q_1 (S/\sqrt{n})$, dimana untuk - q_1 dan q_2 memenuhi

$$P\left[q_1 < (\bar{X} - \mu) / (S/\sqrt{n}) < q_2 \right] = Y .$$

Jadi $(\bar{X} - q_2 (S/\sqrt{n})$, $\bar{X} + q_1 (S/\sqrt{n}))$ adalah suatu interval konfidensi 100 Y persen untuk μ dengan panjang $(q_2 - q_1)(S/\sqrt{n})$.

Dan untuk sebarang sampel maka panjang interval konfidensi itu adalah minimum, jika q_1 dan q_2 dipilih sedemikian hingga $q_2 - q_1$ minimum.

Jadi untuk meminimumkan $L = S/\sqrt{n}\,(q_2 - q_1)$ yang ter gantung dari

$$\int_{q_1}^{q_2} f_T(t)\,dt = Y \text{ .... } (5)$$

dimana $f_T(t)$ adalah density dari distribusi t dengan n -i derajat kebebasan.

Persamaan (3) memberikan q_2 sebagai fungsi dari q_1 dan - deferensial persamaan (3) terhadap q_1 menghasilkan :

$$f_T(q_2)\frac{dq_2}{dq_1} - f_T(q_1) = 0 , \quad \frac{dq_2}{dq_1} = \frac{f_T(q_1)}{f_T(q_2)}$$

Karena L minimum, maka $\frac{dL}{dq_1} = 0$ sehingga

$$\frac{dL}{dq_1} = \frac{S}{\sqrt{n}}\left(\frac{dq_2}{dq_1} - 1\right) = 0$$

Tetapi $S/\sqrt{n}\,(dq_2/dq_1 - 1) = S/\sqrt{n}\left(\frac{f_T(q_1)}{f_T(q_2)} - 1\right) = 0$

Jika dan hanya jika $f_T(q_1) = f_T(q_2)$, maka $q_1 = q_2$

(Untuk $\int^{q_2} f_T(t)\,dt \neq Y$) atau $q_1 = -q_2$

Dan $q_1 = -q_2$ merupakan solusi yang diinginkan, di mana q_1 dan q_2 dapat diperoleh dari tabel distribusi.

2.2. INTERVAL KONFIDENSI UNTUK VARIAN.

Tergantung atas 2 hal yaitu μ diketahui atau tidak diketahui.

Untuk μ tak diketahui, maka pandang

$$Q = \frac{\Sigma (X_j - \bar{X})^2}{\sigma^2} = \frac{(n-1)\, S^2}{\sigma^2}$$

Yang mempunyai distribusi chi kwadrat dengan $n - 1$ derajad kebebasan : Jadi Q merupakan jumlah pivotal yang merupakan fungsi dari $X_1, \ldots, X_n$ dan σ

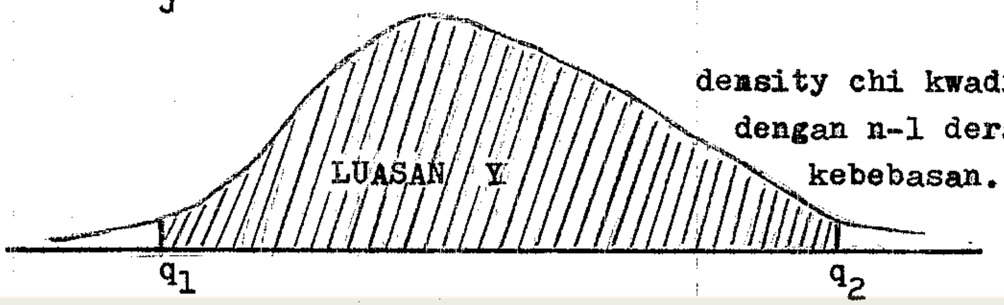
Jadi $q_1 < \frac{(n-1)\, S^2}{\sigma^2} < q_2$ jika dan hanya jika

$$\left\{ \frac{(n-1)S^2}{q_2} < \sigma^2 < \frac{(n-1)S^2}{q_1} \right\} \text{ maka } \left(\frac{(n-1)S^2}{q_2} , \frac{(n-1)S^2}{q_1} \right)$$

adalah interval konfidensi 100 γ persen untuk σ^2, dimana q_1 dan q_2 diberikan oleh $P\left[q_1 < Q < q_2 \right] = \gamma$

(lihat gambar 3)

gb.3



Seringkali q_1 dan q_2 dipilih sehingga

$$P\left[Q < q_1 \right] = P\left[Q > q_2 \right] = (1 - \gamma)/2$$

Dan interval konfidensi demikian dinamakan interval kon fidensi berekor sama untuk σ^2, dimana q_1 dan q_2 dapat - diperoleh dari tabel distribusi chi kwadrat.

Pilih q_1 dan q_2 sehingga panjang interval konfidensi

$$L = (n - 1)\, S^2 \left(\frac{1}{q_1} - \frac{1}{q_2}\right) \text{ minimum}$$

Jika $f_Q(q)$ adalah density chi kwadrat dengan $n - 1$ derajad kebebasan maka defferensial dari.

$$\int_{q_1}^{q_2} f_Q(q)\, dq = Y$$ terhadap q_1 menghasilkan.

$$\frac{dq_2}{dq_1}\, f_Q(q_2) - f_Q(q_1) = 0$$

Dan jika

$$\frac{dL}{dq_1} = (n - 1)\, S^2 \left(-\frac{1}{q_1^2} + \frac{1}{q_2^2}\, \frac{dq_2}{dq_1}\right.$$

$$= (n - 1)\, S^2 \left(-\frac{1}{q_1^2} + \frac{1}{q_2^2}\, \frac{f_Q(q_1)}{f_Q(q_2)}\right) = 0$$

maka $q_1^2 f_Q(q_1) = q_2^2 f_Q(q_2)$

Jadi L adalah minimum jika q_1 dan q_2 dipilih sehingga

$$q_1^2 f_Q(q_1) = q_2^2 f_Q(q_2)$$

dimana tergantung pada

$$\int_{q_1}^{q_2} f_Q(q) = Y$$

Dan untuk setiap q_1 dan q_2 yang memenuhi

$$\int_{q_1}^{q_2} f_Q \ (q) = Y \quad \text{maka}$$

$$\left(\sqrt{\frac{(n-1)\ S^2}{q_2}}\ ,\ \sqrt{\frac{(n-1)\ S^2}{q_1}}\right)$$ adalah interval konfidensi

100 Y persen untuk σ.

2.3. REGION KONFIDENSI SIMULTAN UNTUK MEAN DAN VARIAN.

Untuk suatu estimasi berserikat dari mean μ dan varian σ^2 dari suatu distribusi normal maka dipergunakan - fasal 2.1 dan fasal 2.2.

Sebagai contoh, untuk melukis suatu daerah konfidensi seperti dalam gambar dengan menggunakan dua relasi dibawah ini:

$$P\left[-t_{0.975}\sqrt{S^2/n} < \mu < + t_{0.975}\sqrt{S^2/n}\right] = 0.975 \qquad \text{........... (4)}$$

dan

$$P\left[\frac{(n-1)\ S^2}{X^2_{0.975}} < S^2 < \frac{(n-1)\ S^2}{X^2_{0.025}}\right] = 0.95 \qquad \text{........... (5)}$$

Dimana $t_{0.975}$ adalah titik quanti dari distribusi t dengan derajat kebebasan n - 1, serta $x^2_{0.025}$ dan $x^2_{0.975}$ masing-masing adalah titik quantil ke 0.025 dan ke 0.975dari distribusi chi kwadrat dengan derajat kebebasan n-1.

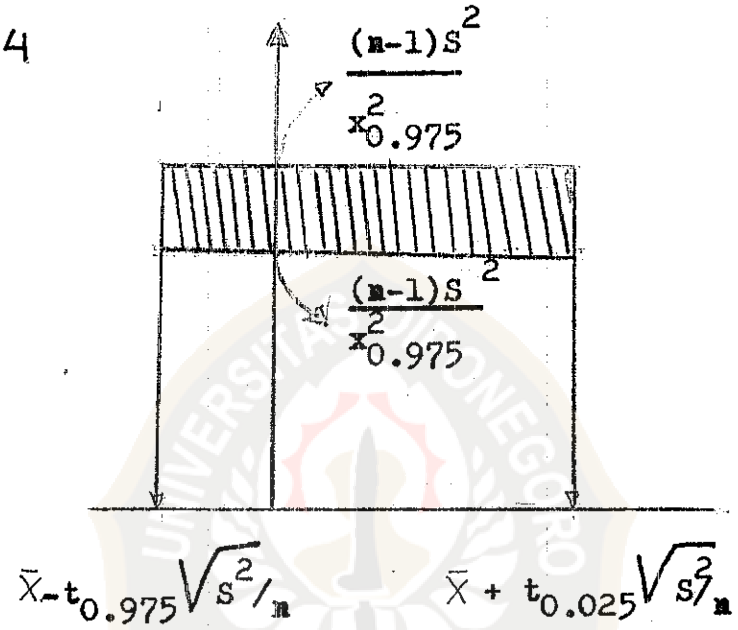
Region dalam gambar 4 adalah daerah konfidensi untuk (μ, σ^2), tetapi koefisien konfidensinya tidak tahu (sebab - koefisien konfidensinya $\neq 0.95^2$, dimana disebabkan oleh 2 peristiwa yang diberikan dalam persamaan (4) dan persamaan (5) yang saling bergantungan).

Suatu daerah konfidensi, yang koefisien konfidensinya dapat dievaluasi, dapat diadakan dengan mengguna -

Oleh karena :   (lihat gambar 4)

$$Q_1 = \frac{\bar{X} - \mu}{\sigma/\sqrt{n}} \qquad \text{dan} \qquad Q_2 = \frac{(n-1)\ S^2}{\sigma^2}$$

gb.4



Untuk setiap jumlah pivotal, maka dapat ditemukan bilangan-bilangan q_1, q_2' dan q_2'' sehingga :

$$P\left[-\ q_1 < \frac{\bar{X} - \mu}{\sigma/\sqrt{n}} < q_1\right] = Y_1 \qquad \ldots\ldots\ (6)$$

dan

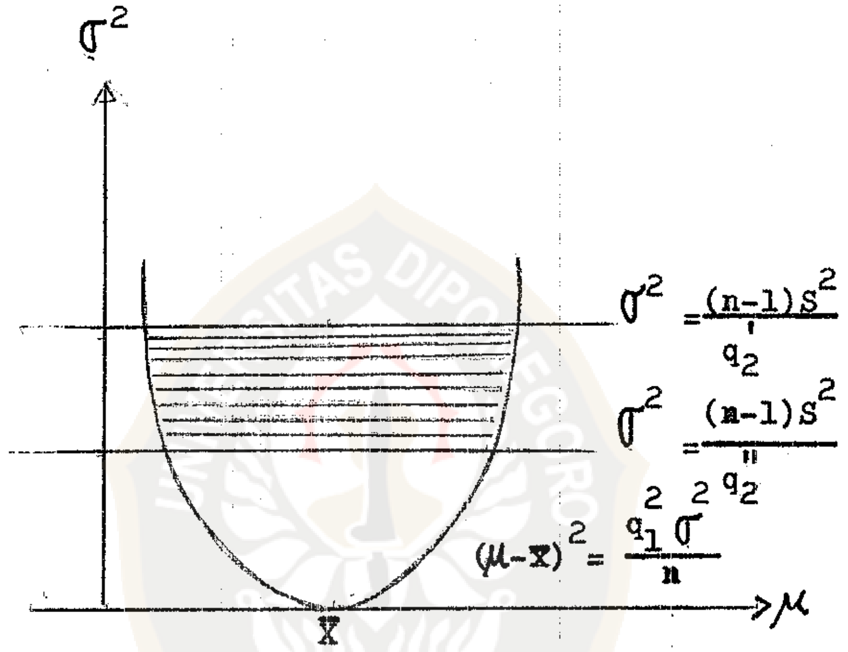
$$P\left[\ q_2' < \frac{(n-1)\ S^2}{2} < q_2''\right] = Y_2 \qquad \ldots\ldots\ (7)$$

Karena Q_1 dan Q_2 tak bergantungan maka probabilitas berserikat.

$$P\left[-q_1 < \frac{\bar{X} - \mu}{\sigma/\sqrt{n}} < q_1;\ q_2' < \frac{(n-1)S^2}{\sigma^2} < q_2''\right] = Y_1\ Y_2 \ \ldots\ldots\ (8)$$

Keempat ketidak samaan dalam persamaan 8 menentukan suatu daerah dalam ruang parameter yang didapat dengan melukis batas-batasnya.

Jika tanda ketidak samaan dalam persamaan 8 diganti dengan tanda kesamaan dan masing-masing persamaan yangdi hasilkan sebagai fungsi dari μ dan σ^2 maka region nya adalah luasan yang diarsir seperti dalam gambar dibawah.



GAMBAR 5.

Sedangkan untuk suatu region konfidensi untuk (μ, σ) dapat diperoleh dengan cara yang sama dimana persamaan diberikan sebagai fungsi dari σ bukannya σ^2, dan parabola pada gambar 5 diatas akan menjadi sepasang garis lurus yang diberikan oleh persamaan:

$$\mu = \bar{x} \pm q_1 \sigma/\sqrt{n} \text{ yang memotong sumbu } \mu \text{ di } \bar{x} .$$

3. METODE UNTUK MENDAPATKAN INTERVAL KONFIDENSI.

Dalam fasal ini dibahas dua metode untuk mendapatkan interval konfidensi yaitu metode jumlah pivotal.

3.1 METODE JUMLAH PIVOTAL.

Sebelumnya dalam fasal 1.3. telah diuraikan sua tu metode untuk mendapatkan interval konfidensi dinamakan metode jumlah pivotal.

Sekarang perhatikan berikut ini.

Jika $X_1, \ldots, X_n$ adalah n sampel random dari suatu fungsi density f (.; θ), dimanafungsi distribusi kumulatif yang bersesuaian dengan f (.;θ) adalah kontinu dalam x, maka dengan probabilitas integral transform $F(X_i;\theta)$ mempunyai distribusi uniform pada interval (0,1).

Oleh karena itu $-\log F(X_i;\theta)$ mempunyai density

$$P\left[-\log F(X_i;\theta) \geq u\right] = P\left[\log F(X_i;\theta) \leq -u\right]$$

$$= P\left[F(X_i;\theta) \leq e^{-u}\right]$$

$$= e^{-u} \text{ untuk } u > 0.$$

Akibatnya $-\sum \log F(X_i;\theta)$ mempunyai distribusi gamma dengan parameter n dan 1.

Sehingga

$$P\left[-\log q_2 < -\sum_{i=1}^{n} \log F(X_i;\theta) < -\log q_1\right] =$$

$$= \int_{-\log q_2}^{-\log q_1} \frac{1}{\Gamma(n)} z^{n-1} e^{-z} \, dz$$

$$= P\left[q_1 < \prod_{i=1}^{n} F(X_i;\theta) < q_2\right]$$

untuk $0 < q_1 < q_2 < 1$.

Jadi $\prod_{i=1}^{n} F(X_i;\theta)$ atau $-\sum_{i=1}^{n} \log F(X_i;\theta)$ ADALAH jumlah pivotal.

Jadi untuk sampel dari suatu populasi yang mempunyaidis tribusi kumulatif yang kontinu maka terdapatlah suatu jumlah pivotal.

Dan jumlah pivotal tersebut belum tentu menjamin untuk mendapatkan suatu interval konfidensi. Tetapi jika $F(x;\theta)$ adalah monoton dalam θ untuk setiap x, maka $\prod_{i=1}^{n} F(x_i;\theta)$ adalah juga monoton dalam θ untuk setiap $x_1, \ldots\ldots, x_n$ dan sifat monoton tersebut diperlukan untuk mendapatkan suatu interval konfidensi untuk θ.

C O N T O H.

Misalkan $X_1, \ldots, X_n$ adalah n sampel random dari $f_X(x) = 1/\theta \; x^{(1-\theta)/\theta} \; I_{(0,1)}$ dimana $\theta > 0$.

Ditanyakan interval konfidensi 100 γ persen untuk θ.

Penyelesaian.

$$F(x) = \int_0^x \frac{1}{\theta} u^{(1-\theta)/\theta} du$$

$$= \frac{1}{\theta}\left[\frac{1}{(1-\theta)/\theta + 1} u^{\frac{1-\theta}{\theta}+1}\right]_0^x$$

$$= x^{1/\theta} \quad \text{untuk} \quad 0 < x < 1.$$

$F_X(x) = 1$ untuk $1 \leq x \leq \infty$

Jadi :

$$F_X(x) = x^{1/\theta} I_{(\theta,1)}(x) + I_{(1,\infty)}(x)$$

$$\gamma = P\left[q_1 \leq \prod_{i=1}^{n} F(X_i;\theta) \leq q_2\right]$$

$$= P\left[q_1 \leq \prod_{i=1}^{n} X_i^{1/\theta} \leq q_2\right]$$

$$= P\left[\log q_1 \leqslant 1/\theta \log \prod_{i=1}^{n} X_i \leqslant \log q_2\right]$$

$$= P\left[\frac{\log q_1}{\log \prod_{i=1}^{n} X_i} \leqslant 1/\theta \leqslant \frac{\log q_2}{\log \prod_{i=1}^{n} X_i}\right]$$

$$= P\left[\frac{\log \prod_{i=1}^{n} X_i}{\log q_2} \leqslant \Theta \leqslant \frac{\log \prod_{i=1}^{n} X_i}{\log q_1}\right]$$

Jadi $\left(\frac{\log \prod_{i=1}^{n} X_i}{\log q_2}, \frac{\log \prod_{i=1}^{n} X_i}{\log q_1}\right)$ adalah interval

konfidensi 100 Y persen untuk o.

4. INTERVAL KONFIDENSI SAMPEL BESAR.

Untuk mendapatkan barisan estimator-estimator T_n dari θ dalam suatu density f $(.;\theta)$ yang didistribusikan secara normal asimptotik mendekati θ, sehingga T_n didistribusikan normal dengan mean θ dan varian $\sigma_n^2(\theta)$, dimana $\sigma_n^2(\theta)$ menunjukan varian yang merupakan fungsi dari θ untuk suatu sampel berukuran n.

Khususnya untuk sampel besar estimator likelihood maximum yaitu $\hat{\theta}_n = \theta(x_1, .. ,x_n)$ untuk parameter θ dari suatu fungsi f $(.;\theta)$ didistribusikan dengan distribusi normal mendekati o dengan varian.

$$\sigma_n^2(\theta) = \frac{1}{n\, E_\theta\left[\left\{(\partial/\partial\theta) \log f(X;\theta)\right\}^2\right]}$$

$$\sigma_n^2(\theta) = \frac{-1}{n\, E_\theta\left[\left\{(\frac{\partial^2}{\partial^2\theta}) \log f(X;\theta)\right]}$$

Jika barisan dari estimator-estimator $|T_n|$ yang berdistribusi normal asimptotik dari θ ada, maka untuk mendapatkan interval konfidensi bisa dicari. Dan $\left(\dfrac{T_n - \theta}{\sigma_n(\theta)} \right)$, dapat diperlakukan sebagai

Jumlah pivotal yang didekati, jadi untuk sampel berukuran besar suatu interval konfidensi dengan suatukoefisien konfidensi γ dapat ditentukan dengan mengubah ketidak samaan.

$$- Z < \frac{T_n - \theta}{\sigma_n(\theta)} < Z$$

dimana $Z = Z_{(1+\gamma)/2}$ didefinisikan dengan $\Phi(Z_{\frac{(1+\gamma)}{2}})$ atau $\Phi(Z) - \Phi(-z) = \gamma$.

Dan sifat optimum dari interval konfidensi sampel besar yang berdasarkan atas estimator-estimator likelihood maximum adalah lebih pendek jika dibandingkan denganinterval-interval yang ditentukan oleh sebarang estimator - yang lain.

5. HIPOTESIS GABUNGAN.

Dalam membicarakan uji hipotesis, perlu diperkenalkan beberapa notasi, istilah dan beberapa definisi. Disini yang dibicarakan hanya sebagian saja - dari bab uji hipotesis yaitu khususnya mengenai hipotesis gabungan dalam uji Generalited likelihood Ratio(uji GLR).

DEPINISI 3 UJI TINGKAT α UNTUK H_0 .

Uji tingkat α untuk H_0.

Suatu uji hipotesis $H_0 : \theta \in H_0$ dikatakan mempunyai tingkat α

Apabila untuk setiap $\theta \in H_o$, $P_\theta \{\text{menolak } H_o\} \leq \alpha$, dan untuk setiap $\delta < \alpha$, terdapat paling sedikit satu $\theta \in H_o$ sedemikian hingga P_θ $[\text{menolak } H_o] > \delta$, Yaitu,

$$\alpha = \sup_{\theta \in H_o} P_o \{\text{menolak } H_o\}$$

Di sini adalah distribusi probabitas yang ditentukan oleh O.

N o t a s i.

$\mathcal{P}_o$ dan $\mathcal{P}_1$ adalah sepasang keluarga parametrik yang saling asing, dimana parameter θ menentukan dengan tunggal elemen $\mathcal{P}_o$ atau $\mathcal{P}_1$,. Selanjutnya, misalkan H_o adalah himpunan harga-harga θ yang berkaitan dengan $\mathcal{P}_o$, sedangkan H_1, berkaitan dengan $\mathcal{P}_1$. P adalah distribusi variabelrandom $X_1, \ldots, X_n$ yang akan diamati.

DEFINISI 4 LIKELIHOOD RATIO TEST SEDERHANA.

Misalkan $X_1, .., X_n$ adalah sampel random salah satu $f_o(.)$ atau $f_1(.)$. Test Y dengan H_o : $x_i \sim f_o(.)$ lawan H_1: $x_i \sim f_1(.)$ didefinisikan sebagai likelihood ratio test sederhana jika λ didefinisikan.

Menolak H_o jika $\lambda < K$

Menerima H_o jika $\lambda > K$

Untuk $\lambda = K$ maka H_o dapat diterima atau dapat ditolak dimana

$$\lambda = \lambda(x_1, .., x_n) = \frac{\prod_{i=1}^{n} f_o(x_i)}{\prod_{i=1}^{n} f_1(x_i)}$$

$$= \frac{L_o(x_1, .., x_n)}{L_1(x_1, .., X_n)} = \frac{L_o}{L_1}$$

Dan K adalah konstante non negatip.

[L_j : L_j (x_1, .. ,X_n) adalah fungsi likelihood untuk sampling dari density f_j (.)]

Untuk K tertentu test dikatakan menolak H_0 jika ratio dari likelihood adalah kecil; yaitu menolak H_0 jika kemungkinan L_1 adalah besar dibandingkan dengan L_0, yaitu sampel yang didapat dari f_1 (.) kemudian dari f_0(.).

4.1. LIKELIHOOD RATIO TEST YANG UMUM (UJI GLR).

Untuk sampel random X_1, .. ,X_n dari density $f(x;\theta)$, $\theta \in \bar{\Theta}$ dilari test dari $H_0 : \theta \in \bar{\Theta}_0$ lawan $H_1 : \theta \in \bar{\Theta}_1$.

DEFINISI 5 LIKELIHOOD RATIO YANG UMUM.

Misalkan L (θ;x_1, .. ,x_n) adalah fungsi likelihod untuk sampel X_1, .. ,X_n mempunyai density bersama $f_{X_1, \ldots, X_n}(x_1, \ldots, x_n)$ dimana $\theta \in \bar{\Theta}$ likelihod ratio yang umum, ditentukan oleh λ atau λ_n, di definisikan sebagai :

$$\lambda = \lambda_n = \lambda(x_1, \ldots, x_n) = \frac{\sup\limits_{\theta \in \bar{\Theta}_0} L(\theta; x_1, \ldots, x_n)}{\sup\limits_{\theta \in \bar{\Theta}} L(\theta; x_1, \ldots, x_n)}$$

Pengamatan yang berkorespondensi dengan random variabel x_1, .. ,x_n dan ditulis dengan Λ untuk λ yaitu $\Lambda = \lambda(x_1, \ldots, x_n)$ fungsi dari random variabel $X_1, \ldots, X_n$

BEBERAPA CATATAN UNTUK LIKELIHOOD RATIO YANG UMUM.

i) Meskipun digunakan simbol λ yang sama untuk menentukan likelihood ratio sederhana, maka likelihood-ratio yang umum tidak mengurangi likelihood ratio sederhana untuk $\bar{\Theta} = \{\theta_0, \theta_1\}$.

ii) λ yang diberikan memenuhi $0 \leq \lambda \leq 1$, apabila $\lambda \geq 0$ maka banyaknya ratio non negatif dan apabila $\lambda \leq 1$ maka supremum diambil pada penyebut dengan himpunan yang lebih besar dari harga parameter dalam pem -bilang. Oleh karena itu penyebut tidak akan lebih kecil dari pada pembilang.

iii) Parameter θ sebagai harga vektor.
Penyebut Λ adalah evaluasi fungsi likelihood padaestimasi maximum likelihood.

iv) Dalam mempertimbangkan likelihood ratio yang umum , seringkali sampel $x_1, .., x_n$ menjadi sampel random dari density $f(x;\theta)$ dimana $\theta \in \bar{\Theta}$.

Harga λ dari statistik Λ digunakan untuk merumuskan test dari $H_0 : \theta \in \bar{\Theta}$ lawan $H_1 : \theta \in \bar{\Theta}_1$, Dengan prinsif likelihood ratio test yang umum maka tolok H_0 jika dan hanya jika $\lambda \leq \lambda_0$, dimana λ_0 konstanta tertentu sehingga $0 \leq \lambda_0 \leq 1$.

CONTOH : LIKELIHOOD RATIO TEST YANG UMUM.

Misalkan $x_1, .. ,x_n$ adalah sampel random dari

$$f(x;\theta) = \theta e^{-\theta x} I_{(0,\infty)}(x)$$

Test $H_0 : \theta \leq \theta_0$ lawan $H_1 : \theta > \theta_0$

$$\sup_{\theta \in \bar{\Theta}} L\left[(o:x_1, .. ,x_n)\right] = \sup_{\theta > 0}\left[\theta^n \exp(-\theta \Sigma x_i)\right] = \left(\frac{n}{\Sigma x_i}\right)^n e^{-n}$$

$$\sup_{\theta \in \bar{\Theta}_0} L\left[(o;x_1,...,x_n)\right] = \sup_{0 < \theta \leq \theta_0}\left[\theta^n \exp(-\theta \Sigma x_i)\right]$$

$$= \begin{cases} \left(\frac{n}{\Sigma x_i}\right)^n e^{-n} & \text{jika } \frac{n}{\Sigma x_i} \leq \theta_0 \\ \theta_0^n \exp(-\theta_0 \Sigma x_i) & \text{jika } \frac{n}{\Sigma x_i} > \theta_0 \end{cases}$$

maka :

$$\lambda = \begin{cases} 1 & \text{jika } \dfrac{n}{\Sigma x_i} \leq \theta_o \\ \dfrac{\theta_o \quad \exp(-\theta_o \Sigma x_i)}{(n/\Sigma x_i)^{n}\, e^{-n}} & \text{jika } \dfrac{n}{\Sigma x_i} > \theta_o \end{cases}$$

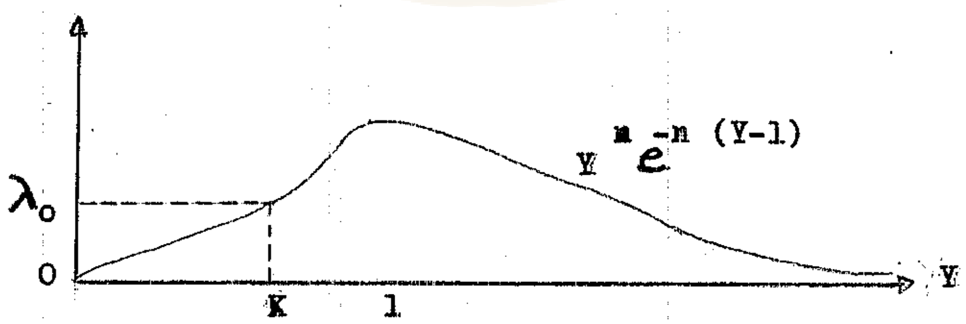
Jika $0 < \lambda_o < 1$, maka likelihood ratio test yang umum diberikan sebagai berikut. Tolok H_o jika $\lambda \leq \lambda_o$, atau tolok H_o jika $\frac{n}{\Sigma x_i} > \theta_o$ dan $\left(\frac{\theta_o \Sigma x_i}{n} \right)^{n} \exp(-e \Sigma x_i + n)$ $\leq \lambda_0$, atau tolok H_o jika $\theta_o\, \bar{x} < 1$ dan $(\theta_o \bar{x})^{n}\, e^{-n}$ $(\theta_o\, \bar{x} - 1) \leq \lambda_o$

Dengan menuliskan $Y = \theta_o \bar{x}$ maka $Y^{n} e^{-n\,(Y-1)} \leq \lambda_o$ jika $Y \leq K$ dimana K konstan yang memenuhi $0 < K < 1$ (lihat gambar 6).

Likelihood ratio test yang umum menjadi tolok H_o jika dan hanya jika.

$\theta_o\, \bar{x} < K$; dimana $0 < K < 1$ (lihat gambar 6)

gb .6



Yaitu, tolok H_o jika $\bar{x}$ kurang dari $1/\theta_o$ jika likelihood ratio test yang umum mempunyai ukuran α yang diinginkan maka K ditentukan dengan menyelesaikan persamaan :

$$\alpha = P_{\theta_o}\left[\theta_o \bar{x} < K\right] = P_{\theta_o}\left[\theta_o \Sigma x_i < nK\right]$$

$$= \int_0^{nK} \frac{1}{\Gamma(n)}\, u^{n-1} e^{-u}\, du$$

$\left(P_\theta\ (\theta_o \bar{x} < K) \leq P_{\theta_o}\ (\theta_o\, \bar{x} < K)\right)$ untuk $\theta \leq \theta_o$.