

**BUKU AJAR  
STATISTIKA MATEMATIKA II  
PAS 306 / 3 SKS**



**Oleh :**

**Dra. TATIK WIDIHARIH, M.Si**

UPT-PUS
No. Deft: 0167/BA/FMIPA/C <sub>1</sub>
Tgl. 22-7-2007

## **KATA PENGANTAR**

Dengan mengucapkan syukur kehadirat Tuhan Yang Maha Esa atas segala limpahan rahmat dan karunianya sehingga Buku Ajar Statistika Matematika 2 ini dapat diselesaikan sesuai dengan waktu yang telah direncanakan. Buku Ajar ini disusun sebagai buku pegangan kuliah untuk mata kuliah Statistika Matematika 2 mahasiswa Program Studi Statistika, Jurusan Matematika, FMIPA UNDIP Semarang semester 4.

Penulis berusaha semaksimal mungkin untuk dapat menyajikan materi secara ringkas dan jelas, sehingga mudah dipelajari para mahasiswa, namun penulis menyadari masih banyak kekurangannya. Penulis berharap kritik dan saran dari para pembaca, teman-teman sekolega demi kesempurnaan buku ajar ini.

Dalam kesempatan ini penulis mengucapkan banyak terima kasih kepada Ketua Program Studi Statistika, Jurusan Matematika, FMIPA UNDIP Semarang, yang telah memberi kesempatan kepada penulis. Ucapan terima kasih juga kepada Universitas Diponegoro yang menyediakan dana untuk kegiatan penulisan buku ajar ini khususnya Alokasi DIPA IDB Proyect. Kepada semua pihak yang tidak dapat penulis sebutkan satu persatu yang telah membantu baik sejak ide penyusunan, saat penyusunan berupa saran dan masukan yang sangat membantu sehingga dapat diselesaikannya buku ajar ini.

Harapan penulis semoga buku ajar ini dapat bermanfaat dan membantu para mahasiswa dalam menenpuh mata kuliah Statistika Matematika 2 serta dapat lebih mengembangkannya sehingga benar-benar

DAFTAR ISI

	Hal
KATA PENGANTAR.....	i
DAFTAR ISI .....	ii
MODUL I      ESTIMATOR TITIK .....	1
1.1 Pengantar .....	1
1.2 Tujuan Instruksional Umum .....	1
1.3 Tujuan Instruksional Khusus .....	1
1.4 Kejiata Belajar .....	1
1.4.1 Kegiatan Belajar 1. METODE ESTIMASI .....	1
1.4.1.1 Uraian dan Contoh .....	1
1. Metode Momen .....	3
2. Metode Maksimum Likelihood .....	6
1.4.1.2 Latihan 1 .....	12
1.4.1.3 Rangkuman .....	13
1.4.1.4 Tes Formatif 1 .....	13
1.4.1.5 Umpan Balik dan Tindak Lanjut .....	14
1.4.2 Kegiatan Belajar 2. EVALUASI ESTIMATOR .....	15
1.4.2.1 Uraian dan Contoh .....	15
1. Sifat Takbias .....	15
2. Sesatan Kuadrat Rata-rata .....	19
3. Konsistensi .....	23
4. Estimator Takbias Terbaik .....	26
1.4.2.2 Latihan 2 .....	36

	2.1 Pengantar .....	39
	2.2 Tujuan Instruksional Umum .....	39
	2.3 Tujuan Instruksional Khusus .....	39
	2.4 Kejiata Belajar KECUKUPAN DAN KELENGKAPAN. ....	40
	2.4.1 Uraian dan Contoh .....	40
	1. Statistik Cukup .....	40
	2. Statistik Lengkap .....	40
	3. Kecukupan, Kelengkapan dan Ketakbiasan .....	50
	2.4.2 Latihan 1 .....	61
	2.4.3 Rangkuman .....	61
	2.4.4 Tes Formatif 1 .....	61
	2.4.5 Umpan Balik dan Tindak Lanjut .....	62
	2.5 Kunci Jawaban Tes Formatif .....	62
	2.6 Referensi .....	63
MODUL III	UJI HIPOTESIS .....	64
	3.1 Pengantar .....	64
	3.2 Tujuan Instruksional Umum .....	64
	3.3 Tujuan Instruksional Khusus .....	64
	3.4 Kejiata Belajar UJI HIPOTESIS .....	65
	3.4.1 Uraian dan Contoh .....	65
	1. Pengertian .....	65
	2. Uji Paling Kuasa .....	68
	3. Uji Paling Kuasa Seragam .....	78
	4. Uji Rasio Likelihood .....	83

MODUL IV	ESTIMASI INTERVAL .....	93
	4.1 Pengantar .....	93
	4.2 Tujuan Instruksional Umum .....	93
	4.3 Tujuan Instruksional Khusus .....	93
	4.4 Kegiatan Belajar . ESTIMASI INTERVAL .....	94
	4.4.1 Uraian dan Contoh .....	94
	1. Pengertian .....	94
	2. Interval Konfindensi .....	96
	3. Metode Inversi Uji Hipotesis .....	97
	4. Metode Besaran Pivot .....	100
	5. Interval Konfindensi Pendekatan .....	110
	4.4.2 Latihan . .....	112
	4.4.3 Rangkuman .....	112
	4.4.4 Tes Formatif .. ..	112
	4.4.5 Umpan Balik dan Tindak Lanjut .....	113
	4.5 Kunci Jawaban Tes Formatif .....	114
	4.6 Referensi .....	115
LAMPIRAN	TABEL STATISTIKA	
	I.    Cumulatif Standart Normal Distribution .....	116
	II.   Percentage Points of the t Distribution .....	121
	III.  Percentage Points of the $\chi^2$ Distribution .....	122

## MODUL I ESTIMATOR TITIK

### 1.1 Pengantar

Parameter populasi biasanya harganya tidak diketahui, sehingga perlu diestimasi berdasarkan pengamatan dari data sampel. Dalam modul ini akan dibahas tentang : metode estimasi yaitu metode momen dan metode likelihood maksimum, evaluasi estimator meliputi sifat takbias, sesatan kuadrat rata-rata, konsistensi dan estimator takbias terbaik.

### 1.2 Tujuan Instruksional Umum:

Setelah mengikuti kuliah dan mempelajari modul ini mahasiswa dapat:

1. Menentukan estimator titik dari suatu parameter.
2. Melakukan evaluasi terhadap estimator titik yang telah diperoleh.

### 1.3 Tujuan Instruksional khusus:

Setelah mengikuti kuliah dan mempelajari modul ini mahasiswa dapat:

1. Menentukan estimator titik dengan metode momen.
2. Menentukan estimator titik dengan metode maksimum likelihood.
3. Membuktikan sifat takbias dari suatu parameter.
4. menghitung sesatan kuadrat rata-rata dari estimator titik.
5. Membuktikan sifat barisan estimator yang konsisten.
6. Menentukan batas bawah Crammer-Rao dari suatu estimator takbias.
7. Membuktikan suatu estimator titik merupakan estimator takbias terbaik.
8. Menentukan estimator takbias terbaik dari suatu parameter.

estimasi titik atau estimasi interval yang biasa dikenal dengan interval konfindensi. Untuk estimasi titik dibedakan menjadi dua pendekatan yaitu pendekatan klasik dan pendekatan teori keputusan.

Dalam pendekatan klasik, inferensi didasarkan sepenuhnya pada informasi yang diperoleh melalui sampel acak yang diambil dari suatu populasi yang berdistribusi tertentu. Metode-metode yang sering digunakan diantaranya metode momen dan metode maksimum likelihood. Estimator yang diperoleh perlu dievaluasi dengan menyelidiki sifat-sifat baik meliputi : sifat takbias, kekonsistenan dari barisan estimator , dan variansi minimum.

Dalam teori keputusan, inferensi didasarkan pada kombinasi dari informasi sampel dengan aspek-aspek lain yang relevan dengan persoalan tertentu agar diperoleh keputusan terbaik. Aspek lain yang dipandang relevan sebagai tambahan informasi sampel salah satunya adalah pengetahuan tentang konsekuensi yang mungkin timbul dari keputusan yang diambil. Pengetahuan ini seringkali diukur dengan menentukan kerugian yang akan terjadi untuk setiap keputusan yang mungkin dan untuk berbagai nilai yang mungkin dari parameter populasi. Fungsi resiko didefinisikan sebagai harga harapan dari fungsi kerugian. Kriteria kebaikan dari estimator ini salah satunya dilihat dari besarnya resiko dari estimator tersebut.

Misalkan  $X$  adalah variabel random dengan fungsi densitas  $f(x|\theta)$  dimana  $\theta$  adalah parameter yang harganya tidak diketahui. Dalam estimasi titik ini berarti mengestimasi (menduga) harga parameter  $\theta$  (dimungkinkan juga fungsi dari  $\theta$  katakana  $g(\theta)$ ) melalui / berdasarkan data observasi dari sampel. Secara matematika data observasi ini biasa disebut sampel random.

Bila  $X_1, X_2, \dots, X_n$  variabel random independent masing-masing dengan

merupakan himpunan bagian dari ruang Euclid dengan dimensi yang bersesuaian dengan dimensi  $\theta$ .

### 1. Metode Momen (Method of Moment Estimator / MME).

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

#### Definisi 1.1

Statistik  $T\left(\underset{\sim}{X}\right)$  adalah fungsi dari data observasi  $X_1, X_2, \dots, X_n$  yang digunakan untuk mengestimasi dari harga  $g(\theta)$  disebut estimator dari  $g(\theta)$  dan harga terobservasi dari statistik  $t\left(\underset{\sim}{x}\right)$  disebut estimasi (dugaan) dari  $g(\theta)$ .

Kadang-kadang estimator dari  $\theta$  juga dinotasikan dengan  $\hat{\theta}$ . Kita bias menghitung harga  $\hat{\theta} = T\left(\underset{\sim}{X}\right)$  langsung dari data observasi  $x_1, x_2, \dots, x_n$  karena fungsi tersebut tidak tergantung dari harga parameter  $\theta$  yang tidak diketahui.

#### Definisi 1.2.

1. Misalkan suatu populasi dengan fungsi densitas  $f(x|\theta)$ , maka momen populasi ke  $k$  didefinisikan sebagai  $\mu_k = E(X^k)$
2. Jika  $X_1, X_2, \dots, X_n$  adalah sampel random dari populasi dengan fungsi densitas  $f(x|\theta)$ , maka momen sampel ke  $k$  didefinisikan

persamaan simultan yang dihasilkan. Momen populasi ke  $j$  ( $\mu_j$ ) biasanya merupakan fungsi dari  $\theta = (\theta_1, \theta_2, \dots, \theta_k)$ . Estimator metode momen  $(\hat{\theta}_1, \hat{\theta}_2, \dots, \hat{\theta}_k)$  dari  $(\theta_1, \theta_2, \dots, \theta_k)$  didapatkan dengan menyelesaikan system persamaan  $(\theta_1, \theta_2, \dots, \theta_k)$  dalam bentuk  $m_1, m_2, \dots, m_k$  yaitu :

$$\mu_1 = m_1$$

$$\mu_2 = m_2$$

.....

$$\mu_k = m_k$$

### Contoh 1.1

Misalkan  $X_1, X_2, \dots, X_n$  adalah sampel random dari populasi yang berdistribusi Poisson dengan parameter  $\lambda$ . Dengan metode momen tentukan estimator titik untuk  $\lambda$ .

**Penyelesaian:**

$$X \sim P(\lambda), \text{ maka fungsi densitasnya : } f(x | \lambda) = e^{-\lambda} \frac{\lambda^x}{x!} \quad x=0,1,2,\dots$$

$$E(X) = \sum_{x=0}^{\infty} x \cdot e^{-\lambda} \frac{\lambda^x}{x!} = \lambda \sum_{x=1}^{\infty} e^{-\lambda} \frac{\lambda^{x-1}}{(x-1)!} = \lambda$$

Karena hanya satu parameter yang akan diestimasi maka hanya diperlukan satu persamaan dan langsung diperoleh estimator titiknya.

$$E(X) = \frac{1}{n} \sum_{i=1}^n X_i$$

**Contoh 1.2**

Misalkan  $X_1, X_2, \dots, X_n$  adalah sampel random dari populasi yang berdistribusi Normal dengan rata-rata  $\mu$  dan variansi  $\sigma^2$ . Dengan menggunakan metode momen tentukan estimator titik untuk  $\mu$  dan  $\sigma^2$ .

**Penyelesaian:**

$X \sim N(\mu, \sigma^2)$  berarti  $E(X) = \mu$  dan  $\text{Var}(Z) = \sigma^2$

$$\text{Var}(X) = E(X^2) - (E(X))^2$$

$$\sigma^2 = E(X^2) - \mu^2$$

$$E(X^2) = \mu^2 + \sigma^2$$

Sehingga diperoleh persamaan sebagai berikut :

$$E(X) = \frac{1}{n} \sum_{i=1}^n X_i \quad \text{sehingga} \quad \hat{\mu} = \frac{1}{n} \sum_{i=1}^n X_i$$

$$E(X^2) = \frac{1}{n} \sum_{i=1}^n X_i^2 \quad \text{maka} \quad \hat{\mu}^2 + \hat{\sigma}^2 = \frac{1}{n} \sum_{i=1}^n X_i^2$$

$$\hat{\sigma}^2 = \frac{1}{n} \sum_{i=1}^n X_i^2 - \left( \frac{1}{n} \sum_{i=1}^n X_i \right)^2$$

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

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

**Contoh 1.3**

Misalkan  $X_1, X_2, \dots, X_n$  adalah sampel random dari populasi yang berdistribusi  $\text{Exp}(1, \theta)$ . Dengan menggunakan metode momen tentukan estimator titik

$$\begin{aligned}
 E(X) &= \int_{\theta} x \cdot \exp(-x + \theta) dx \\
 &= - \int_{\theta} x \cdot d(\exp(-x + \theta)) \\
 &= -x \cdot \exp(-x + \theta) \Big|_{\theta}^{\infty} + \int_{\theta} \exp(-x + \theta) dx \\
 &= \theta - \exp(-x + \theta) \Big|_{\theta}^{\infty} \\
 &= \theta + 1
 \end{aligned}$$

Diperoleh persamaan :

$$E(X) = \frac{1}{n} \sum_{i=1}^n X_i$$

$$\hat{\theta} + 1 = \frac{1}{n} \sum_{i=1}^n X_i$$

$$\hat{\theta} = \frac{1}{n} \sum_{i=1}^n X_i - 1$$

## 2. Metode Maksimum Likelihood (maximum Likelihood Estimator / MLE)

Sejauh ini metode maksimum likelihood merupakan metode yang paling populer dalam menghasilkan estimator. Misalkan  $X_1, X_2, \dots, X_n$  adalah sampel random dari populasi dengan densitas  $f(x | \theta_1, \theta_2, \dots, \theta_k)$ , fungsi likelihood didefinisikan dengan :  $L(\theta_1, \theta_2, \dots, \theta_k | X) = \prod_{i=1}^n f(x_i | \theta_1, \theta_2, \dots, \theta_k)$ .

Bila fungsi likelihood ini terdiferensialkan dalam  $\theta_1, \theta_2, \dots, \theta_k$  maka calon estimator maksimum likelihood yang mungkin adalah harga-harga

$\left( \hat{\theta}_1, \hat{\theta}_2, \dots, \hat{\theta}_k \right)$  sedemikian sehingga :

$$\partial L(\theta_1, \theta_2, \dots, \theta_k | X)$$

$$\frac{\partial^2 L(\theta_1, \theta_2, \dots, \theta_k | \underline{X})}{\partial \theta_i^2} \Big|_{\theta_i = \hat{\theta}_i} < 0$$

Dalam banyak kasus dimana diferensi digunakan, akan lebih mudah bekerja pada logaritma alam (ln) dari  $L(\theta_1, \theta_2, \dots, \theta_k | \underline{X})$  yaitu :

$l = \log L(\theta_1, \theta_2, \dots, \theta_k | \underline{X})$ . Hal ini dimungkinkan karena fungsi logaritma

naik tegas pada  $(0, \infty)$ , yang berarti bahwa  $L(\theta_1, \theta_2, \dots, \theta_k | \underline{X})$  dan

$l = \log L(\theta_1, \theta_2, \dots, \theta_k | \underline{X})$  mempunyai ekstrem yang sama.

Jelasnya untuk menentukan estimator maksimum likelihood dari  $\theta = (\theta_1, \theta_2, \dots, \theta_k)$  dengan langkah-langkah sebagai berikut :

1. Tentukan fungsi likelihood :  $L(\theta_1, \theta_2, \dots, \theta_k | \underline{X}) = \prod_{i=1}^n f(x_i | \theta_1, \theta_2, \dots, \theta_k)$
2. Bentuk log likelihood :  $l = \log L(\theta_1, \theta_2, \dots, \theta_k | \underline{X})$
3. Tentukan turunan dari  $l = \log L(\theta_1, \theta_2, \dots, \theta_k | \underline{X})$  terhadap  $\theta_1, \theta_2, \dots, \theta_k$  :

$$\frac{\partial \log \left( L(\theta_1, \theta_2, \dots, \theta_k | \underline{X}) \right)}{\partial \theta_i} \quad \text{untuk } i = 1, 2, \dots, k$$

4. Bentuk persamaan likelihood :

$$\frac{\partial \log \left( L(\theta_1, \theta_2, \dots, \theta_k | \underline{X}) \right)}{\partial \theta_i} = 0 \quad \text{untuk } i = 1, 2, \dots, k$$

**Contoh 2.1**

Misalkan  $X_1, X_2, \dots, X_n$  adalah sampel random dari populasi yang berdistribusi Poisson dengan parameter  $\lambda$ . Dengan metode maksimum likelihood tentukan estimator titik untuk  $\lambda$ .

**Penyelesaian:**

$X_i \sim P(\lambda)$ , maka fungsi densitasnya :  $f(x_i | \lambda) = e^{-\lambda} \frac{\lambda^{x_i}}{x_i!}$   $x=0,1,2,\dots$

Fungsi likelihood :

$$\begin{aligned} L(\lambda | \underline{X}) &= \prod_{i=1}^n f(x_i | \lambda) \\ &= e^{-\lambda} \frac{\lambda^{x_1}}{x_1!} \cdot e^{-\lambda} \frac{\lambda^{x_2}}{x_2!} \cdots e^{-\lambda} \frac{\lambda^{x_n}}{x_n!} \\ &= \frac{1}{x_1! \cdot x_2! \cdots x_n!} e^{-n\lambda} \lambda^{\sum_{i=1}^n x_i} \end{aligned}$$

Log likelihood :

$$\begin{aligned} l &= \log \left( L(\lambda | \underline{X}) \right) \\ &= \log \left( \prod_{i=1}^n x_i \right)^{-1} - n\lambda + \sum_{i=1}^n x_i \log \lambda \\ &= -\sum_{i=1}^n \log(x_i) - n\lambda + \sum_{i=1}^n x_i \log \lambda \\ \frac{\partial l}{\partial \lambda} &= -n + \frac{1}{\lambda} \sum_{i=1}^n x_i \end{aligned}$$

Persamaan likelihood:

( ( \ ) )

**Contoh 2.2**

Misalkan  $X_1, X_2, \dots, X_n$  adalah sampel random dari populasi yang berdistribusi normal dengan rata-rata  $\mu$  dan variansi  $\sigma^2$ . Dengan menggunakan metode maksimum likelihood tentukan estimator titik untuk  $\mu$  dan  $\sigma^2$ .

**Penyelesaian :**

$X_i \sim N(\mu, \sigma^2)$  , maka densitasnya adalah :

$$f(x_i | \mu, \sigma^2) = \frac{1}{\sqrt{2\pi\sigma}} \cdot \exp\left(-\frac{1}{2\sigma^2}(x_i - \mu)^2\right)$$

Fungsi likelihood :

$$\begin{aligned} L(\mu, \sigma^2 | \underline{X}) &= \prod_{i=1}^n f(x_i | \mu, \sigma^2) \\ &= \frac{1}{\sqrt{2\pi\sigma}} \exp\left(-\frac{1}{2\sigma^2}(x_1 - \mu)^2\right) \dots \frac{1}{\sqrt{2\pi\sigma}} \exp\left(-\frac{1}{2\sigma^2}(x_n - \mu)^2\right) \\ &= (2\pi)^{-n/2} (\sigma^2)^{-n/2} \exp\left(-\frac{1}{2\sigma^2} \sum_{i=1}^n (x_i - \mu)^2\right) \end{aligned}$$

Log likelihood :

$$\begin{aligned} l &= \log\left(L(\mu, \sigma^2 | \underline{X})\right) \\ &= -\frac{n}{2} \log(2\pi) - \frac{n}{2} \log \sigma^2 - \frac{1}{2\sigma^2} \sum_{i=1}^n (x_i - \mu)^2 \end{aligned}$$

$$\frac{\partial l}{\partial \mu} = -\frac{1}{2\sigma^2} \cdot 2 \cdot \sum_{i=1}^n (x_i - \mu) \cdot (-1)$$

$$= \frac{1}{\sigma^2} \sum_{i=1}^n (x_i - \mu)$$

Persamaan likelihood :

$$\begin{aligned} \frac{\partial l}{\partial \hat{\mu}} &= 0 \\ \frac{1}{\hat{\sigma}^2} \sum_{i=1}^n (x_i - \hat{\mu}) &= 0 \\ \sum_{i=1}^n (x_i - \hat{\mu}) &= 0 \\ \sum_{i=1}^n x_i - n \hat{\mu} &= 0 \\ \hat{\mu} &= \frac{1}{n} \sum_{i=1}^n x_i = \bar{X} \\ \frac{\partial l}{\partial \hat{\sigma}^2} &= 0 \\ -\frac{n}{2 \hat{\sigma}^2} + \frac{1}{2 \hat{\sigma}^4} \sum_{i=1}^n (x_i - \hat{\mu})^2 &= 0 \\ -n \hat{\sigma}^2 + \sum_{i=1}^n (x_i - \bar{X})^2 &= 0 \\ \hat{\sigma}^2 &= \frac{1}{n} \sum_{i=1}^n (x_i - \bar{X})^2 \end{aligned}$$

### Contoh 2.3

Misalkan  $X_1, X_2, \dots, X_n$  adalah sampel random dari populasi yang berdistribusi Gamma( $\alpha, \beta$ ). Dengan menggunakan metode maksimum likelihood tentukan estimator titik  $\alpha$  dan  $\beta$ .

**Penyelesaian:**

$X_i \sim \text{Gamma}(\alpha, \beta)$ , maka densitasnya adalah :

$$f(x_i | \alpha, \beta) = \frac{1}{\Gamma(\alpha) \beta^\alpha} \cdot x_i^{\alpha-1} \exp\left(-\frac{x}{\beta}\right) \quad x_i > 0 \quad \alpha, \beta > 0$$

Fungsi likelihood :

$$\begin{aligned} L(\alpha, \beta | \underline{X}) &= \prod_{i=1}^n f(x_i | \alpha, \beta) \\ &= \frac{1}{\Gamma(\alpha)\beta^\alpha} x_1^{\alpha-1} \exp\left(-\frac{x_1}{\beta}\right) \cdots \frac{1}{\Gamma(\alpha)\beta^\alpha} x_n^{\alpha-1} \exp\left(-\frac{x_n}{\beta}\right) \\ &= (\Gamma(\alpha))^{-n} \beta^{-n\alpha} \prod_{i=1}^n x_i^{\alpha-1} \exp\left(-\frac{1}{\beta} \sum_{i=1}^n x_i\right) \end{aligned}$$

Log likelihood :

$$\begin{aligned} l &= \log\left(L(\alpha, \beta | \underline{X})\right) \\ &= -n \log(\Gamma(\alpha)) - n\alpha \log \beta - \frac{1}{\beta} \sum_{i=1}^n x_i \\ \frac{\partial l}{\partial \alpha} &= -n \frac{\Gamma'(\alpha)}{\Gamma(\alpha)} - n \log \beta + \sum_{i=1}^n \log(x_i) \\ \frac{\partial l}{\partial \beta} &= -\frac{n\alpha}{\beta} + \frac{1}{\beta^2} \sum_{i=1}^n x_i \end{aligned}$$

Persamaan likelihood :

$$\begin{aligned} \frac{\partial l}{\partial \beta} &= 0 \\ -\frac{n\hat{\alpha}}{\hat{\beta}} + \frac{1}{\hat{\beta}^2} \sum_{i=1}^n x_i &= 0 \\ \hat{\beta} &= \frac{\sum_{i=1}^n x_i}{n\hat{\alpha}} = \frac{\bar{X}}{\hat{\alpha}} \\ \frac{\partial l}{\partial \alpha} &= 0 \\ -n\varphi'(\hat{\alpha}) - n \log \hat{\beta} + \sum_{i=1}^n \log(x_i) &= 0 \quad \text{dengan} \quad \varphi'(\alpha) = \frac{\Gamma'(\alpha)}{\Gamma(\alpha)} \\ -n \log \frac{\bar{X}}{\hat{\alpha}} - n\varphi'(\hat{\alpha}) + \log \frac{\left(\prod_{i=1}^n x_i\right)^{1/n}}{\bar{X}} &= 0 \end{aligned}$$

Bentuk tersebut diatas tidak "closed form" sehingga diselesaikan secara numeric. Menurut Greenwood dan Durand (1960) dalam Bain dkk(1992) :

$$\hat{\alpha} = \begin{cases} (0,5000876 + 0,1648852M - 0,0544274 M^2) / M & \text{untuk } 0 < M \leq 0,5772 \\ \frac{8,898919 + 9,059950M + 0,9775373 M^2}{M(17,79758 + 11,968477M + M^2)} & \text{untuk } 0,5772 < M \leq 17 \\ M^{-1} & \text{untuk } M > 17 \end{cases}$$

dengan  $M = \ln \left( \frac{\left( \sum_{i=1}^n x_i \right)^{1/n}}{\bar{X}} \right)$

Salah satu sifat yang paling berguna dari estimator maksimum likelihood ini adalah sifat invariannya. Misalkan dalam hal ini tertarik pada estimasi  $g(\theta)$  dengan  $\theta \rightarrow g(\theta)$  fungsi satu-satu. Sifat invariant ini mengatakan bahwa jika  $\hat{\theta}$  adalah estimator maksimum likelihood untuk  $\theta$ , maka  $g(\hat{\theta})$  juga estimator maksimum likelihood untuk  $g(\theta)$ .

**Contoh 2.4**

Berdasarkan contoh 1.3.2 telah diperoleh :  $\hat{\sigma}^2 = \frac{1}{n} \sum_{i=1}^n (x_i - \bar{X})^2$ , dengan sifat

invariant maka  $\hat{\sigma} = \sqrt{\frac{1}{n} \sum_{i=1}^n (x_i - \bar{X})^2}$

**1.4.1.2 Latihan 1**

1. Apa perbedaan antara momen populasi dengan momen sampel
2. Mengapa dalam metode likelihood maksimum disyaratkan bahwa  $X_1, X_2, \dots, X_n$  peubah acak independen dan berdistribusi identik, bagaimana jika syarat tersebut tidak dipenuhi?
3. Apakah pendugaan dengan menggunakan metode momen dan metode likelihood maksimum selalu menghasilkan estimator titik yang sama? Jelaskan dan berikan contoh.

### 1.4.1.3 Rangkuman

1. Dalam pendugaan parameter dengan metode momen yang terpenting adalah menentukan momen ke - k dari populasi maupun sampel dan membentuk k buah persamaan beserta penyelesaiannya.
2. Dalam pendugaan parameter dengan metode likelihood maksimum, memaksimumkan fungsi likelihood dalam dilakukan dengan memaksimumkan log likelihood.
3. Sifat yang terpenting dalam estimator likelihood maksimum adalah sifat invarian.

### 1.4.1.4 Test Formatif 1

1. Tentukan estimator dengan metode momen berdasarkan sampel random  $X_1, X_2, \dots, X_n$  dari populasi masing-masing dengan densitas :

$$a. f(x | \theta) = \theta x^{\theta-1} \quad 0 < x < 1 \quad \theta > 0$$

$$b. f(x | \theta) = (\theta + 1) x^{-\theta-2} \quad x > 1 \quad \theta > 0$$

$$c. f(x | \theta) = \theta^2 x \cdot \exp(-\theta x) \quad x > 0 \quad \theta > 0$$

$$d. X_i \sim \text{Gamma}(2, \theta)$$

$$e. X_i \sim \text{Exp}(\theta)$$

2. Tentukan estimator dengan metode maksimum likelihood berdasarkan sampel random  $X_1, X_2, \dots, X_n$  dari populasi masing-masing dengan densitas :

$$a. f(x | \theta) = \theta x^{\theta-1} \quad 0 < x < 1 \quad \theta > 0$$

$$b. f(x | \theta) = (\theta + 1) x^{-\theta-2} \quad x > 1 \quad \theta > 0$$

$$c. f(x | \theta) = \theta^2 x \cdot \exp(-\theta x) \quad x > 0 \quad \theta > 0$$

$$d. X_i \sim \text{Gamma}(2, \theta)$$

$$e. X_i \sim \text{Exp}(\theta)$$

#### 1.4.1.5 Umpan Balik dan tidak Lanjut

Cocokkanlah jawaban anda dengan kunci jawaban dari modul ini. Hitunglah jawaban anda yang benar, kemudian gunakanlah rumus dibawah ini untuk mengetahui tingkat penguasaan anda terhadap materi ini.

$$\text{Tingkat penguasaan} = (\text{jumlah jawaban benar}) / 10 \times 100\%$$

Arti tingkat penguasaan yang anda capai :

90% - 100% : baik sekali

80% - 89% : baik

70% - 79% : sedang

< 70% : kurang

Bila anda telah mencapai tingkat penguasaan 80% atau lebih, anda dapat meneruskan kegiatan belajar selanjutnya. Tetapi apabila tingkat penguasaan anda kurang dari 80% , sebaiknya anda mengulangi materi ini terutama bagian yang belum anda kuasai sebelum meneruskan kegiatan belajar selanjutnya.

## 1.4.2 Kegiatan Belajar 2:

### EVALUASI ESTIMATOR

#### 1.4 .2.1 Uraian dan Contoh.

Estimator titik untuk parameter  $\theta$  melalui pendekatan klasik yaitu metode momen dan metode maksimum likelihood mungkin diperoleh estimator yang berbeda. Masalahnya sekarang adalah bagaimana memilih salah satu estimator terbaik yang memenuhi sifat-sifat kebaikan suatu estimator. Dalam submodul ini akan diperkenalkan patokan dasar untuk mengevaluasi estimator dan menyelidiki kelakuan beberapa estimator terhadap kriteria tertentu.

#### 1. Sifat Takbias (unbias)

Sifat takbias ini merupakan sifat baik dari suatu estimator yang diperoleh melalui pendekatan klasik, dalam pemilihan estimator terbaik salah satunya harus memenuhi sifat takbias ini.

#### Definisi 1.1

Misalkan  $T\left(\underset{\sim}{X}\right)$  suatu estimator titik untuk parameter  $\theta$ , maka

$T\left(\underset{\sim}{X}\right)$  disebut estimator takbias untuk  $\theta$  jika  $E\left(T\left(\underset{\sim}{X}\right)\right) = \theta$ .

#### Contoh 1.1

Misalkan  $X_1, X_2, \dots, X_n$  adalah sampel random independen dari populasi berdistribusi Poisson dengan parameter  $\lambda$ . Telah diperoleh dari contoh 1.2.1

$\hat{\lambda} = \frac{1}{n} \sum_{i=1}^n X_i$ , apakah estimator ini merupakan estimator yang takbias untuk  $\lambda$ .

#### Penyelesaian:

Akan ditunjukkan bahwa  $E\left(\hat{\lambda}\right) = \lambda$

$X_i \sim \text{Poisson}(\lambda)$ , maka  $E(X_i) = \lambda$

$$\begin{aligned}
 E(\hat{\lambda}) &= E\left(\frac{1}{n} \sum_{i=1}^n X_i\right) \\
 &= \frac{1}{n} \sum_{i=1}^n E(X_i) \\
 &= \frac{1}{n} (E(X_1) + E(X_2) + \dots + E(X_n)) \\
 &= \frac{1}{n} (\lambda + \lambda + \dots + \lambda) \\
 &= \frac{1}{n} (n\lambda) \\
 &= \lambda
 \end{aligned}$$

Sehingga  $\hat{\lambda} = \frac{1}{n} \sum_{i=1}^n X_i$  adalah estimator yang takbias untuk  $\lambda$ .

### Contoh 1.2

Misalkan  $X_1, X_2, \dots, X_n$  adalah sampel random independen dari populasi berdistribusi normal dengan rata-rata  $\mu$  dan variansi  $\sigma^2$ . Dari contoh 1.3.2

telah diperoleh :  $\hat{\mu} = \frac{1}{n} \sum_{i=1}^n X_i$  dan  $\hat{\sigma}^2 = \frac{1}{n} \sum_{i=1}^n (x_i - \bar{X})^2$ , apakah estimator

tersebut merupakan estimator yang takbias untuk  $\mu$  dan  $\sigma^2$ .

### Penyelesaian :

$X_i \sim N(\mu, \sigma^2)$ , maka  $E(X_i) = \mu$

$$\sum_{i=1}^n \left( \frac{x_i - \bar{X}}{\sigma} \right)^2 \sim \chi_{n-1}^2$$

$$E \left( \sum_{i=1}^n \left( \frac{x_i - \bar{X}}{\sigma} \right)^2 \right) = n-1$$

Akan ditunjukkan :  $E(\hat{\mu}) = \mu$  dan  $E(\hat{\sigma}^2) = \sigma^2$

$$\begin{aligned}
 E(\hat{\mu}) &= E\left(\frac{1}{n}\sum_{i=1}^n X_i\right) \\
 &= \frac{1}{n}\sum_{i=1}^n E(X_i) \\
 &= \frac{1}{n}(E(X_1) + E(X_2) + \dots + E(X_n)) \\
 &= \frac{1}{n}(\mu + \mu + \dots + \mu) \\
 &= \frac{1}{n}(n\mu) \\
 &= \mu
 \end{aligned}$$

Sehingga  $\hat{\mu} = \frac{1}{n}\sum_{i=1}^n X_i$  merupakan estimator yang takbias untuk  $\mu$ .

$$\begin{aligned}
 E(\hat{\sigma}^2) &= E\left(\frac{1}{n}\sum_{i=1}^n (x_i - \bar{X})^2\right) \\
 &= \frac{\sigma^2}{n} E\left(\sum_{i=1}^n \left(\frac{x_i - \bar{X}}{\sigma}\right)^2\right) \\
 &= \frac{\sigma^2}{n} \cdot (n-1)
 \end{aligned}$$

Karena  $E(\hat{\sigma}^2) \neq \sigma^2$ , maka  $\hat{\sigma}^2 = \frac{1}{n}\sum_{i=1}^n (x_i - \bar{X})^2$  merupakan estimator

yang bias. Tetapi apabila diambil estimator lain untuk  $\sigma^2$  katakan  $S^2$

dengan :  $S^2 = \frac{1}{n-1}\sum_{i=1}^n (x_i - \bar{X})^2$  dapat ditunjukkan bahwa  $S^2$  merupakan

estimator yang takbias untuk  $\sigma^2$  sebagai berikut :

$$\begin{aligned}
 E(S^2) &= E\left(\frac{1}{n-1}\sum_{i=1}^n (x_i - \bar{X})^2\right) \\
 &= \frac{\sigma^2}{n-1} E\left(\sum_{i=1}^n \left(\frac{x_i - \bar{X}}{\sigma}\right)^2\right) \\
 &= \frac{\sigma^2}{n-1} \cdot (n-1) \\
 &= \sigma^2
 \end{aligned}$$

Karena  $E(S^2) = \sigma^2$  maka  $S^2$  merupakan estimator yang takbias untuk  $\sigma^2$ .

Apabila  $T\left(\underset{\sim}{X}\right)$  merupakan estimator yang bias untuk  $\theta$ , maka :

$$\text{Bias}\left(T\left(\underset{\sim}{X}\right)\right) = E\left(T\left(\underset{\sim}{X}\right)\right) - \theta.$$

Berdasarkan contoh diatas maka :

$$\begin{aligned} \text{Bias}\left(\hat{\sigma}^2\right) &= E\left(\hat{\sigma}^2\right) - \sigma^2 \\ &= \frac{(n-1)\sigma^2}{n} - \sigma^2 \\ &= -\frac{\sigma^2}{n} \end{aligned}$$

### Contoh 1.3

Misalkan  $X_1, X_2, \dots, X_n$  adalah sampel random independen dari populasi berdistribusi Uniform( $\theta-0,5$ ,  $\theta+0,5$ ). Dengan menggunakan metode momen tentukan estimator untuk  $\theta$  dan selidiki apakah estimator tersebut bersifat takbias.

**Penyelesaian :**

$X \sim U(\theta-0,5, \theta+0,5)$ , maka densitasnya adalah :

$$f(x|\theta) = \frac{1}{(\theta+0,5) - (\theta-0,5)} = 1$$

$$\begin{aligned} E(X) &= \int_x x \cdot f(x|\theta) dx \\ &= \int_{\theta-0,5}^{\theta+0,5} x \cdot dx \\ &= \frac{x^2}{2} \Big|_{\theta-0,5}^{\theta+0,5} \\ &= \frac{1}{2} \left( (\theta+0,5)^2 - (\theta-0,5)^2 \right) \\ &= \frac{1}{2} \left( \theta^2 + \theta + 0,25 - \theta^2 + \theta - 0,25 \right) \\ &= \theta \end{aligned}$$

Sehingga diperoleh persamaan :

$$E(X) = \frac{1}{n} \sum_{i=1}^n X_i$$

$$\hat{\theta} = \frac{1}{n} \sum_{i=1}^n X_i$$

Selanjutnya untuk membuktikan sifat takbias akan ditunjukkan :  $E(\hat{\theta}) = \theta$ .

$$\begin{aligned} E(\hat{\theta}) &= E\left(\frac{1}{n} \sum_{i=1}^n X_i\right) \\ &= \frac{1}{n} \sum_{i=1}^n E(X_i) \\ &= \frac{1}{n} (E(X_1) + E(X_2) + \dots + E(X_n)) \\ &= \frac{1}{n} (\theta + \theta + \dots + \theta) \\ &= \frac{1}{n} (n\theta) \\ &= \theta \end{aligned}$$

Sehingga terbukti  $\hat{\theta} = \frac{1}{n} \sum_{i=1}^n X_i$  merupakan estimator yang takbias untuk  $\theta$ .

## 2. Sesatan Kuadrat Rata-rata (Mean Square Error / MSE)

Ukuran kualitas estimator yang pertama kali dibahas adalah sesatan kuadrat rata-rata (MSE = mean square error) yang didefinisikan sebagai berikut :

### Definisi 2.1

Sesatan kuadrat rata-rata estimator  $T(\underline{X})$  untuk parameter  $\theta$  dinotasikan dengan  $MSE(T(\underline{X}))$  didefinisikan dengan :

$$MSE\left(T(\underline{X})\right) = E\left(T(\underline{X}) - \theta\right)^2$$

Bentuk ekspektasi dari definisi tersebut dapat disederhanakan sebagai berikut

$$\begin{aligned}
 \text{MSE}\left(T\left(\underset{\sim}{X}\right)\right) &= E\left(T\left(\underset{\sim}{X}\right) - \theta\right)^2 \\
 &= E\left(T\left(\underset{\sim}{X}\right) - E\left(T\left(\underset{\sim}{X}\right)\right) + E\left(T\left(\underset{\sim}{X}\right)\right) - \theta\right)^2 \\
 &= E\left(T\left(\underset{\sim}{X}\right) - E\left(T\left(\underset{\sim}{X}\right)\right)\right)^2 + E\left(E\left(T\left(\underset{\sim}{X}\right)\right) - \theta\right)^2 \\
 &\quad + 2E\left(T\left(\underset{\sim}{X}\right) - E\left(T\left(\underset{\sim}{X}\right)\right)\right) \cdot E\left(E\left(T\left(\underset{\sim}{X}\right)\right) - \theta\right) \\
 &= \text{Var}\left(T\left(\underset{\sim}{X}\right)\right) + \left(\text{bias}\left(T\left(\underset{\sim}{X}\right)\right)\right)^2
 \end{aligned}$$

Apabila  $T\left(\underset{\sim}{X}\right)$  merupakan estimator yang takbias untuk  $\theta$  maka :

$$\text{MSE}\left(T\left(\underset{\sim}{X}\right)\right) = \text{Var}\left(T\left(\underset{\sim}{X}\right)\right).$$

Secara umum MSE mempunyai dua komponen, yaitu variansi yang mengukur variabilitas (precision) dan bias yang mengukur keakuratan (accuracy) dari estimator.

### Contoh 2.1

Misalkan  $X_1, X_2, \dots, X_n$  adalah sampel random independen dari populasi berdistribusi Eksponensial dengan rata-rata  $\theta$ . Dengan menggunakan metode momen tentukan estimator untuk  $\theta$  dan hitunglah sesatan kuadrat rata-rata dari estimator tersebut.

### Penyelesaian :

$X \sim \text{Exp}(\theta)$  , maka densitasnya adalah :

$$\begin{aligned}
 f(x|\theta) &= \frac{1}{\theta} \cdot \exp\left(-\frac{x}{\theta}\right) \quad x > 0 \quad \theta > 0 \\
 E(X) &= \int_0^{\infty} x \cdot \frac{1}{\theta} \cdot \exp\left(-\frac{x}{\theta}\right) dx \\
 &= \theta \int_0^{\infty} \left(\frac{x}{\theta}\right)^{2-1} \exp\left(-\frac{x}{\theta}\right) d\frac{x}{\theta} \\
 &= \theta \cdot \Gamma(2) \\
 &= \theta
 \end{aligned}$$

Sehingga diperoleh persamaan :

$$E(X) = \frac{1}{n} \sum_{i=1}^n X_i$$

$$\hat{\theta} = \frac{1}{n} \sum_{i=1}^n X_i$$

$$MSE(\hat{\theta}) = \text{var}(\hat{\theta}) + \left( \text{bias}(\hat{\theta}) \right)^2$$

Perlu diselidiki sifat takbias dari  $\hat{\theta}$  untuk dapat menghitung  $\text{bias}(\hat{\theta})$ .

$$\begin{aligned} E(\hat{\theta}) &= E\left(\frac{1}{n} \sum_{i=1}^n X_i\right) \\ &= \frac{1}{n} \sum_{i=1}^n E(X_i) \\ &= \frac{1}{n} (E(X_1) + E(X_2) + \dots + E(X_n)) \\ &= \frac{1}{n} (\theta + \theta + \dots + \theta) \\ &= \frac{1}{n} (n\theta) \\ &= \theta \end{aligned}$$

Terlihat bahwa  $\hat{\theta}$  merupakan estimator yang takbias untuk  $\theta$ , sehingga  $\text{bias}(\hat{\theta}) = 0$ .

$$\begin{aligned} E(X^2) &= \int_0^{\infty} x^2 \cdot \frac{1}{\theta} \cdot \exp\left(-\frac{x}{\theta}\right) dx \\ &= \theta^2 \int_0^{\infty} \left(\frac{x}{\theta}\right)^{3-1} \exp\left(-\frac{x}{\theta}\right) d\frac{x}{\theta} \\ &= \theta^2 \Gamma(3) \\ &= 2\theta^2 \end{aligned}$$

$$\begin{aligned} \text{Var}(X) &= E(X^2) - (E(X))^2 \\ &= 2\theta^2 - \theta^2 \\ &= \theta^2 \end{aligned}$$

$$\begin{aligned}
 MSE(\hat{\theta}) &= \text{var}(\hat{\theta}) + \left( \text{bias}(\hat{\theta}) \right)^2 \\
 &= \text{var}\left(\frac{1}{n} \sum_{i=1}^n X_i\right) \\
 &= \frac{1}{n^2} \cdot \sum_{i=1}^n \text{var}(X_i) \\
 &= \frac{1}{n^2} \cdot \sum_{i=1}^n \theta^2 \\
 &= \frac{1}{n^2} \cdot n \theta^2 = \frac{\theta^2}{n}
 \end{aligned}$$

**Contoh 2.2**

Misalkan  $X_1, X_2, \dots, X_n$  adalah sampel random independen dari populasi berdistribusi Normal dengan rata-rata  $\mu$  dan variansi  $\sigma^2$ . Telah diperoleh :

$$\hat{\mu} = \frac{1}{n} \sum_{i=1}^n X_i \quad \text{dan} \quad \hat{\sigma}^2 = \frac{1}{n} \sum_{i=1}^n (x_i - \bar{X})^2, \text{ hitunglah sesatan kuadrat rata-}$$

rata dari kedua estimator tersebut.

**Penyelesaian :**

Berdasarkan contoh 1.4.2 telah diperoleh bahwa  $\hat{\mu} = \frac{1}{n} \sum_{i=1}^n X_i$  merupakan estimator yang takbias untuk  $\mu$ , sehingga :

$$\begin{aligned}
 MSE(\hat{\mu}) &= \text{var}(\hat{\mu}) \\
 &= \text{var}\left(\frac{1}{n} \sum_{i=1}^n X_i\right) \\
 &= \frac{1}{n^2} \sum_{i=1}^n \text{var}(X_i) \\
 &= \frac{1}{n^2} \sum_{i=1}^n \sigma^2 \\
 &= \frac{1}{n^2} \cdot n \sigma^2 \\
 &= \frac{\sigma^2}{n}
 \end{aligned}$$

Sedangkan :  $\hat{\sigma}^2 = \frac{1}{n} \sum_{i=1}^n (x_i - \bar{X})^2$  merupakan estimator yang bias untuk  $\sigma^2$  dengan  $bias\left(\hat{\sigma}^2\right) = -\frac{\sigma^2}{n}$ .

Karena :

$$\sum_{i=1}^n \left( \frac{X_i - \bar{X}}{\sigma} \right)^2 \sim \chi_{n-1}^2 \quad \text{maka}$$

$$n \frac{\hat{\sigma}^2}{\sigma^2} \sim \chi_{n-1}^2$$

$$\text{Var} \left( n \frac{\hat{\sigma}^2}{\sigma^2} \right) = 2(n-1)$$

$$\frac{n^2}{\sigma^4} \text{var} \left( \hat{\sigma}^2 \right) = 2(n-1)$$

$$\text{var} \left( \hat{\sigma}^2 \right) = 2(n-1) \cdot \frac{\sigma^4}{n}$$

$$\begin{aligned} \text{MSE} \left( \hat{\sigma}^2 \right) &= \text{var} \left( \hat{\sigma}^2 \right) + \left( \text{bias} \left( \hat{\sigma}^2 \right) \right)^2 \\ &= 2(n-1) \cdot \frac{\sigma^4}{n} + \frac{\sigma^4}{n} \\ &= (2n-1) \frac{\sigma^4}{n} \end{aligned}$$

### 3. Konsistensi

Sejauh ini kriteria yang dibicarakan adalah kriteria untuk sampel berhingga. Sebaliknya konsistensi adalah sifat asimtotis, yaitu menggambarkan sifat estimator bila ukuran sampel menjadi takhingga. Konsistensi adalah sifat dari barisan estimator bukan dari estimator tunggal, walaupun biasanya disebut suatu estimator konsisten. Bila dipunyai

observasi  $X_1, X_2, \dots$  dengan densitas  $f(x_i|\theta)$ , dapat dikonstruksikan barisan estimator  $T_n(\underline{X}) = T_n(X_1, X_2, \dots, X_n)$  dengan melakukan prosedur estimasi yang sama untuk setiap ukuran sampel  $n$ . Sebagai contoh :

$$\bar{X}_1 = X_1 \quad \bar{X}_2 = \frac{1}{2}(X_1 + X_2) \quad \bar{X}_3 = \frac{1}{3}(X_1 + X_2 + X_3) \quad \dots \text{dan seterusnya.}$$

Sehingga dapat didefinisikan barisan estimator konsisten sebagai berikut :

**Definisi 3.1**

Barisan estimator  $T_n(\underline{X}) = T_n(X_1, X_2, \dots, X_n)$  adalah barisan estimator konsisten untuk parameter  $\theta$ , bila untuk setiap  $\varepsilon > 0$  dan untuk setiap  $\theta \in \Omega$  berlaku :  $\lim_{n \rightarrow \infty} P\left(\left|T_n(\underline{X}) - \theta\right| < \varepsilon\right) = 1$  atau ekuivalen dengan :  $\lim_{n \rightarrow \infty} P\left(\left|T_n(\underline{X}) - \theta\right| \geq \varepsilon\right) = 0$

Definisi diatas mengatakan bahwa barisan estimator yang konsisten konvergen dalam promodulilitas pada parameter  $\theta$ .

**Contoh 3.1**

Misalkan  $X_1, X_2, \dots, X_n$  variabel random independen berdistribusi normal dengan rata-rata  $\theta$  dan variansi 1. Telah diketahui bahwa  $\hat{\theta} = \bar{X}_n = \frac{1}{n} \sum_{i=1}^n X_i$ .

Akan ditunjukkan bahwa estimator untuk  $\theta$  tersebut merupakan barisan estimator yang konsisten.

**Penyelesaian :**

Karena :  $X_i \sim N(\theta, 1)$  maka  $\bar{X}_n \sim N\left(\theta, \frac{1}{n}\right)$

$$\begin{aligned}
 P(|\bar{X}_n - \theta| < \varepsilon) &= \int_{\theta - \varepsilon}^{\theta + \varepsilon} \left(\frac{n}{2\pi}\right)^{1/2} \exp\left(-\frac{n}{2}(\bar{x}_n - \theta)^2\right) d\bar{x}_n \\
 &= \int_{-\varepsilon}^{\varepsilon} \left(\frac{n}{2\pi}\right)^{1/2} \exp\left(-\frac{n}{2}y^2\right) dy \\
 &= \int_{-\varepsilon\sqrt{n}}^{\varepsilon\sqrt{n}} \left(\frac{1}{2\pi}\right)^{1/2} \exp\left(-\frac{1}{2}t^2\right) dt \\
 &= P(-\varepsilon\sqrt{n} < Z < \varepsilon\sqrt{n}) \rightarrow 1 \quad \text{untuk } n \rightarrow \infty
 \end{aligned}$$

Sehingga  $\hat{\theta} = \bar{X}_n = \frac{1}{n} \sum_{i=1}^n X_i$  adalah barisan estimator konsisten untuk  $\theta$ .

Secara umum, perhitungan mendetail seperti diatas tidak perlu dilakukan untuk membuktikan konsistensi. Ketaksamaan Chebychev lebih

mudah digunakan yaitu : 
$$P\left(\left|T_n\left(\underset{\sim}{X}\right) - \theta\right| \geq \varepsilon\right) = \frac{1}{\varepsilon^2} E\left(T_n\left(\underset{\sim}{X}\right) - \theta\right)^2$$

Sehingga ;  $\lim_{n \rightarrow \infty} E\left(T_n\left(\underset{\sim}{X}\right) - \theta\right)^2 = 0$  adalah syarat cukup agar suatu

barisan estimator  $T_n\left(\underset{\sim}{X}\right) = T_n(X_1, X_2, \dots, X_n)$  merupakan barisan

estimator yang konsisten. Selanjutnya dapat ditulis :

$$\begin{aligned}
 E\left(T_n\left(\underset{\sim}{X}\right) - \theta\right)^2 &= E\left(T_n\left(\underset{\sim}{X}\right) - E\left(T_n\left(\underset{\sim}{X}\right)\right) + E\left(T_n\left(\underset{\sim}{X}\right)\right) - \theta\right)^2 \\
 &= E\left(T_n\left(\underset{\sim}{X}\right) - E\left(T_n\left(\underset{\sim}{X}\right)\right)\right)^2 + \left(E\left(T_n\left(\underset{\sim}{X}\right)\right) - \theta\right)^2 \\
 &= \text{Var}\left(T_n\left(\underset{\sim}{X}\right)\right) + \left(\text{bias}\left(T_n\left(\underset{\sim}{X}\right)\right)\right)^2
 \end{aligned}$$

Berdasarkan persamaan diatas dapat disimpulkan bahwa  $T_n(\underline{X}) = T_n(X_1, X_2, \dots, X_n)$  merupakan barisan estimator konsisten bila memenuhi :

1.  $\lim_{n \rightarrow \infty} Var\left(T_n\left(\underline{X}\right)\right) = 0$
2.  $\lim_{n \rightarrow \infty} \left(bias\left(T_n\left(\underline{X}\right)\right)\right)^2 = 0$

### Contoh 3.2

Berdasarkan contoh 1.4.6 untuk menunjukkan bahwa  $\hat{\theta} = \overline{X}_n = \frac{1}{n} \sum_{i=1}^n X_i$  merupakan barisan estimator konsisten untuk  $\theta$  cukup dengan menunjukkan syarat 1 dan 2 diatas dipenuhi.

Karena  $\overline{X}_n \sim N\left(\theta, \frac{1}{n}\right)$ , berarti  $E(\overline{X}_n) = \theta$  atau  $\hat{\theta} = \overline{X}_n = \frac{1}{n} \sum_{i=1}^n X_i$  adalah estimator yang takbias untuk  $\theta$ , sehingga  $bias(\overline{X}_n) = 0$ , dan  $Var(\overline{X}_n) = \frac{1}{n}$ .

Dengan mengambil limit untuk  $n \rightarrow \infty$  maka syarat 1 dan 2 dipenuhi, berarti  $\hat{\theta} = \overline{X}_n = \frac{1}{n} \sum_{i=1}^n X_i$  merupakan barisan estimator yang konsisten.

## 4. Estimator Takbias Terbaik

Dalam beberapa kasus mungkin bisa ditemukan lebih dari satu estimator takbias untuk parameter  $\theta$ . Masalahnya sekarang harus dipilih satu estimator takbias terbaiknya. Sebagai ilustrasi misalkan  $X_1, X_2, \dots, X_n$  variabel random independen berdistribusi Poisson dengan parameter  $\lambda$ . Diambil  $\overline{X}$  dan  $S^2$  masing-masing adalah rata-rata dan variansinya. Telah diperoleh bahwa  $E(\overline{X}) = \lambda$  dan  $E(S^2) = \lambda$ , sehingga keduanya merupakan estimator yang takbias untuk  $\lambda$ . Bila diambil  $\alpha$  suatu konstanta sembarang, dibentuk

$T(\bar{X}, S^2) = \alpha \bar{X} + (1 - \alpha)S^2$ , maka  $E(T(\bar{X}, S^2)) = \lambda$ . Sehingga dapat diperoleh takberhingga banyak estimator takbias untuk  $\lambda$ . Masalahnya sekarang adalah manakah yang merupakan estimator takbias terbaik.

**Definisi 4.1**

Estimator  $T^*(\underline{X})$  disebut estimator takbias terbaik untuk  $g(\theta)$  bila  $E\left(T^*(\underline{X})\right) = g(\theta)$  dan untuk sebarang estimator  $T(\underline{X})$  maka berlaku  $Var\left(T^*(\underline{X})\right) \leq Var\left(T(\underline{X})\right)$  untuk setiap  $\theta$ . Dalam hal ini  $T^*(\underline{X})$  juga disebut estimator takbias variansi minimum seragam atau " a uniform minimum variance unbiased estimator (UMVUE)" dari  $g(\theta)$ .

Salah satu cara menentukan UMVUE dengan melalui batas bawah Cramer Rao sebagai berikut : misalkan dalam melakukan estimasi  $g(\theta)$  dari sampel random suatu populasi dengan densitas  $f(x|\theta)$  dapat ditentukan batas bawah variansi (misalkan  $B(\theta)$ ) dari setiap estimator takbias. Selanjutnya bila dapat ditemukan estimator takbias  $T^*(\underline{X})$  sedemikian sehingga  $var\left(T^*(\underline{X})\right) = B(\theta)$ , maka  $T^*(\underline{X})$  merupakan UMVUE dari  $g(\theta)$ .

**Theorema 4.1 (Cramer Rao)**

Misalkan  $X_1, X_2, \dots, X_n$  sampel random independen dari suatu populasi dengan densitas  $f(x|\theta)$ ,  $T(\underline{X})$  sebarang estimator dengan  $E(T(\underline{X}))$  terdifferensialkan dalam  $\theta$ . Andaikan densitas bersama  $f(\underline{x}|\theta)$  memenuhi :

$$\frac{d}{d\theta} \int \dots \int h(\underline{x}) \cdot f(\underline{x} | \theta) d\underline{x} = \int \dots \int h(\underline{x}) \frac{\partial}{\partial \theta} f(\underline{x} | \theta) d\underline{x} \quad \text{untuk}$$

setiap fungsi  $h(\underline{x})$  dengan  $E(h(\underline{x})) < \infty$ , maka :

$$\text{var}\left(T(\underline{X})\right) \geq \frac{\left(\frac{d}{d\theta} E\left(T(\underline{X})\right)\right)^2}{E\left(\left(\frac{\partial}{\partial \theta} \log f(\underline{x} | \theta)\right)^2\right)}$$

dengan  $E\left(\left(\frac{\partial}{\partial \theta} \log f(\underline{x} | \theta)\right)^2\right)$  disebut informasi Fisher

**Bukti :**

Dipunyai ketaksamaan Schwarz sebagai berikut :

$$[\text{cov}(X, Y)]^2 \leq \text{Var}(X) \cdot \text{Var}(Y)$$

$$\text{Var}(X) \geq \frac{[\text{cov}(X, Y)]^2}{\text{Var}(Y)}$$

dengan mengambil  $X = T(\underline{X})$  dan  $Y = \frac{\partial}{\partial \theta} \log f(\underline{x} | \theta)$

$$\text{cov}(X, Y) = E[(X - E(X))(Y - E(Y))]$$

$$E(Y) = E\left(\frac{\partial}{\partial \theta} \log f(\underline{x} | \theta)\right)$$

$$= E\left[\frac{\frac{\partial}{\partial \theta} f(\underline{x} | \theta)}{f(\underline{x} | \theta)}\right]$$

$$= \int \dots \int \left[\frac{\frac{\partial}{\partial \theta} f(\underline{x} | \theta)}{f(\underline{x} | \theta)}\right] f(\underline{x} | \theta) d\underline{x}$$

$$= \frac{\partial}{\partial \theta} \int \dots \int f(\underline{x} | \theta) d\underline{x}$$

$$= \frac{\partial}{\partial \theta} 1 = 0$$

sehingga:

$$\begin{aligned} \text{Cov}\left(T(\underline{X}), \frac{\partial}{\partial \theta} \log f(\underline{x} | \theta)\right) &= E\left\{T(\underline{X}) \cdot \frac{\partial}{\partial \theta} \log f(\underline{x} | \theta)\right\} - \frac{\partial}{\partial \theta} \log f(\underline{x} | \theta) \cdot E\left(T(\underline{X})\right) - \\ &\quad T(\underline{X}) \cdot E\left(\frac{\partial}{\partial \theta} \log f(\underline{x} | \theta)\right) + E\left(T(\underline{X})\right) \cdot E\left(\frac{\partial}{\partial \theta} \log f(\underline{x} | \theta)\right) \\ &= E\left(T(\underline{X}) \cdot \frac{\partial}{\partial \theta} \log f(\underline{x} | \theta)\right) - E\left(T(\underline{X})\right) \cdot E\left(\frac{\partial}{\partial \theta} \log f(\underline{x} | \theta)\right) \\ &= E\left(T(\underline{X}) \cdot \frac{\partial}{\partial \theta} \log f(\underline{x} | \theta)\right) \\ &= \int \dots \int \left[T(\underline{X}) \cdot \frac{\partial}{\partial \theta} \log f(\underline{x} | \theta)\right] d\underline{x} \\ &= \int \dots \int T(\underline{X}) \frac{\frac{\partial}{\partial \theta} f(\underline{x} | \theta)}{f(\underline{x} | \theta)} d\underline{x} \\ &= \frac{d}{d\theta} \int \dots \int T(\underline{X}) f(\underline{x} | \theta) d\underline{x} \\ &= \frac{d}{d\theta} E\left(T(\underline{X})\right) \end{aligned}$$

$$\begin{aligned} \text{var}\left(\frac{\partial}{\partial \theta} \log f(\underline{x} | \theta)\right) &= E\left(\left(\frac{\partial}{\partial \theta} \log f(\underline{x} | \theta)\right)^2\right) - \left(E\left(\frac{\partial}{\partial \theta} \log f(\underline{x} | \theta)\right)\right)^2 \\ &= E\left(\left(\frac{\partial}{\partial \theta} \log f(\underline{x} | \theta)\right)^2\right) \end{aligned}$$

$$\text{sehingga: } \text{var}\left(T(\underline{X})\right) \geq \frac{\left(\frac{d}{d\theta} E\left(T(\underline{X})\right)\right)^2}{E\left(\left(\frac{\partial}{\partial \theta} \log f(\underline{x} | \theta)\right)^2\right)}$$

#### Contoh 4.1

Misalkan  $X_1, X_2, \dots, X_n$  adalah sampel random independen dari populasi Poisson dengan parameter  $\lambda$ . Bila diambil  $g(\lambda) = \lambda$ , tentukan UMVUE untuk  $g(\lambda) = \lambda$ .

**Penyelesaian :**

Akan dicari  $T(\underline{X})$  yang merupakan UMVUE untuk  $g(\lambda) = \lambda$ .

Berarti  $E(T(\underline{X})) = \lambda$ , sehingga  $\frac{d}{d\lambda} E(T(\underline{X})) = 1$

$$f(x_i | \lambda) = e^{-\lambda} \cdot \frac{\lambda^{x_i}}{x_i!}$$

$$f(\underline{x} | \lambda) = f(x_1 | \lambda) \cdot f(x_2 | \lambda) \dots f(x_n | \lambda)$$

$$= e^{-\lambda} \cdot \frac{\lambda^{x_1}}{x_1!} \cdot e^{-\lambda} \cdot \frac{\lambda^{x_2}}{x_2!} \dots e^{-\lambda} \cdot \frac{\lambda^{x_n}}{x_n!}$$

$$= \frac{1}{x_1! \cdot x_2! \dots x_n!} e^{-n\lambda} \cdot \lambda^{\sum_{i=1}^n x_i}$$

$$\log(f(\underline{x} | \lambda)) = -\sum_{i=1}^n \log x_i - n\lambda + \sum_{i=1}^n x_i \cdot \log \lambda$$

$$\frac{\partial}{\partial \lambda} \log(f(\underline{x} | \lambda)) = -n + \frac{1}{\lambda} \sum_{i=1}^n x_i$$

$$\left( \frac{\partial}{\partial \lambda} \log(f(\underline{x} | \lambda)) \right)^2 = n^2 - \frac{2n}{\lambda} \sum_{i=1}^n x_i + \frac{1}{\lambda^2} \left( \sum_{i=1}^n x_i \right)^2$$

$$E\left( \left( \frac{\partial}{\partial \lambda} \log(f(\underline{x} | \lambda)) \right)^2 \right) = n^2 - \frac{2n}{\lambda} E\left( \sum_{i=1}^n x_i \right) + \frac{1}{\lambda^2} E\left( \left( \sum_{i=1}^n x_i \right)^2 \right)$$

$X_i \sim P(\lambda)$  maka fungsi pembangkit momemnya adalah :

$$M_{X_i}(t) = \exp(\lambda(e^t - 1))$$

$$M_{\sum_{i=1}^n X_i}(t) = M_{X_1}(t) \cdot M_{X_2}(t) \dots M_{X_n}(t)$$

$$= \exp(\lambda(e^t - 1)) \cdot \exp(\lambda(e^t - 1)) \dots \exp(\lambda(e^t - 1))$$

$$= \exp(n\lambda(e^t - 1))$$

Maka :

$$\sum_{i=1}^n X_i \sim P(n\lambda)$$

$$E\left(\sum_{i=1}^n X_i\right) = n\lambda$$

$$\text{var}\left(\sum_{i=1}^n X_i\right) = n\lambda$$

$$E\left(\left(\sum_{i=1}^n X_i\right)^2\right) - \left(E\left(\sum_{i=1}^n X_i\right)\right)^2 = n\lambda$$

$$E\left(\left(\sum_{i=1}^n X_i\right)^2\right) = n\lambda + (n\lambda)^2$$

$$E\left[\left(\frac{\partial}{\partial \lambda} \log f\left(\underset{\sim}{x}|\lambda\right)\right)^2\right] = n^2 - \frac{2n}{\lambda} \cdot n\lambda + \frac{1}{\lambda^2} \cdot (n\lambda + (n\lambda)^2)$$

$$= \frac{n}{\lambda}$$

Sehingga :

$$\text{var}\left(T\left(\underset{\sim}{X}\right)\right) \geq \frac{\left(\frac{d}{d\lambda} E\left(T\left(\underset{\sim}{X}\right)\right)\right)^2}{E\left[\left(\frac{\partial}{\partial \lambda} \log f\left(\underset{\sim}{x}|\lambda\right)\right)^2\right]}$$

$$\text{var}\left(T\left(\underset{\sim}{X}\right)\right) \geq \frac{1}{\frac{n}{\lambda}} = \frac{\lambda}{n}$$

Ambil :  $T\left(\underset{\sim}{X}\right) = \frac{1}{n} \sum_{i=1}^n X_i$ , maka :

$$\text{Var}\left(\frac{1}{n} \sum_{i=1}^n X_i\right) = \frac{1}{n^2} \text{var}\left(\sum_{i=1}^n X_i\right)$$

$$= \frac{1}{n^2} \cdot n\lambda = \frac{\lambda}{n}$$

Sehingga  $\text{Var}(T(\underline{X}))$  dapat mencapai batas bawah Cramer Rao, jadi dapat disimpulkan  $T(\underline{X}) = \frac{1}{n} \sum_{i=1}^n X_i$  adalah UMVUE dari  $g(\lambda) = \lambda$ .

**Akibat 4.1**

Misalkan  $X_1, X_2, \dots, X_n$  sampel random independen dari suatu populasi dengan densitas bersama  $f(\underline{x} | \theta)$ , dan misalkan  $T(\underline{X})$  sebarang estimator dengan  $E(T(\underline{X}))$  terdifferensialkan dalam  $\theta$ . Bila densitas bersama  $f(\underline{x} | \theta) = \prod_{i=1}^n f(x_i | \theta)$  memenuhi syarat theorema Cramer Rao maka:

$$\text{var}\left(T(\underline{X})\right) \geq \frac{\left(\frac{d}{d\theta} E\left(T(\underline{X})\right)\right)^2}{n \cdot E\left(\left(\frac{\partial}{\partial \theta} \log f(x|\theta)\right)^2\right)}$$

**Contoh 4.2**

Misalkan  $X_1, X_2, \dots, X_n$  sampel random independen dari suatu populasi berdistribusi Poisson dengan parameter  $\lambda$ . Bila  $g(\lambda) = \lambda$ , tentukan UMVUE dari  $g(\lambda) = \lambda$ .

**Penyelesaian:**

Akan dicari  $T(\underline{X})$  yang merupakan UMVUE dari  $g(\lambda) = \lambda$ ,

Berarti  $E(T(\underline{X})) = \lambda$

Sehingga ;  $\frac{d}{d\lambda} E\left(T(\underline{X})\right) = 1$

$$f(x|\lambda) = e^{-\lambda} \frac{\lambda^x}{x!}$$

$$\log f(x|\lambda) = -\log x! + x \log \lambda - \lambda$$

$$\frac{\partial}{\partial \lambda} \log f(x|\lambda) = \frac{x}{\lambda} - 1$$

$$\begin{aligned} E\left(\left(\frac{\partial}{\partial \lambda} \log f(x|\lambda)\right)^2\right) &= n \cdot E\left(\left(\frac{\partial}{\partial \lambda} \log f(x|\lambda)\right)^2\right) \\ &= n \cdot E\left(\left(\frac{x}{\lambda} - 1\right)^2\right) \\ &= n \cdot E\left(\frac{x^2}{\lambda^2} - 2 \cdot \frac{x}{\lambda} + 1\right) \\ &= n \left(\frac{\lambda + \lambda^2}{\lambda^2} - 2 \cdot \frac{\lambda}{\lambda} + 1\right) \\ &= \frac{n}{\lambda} \end{aligned}$$

Sehingga :

$$\text{var}\left(T\left(\underline{X}\right)\right) \geq \frac{\left(\frac{d}{d\lambda} E\left(T\left(\underline{X}\right)\right)\right)^2}{n \cdot E\left(\left(\frac{\partial}{\partial \lambda} \log f(x|\lambda)\right)^2\right)}$$

$$\text{var}\left(T\left(\underline{X}\right)\right) \geq \frac{1}{\frac{n}{\lambda}} = \frac{\lambda}{n}$$

Ambil  $T\left(\underline{X}\right) = \frac{1}{n} \sum_{i=1}^n X_i$ , maka :

$$\begin{aligned} \text{var}\left(T\left(\underline{X}\right)\right) &= \text{var}\left(\frac{1}{n} \sum_{i=1}^n X_i\right) \\ &= \frac{1}{n^2} \sum_{i=1}^n \text{var}\left(X_i\right) \\ &= \frac{1}{n^2} \sum_{i=1}^n \lambda \\ &= \frac{1}{n^2} \cdot n \cdot \lambda = \frac{\lambda}{n} \end{aligned}$$

Jadi  $\text{var}\left(T\left(\underline{X}\right) = \frac{1}{n} \sum_{i=1}^n X_i\right)$  dapat mencapai batas bawah

Cramer Rao, jadi dapat disimpulkan bahwa  $T\left(\underset{\sim}{X}\right) = \frac{1}{n} \sum_{i=1}^n X_i$  adalah UMVUE dari  $g(\lambda) = \lambda$

Sangat penting untuk kita ingat bahwa andaikan kunci dalam theorema Cramer Rao adalah kemampuan untuk mendifferensialkan dibawah tanda integral yang dengan sendirinya agak terbatas penerapannya. Untuk densitas dalam keluarga eksponensial satu parameter memenuhi persyaratan theorema Cramer Rao dan bila estimator takbias dari  $g(\theta)$  ada, maka pasti akan ada estimator takbias yang mencapai batas bawah Cramer Rao.

**Akibat 4.2**

Misalkan  $X_1, X_2, \dots, X_n$  sampel random independen dari suatu populasi dengan densitas  $f(x|\theta)$ , yang memenuhi syarat theorema Cramer Rao.

Misalkan  $L\left(\theta | \underset{\sim}{x}\right) = \prod_{i=1}^n f(x_i | \theta)$  menyatakan fungsi likelihood. Bila

$T\left(\underset{\sim}{X}\right)$  sebarang estimator takbias dari  $g(\theta)$ , maka  $T\left(\underset{\sim}{X}\right)$  mencapai batas bawah Cramer Rao (Cramer Rao Lower Bound / CRLB) bila dan hanya bila :

$$\frac{\partial}{\partial \theta} \log L\left(\theta | \underset{\sim}{X}\right) = a(\theta) \left[ T\left(\underset{\sim}{X}\right) - g(\theta) \right]$$

untuk suatu fungsi  $g(\theta)$ .

**Contoh 4.3**

Misalkan  $X_1, X_2, \dots, X_n$  sampel random independen dari suatu populasi berdistribusi Gamma(2,  $\theta$ ). Tentukan UMVUE untuk  $g(\theta) = \theta$

**Penyelesaian :**

$X_i \sim \text{Gamma}(2, \theta)$  maka densitasnya adalah :

$$f(x_i | \theta) = \theta^{-2} \cdot x_i \cdot \exp\left(-\frac{x_i}{\theta}\right) \quad x_i > 0 \quad \theta > 0$$

$$\begin{aligned} L(\theta | \underline{x}) &= f(x_1 | \theta) \cdot f(x_2 | \theta) \cdots f(x_n | \theta) \\ &= \theta^{-2} \cdot x_1 \cdot \exp\left(-\frac{x_1}{\theta}\right) \cdot \theta^{-2} \cdot x_2 \cdot \exp\left(-\frac{x_2}{\theta}\right) \cdots \theta^{-2} \cdot x_n \cdot \exp\left(-\frac{x_n}{\theta}\right) \\ &= \theta^{-2n} (x_1 \cdot x_2 \cdots x_n) \cdot \exp\left(-\frac{1}{\theta} \sum_{i=1}^n x_i\right) \end{aligned}$$

$$\log\left(L(\theta | \underline{x})\right) = -2n \log \theta + \sum_{i=1}^n \log x_i - \frac{1}{\theta} \sum_{i=1}^n x_i$$

$$\begin{aligned} \frac{\partial}{\partial \theta} \log\left(L(\theta | \underline{x})\right) &= -\frac{2n}{\theta} + \frac{1}{\theta^2} \sum_{i=1}^n x_i \\ &= \frac{2n}{\theta^2} \left[ \frac{\sum_{i=1}^n x_i}{2n} - \theta \right] \end{aligned}$$

Dari bentuk terakhir dan dengan menggunakan akibat 1.5.2 diperoleh :

$$a(\theta) = \frac{2n}{\theta^2} \quad \text{dan} \quad T(\underline{X}) = \frac{\sum_{i=1}^n X_i}{2n}$$

Apabila  $T(\underline{X})$  merupakan estimator takbias, maka  $T(\underline{X})$  merupakan UMVUE untuk  $\theta$ .

$$\begin{aligned} E(X_i) &= \int_0^{\infty} x_i \cdot \theta^{-2} \cdot x_i \cdot \exp\left(-\frac{x_i}{\theta}\right) dx_i \\ &= \theta \int_0^{\infty} \left(\frac{x_i}{\theta}\right)^2 \cdot \exp\left(-\frac{x_i}{\theta}\right) d\frac{x_i}{\theta} \\ &= \theta \cdot \Gamma(3) = 2\theta \end{aligned}$$

$$E\left(T(\underline{X})\right) = E\left(\frac{1}{2n} \sum_{i=1}^n x_i\right) = \frac{1}{2n} \cdot \sum_{i=1}^n E(x_i) = \frac{1}{2n} \sum_{i=1}^n 2\theta = \frac{1}{2n} \cdot n\theta = \theta$$

Berarti  $T(\underline{X})$  merupakan estimator takbias untuk  $\theta$ , sehingga merupakan UMVUE untuk  $\theta$ .

#### 1.4.2.2 Latihan

1. Apa hubungannya definisi sesatan kuadrat rata-rata dengan sifat barisan estimator titik yang konsisten?
2. Apakah setiap estimator takbias selalu dapat mencapai batas bawah Crammer-Rao?
3. Apakah estimator takbias terbaik dari setiap parameter selalu ada? Mengapa?

#### 1.4.2.3 Rangkuman

1. Sifat estimator takbias sangat penting dalam hal akan menghitung sesatan kuadrat rata-rata, demikian juga untuk menentukan estimator takbias terbaik.
2. Membuktikan sifat barisan estimator konsisten lebih mudah melalui sesatan kuadrat rata-rata. Bila limit untuk  $n \rightarrow \infty$  dari bias dan sesatan kuadrat rata-rata dari estimator tersebut adalah nol, maka barisan estimator tersebut bersifat konsisten.
3. Menentukan estimator takbias terbaik dapat melalui batas bawah Crammer-Rao atau melalui akibat 4.2.

#### 1.4.2.4 Tes Formatif 2

1. Misalkan  $X_1, X_2, \dots, X_n$  adalah sampel random independen dari populasi berdistribusi Gamma(7,  $\theta$ ).
  - a. Tentukan estimator titik untuk  $\theta$  dengan metode likelihood maksimum.
  - b. Apakah estimator dalam a) bersifat takbias? Buktikan.
  - c. Hitunglah sesatan kuadrat rata-rata dari estimator dalam a).
  - d. Apakah estimator dalam a) merupakan barisan estimator yang konsisten? Buktikan.
  - e. Tentukan UMVUE untuk  $\theta$ .
2. Misalkan  $X_1, X_2, \dots, X_n$  adalah sampel random independen dari populasi berdistribusi Exponensial dengan rata-rata  $\theta$ .

- a. Tentukan estimator titik untuk  $\theta$  dengan metode likelihood maksimum.
- b. Apakah estimator dalam a) bersifat takbias? Buktikan.
- c. Hitunglah sesatan kuadrat rata-rata dari estimator dalam a).
- d. Apakah estimator dalam a) merupakan barisan estimator yang konsisten? Buktikan.
- e. Tentukan UMVUE untuk  $\theta$ .

#### 1.4.2.5 Umpan Balik dan tidak Lanjut

Cocokkanlah jawaban anda dengan kunci jawaban dari modul ini. Hitunglah jawaban anda yang benar, kemudian gunakanlah rumus dibawah ini untuk mengetahui tingkat penguasaan anda terhadap materi ini.

$$\text{Tingkat penguasaan} = (\text{jumlah jawaban benar}) / 10 \times 100\%$$

Arti tingkat penguasaan yang anda capai :

90% - 100% : baik sekali

80% - 89% : baik

70% - 79% : sedang

< 70% : kurang

Bila anda telah mencapai tingkat penguasaan 80% atau lebih, anda dapat meneruskan kegiatan belajar selanjutnya. Tetapi apabila tingkat penguasaan anda kurang dari 80% , sebaiknya anda mengulangi materi ini terutama bagian yang belum anda kuasai sebelum meneruskan kegiatan belajar selanjutnya.

### 1.5 Kunci Jawaban Tes Formatif

#### Tes Formatif 1

$$1. \frac{\bar{X}}{1-\bar{X}} \quad 2. \frac{1}{\bar{X}-1} \quad 3. \frac{2}{\bar{X}} \quad 4. \frac{\bar{X}}{2} \quad 5. \bar{X}$$

$$6. \frac{-n}{\sum_{i=1}^n \log X_i} \quad 7. 1 - \frac{n}{\sum_{i=1}^n \log X_i} \quad 8. \frac{2}{\bar{X}} \quad 9. \frac{\bar{X}}{2} \quad 10. \bar{X}$$

#### Tes Formatif 2

$$1. a. \frac{\bar{X}}{7} \quad b. ya \quad c. \frac{\theta^2}{7n} \quad d. ya \quad e. \frac{\bar{X}}{7}$$

$$2. a. \bar{X} \quad b. ya \quad c. \frac{\theta^2}{n} \quad d. ya \quad e. \bar{X}$$

### 1.6 Referensi

1. Bain, L.J. and Engelhart, M. (1992); Introduction to Promodulility and Mathematical Statistics, Second Edition. Duxbury Press, Belmont, California.
2. Casella, G and Berger, R.L. (1990); Statistical Inference. Wadsworth Inc. Belmont, California.
3. Dudewicz, E.J and Mishra, S.N. (1998); Modern Mathematical Statistics. John Willey and Sons, Singapore.
4. Roussas, G.G. (1976); A First Course in Mathematical Statistics. Mei Ya Publication Inc. Taipei, Taiwan.
5. ....; (2007). Estimation Theory. "[http://en.wikipedia.org/wiki/Point\\_estimation](http://en.wikipedia.org/wiki/Point_estimation)" This page was last modified 00:26, 29 June 2007

## MODUL II KECUKUPAN DAN KELENGKAPAN

### 2.1. Pengantar

Dalam modul ini akan dipelajari cara penentuan statistik cukup  $T(\underline{X})$  untuk parameter  $\theta$ , menyelidiki sifat kelengkapan suatu statistik  $T(\underline{X})$  dan hubungan antara sifat kecukupan, kelengkapan dan ketakbiasan. Selanjutnya juga akan dipelajari penentuan estimator takbias terbaik melalui statistik cukup, lengkap dan takbias. Cara ini dilakukan bila penentuan batas bawah Cramer\_Rao menemui kesulitan.

### 2.2. Tujuan Instruksional Umum

Setelah mengikuti mempelajari modul ini mahasiswa dapat :

1. Mendefinisikan statistik cukup dan cara penentuan statistik cukup.
2. Membuktikan sifat lengkap suatu statistik cukup melalui definisi statistik lengkap.
3. Menentukan estimator takbias terbaik dari suatu parameter.

### 2.3. Tujuan Instruksional Khusus.

Setelah mengikuti mempelajari modul ini mahasiswa dapat :

1. Membuktikan suatu statistik merupakan statistik cukup melalui definisi statistik cukup.
2. Menentukan statistik cukup dengan menggunakan teorema faktorisasi Neyman-Pearson.
3. Membuktikan suatu keluarga distribusi merupakan anggota keluarga eksponensial.
4. Menentukan statistik cukup melalui keluarga eksponensial.
5. Membuktikan suatu statistik cukup bersifat lengkap melalui definisi statistik lengkap.

6. Menentukan estimator takbias terbaik dari suatu parameter dengan menggunakan sifat cukup lengkap dan takbias.

#### 2.4. Kegiatan Belajar :

### KECUKUPAN DAN KELENGKAPAN

#### 2.4.1. Uraian dan Contoh

Informasi dalam sampel  $\underline{X} = (X_1, X_2, \dots, X_n)$  akan digunakan untuk melakukan inferensi tentang  $\theta$ . Data terobservasi  $\underline{x} = (x_1, x_2, \dots, x_n)$  merupakan daftar bilangan yang sukar untuk diinterpretasikan. Sehingga akan dilakukan pendugaan melalui penghitungan statistik  $T(\underline{X})$  yang merupakan fungsi dari sampel. Arti *statistic* sendiri adalah *penduga parameter*, sehingga ilmu yang mempelajari pendugaan parameter dikenal dengan *statistika*. Apabila distribusi dari sampel diketahui maka untuk keperluan inferensi akan ditentukan statistic terbaik, yaitu statistic yang mempunyai sifat-sifat baik (takbias, variansi minimum, atau sifat lain sesuai inferensi yang dilakukan).

#### 1. Statistik Cukup.

Statistik cukup untuk parameter  $\theta$  adalah statistic dalam arti tertentu dapat menyerap semua informasi tentang  $\theta$  yang termuat dalam sampel. Bila  $T(\underline{X})$  adalah statistik cukup untuk  $\theta$  maka setiap inferensi tentang  $\theta$  harus tergantung pada sampel  $\underline{X} = (X_1, X_2, \dots, X_n)$  hanya melalui  $T(\underline{X})$ . Adapun definisi dari statistik cukup sebagai berikut :

**Definisi 1.1.**

Statistik  $T(\underline{X})$  disebut statistik cukup untuk  $\theta$  bila distribusi bersyarat sampel  $\underline{X} = (X_1, X_2, \dots, X_n)$  diberikan harga  $T(\underline{X})$  tidak tergantung  $\theta$ .

Maksud dari definisi tersebut dapat dijelaskan sebagai berikut : bila  $f(\underline{x}|\theta)$  adalah densitas dari  $\underline{X} = (X_1, X_2, \dots, X_n)$  dan  $g(T(\underline{x})|\theta)$  adalah densitas dari  $T(\underline{X})$ , maka  $T(\underline{X})$  adalah statistik cukup untuk  $\theta$  bila untuk setiap

$\underline{X} = (X_1, X_2, \dots, X_n)$  dalam ruang sampel, rasio :  $\frac{f(\underline{x}|\theta)}{g(T(\underline{x})|\theta)}$  tidak

bergantung pada  $\theta$ .

**Contoh 1.1.**

Misalkan  $X_1, X_2, \dots, X_n$  variabel random independent berdistribusi Bernoulli dengan parameter  $\theta$ ,  $0 < \theta < 1$ . Akan ditunjukkan bahwa :

$T(\underline{X}) = X_1 + X_2 + \dots + X_n = \sum_{i=1}^n X_i$  adalah statistik cukup untuk  $\theta$ .

**Penyelesaian :**

$X_i \sim B(1, \theta)$  densitas dari  $X_i$  adalah :  $f(x_i|\theta) = \theta^{x_i}(1-\theta)^{1-x_i}$  dengan  $i=0,1$

$$\begin{aligned} f(\underline{x}|\theta) &= f(x_1|\theta) \cdot f(x_2|\theta) \cdot \dots \cdot f(x_n|\theta) \\ &= \theta^{x_1}(1-\theta)^{1-x_1} \cdot \theta^{x_2}(1-\theta)^{1-x_2} \cdot \dots \cdot \theta^{x_n}(1-\theta)^{1-x_n} \\ &= \theta^{\sum_{i=1}^n x_i} (1-\theta)^{n-\sum_{i=1}^n x_i} \end{aligned}$$

Fungsi pembangkit momen dari  $X_i$  adalah  $M_{X_i}(t) = \theta e^t + (1-\theta)$

$$M_{\sum_{i=1}^n X_i}(t) = M_{X_1}(t) \cdot M_{X_2}(t) \cdot \dots \cdot M_{X_n}(t)$$

$$\begin{aligned}
 &= \theta e^{t + (1-\theta)} \cdot \theta e^{t + (1-\theta)} \dots \theta e^{t + (1-\theta)} \\
 &= \left( \theta e^{t + (1-\theta)} \right)^n \text{ (yang merupakan fungsi pembangkit momen dari} \\
 &\quad \text{distribusi Binomial (n, } \theta))
 \end{aligned}$$

Sehingga :  $T\left(\underset{\sim}{X}\right) = \sum_{i=1}^n X_i$  berdistribusi Binomial (n,  $\theta$ )

$$g(T(\underset{\sim}{x}) | \theta) = \binom{n}{\sum_{i=1}^n x_i} \cdot \theta^{\sum_{i=1}^n x_i} (1-\theta)^{n - \sum_{i=1}^n x_i}$$

$$\text{Maka : } \frac{f(\underset{\sim}{x} | \theta)}{g(T(\underset{\sim}{x}) | \theta)} = \frac{\theta^{\sum_{i=1}^n x_i} \cdot (1-\theta)^{n - \sum_{i=1}^n x_i}}{\binom{n}{\sum_{i=1}^n x_i} \cdot \theta^{\sum_{i=1}^n x_i} \cdot (1-\theta)^{n - \sum_{i=1}^n x_i}} = \frac{1}{\binom{n}{\sum_{i=1}^n x_i}}$$

yang tidak bergantung pada  $\theta$

Jadi  $T\left(\underset{\sim}{X}\right) = \sum_{i=1}^n X_i$  adalah statistik cukup untuk  $\theta$ .

**Contoh 1.2.**

Misalkan  $X_1, X_2, \dots, X_n$  variabel random independent berdistribusi normal dengan rata-rata  $\mu$  dan variansi 9. Akan ditunjukkan bahwa rata-rata sampel :

$$T\left(\underset{\sim}{X}\right) = \bar{X} = \frac{\sum_{i=1}^n X_i}{n} \text{ adalah statistic cukup untuk } \mu.$$

**Penyelesaian.**

$X_i \sim N(\mu, 9)$  maka densitas dari  $X_i$  adalah :

$$f(x_i | \mu) = (2\pi)^{-1/2} 9^{-1/2} \exp\left(-\frac{1}{2 \cdot 9} (x_i - \mu)^2\right)$$

$$f(\underset{\sim}{x} | \mu) = f(x_1 | \mu) f(x_2 | \mu) \dots f(x_n | \mu) =$$

$$\begin{aligned}
 &(2\pi)^{-1/2} 9^{-1/2} \exp\left(-\frac{1}{2 \cdot 9} (x_1 - \mu)^2\right) \cdot (2\pi)^{-1/2} 9^{-1/2} \exp\left(-\frac{1}{2 \cdot 9} (x_2 - \mu)^2\right) \dots \\
 &(2\pi)^{-1/2} 9^{-1/2} \exp\left(-\frac{1}{2 \cdot 9} (x_n - \mu)^2\right)
 \end{aligned}$$

$$= (2\pi)^{-n/2} \cdot 9^{-n/2} \exp\left(-\frac{1}{18} \sum_{i=1}^n (x_i - \mu)^2\right)$$

Selanjutnya akan ditentukan densitas dari  $T(\underline{X})$  yaitu  $g(T(\underline{X}) | \mu)$

Fungsi pembangkit momen untuk  $X_i$  adalah :  $M_{X_i}(t) = \exp\left(\mu t + \frac{1}{2} \cdot 9 \cdot t^2\right)$

$$M_{\sum_{i=1}^n X_i}(t) = M_{X_1}(t) \cdot M_{X_2}(t) \cdot \dots \cdot M_{X_n}(t)$$

$$= \exp\left(\mu t + \frac{1}{2} \cdot 9 \cdot t^2\right) \cdot \exp\left(\mu t + \frac{1}{2} \cdot 9 \cdot t^2\right) \cdot \dots \cdot \exp\left(\mu t + \frac{1}{2} \cdot 9 \cdot t^2\right)$$

$$= \exp\left(n \cdot \mu t + n \cdot \frac{1}{2} \cdot 9 \cdot t^2\right)$$

$$M_{\sum_{i=1}^n X_i/n}(t) = \exp\left(n \cdot \mu \cdot \frac{t}{n} + n \cdot \frac{1}{2} \cdot 9 \cdot \left(\frac{t}{n}\right)^2\right) = \exp\left(\mu t + \frac{1}{2} \cdot \frac{9}{n} \cdot t^2\right) \quad \text{yang merupakan}$$

fungsi pembangkit momen untuk distribusi normal dengan rata-rata  $\mu$  dan

variansi  $9/n$ . Sehingga :  $\bar{X} = \frac{\sum_{i=1}^n X_i}{n}$  berdistribusi normal dengan rata-rata  $\mu$  dan variansi  $9/n$

$$g(\bar{X} | \mu) = (2\pi)^{-1/2} \cdot \left(\frac{9}{n}\right)^{-1/2} \cdot \exp\left(-\frac{1}{2} \cdot \frac{n}{9} (\bar{x} - \mu)^2\right)$$

$$= (2\pi)^{-1/2} \cdot \left(\frac{9}{n}\right)^{-1/2} \cdot \exp\left(-\frac{1}{2 \cdot 9} \cdot n \cdot (\bar{x} - \mu)^2\right)$$

$$\begin{aligned} \frac{f(\underline{x} | \mu)}{g(T(\underline{X}) | \mu)} &= \frac{(2\pi)^{-n/2} \cdot 9^{-n/2} \cdot \exp\left(-\frac{1}{2 \cdot 9} \sum_{i=1}^n (x_i - \mu)^2\right)}{(2\pi)^{-1/2} \cdot \left(\frac{9}{n}\right)^{-1/2} \cdot \exp\left(-\frac{1}{2 \cdot 9} \cdot n \cdot (\bar{x} - \mu)^2\right)} \\ &= \frac{(2\pi)^{-(n-1)/2} \cdot 9^{-(n-1)/2} \cdot \exp\left(-\frac{1}{2 \cdot 9} \cdot \sum_{i=1}^n \left((x_i - \bar{x}) - (\bar{x} - \mu)\right)^2\right)}{n^{1/2} \cdot \exp\left(-\frac{1}{2 \cdot 9} \cdot n \cdot (\bar{x} - \mu)^2\right)} \end{aligned}$$

$$\begin{aligned}
 &= \frac{(2\pi)^{-(n-1)/2} \cdot 9^{-(n-1)/2} \cdot \exp\left(-\frac{1}{2 \cdot 9} \left(\sum_{i=1}^n (x_i - \bar{x})^2 + n \cdot (\bar{x} - \mu)^2\right)\right)}{n^{1/2} \cdot \exp\left(-\frac{1}{2 \cdot 9} \cdot n \cdot (\bar{x} - \mu)^2\right)} \\
 &= n^{-1/2} (2\pi)^{-(n-1)/2} 9^{-(n-1)/2} \cdot \exp\left(-\frac{1}{2 \cdot 9} \sum_{i=1}^n (x_i - \bar{x})^2\right)
 \end{aligned}$$

Bentuk terakhir ini tidak mengandung  $\mu$  sehingga  $T(\underline{X}) = \bar{X} = \frac{\sum_{i=1}^n X_i}{n}$  statistic cukup untuk  $\mu$ .

Dengan menggunakan definisi 2.2.1. menentukan statistik cukup merupakan hal yang sulit, karena harus menduga bentuk  $T(\underline{X})$  yang kemudian harus dibuktikan bahwa  $T(\underline{X})$  tersebut benar-benar merupakan statistik cukup. Cara yang lebih mudah untuk menentukan statistik cukup dengan menggunakan teorema faktorisasi Neyman berikut.

**Teorema 1.1.**

Misalkan  $f(\underline{x}|\theta)$  adalah densitas bersama dari sampel

$\underline{X} = (X_1, X_2, \dots, X_n)$ ,  $T(\underline{X})$  adalah statistik cukup untuk  $\theta$  bila dan

hanya bila terdapat fungsi  $g(T(\underline{x})|\theta)$  dan  $h(\underline{x})$  sedemikian sehingga untuk semua titik sampel  $\underline{X} = (X_1, X_2, \dots, X_n)$  dan semua titik

parameter  $\theta$  berlaku :  $f(\underline{x}|\theta) = h(\underline{x}) \cdot g(T(\underline{x})|\theta)$ , dengan  $h(\underline{x})$

adalah fungsi dari  $\underline{x}$  yang tidak tergantung  $\theta$  dan  $g(T(\underline{x})|\theta)$

adalah fungsi yang tergantung  $\theta$  hanya melalui  $T(\underline{X})$ .

**Bukti :**

$\Rightarrow$  Misalkan  $T(\underline{X})$  statistic cukup untuk  $\theta$ , akan dibuktikan berlaku :

$$f(\underline{x}|\theta) = h(\underline{x}) \cdot g(T(\underline{x})|\theta)$$

Ambil :

$$h(\underline{x}) = P(\underline{X} = \underline{x} | T(\underline{X}) = t)$$

$$g(T(\underline{x})|\theta) = P(T(\underline{X}) = t)$$

Sehingga :

$$\begin{aligned} f(\underline{x}|\theta) &= P(\underline{X} = \underline{x}) \\ &= P(\underline{X} = \underline{x} \text{ dan } T(\underline{X}) = t) \text{ karena } \underline{X} = \underline{x} \subset T(\underline{X}) = t \\ &= P(T(\underline{X}) = t) P(\underline{X} = \underline{x} | T(\underline{X}) = t) \\ &= g(T(\underline{x})|\theta) \cdot h(\underline{x}) \end{aligned}$$

$\Leftarrow$  Bila  $f(\underline{x}|\theta) = h(\underline{x}) \cdot g(T(\underline{x})|\theta)$  akan dibuktikan  $T(\underline{X})$  statistik cukup untuk  $\theta$ .

Andaikan :  $q(T(\underline{x})|\theta)$  adalah densitas dari  $T(\underline{X})$ , akan ditunjukkan bahwa

$$: \frac{f(\underline{x}|\theta)}{q(T(\underline{x})|\theta)} \text{ tidak tergantung pada } \theta$$

Didefinisikan :  $A_{T(\underline{x})} = \{ \underline{Y} : T(\underline{Y}) = T(\underline{X}) \}$

$$\begin{aligned} \frac{f(\underline{x}|\theta)}{q(T(\underline{x})|\theta)} &= \frac{h(\underline{x}) \cdot g(T(\underline{x})|\theta)}{q(T(\underline{x})|\theta)} \\ &= \frac{h(\underline{x}) \cdot g(T(\underline{x})|\theta)}{\sum_{A_T(\underline{x})} h(\underline{y}) \cdot g(T(\underline{y})|\theta)} \\ &= \frac{h(\underline{x}) \cdot g(T(\underline{x})|\theta)}{g(T(\underline{x})|\theta) \sum_{A_T(\underline{x})} h(\underline{y})} \\ &= \frac{h(\underline{x})}{\sum_{A_T(\underline{x})} h(\underline{y})} \end{aligned}$$

Terlihat bahwa  $\frac{f(\underline{x}|\theta)}{q(T(\underline{x})|\theta)}$  tidak tergantung pada  $\theta$ , sehingga  $T(\underline{X})$

adalah statistic cukup untuk  $\theta$

### Contoh 1.3.

Misalkan  $X_1, X_2, \dots, X_n$  variabel random independent berdistribusi Poisson dengan parameter  $\lambda$ . Tentukan statistik cukup untuk  $\lambda$ .

**Penyelesaian:**

$X_i \sim P(\lambda)$  maka densitasnya adalah:  $f(x_i|\lambda) = e^{-\lambda} \cdot \frac{\lambda^{x_i}}{x_i!}$  dengan  $x_i = 0, 1, 2, 3, \dots$

$$f(\underline{x}|\lambda) = f(x_1|\lambda) \cdot f(x_2|\lambda) \cdot \dots \cdot f(x_n|\lambda)$$

$$= e^{-\lambda} \frac{\lambda^{x_1}}{x_1!} \cdot e^{-\lambda} \frac{\lambda^{x_2}}{x_2!} \cdots e^{-\lambda} \frac{\lambda^{x_n}}{x_n!} = \frac{1}{x_1! \cdot x_2! \cdots x_n!} e^{-n\lambda} \lambda^{\sum_{i=1}^n x_i}$$

Ambil:  $h(\underline{x}) = \frac{1}{x_1! \cdot x_2! \cdots x_n!}$

$$g\left(T(\underline{X}) \mid \lambda\right) = e^{-n\lambda} \cdot \lambda^{\sum_{i=1}^n x_i}$$

Sehingga:  $T(\underline{X}) = \sum_{i=1}^n X_i$  adalah statistik cukup  $\lambda$ .

Untuk keluarga eksponensial penentuan statistik cukup dengan menggunakan teorema faktorisasi Neyman sangat mudah dilakukan. Untuk itu akan dibahas dahulu mengenai keluarga eksponensial.

**Definisi 1.2.**

Keluarga densitas disebut anggota keluarga eksponensial k parameter bila densitas tersebut dapat dinyatakan dalam bentuk :

$$f(\underline{x} \mid \underline{\theta}) = h(\underline{x}) \cdot c(\underline{\theta}) \cdot \exp\left(\sum_{i=1}^k w_i(\underline{\theta}) \cdot t_i(\underline{x})\right),$$

dengan:  $h(\underline{x})$  fungsi nonnegatif dari  $\underline{x}$

$t_i(\underline{x})$  fungsi berharga nyata dari  $\underline{x}$

$c(\underline{\theta})$  fungsi nonnegatif dari  $\underline{\theta}$

$w_i(\underline{\theta})$  fungsi berharga nyata dari  $\underline{\theta}$

**Contoh 1.4**

Misalkan  $X_1, X_2, \dots, X_n$  variabel random independent berdistribusi Poisson dengan parameter  $\lambda$ . Selidikilah apakah merupakan anggota keluarga eksponensial satu parameter.

**Penyelesaian :**

$X_i \sim P(\lambda)$  maka densitasnya adalah :  $f(x_i | \lambda) = e^{-\lambda} \cdot \frac{\lambda^{x_i}}{x_i!}$  dengan  $x_i = 0, 1, 2, 3, \dots$

$$\begin{aligned} f(\underline{x} | \lambda) &= f(x_1 | \lambda) f(x_2 | \lambda) \dots f(x_n | \lambda) \\ &= e^{-\lambda} \frac{\lambda^{x_1}}{x_1!} \cdot e^{-\lambda} \frac{\lambda^{x_2}}{x_2!} \dots e^{-\lambda} \frac{\lambda^{x_n}}{x_n!} = \frac{1}{x_1! \cdot x_2! \dots x_n!} e^{-n\lambda} \lambda^{\sum_{i=1}^n x_i} \\ &= \frac{1}{x_1! \cdot x_2! \dots x_n!} e^{-n\lambda} \cdot \exp\left(\sum_{i=1}^n x_i \cdot \ln \lambda\right) \end{aligned}$$

dengan mengambil :  $h(\underline{x}) = \frac{1}{x_1! \cdot x_2! \dots x_n!}$

$$c(\lambda) = e^{-n\lambda} \quad , \quad w(\lambda) = \ln \lambda \quad \text{dan} \quad t(\underline{x}) = \sum_{i=1}^n x_i \quad , \quad \text{maka dapat}$$

disimpulkan keluarga densitas tersebut merupakan anggota keluarga eksponensial satu parameter.

### Contoh 1.5

Misalkan  $X_1, X_2, \dots, X_n$  variabel random independent berdistribusi Normal dengan rata-rata  $\mu$  dan variansi  $\sigma^2$ . Selidiki apakah merupakan anggota keluarga eksponensial dua parameter.

**Penyelesaian:**

$X_i \sim N(\mu, \sigma^2)$  maka densitas dari  $X_i$  adalah :

$$f(x_i | \mu, \sigma^2) = (2\pi)^{-1/2} (\sigma^2)^{-1/2} \exp\left(-\frac{1}{2\sigma^2} (x_i - \mu)^2\right)$$

$$f(\underline{x} | \mu, \sigma^2) = f(x_1 | \mu, \sigma^2) f(x_2 | \mu, \sigma^2) \dots f(x_n | \mu, \sigma^2) =$$

$$\begin{aligned}
 & (2\pi)^{-1/2}(\sigma^2)^{-1/2} \exp\left(-\frac{1}{2\sigma^2}(x_1-\mu)^2\right) \cdot (2\pi)^{-1/2}(\sigma^2)^{-1/2} \exp\left(-\frac{1}{2\sigma^2}(x_2-\mu)^2\right) \\
 & \dots\dots\dots (2\pi)^{-1/2}(\sigma^2)^{-1/2} \exp\left(-\frac{1}{2\sigma^2}(x_n-\mu)^2\right) \\
 & = (2\pi)^{-n/2}(\sigma^2)^{-n/2} \exp\left(-\frac{1}{2\sigma^2}\sum_{i=1}^n(x_i-\mu)^2\right) \\
 & = (2\pi)^{-n/2}(\sigma^2)^{-n/2} \exp\left(-\frac{1}{2\sigma^2}\left(\sum_{i=1}^n x_i^2 - 2\mu\sum_{i=1}^n x_i + n\mu^2\right)\right) \\
 & = (2\pi)^{-n/2}(\sigma^2)^{-n/2} \cdot \exp\left(-\frac{n\mu^2}{2\sigma^2}\right) \cdot \exp\left(-\frac{1}{2\sigma^2}\sum_{i=1}^n x_i^2 + \frac{\mu}{\sigma^2}\sum_{i=1}^n x_i\right)
 \end{aligned}$$

dengan mengambil :  $h(\underline{x}) = (2\pi)^{-n/2}$

$$c(\mu, \sigma^2) = (\sigma^2)^{-n/2} \cdot \exp\left(-\frac{n\mu^2}{2\sigma^2}\right)$$

$$w_1(\mu, \sigma^2) = -\frac{1}{2\sigma^2} \quad \text{dan} \quad w_2(\mu, \sigma^2) = \frac{\mu}{\sigma^2}$$

$$t_1(\underline{x}) = \sum_{i=1}^n x_i^2 \quad \text{dan} \quad t_2(\underline{x}) = \sum_{i=1}^n x_i$$

maka dapat disimpulkan keluarga densitas tersebut merupakan anggota keluarga eksponensial dua parameter.

Misalkan  $X_1, X_2, \dots, X_n$  variabel random independent dengan densitas bersama :

$$f(\underline{x} | \underline{\theta}) = h(\underline{x}) \cdot c(\underline{\theta}) \cdot \exp\left(\sum_{i=1}^k w_i(\underline{\theta}) \cdot t_i(\underline{x})\right), \text{ maka}$$

$t(\underline{x}) = \left(t_1(\underline{x}), t_2(\underline{x}), \dots, t_k(\underline{x})\right)$  merupakan statistik cukup untuk  $\underline{\theta}$ .

Berdasarkan contoh 2.2.5 tersebut maka statistik cukup untuk  $\mu$  adalah  $\sum_{i=1}^n x_i$

dan statistik cukup untuk  $\sigma^2$  adalah  $\sum_{i=1}^n x_i^2$ .

Statistik cukup untuk  $\theta$  tidak tunggal, setiap fungsi satu-satu dari statistik cukup untuk  $\theta$  adalah juga statistik cukup untuk  $\theta$ . Sebagai contoh yang paling sederhana bila  $\sum_{i=1}^n x_i$  adalah statistik cukup untuk  $\theta$ , maka  $\sum_{i=1}^n x_i/n$  juga merupakan statistik cukup untuk  $\theta$ .

## 2. Statistik Lengkap.

Sifat kelengkapan suatu statistik  $T(\underline{X})$  adalah dilihat dari distribusi atau densitas dari  $T(\underline{X})$ . Bila densitas dari  $T(\underline{X})$  lengkap maka  $T(\underline{X})$  adalah statistik lengkap.

### Definisi 2.1.

Misalkan  $f(t|\theta)$  adalah keluarga densitas untuk  $T(\underline{X})$ . Keluarga densitas disebut lengkap bila  $E(g(T)) = 0$  untuk setiap  $\theta$  dalam ruang parameter, maka  $g(T) = 0$ . Secara ekuivalen  $T(\underline{X})$  disebut statistik lengkap.

### Contoh 2.1.

Misalkan  $T(\underline{X})$  berdistribusi Binomial  $(n, \theta)$ ,  $0 < \theta < 1$ . Akan ditunjukkan bahwa  $T(\underline{X})$  adalah statistik lengkap.

**Penyelesaian:**

Andaikan  $f(t|\theta)$  adalah densitas dari  $T\left(\underline{X}\right)$ , maka  $f(t|\theta) = \binom{n}{t} \theta^t (1-\theta)^{n-t}$

Misal  $g$  suatu fungsi sedemikian sehingga  $E(g(T)) = 0$

$$\sum_t g(t) \binom{n}{t} \theta^t (1-\theta)^{n-t} = 0$$

$$(1-\theta)^n \sum_t g(t) \binom{n}{t} \left(\frac{\theta}{1-\theta}\right)^t = 0$$

Ambil:  $\frac{\theta}{1-\theta} = r$  dengan  $0 < r < \infty$ ,  $(1-\theta)^n \sum_t g(t) \binom{n}{t} r^t = 0$  merupakan suatu polynomial dalam  $r$  dengan  $r^t > 0$  dan jumlahnya sama dengan nol. Sehingga  $g(t) \binom{n}{t} = 0$ , padahal  $\binom{n}{t}$  selalu positif, maka  $g(t) = 0$ . Jadi  $T\left(\underline{X}\right)$  adalah statistik lengkap.

**Contoh 2.2**

Misalkan  $T\left(\underline{X}\right)$  berdistribusi Poisson dengan parameter  $\lambda$ . Akan ditunjukkan bahwa  $T\left(\underline{X}\right)$  adalah statistik lengkap.

**Penyelesaian :**

Andaikan  $f(t|\lambda)$  adalah densitas dari  $T\left(\underline{X}\right)$ , maka  $f(t|\lambda) = e^{-\lambda} \frac{\lambda^t}{t!}$ .

Misal  $g$  suatu fungsi sedemikian sehingga  $E(g(T)) = 0$

$$\sum_t g(t) e^{-\lambda} \frac{\lambda^t}{t!} = 0$$

$$e^{-\lambda} \sum_t g(t) \frac{\lambda^t}{t!} = 0$$

karena  $e^{-\lambda} > 0$  maka  $g(t) \cdot \frac{\lambda^t}{t!} = 0$ , padahal  $\frac{\lambda^t}{t!}$  selalu positif, maka  $g(t) = 0$ .

Jadi  $T(\underline{X})$  adalah statistik lengkap.

### 3. Kecukupan, Kelengkapan dan Ketakbiasan.

Pada submodul ini, konsep kecukupan dan kelengkapan akan digunakan untuk menentukan estimator takbias terbaik. Sebelumnya akan dipelajari terlebih dahulu konsep kecukupan dengan estimator likelihood maksimum seperti pada teorema berikut :

#### Teorema 3.1.

Jika  $T(\underline{X})$  adalah statistik cukup untuk  $\theta$  dan  $\hat{\theta}$  adalah satu-satunya estimator maksimum likelihood dari  $\theta$ , maka  $\hat{\theta}$  merupakan fungsi dari  $T(\underline{X})$

**Bukti :**

Telah diperoleh bahwa :  $L(\theta | \underline{x}) = f(\underline{x} | \theta)$

Dengan teorema faktorisasi :  $f(\underline{x} | \theta) = h(\underline{x}) \cdot q(T(\underline{x}) | \theta)$

Berarti bahwa nilai yang memaksimumkan fungsi likelihood harus tergantung  $T(\underline{X})$ , katakan :  $\hat{\theta} = g(T(\underline{X}))$ . Jika estimator likelihood maksimum tunggal, akan merupakan fungsi dari  $T(\underline{X})$ .

**Contoh 3.1.**

Misalkan  $X_1, X_2, \dots, X_n$  sampel random independen dari populasi yang berdistribusi Bernoulli dengan parameter  $\theta$ . Akan ditunjukkan bahwa estimator maksimum likelihood dari  $\theta$  merupakan fungsi dari statistic cukup untuk  $\theta$ .

**Penyelesaian :**

Berdasarkan contoh 2.2.1. telah diperoleh bahwa :  $T(\underline{X}) = \sum_{i=1}^n X_i$  merupakan statistik cukup untuk  $\theta$  dan  $\sum_{i=1}^n X_i$  berdistribusi Binomial ( $n, \theta$ ).

Akan dicari MLE untuk  $\theta$  sebagai berikut :

$X_i \sim B(1, \theta)$  , maka fungsi densitasnya :  $f(x|\theta) = \theta^{x_i} \cdot (1-\theta)^{1-x_i}$  ,  $x_i = 0, 1$ .

$$L(\theta | \underline{X}) = f(x_1|\theta) \cdot f(x_2|\theta) \cdot \dots \cdot f(x_n|\theta)$$

$$= \theta^{x_1} (1-\theta)^{1-x_1} \cdot \theta^{x_2} (1-\theta)^{1-x_2} \cdot \dots \cdot \theta^{x_n} (1-\theta)^{1-x_n} = \theta^{\sum_{i=1}^n x_i} (1-\theta)^{n-\sum_{i=1}^n x_i}$$

$$\log(L(\theta | \underline{X})) = \log\left(\theta^{\sum_{i=1}^n x_i} \cdot (1-\theta)^{n-\sum_{i=1}^n x_i}\right) = \sum_{i=1}^n x_i \log \theta + \left(n - \sum_{i=1}^n x_i\right) \log(1-\theta)$$

$$\frac{\partial \log(L(\theta | \underline{X}))}{\partial \theta} = \frac{\sum_{i=1}^n x_i}{\theta} - \frac{n - \sum_{i=1}^n x_i}{(1-\theta)}$$

Persamaan likelihood :

$$\frac{\partial \log \left( L(\theta | \underline{X}) \right)}{\partial \hat{\theta}} = 0$$

$$\frac{\sum_{i=1}^n x_i}{\hat{\theta}} - \frac{n - \sum_{i=1}^n x_i}{(1 - \hat{\theta})} = 0$$

$$(1 - \hat{\theta}) \sum_{i=1}^n x_i = \hat{\theta} \left( n - \sum_{i=1}^n x_i \right)$$

$$\sum_{i=1}^n x_i = n \cdot \hat{\theta}$$

$$\hat{\theta} = \frac{\sum_{i=1}^n x_i}{n}$$

Dari baris terakhir terlihat bahwa  $\hat{\theta} = \frac{T(\underline{X})}{n}$  yang berarti bahwa  $\hat{\theta}$  (MLE) merupakan fungsi dari  $T(\underline{X})$  (statistik cukup)

Dalam pembahasan terdahulu konsep kecukupan tidak digunakan dalam pencarian estimator takbias. Sekarang akan dilihat bahwa konsep kecukupan merupakan alat yang berguna untuk menentukan estimator takbias. Perlu diingat bahwa jika X dan Y adalah dua peubah acak sedemikian sehingga harga harapannya ada, maka berlaku :

$$E(X) = E(E[X | Y]) \text{ dan } \text{Var}(X) = \text{Var}(E[X | Y]) + E(\text{Var}[X | Y]) \dots \dots \dots (2.1)$$

Sifat ini akan digunakan untuk membuktikan teorema berikut.

**Teorema 3.2.**

Misalkan W adalah estimator takbias untuk  $g(\theta)$ , dan T adalah statistic cukup untuk  $\theta$ . Didefinisikan :  $\varphi(t) = E(W | T)$ , maka  $E(\varphi(t)) = g(\theta)$ , dan  $\text{var}(\varphi(t)) \leq \text{Var}(W)$  untuk setiap  $\theta$ , yaitu  $\varphi(t)$  secara seragam (uniform) merupakan estimator dari  $g(\theta)$  yang takbias dan lebih baik.

**Bukti :**

Dengan menggunakan persamaan (2.1) :  $E(W) = E(E[W | T]) = E(\varphi(t))$ . Karena  $W$  estimator takbias untuk  $g(\theta)$  maka  $E(W) = g(\theta)$ . Sehingga  $E(\varphi(t)) = g(\theta)$  yang berarti  $\varphi(t)$  estimator takbias untuk  $g(\theta)$ .

$$\begin{aligned}\text{Var}(W) &= \text{Var}(E(W | T)) + E(\text{Var}(W | T)) \\ &= \text{Var}(\varphi(t)) + E(\text{Var}(W | T)) \\ &\geq \text{Var}(\varphi(t)) \text{ karena } \text{Var}(W | T) \geq 0.\end{aligned}$$

Karena itu  $\varphi(t)$  secara seragam lebih baik dari  $W$ , dan tinggal membuktikan bahwa  $\varphi(t)$  benar-benar estimator, yaitu  $\varphi(t) = E(W | T)$  hanya merupakan fungsi sampel dan khususnya independen dengan  $\theta$ . Dari definisi statistik cukup, distribusi  $W | T$  independen dengan  $\theta$  dan dari kenyataan  $W$  fungsi sampel saja maka  $\varphi(t)$  adalah statistik (merupakan fungsi sampel saja).

Teorema diatas mengatakan bahwa dalam mencari estimator takbias terbaik kita hanya memandangi statistik yang merupakan fungsi dari statistik cukup. Sekarang kita mengetahui bahwa dalam mencari estimator takbias terbaik untuk  $g(\theta)$  kita hanya perlu memandangi estimator yang berdasar pada statistik cukup. Pertanyaan yang muncul sekarang adalah bila kita mempunyai  $E(\varphi(t))=g(\theta)$  dan  $\varphi(t)$  berdasarkan statistik cukup  $T$ , bagaimana kita mengetahui bahwa  $\varphi(t)$  adalah takbias terbaik?. Memang, bila  $\varphi(t)$  mencapai batas bawah Cramer-Rao maka ia takbias terbaik, tetapi bila tidak, apakah kita mendapatkan sesuatu?. Dalam beberapa kasus, hanya ada satu fungsi dari statistik cukup  $T$  katakan  $f(T)$  yang takbias untuk  $g(\theta)$ , maka  $f(T)$  akan merupakan estimator takbias terbaik untuk  $g(\theta)$ .

**Teorema 3.3**

Bila  $W$  estimator takbias terbaik untuk  $g(\theta)$ , maka  $W$  tunggal.

**Bukti :**

Misalkan  $W'$  juga merupakan estimator takbias terbaik dari  $g(\theta)$ , maka

$\text{Var}(W') = \text{Var}(W)$ . Pandang estimator  $W^* = \frac{1}{2} (W + W')$ , maka  $W^*$  juga merupakan estimator takbias dan :

$$\begin{aligned}\text{Var}(W^*) &= \text{Var}\left(\frac{1}{2} (W + W')\right) \\ &= \frac{1}{4} \text{Var}(W) + \frac{1}{4} \text{Var}(W') + \frac{1}{2} \text{Cov}(W, W')\end{aligned}$$

Dengan menggunakan ketaksamaan Cauchy-schwarz :

$[\text{Cov}(W, W')]^2 \leq \text{Var}(W) \cdot \text{Var}(W')$ , maka :

$$\text{Var}(W^*) \leq \frac{1}{4} \text{Var}(W) + \frac{1}{4} \text{Var}(W') + \frac{1}{2} [\text{Var}(W) \cdot \text{Var}(W')]^{1/2} = \text{Var}(W)$$

Tetapi bila ketaksamaan diatas tegas, maka ketakbiasan terbaik dari  $W$  diingkari, sehingga kita harus mempunyai kesamaan untuk setiap  $\theta$ . Karena kesamaan adalah terapan dari Cauchy-Schwarz, maka kita mempunyai kesamaan hanya bila  $W' = a(\theta) W + b(\theta)$ . Sekarang dengan menggunakan sifat dari kovariansi didapat :

$$\begin{aligned}\text{Cov}(W, W') &= \text{Cov}(W, a(\theta) W + b(\theta)) \\ &= \text{Cov}(W, a(\theta) W) \\ &= a(\theta) \text{Var}(W)\end{aligned}$$

Karena  $\text{Cov}(W, W') = \text{Var}(W)$ , sehingga  $a(\theta) = 1$  dan karena  $E(W') = g(\theta)$  akibatnya  $b(\theta) = 0$  dan  $W = W'$  yang berarti  $W$  tunggal.

Hubungan antara kecukupan dan kelengkapan statistic  $T$  dengan estimator takbias terbaik dinyatakan dengan teorema berikut :

#### **Teorema 3.4**

Misalkan  $T$  adalah statistic cukup dan lengkap untuk parameter  $\theta$ , dan misalkan  $\varphi(t)$  adalah sebarang estimator yang hanya bergantung pada  $T$ , maka  $\varphi(t)$  adalah estimator takbias terbaik tunggal dari harga harapannya.

Dalam beberapa situasi, calon untuk estimator takbias terbaik sangat sukar untuk dilihat. Tetapi dalam hal adanya kelengkapan (completeness), teori yang kita bicarakan mengatakan bahwa bila kita dapat mencari estimator takbias maka kita bisa mendapatkan estimator takbias terbaik. Bila  $T$  statistic cukup dan lengkap

untuk parameter  $\theta$ , dan  $H(X_1, X_2, \dots, X_n)$  sebarang estimator takbias untuk  $g(\theta)$ , maka :  $\varphi(t) = E(H(X_1, X_2, \dots, X_n) | T)$  adalah estimator takbias terbaik untuk  $g(\theta)$ .

**Contoh 3.2**

Misalkan  $X_1, X_2, \dots, X_n$  adalah peubah acak independen berdistribusi Binomial( $k, \theta$ ). Ambil  $g(\theta) = P(X=1) = k\theta(1-\theta)^{k-1}$ , tentukan estimator takbias terbaik untuk  $g(\theta)$ .

**Penyelesaian :**

$$X_i \sim B(k, \theta) \text{ maka densitas dari } X_i \text{ adalah ; } f(x_i | \theta) = \binom{k}{x_i} \theta^{x_i} (1-\theta)^{k-x_i}$$

Karena  $X_1, X_2, \dots, X_n$  independen maka :

$$\begin{aligned} f(\underline{x} | \theta) &= f(x_1 | \theta) \cdot f(x_2 | \theta) \cdot \dots \cdot f(x_n | \theta) \\ &= \binom{k}{x_1} \theta^{x_1} (1-\theta)^{k-x_1} \cdot \binom{k}{x_2} \theta^{x_2} (1-\theta)^{k-x_2} \cdot \dots \cdot \binom{k}{x_n} \theta^{x_n} (1-\theta)^{k-x_n} \\ &= \prod_{i=1}^n \binom{k}{x_i} \cdot \theta^{\sum_{i=1}^n x_i} (1-\theta)^{n \cdot k - \sum_{i=1}^n x_i} \end{aligned}$$

Sehingga  $T(\underline{X}) = \sum_{i=1}^n X_i$  adalah statistik cukup untuk  $\theta$ , dan  $T(\underline{X}) = \sum_{i=1}^n X_i$

berdistribusi Binomial( $n \cdot k, \theta$ ), maka  $T(\underline{X}) = \sum_{i=1}^n X_i$  merupakan statistik lengkap.

Pandang :

$$H(\underline{X}) = \begin{cases} 1, \dots, \dots, X_i = 1 \\ 0, \dots, \dots, \text{yanglain} \end{cases}$$

$$E(H(\underline{X})) = \sum_{x_i} h(x_i) \cdot \binom{k}{x_i} \theta^{x_i} (1-\theta)^{k-x_i} = 1 \cdot \binom{k}{1} \theta^1 (1-\theta)^{k-1} = k\theta(1-\theta)^{k-1}$$

Sehingga  $H(\underline{X})$  merupakan estimator takbias untuk  $g(\theta) = P(X=1) = k\theta(1-\theta)^{k-1}$

$\varphi(T) = E(H(\underline{X} | T))$  akan merupakan estimator takbias terbaik untuk

$g(\theta) = P(X=1) = k\theta(1-\theta)^{k-1}$ . Misalkan telah diobservasi  $\sum_{i=1}^n X_i = t$  maka :

$$\begin{aligned}
 \varphi(T) &= E(H(\underline{X} | \sum_{i=1}^n X_i = t)) \\
 &= \sum_{x_i} H(\underline{x}) \cdot P(H(\underline{x}) | \sum_{i=1}^n X_i = t) \\
 &= 1 \cdot P(H(\underline{x}) | \sum_{i=1}^n X_i = t) \\
 &= \frac{P(H(\underline{x}), \sum_{i=1}^n X_i = t)}{P(\sum_{i=1}^n X_i = t)} \\
 &= \frac{P(X_1 = 1, X_1 + X_2 + \dots + X_n = t)}{P(\sum_{i=1}^n X_i = t)} \\
 &= \frac{P(X_1 = 1, X_2 + X_3 + \dots + X_n = t-1)}{P(\sum_{i=1}^n X_i = t)} \\
 &= \frac{k\theta(1-\theta)^{k-1} \cdot \binom{k(n-1)}{t-1} \theta^{t-1} (1-\theta)^{k(n-1)-(t-1)}}{\binom{kn}{t} \theta^t (1-\theta)^{kn-t}} \\
 &= \frac{k \binom{k(n-1)}{t-1}}{\binom{kn}{t}}
 \end{aligned}$$

Sehingga : 
$$\varphi(T) = \frac{k \binom{k(n-1)}{\sum_{i=1}^n X_i - 1}}{\binom{k.n}{\sum_{i=1}^n X_i}}$$
 merupakan estimator takbias terbaik

untuk  $g(\theta) = P(X=1) = k\theta(1-\theta)^{k-1}$ .

**Teorema 3.5. (Lehmann-Scheffe)**

Misalkan  $X_1, X_2, \dots, X_n$  mempunyai densitas bersama  $f(x | \theta)$  dan misalkan  $T$  statistik cukup dan lengkap untuk  $\theta$ . Jika  $T^* = f(T)$  yaitu  $T^*$  merupakan fungsi dari  $T$  adalah statistik yang takbias untuk  $g(\theta)$  maka  $T^*$  merupakan estimator takbias terbaik untuk  $g(\theta)$ .

**Bukti :**

Dengan menggunakan sifat kelengkapan dari sebarang statistik yang merupakan fungsi dari  $T$  dan takbias untuk  $g(\theta)$  harus sama dengan  $T^*$  dengan peluang 1. jika  $W$  sebarang estimator takbias untuk  $g(\theta)$ , dengan menggunakan teorema Rao-Blackwell  $E(W | T)$  fungsi dari  $T$  yang takbias untuk  $g(\theta)$ , sehingga dengan teorema ketunggalan dari estimator takbias terbaik maka  $T^* = E(W | T)$  dengan peluang 1. Selanjutnya  $Var(T^*) \leq Var(W)$  untuk setiap  $\theta$ . Jadi  $T^*$  estimator takbias terbaik untuk  $g(\theta)$ .

**Contoh 3.3**

Misalkan  $X_1, X_2, \dots, X_n$  peubah acak independen dengan densitas :  $f(x | \theta) = \theta \cdot x^{\theta-1}$  untuk  $0 < x < 1$  dan  $\theta > 0$ . Tentukan estimator takbias terbaik untuk  $g(\theta) = \theta$ .

**Penyelesaian :**

$$f(x | \theta) = \theta \cdot x^{\theta-1}$$

misalkan  $y = -\ln(x)$ , karena  $0 < x < 1$ , maka  $y > 0$ . berarti  $x = e^{-y}$ ,  $dx/dy = -e^{-y}$  sehingga  $|dx/dy| = e^{-y}$

Densitas dari Y adalah:  $g(y | \theta) = \theta \cdot (e^{-y})^{\theta-1} \cdot e^{-y} = \theta \cdot e^{-\theta y}$ ,  $y > 0$

Terlihat bahwa Y berdistribusi Eksponensial dengan rata-rata  $1/\theta$ .

Fungsi pembangkit moment dari Y adalah:  $M_Y(t) = \left(1 - \frac{t}{\theta}\right)^{-1}$

Bila  $Y_1, Y_2, \dots, Y_n$  independen dan berdistribusi Eksponensial dengan rata-rata  $1/\theta$

maka  $M_{\sum_{i=1}^n Y_i}(t) = \prod_{i=1}^n M_{Y_i}(t) = \left(1 - \frac{t}{\theta}\right)^{-n}$  yang merupakan fungsi pembangkit momen dari distribusi Gamma( $n, 1/\theta$ ).

Notasikan  $S = \sum_{i=1}^n Y_i$  maka densitas dari S adalah:

$$k(s | \theta) = \frac{1}{\Gamma(n)} \cdot \theta^n s^{n-1} \cdot e^{-\theta s} \text{ dengan } s > 0 \text{ dan } \theta > 0$$

$$\begin{aligned} E\left(\frac{1}{S}\right) &= \int_0^{\infty} \frac{1}{s} \frac{1}{\Gamma(n)} \theta^n s^{n-1} \cdot e^{-\theta s} ds \\ &= \frac{\theta}{(n-1)!} \int_0^{\infty} (\theta s)^{n-2} \cdot e^{-\theta s} d(\theta s) \\ &= \frac{\theta}{(n-1)!} \Gamma(n-1) \\ &= \frac{\theta}{(n-1)} \end{aligned}$$

Sehingga bila diambil  $H(S) = \frac{n-1}{S}$  maka  $E(H(S)) = \theta$

Jadi  $T(X) = \frac{n-1}{-\sum_{i=1}^n \log X_i} = \frac{1-n}{\sum_{i=1}^n \log X_i}$  adalah estimator takbias terbaik untuk  $\theta$ .

### 2.4.2 Latihan

1. Apakah hubungannya antara statistik cukup dengan estimator titik.
2. Jelaskan perbedaan menentukan estimator takbias terbaik melalui batas bawah Cramer - Rao dengan melalui hubungan antara ketakbiasan, kecukupan dan kelengkapan.

### 2.4.3 Rangkuman

1. Menentukan statistik cukup dapat melalui dua cara yaitu dengan menggunakan faktorisasi Neyman-Pearson dan dengan melalui keluarga eksponensial.
2. Membuktikan sifat statistik lengkap paling sederhana dengan membuktikan bahwa densitas dari statistik tersebut merupakan anggota keluarga eksponensial, karena keluarga eksponensial bersifat lengkap.
3. Menentukan estimator takbias terbaik melalui hubungan antara ketakbiasan, kecukupan dan kelengkapan harus dilakukan dengan cermat, teliti dan hati-hati, karena perhitungannya relatif rumit.

### 2.4.4 Tes Formatif

1. Tunjukkan bahwa keluarga distribusi berikut merupakan anggota keluarga eksponensial dan untuk setiap kasus tentukan statistik cukup dan lengkap berdasarkan sampel random  $X_1, X_2, \dots, X_n$ 
  - a. Binomial  $(n, \theta)$  dengan  $0 < \theta < 1$
  - b. Gamma  $(2, \theta)$  dengan  $\theta > 0$
  - c. Eksponensial  $(\theta)$  dengan  $\theta > 0$
2. Misalkan  $X_1, X_2, \dots, X_n$  sampel random dari populasi berdistribusi Bernoulli  $(\theta)$ , tentukan estimator takbias terbaik untuk :
  - a.  $\theta(1-\theta)$
  - b.  $\theta^2$

3. Suatu sampel random berukuran  $n$  diambil dari populasi dengan densitas

$f(x | \theta) = \theta \cdot x^{\theta-1}$ ,  $0 < x < 1$ ,  $\theta > 0$ . Tentukan estimator takbias terbaik untuk  $1/\theta$ .

4.  $X_1, X_2, \dots, X_n$  sampel random dari populasi dengan densitas :

$$f(x | \theta) = \theta(1-x)^{-(1+\theta)}, x > 0, \theta > 0$$

- Tentukan statistik cukup untuk  $\theta$
- Tentukan batas bawah Cramer-Rao untuk  $1/\theta$
- Tentukan estimator takbias terbaik untuk  $1/\theta$
- Tentukan estimator takbias terbaik untuk  $\theta$

#### 2.4.5 Umpan Balik dan Tindak Lanjut.

Cocokkanlah jawaban anda dengan kunci jawaban dari modul ini. Hitunglah jawaban anda yang benar, kemudian gunakanlah rumus dibawah ini untuk mengetahui tingkat penguasaan anda terhadap materi ini.

$$\text{Tingkat penguasaan} = (\text{jumlah jawaban benar}) / 10 \times 100\%$$

Arti tingkat penguasaan yang anda capai :

90% - 100% : baik sekali

80% - 89% : baik

70% - 79% : sedang

< 70% : kurang

Bila anda telah mencapai tingkat penguasaan 80% atau lebih, anda dapat meneruskan kegiatan belajar selanjutnya. Tetapi apabila tingkat penguasaan anda kurang dari 80% , sebaiknya anda mengulangi materi ini terutama bagian yang belum anda kuasai sebelum meneruskan kegiatan belajar selanjutnya.

#### 2.5 Kunci Jawaban Tes Formatif

1. a  $\sum_{i=1}^n X_i$

$$\text{b. } \sum_{i=1}^n X_i$$

$$\text{c. } \sum_{i=1}^n X_i$$

$$2. \text{ a. } \left( \sum_{i=1}^n X_i - \left( \sum_{i=1}^n X_i \right)^2 / n \right) / (n-1)$$

$$\text{b. } \left[ \left( \sum_{i=1}^n X_i \right)^2 - \sum_{i=1}^n X_i \right] / (n(n-1))$$

$$3. -\frac{1}{n} \sum_{i=1}^n \ln X_i$$

$$4. \text{ a. } \sum_{i=1}^n \ln(1 + X_i)$$

$$\text{b. } \frac{\theta^2}{n}$$

$$\text{c. } \frac{1}{n} \sum_{i=1}^n \ln X_i$$

$$\text{d. } \frac{n-1}{\sum_{i=1}^n \ln X_i}$$

## 2.6 Referensi :

1. Bain, L.J. and Engelhart, M. (1992); Introduction to Promodulility and Mathematical Statistics, Second Edition. Duxbury Press, Belmont, California.
2. Casella, G and Berger, R.L. (1990); Statistical Inference. Wadsworth Inc. Belmont, California.
3. Dudewicz, E.J and Mishra, S.N. (1998); Modern Mathematical Statistics. John Willey and Sons, Singapore.
4. Roussas, G.G. (1976); A First Course in Mathematical Statistics. Mei Ya Publication Inc. Taipei, Taiwan.

## MODUL III UJI HIPOTESIS

### 3.1 Pengantar

Pada bagian ini anda akan mempelajari konsep uji hipotesis, dimulai dari pengertian uji hipotesis , hipotesis sederhana (simple hipotesis) dan hipotesis composite. Pokok permasalahannya adalah menentukan statistik uji (statistik hitung) dan statistik tabel serta kriteria ujinya (menentukan daerah penolakan). Ada beberapa macam metode penentuan kriteria uji tergantung dari bentuk hipotesis yang diambil. Dalam modul ini akan dibahas tentang uji paling kuasa (most powerful tes = MP tes), uji paling kuasa seragam (uniformly most powerful tes = UMP tes) dan uji ratio likelihood (generalized likelihood ratio tes = GLRT)

### 3.2 Tujuan Instruksional Umum :

Setelah mempelajari bagian ini mahasiswa dapat memilih uji yang sesuai, menentukan statistik uji dan statistik tabel sehingga dapat ditentukan kriteria ujinya.

### 3.3 Tujuan Instruksional Khusus :

Setelah mempelajari bagian ini mahasiswa dapat :

1. Mendefinisikan suatu hipotesis.
2. Membedakan antara hipotesis sederhana dan hipotesis composite.
3. Menggunakan uji paling kuasa untuk menentukan statistik uji, statistik tabel dan kriteria ujinya. Selanjutnya berdasarkan data terobservasi mahasiswa dapat melakukan inferensi.
4. Menggunakan uji paling kuasa uniform uji paling kuasa untuk menentukan statistik uji, statistik tabel dan kriteria ujinya. Selanjutnya berdasarkan data terobservasi mahasiswa dapat melakukan inferensi.
5. Menggunakan uji rasio likelihood untuk menentukan statistik uji, statistik tabel dan kriteria ujinya. Selanjutnya berdasarkan data terobservasi mahasiswa dapat melakukan inferensi.

### 3.4 Kegiatan Belajar :

## UJI HIPOTESIS

### 3.4.1. Uraian dan Contoh.

#### 1. Pengertian.

Masalah pada estimasi titik adalah mendapatkan suatu harga dugaan dari suatu parameter. Pada beberapa situasi sering kita dihadapkan dengan masalah memilih satu tindakan diantara dua pilihan. Misalnya, setelah suatu jenis obat penurun tekanan darah bagi penderita hipertensi dikembangkan, uji coba yang ekstensif akan dilakukan untuk membuktikan apakah obat tersebut benar-benar dapat menurunkan tekanan darah bagi penderita hipertensi. Dalam hal ini diperlukan suatu uji yang cermat untuk menentukan satu pilihan apakah obat tersebut seyogyanya dipergunakan ataukah sebaliknya yaitu obat tersebut tidak dipergunakan. Pemilihan keputusan seperti ini diperlukan teori uji hipotesis.

#### Definisi 1.1

Hipotesis adalah pernyataan (dugaan) tentang parameter populasi.

Definisi tersebut agak umum, tetapi hal penting yang harus diketahui adalah hipotesis membuat suatu pernyataan tentang populasi. Tujuan dari uji hipotesis adalah memutuskan berdasarkan sampel dari populasi, mana dari dua hipotesis yang saling asing yang benar.

#### Definisi 1.2

Dua hipotesis yang saling asing dalam persoalan uji hipotesis disebut hipotesis nol dinotasikan  $H_0$  dan hipotesis alternatif dinotasikan  $H_1$ .

Kesimpulan yang diperoleh dari uji hipotesis hanya menolak  $H_0$  (berarti menerima  $H_1$ ) atau menerima  $H_0$  (berarti menolak  $H_1$ ). Bila kita

mempunyai suatu hipotesis (dugaan), biasanya kita berkeyakinan dugaan tersebut benar, sehingga harapannya diterima, maka kebiasaan ilmiah dugaan yang dianggap benar diletakkan sebagai hipotesis alternatif ( $H_1$ ), dan dugaan yang diharapkan untuk ditolak diletakkan sebagai hipotesis nol ( $H_0$ ).

Bila  $\theta$  menyatakan parameter populasi,  $\Omega$  menyatakan ruang parameter,  $\Omega_0 \subset \Omega$ , format umum dari hipotesis nol dan hipotesis alternatif adalah :  $H_0 : \theta \in \Omega_0$  dan  $H_1 : \theta \in \Omega_0^c$  dengan  $\Omega_0^c$  komplemen dari  $\Omega_0$ . jika  $\Omega_0$  hanya memuat satu harga  $\theta$ , yaitu  $\Omega_0 = \{\theta\}$  maka hipotesis nol ( $H_0$ ) disebut hipotesis sederhana (simple hypothesis) dan selain itu disebut hipotesis composite. Hal yang sama juga berlaku untuk  $H_1$ . Keuntungan bentuk  $H_0$  dan  $H_1$  yang dapat diuji adalah :

1.  $H_0$  sederhana dan  $H_1$  sederhana
2.  $H_0$  sederhana dan  $H_1$  composite
3.  $H_0$  composite dan  $H_1$  composite

Dalam persoalan uji hipotesis, sesudah mengobservasi sampel langkah selanjutnya menentukan apakah menerima  $H_0$  yang berarti menolak  $H_1$  atau menerima  $H_1$  yang berarti menolak  $H_0$ . Himpunan bagian ruang sampel dimana  $H_0$  ditolak disebut daerah penolakan atau daerah kritis dinotasikan  $R$  dan komplemen daerah penolakan disebut daerah penerimaan dinotasikan  $R^c$ .

### Definisi 1.3

Fungsi uji (test function)  $\phi(\underline{X})$  adalah fungsi pada ruang sampel dengan harga 1 bila  $\underline{X} = (X_1, X_2, \dots, X_n)$  dalam daerah penolakan dan 0 bila  $\underline{X} = (X_1, X_2, \dots, X_n)$  dalam daerah penerimaan.

$$\phi(\underline{X}) = \begin{cases} 1, & \underline{X} \in R \\ 0, & \underline{X} \in R^c \end{cases}$$

Biasanya, uji hipotesis dinyatakan dalam bentuk statistik uji  $T(\underline{X}) = T(X_1, X_2, \dots, X_n)$  yaitu fungsi dari sampel. Sebagai contoh, uji yang menyatakan  $H_0$  ditolak bila rata-rata sampel  $(\bar{X})$  lebih besar atau sama dengan 3. Dalam hal ini  $T(\underline{X}) = \bar{X}$  adalah statistik uji dan daerah penolakannya adalah  $R = \{(X_1, X_2, \dots, X_n) \mid \bar{X} \geq 3\}$  dinotasikan

dalam fungsi uji : 
$$\phi(\underline{X}) = \begin{cases} 1, & \bar{X} \geq 3 \\ 0, & \bar{X} < 3 \end{cases}$$

Dalam memutuskan untuk menerima atau menolak  $H_0$  kita bisa membuat suatu kesalahan. Ada dua jenis kesalahan dalam uji hipotesis yang dikenal dengan kesalahan tipe I dinotasikan  $\alpha$ , dan kesalahan tipe II dinotasikan  $\beta$ . Untuk lebih jelasnya perhatikan ilustrasi berikut :

Keputusan	Hipotesis nol ( $H_0$ )	
	benar	salah
Menolak $H_0$	Kesalah tipe I ( $\alpha$ )	Keputusan benar
Menerima $H_0$	Keputusan benar	Kesalahan tipe II ( $\beta$ )

Berdasarkan ilustrasi tersebut diperoleh definisi berikut :

**Definisi 1.4**

Kesalahan tipe I ( $\alpha$ ) dan kesalahan tipe II ( $\beta$ ) didefinisikan sebagai :

$$\alpha = P(\text{menolak } H_0 \mid H_0 \text{ benar})$$

$$\beta = P(\text{menerima } H_0 \mid H_0 \text{ salah}) = P(\text{menerima } H_0 \mid H_1 \text{ benar})$$

Pandang uji hipotesis  $H_0 : \theta \in \Omega_0$  vs  $H_1 : \theta \in \Omega_0^c$ ,  $R$  menyatakan daerah penolakan, maka untuk  $\theta \in \Omega_0$  akan membuat kesalahan tipe I bila  $\underline{X} \in R$ , sehingga promodulilitas kesalahan tipe I adalah  $P(\underline{X} \in R)$ .

Untuk  $\theta \in \Omega_0^c$  promodulilitas kesalahan tipe II adalah  $P(X \in R^c)$ .

$$\text{Dapat diringkaskan: } P(X \in R) = \begin{cases} \alpha, & \text{bila } \theta \in \Omega_0 \\ 1 - \beta, & \text{bila } \theta \in \Omega_0^c \end{cases}$$

### Definisi 1.5

Fungsi kuasa (power function)  $\pi(\theta)$  dari uji  $H_0$  adalah promodulilitas menolak  $H_0$  bila harga parameter yang benar adalah  $\theta$ .

Sebagai contoh untuk hipotesis  $H_0 : \theta = \theta_0$  vs  $H_1 : \theta = \theta_1$  maka :

$$\pi(\theta_0) = P(\text{menolak } H_0 \mid \theta = \theta_0) = \alpha$$

$$\pi(\theta_1) = P(\text{menolak } H_0 \mid \theta = \theta_1) = 1 - P(\text{menerima } H_0 \mid \theta = \theta_1) = 1 - \beta.$$

Untuk hipotesis composite  $H_0 : \theta \in \Omega_0$  vs  $H_1 : \theta \in \Omega_0^c$  ukuran test (daerah

kritis);  $\alpha = \max_{\theta \in \Omega_0} \pi(\theta)$

Setelah uji hipotesis dilaksanakan, kesimpulan harus dilaporkan dalam bentuk yang secara statistik mempunyai arti. Salah satu bentuknya adalah ukuran  $\alpha$  dari uji yang digunakan. Bila  $\alpha$  kecil, keputusan untuk menolak  $H_0$  sangat meyakinkan, tetapi apabila  $\alpha$  besar keputusan untuk menolak  $H_0$  tidak begitu meyakinkan karena uji mempunyai promodulilitas yang besar untuk membuat keputusan yang salah.

Cara lain untuk melaporkan hasil uji hipotesis adalah melaporkan harga  $-p$  (p-value). P-value untuk titik sampel adalah harga  $\alpha$  terkecil sampai titik sampel ini membawa pada keputusan menolak  $H_0$ . Makin kecil p-value makin kuat dukungan sampel bahwa  $H_1$  benar.

## 2. Uji Paling Kuasa (Most Powerful Test / MP Test)

Dalam uji hipotesis kesalahan tipe I paling besar sama dengan  $\alpha$  untuk  $\theta \in \Omega_0$ . Uji yang baik juga akan mempunyai kesalahan tipe II yang kecil, sehingga mempunyai fungsi kuasa yang besar untuk  $\theta \in \Omega_0^c$ . Bila mempunyai

uji dengan kesalahan tipe II yang lebih kecil maka akan merupakan calon uji terbaik. Jika  $\underline{X} = (X_1, X_2, \dots, X_n)$  mempunyai densitas bersama  $f(\underline{x} | \theta)$  dan daerah penolakan  $R$ , notasi untuk fungsi kuasa yang bersesuaian dengan  $R$  adalah:  $\pi_R(\theta) = P((X_1, X_2, \dots, X_n) \in R | \theta)$

**Definisi 2.1**

Suatu uji  $H_0 : \theta = \theta_0$  vs  $H_1 : \theta = \theta_1$  dengan daerah penolakan  $R^*$  disebut uji paling kuasa ukuran  $\alpha$  jika:  $\pi_{R^*}(\theta_0) = \alpha$  dan  $\pi_{R^*}(\theta_1) \geq \pi_R(\theta_1)$  untuk sebarang daerah penolakan  $R$  ukuran  $\alpha$ .

Lemma berikut digunakan untuk menentukan daerah penolakan uji paling kuasa dari hipotesis sederhana vs hipotesis sederhana dengan bentuk:

$$H_0 : \theta = \theta_0 \text{ vs } H_1 : \theta = \theta_1$$

**Lemma 2.1**

Andaikan  $\underline{X} = (X_1, X_2, \dots, X_n)$  mempunyai densitas bersama

$$f(\underline{x} | \theta), \quad \text{ambil} \quad : \quad \lambda(\underline{x}, \theta_0, \theta_1) = \frac{f(\underline{x} | \theta_0)}{f(\underline{x} | \theta_1)},$$

$$R^* = \{(X_1, X_2, \dots, X_n) | \lambda(\underline{x}, \theta_0, \theta_1) \leq k\} \quad \text{dengan } k \text{ suatu}$$

konstanta sehingga  $P((X_1, X_2, \dots, X_n) \in R^* | \theta_0) = \alpha$ , maka  $R^*$  adalah daerah penolakan uji paling kuasa ukuran  $\alpha$  untuk uji:

$$H_0 : \theta = \theta_0 \text{ vs } H_1 : \theta = \theta_1.$$

**Bukti:**

$$P(\underline{X} \in A | \theta) = \int_A f(\underline{x} | \theta) d\underline{x}$$

$$R^* = \left\{ (X_1, X_2, \dots, X_n) \mid \lambda(\underline{x}, \theta_0, \theta_1) \leq k \right\}$$

Dalam  $R^*$ ,  $f(\underline{x} \mid \theta_0) \leq f(\underline{x} \mid \theta_1)$

sehingga:  $\int_R f(\underline{x} \mid \theta_0) d\underline{x} \leq k \cdot \int_R f(\underline{x} \mid \theta_1) d\underline{x}$

Jika  $A \subset R^*$  maka:  $P(\underline{X} \in A \mid \theta_0) \leq k \cdot P(\underline{X} \in A \mid \theta_1)$  .....3.2.1

Jika  $A \subset R^{*c}$  maka:  $P(\underline{X} \in A \mid \theta_0) \geq k \cdot P(\underline{X} \in A \mid \theta_1)$  .....3.2.1

$R^* = (R^* \cap R) \cup (R^* \cap R^c)$  dan  $R = (R \cap R^*) \cup (R \cap R^{*c})$

$$\pi_{R^*}(\theta) = P(\underline{X} \in R^* \cap R \mid \theta) + P(\underline{X} \in R^* \cap R^c \mid \theta)$$

$$\pi_R(\theta) = P(\underline{X} \in R \cap R^* \mid \theta) + P(\underline{X} \in R \cap R^{*c} \mid \theta)$$

$$\pi_{R^*}(\theta) - \pi_R(\theta) = P(\underline{X} \in R^* \cap R^c \mid \theta) - P(\underline{X} \in R \cap R^{*c} \mid \theta)$$

$$\pi_{R^*}(\theta_1) - \pi_R(\theta_1) = P(\underline{X} \in R^* \cap R^c \mid \theta_1) - P(\underline{X} \in R \cap R^{*c} \mid \theta_1)$$

dengan menggunakan pertidaksamaan (3.2.1) dan (3.2.2) diperoleh :

$$\pi_{R^*}(\theta_1) - \pi_R(\theta_1) \geq k \cdot [P(\underline{X} \in R^* \cap R^c \mid \theta_0) - P(\underline{X} \in R \cap R^{*c} \mid \theta_0)]$$

$$\pi_{R^*}(\theta_1) - \pi_R(\theta_1) \geq k \pi_{R^*}(\theta_0) - \pi_R(\theta_0)$$

$$\pi_{R^*}(\theta_1) - \pi_R(\theta_1) \geq k(\alpha - \alpha)$$

$$\pi_{R^*}(\theta_1) \geq \pi_R(\theta_1)$$

Dari baris terakhir terlihat memenuhi syarat definisi 3.2.1 , sehingga  $R^*$  merupakan daerah penolakan uji paling kuasa ukuran  $\alpha$  untuk :

$$H_0 : \theta = \theta_0 \text{ vs } H_1 : \theta = \theta_1.$$

**Contoh 2.1**

Misalkan  $X_1, X_2, \dots, X_{10}$  Sampel random dari populasi berdistribusi normal dengan rata-rata  $\theta$  dan variansi 9.

Diambil hipotesis :  $H_0 : \theta = 0.25$  vs  $H_1 : \theta = 1.95$

- a. Tentukan uji paling kuasa dari hipotesis tersebut dengan  $\alpha = 5\%$ .

- b. Bila dipunyai data : -1.82124 5.42107 5.51125 2.70837 0.08660 5.81389  
 -1.41054 0.83611 0.88826 0.53973 bagaimana kesimpulan anda.
- c. Tentukan kuasa ujinya jika  $\theta = 1.95$

**Penyelesaian:**

a. Karena :  $X_i \sim N(\theta, 9)$

maka densitasnya :  $f(x_i | \theta) = \frac{1}{\sqrt{2\pi} \cdot \sqrt{9}} \exp\left(-\frac{1}{2 \cdot 9} (x - \theta)^2\right)$

$$f(x_i | \theta) = (18\pi)^{-1/2} \cdot \exp\left(-\frac{1}{18} (x - \theta)^2\right)$$

Densitas bersama dari  $X_1, X_2, \dots, X_{10}$  :

$$f(\underline{x} | \theta) = f(x_1 | \theta) \cdot f(x_2 | \theta) \cdot \dots \cdot f(x_{10} | \theta)$$

$$f(\underline{x} | \theta) = (18\pi)^{-1/2} \exp\left(-\frac{(x_1 - \theta)^2}{18}\right) \cdot \dots \cdot (18\pi)^{-1/2} \exp\left(-\frac{(x_{10} - \theta)^2}{18}\right)$$

$$f(\underline{x} | \theta) = (18\pi)^{-10/2} \cdot \exp\left(-\sum_{i=1}^{10} \frac{(x_i - \theta)^2}{18}\right)$$

$$\lambda(\underline{x}, 0.25, 1.95) = \frac{f(\underline{x} | 0.25)}{f(\underline{x} | 1.95)}$$

$$= \frac{(18\pi)^{-10/2} \cdot \exp\left(-\sum_{i=1}^{10} \frac{(x_i - 0.25)^2}{18}\right)}{(18\pi)^{-10/2} \cdot \exp\left(-\sum_{i=1}^{10} \frac{(x_i - 1.95)^2}{18}\right)}$$

$$= \exp\left(\left[\sum_{i=1}^{10} \frac{(x_i - 1.95)^2}{18} - \sum_{i=1}^{10} \frac{(x_i - 0.25)^2}{18}\right]\right)$$

$$= \exp\left(\frac{(0.25 - 1.95)2 \sum_{i=1}^{10} x_i}{18} + \frac{(1.95^2 - 0.25^2)10}{18}\right)$$

$$= \exp\left(-0.1889 \sum_{i=1}^{10} x_i\right) \cdot \exp(0.2078)$$

$$= 1.2309 \cdot \exp\left(-0.1889 \sum_{i=1}^{10} x_i\right)$$

Tolak  $H_0$  jika :  $\lambda\left(\bar{x}, 0.25, 1.95\right) \leq k$

$$1.2309 \cdot \exp\left(-0.1889 \sum_{i=1}^{10} x_i\right) \leq k$$

$$\exp\left(-0.1889 \sum_{i=1}^{10} x_i\right) \leq k_1 \dots \text{dengan} \dots k_1 = k \cdot (1.2309)^{-1}$$

$$-0.1889 \sum_{i=1}^{10} x_i \leq k_2 \dots \text{dengan} \dots k_2 = \ln k_1$$

$$\sum_{i=1}^{10} x_i \geq k_3 \dots \text{dengan} \dots k_3 = k_2 \cdot (-0.1889)^{-1}$$

$$\bar{x} \geq c \dots \text{dengan} \dots c = k_3 / 10$$

Diperoleh fungsi uji berikut :

$$\phi\left(\bar{X}\right) = \begin{cases} 1, & \bar{X} \geq c \\ 0, & \bar{X} < c \end{cases}$$

Harga C ditentukan melalui definisi kesalahan tipe I ( $\alpha$ ) , dalam masalah ini diambil  $\alpha = 5\%$  , sebagai berikut :

$$\alpha = P(\text{menolak } H_0 \mid H_0 \text{ benar})$$

$$0.05 = P\left(\bar{X} \geq c \mid \theta = 0.25\right)$$

$$0.05 = P\left(\frac{\bar{X} - 0.25}{3/\sqrt{10}} \geq \frac{c - 0.25}{3/\sqrt{10}}\right)$$

$$0.05 = P\left(Z \geq \frac{c - 0.25}{3/\sqrt{10}}\right)$$

$$\frac{c - 0.25}{3/\sqrt{10}} = 1.645$$

$$c = 1.8106$$

Sehingga diperoleh uji paling kuasa dengan  $\alpha = 5\%$  sebagai berikut :

$$\phi\left(\bar{X}\right) = \begin{cases} 1, & \bar{X} \geq 1.8106 \\ 0, & \bar{X} < 1.8106 \end{cases}$$

b. Berdasarkan data sampel diperoleh :  $\bar{X} = 1.858$  sehingga disimpulkan tolak  $H_0$ .

c. Kuasa uji jika  $\theta = 1.95$  adalah :

$$\begin{aligned} 1 - \beta &= 1 - P(\bar{X} < 1.8106 | \theta = 1.95) \\ &= 1 - P\left(\frac{\bar{X} - 1.95}{3/\sqrt{10}} < \frac{1.8106 - 1.95}{3/\sqrt{10}}\right) \\ &= 1 - P(Z < -0.15) \\ &= 1 - 0.4404 = 0.5596 \end{aligned}$$

### Contoh 2.2

Misalkan  $X_1, X_2, \dots, X_{10}$  Sampel random dari populasi berdistribusi Eksponensial dengan rata-rata  $\theta$ .

Diambil hipotesis :  $H_0 : \theta = 5$  vs  $H_1 : \theta = 1.5$

- Tentukan uji paling kuasa dari hipotesis tersebut dengan  $\alpha = 2.5\%$ .
- Bila dipunyai data : 1.34265 1.61195 2.91253 1.20288 3.21883 6.19548 0.36283 2.96823 1.67552 1.57278 bagaimana kesimpulan anda.
- Tentukan kuasa ujinya jika  $\theta = 1.5$

Penyelesaian :

a.  $X_i \sim \text{Exp}(\theta)$  maka densitasnya adalah :  $f(x_i | \theta) = \frac{1}{\theta} \exp\left(-\frac{x_i}{\theta}\right)$

Densitas bersama dari  $X_1, X_2, \dots, X_{10}$  adalah :

$$\begin{aligned} f(\underline{x} | \theta) &= f(x_1 | \theta).f(x_2 | \theta). \dots .f(x_{10} | \theta) \\ &= \frac{1}{\theta} \exp\left(-\frac{x_1}{\theta}\right). \frac{1}{\theta} \exp\left(-\frac{x_2}{\theta}\right). \dots . \frac{1}{\theta} \exp\left(-\frac{x_{10}}{\theta}\right) \\ &= \theta^{-10} \cdot \exp\left(-\frac{1}{\theta} \sum_{i=1}^{10} x_i\right) \end{aligned}$$

$$\lambda(\underline{x}, 5, 1.5) = \frac{f(\underline{x} | 5)}{f(\underline{x} | 1.5)}$$

$$\begin{aligned}
 &= \frac{5^{-10} \cdot \exp\left(-\frac{1}{5} \sum_{i=1}^{10} x_i\right)}{1.5^{-10} \cdot \exp\left(-\frac{1}{1.5} \sum_{i=1}^{10} x_i\right)} \\
 &= 0.0000059 \cdot \exp\left(\left(\frac{1}{1.5} - \frac{1}{5}\right) \sum_{i=1}^{10} x_i\right) \\
 &= 0.0000059 \cdot \exp\left(0.4667 \sum_{i=1}^{10} x_i\right)
 \end{aligned}$$

Tolak  $H_0$  jika  $\lambda(\underline{x}, 5, 1.5) \leq k$

$$0.0000059 \cdot \exp\left(0.4667 \sum_{i=1}^{10} x_i\right) \leq k$$

$$\exp\left(0.4667 \sum_{i=1}^{10} x_i\right) \leq k_1 \dots \text{dengan : } k_1 = k / 0.0000059$$

$$0.4667 \sum_{i=1}^{10} x_i \leq k_2 \dots \text{dengan : } k_2 = \ln k_1$$

$$\sum_{i=1}^{10} x_i \leq c \dots \text{dengan : } c = k_2 / 0.4667$$

Diperoleh fungsi uji : 
$$\phi(\underline{X}) = \begin{cases} 1, & \sum_{i=1}^{10} X_i \leq c \\ 0, & \sum_{i=1}^{10} X_i > c \end{cases}$$

Harga C ditentukan melalui definisi kesalahan tipe I ( $\alpha$ ) , dalam masalah ini diambil  $\alpha = 2.5\%$  , sebagai berikut :

$$P\left(\sum_{i=1}^{10} X_i \leq c \mid \theta = 5\right) = 0.025$$

Akan dicari bentuk transformasi yang memuat  $\sum_{i=1}^{10} X_i$  dan  $\theta$  sehingga distribusinya mudah (telah ditabelkan) dengan menggunakan fungsi pembangkit momen berikut :

$X \sim \text{Eks}(\theta)$

$$M_x(t) = \int_0^{\infty} e^{tx} \frac{1}{\theta} e^{-\frac{x}{\theta}} dx$$

$$\begin{aligned}
 &= \frac{1}{(1-\theta t)} \int_0^{\infty} e^{-\frac{x}{\theta}(1-\theta)} d \frac{x}{\theta} (1-\theta t) \\
 &= (1-\theta t)^{-1} \cdot \Gamma(1) = (1-\theta t)^{-1}
 \end{aligned}$$

Karena :  $X_1, X_2, \dots, X_{10}$  independen maka :

$$\begin{aligned}
 M_{\sum_{i=1}^{10} X_i}(t) &= M_{X_1}(t) M_{X_2}(t) \dots M_{X_{10}}(t) \\
 &= (1-\theta t)^{-1} \cdot (1-\theta t)^{-1} \dots (1-\theta t)^{-1} = (1-\theta t)^{-10}
 \end{aligned}$$

$M_{2 \sum_{i=1}^{10} X_i / \theta}(t) = (1-2t)^{-20/2}$  (merupakan fungsi pembangkit momen dari

distribusi Chi-kuadrat dengan derajat bebas 20)

sehingga :  $\frac{2 \sum_{i=1}^{10} X_i}{\theta} \sim \chi_{20}^2$

$$P\left(\sum_{i=1}^{10} X_i \leq c \mid \theta = 5\right) = 0.025$$

$$P\left(\frac{2 \sum_{i=1}^{10} X_i}{5} \leq \frac{2c}{5}\right) = 0.025$$

$$P\left(\chi_{20}^2 \leq \frac{2c}{5}\right) = 0.025$$

Sehingga :  $\frac{2c}{5} = 9.59; \dots$  diperoleh :  $c = 23.975$

Sehingga diperoleh uji paling kuasa dengan  $\alpha = 2.5\%$  sebagai berikut :

$$\phi(\tilde{X}) = \begin{cases} 1, & \sum_{i=1}^{10} X_i \leq 23.975 \\ 0, & \sum_{i=1}^{10} X_i > 23.975 \end{cases}$$

b. Dari data diperoleh :  $\sum_{i=1}^{10} X_i = 22.76368$  , sehingga disimpulkan tolak  $H_0$ .

c. Kuasa uji jika  $\theta=1.95$

$$1 - \beta = P\left(\sum_{i=1}^{10} X_i > 23.975 \mid \theta = 1.5\right)$$

$$\begin{aligned}
 &= 1 - P\left(\frac{2\sum_{i=1}^{10} X_i}{1.5} > \frac{2 \times 23.975}{1.5}\right) \\
 &= 1 - P(\chi_{20}^2 > 31.9667) \\
 &= 1 - 0.0449 = 0.9551
 \end{aligned}$$

Penggunaan dari lemma Neyman-Pearson kebanyakan pada kasus uji dengan  $\tilde{X}$  sampel random dari suatu populasi tertentu (densitasnya tertentu) dengan harga parameter yang berbeda, yaitu dengan bentuk  $H_0 : f(x|\theta_0)$  vs  $H_1 : f(x|\theta_1)$ . Perhatikan contoh 3.2.1 mempunyai bentuk  $H_0 : N(\theta_0, 9)$  vs  $H_1 : N(\theta_1, 9)$  dengan  $\theta_0 = 0.25$  dan  $\theta_1 = 1.95$ , sedangkan pada contoh 3.2.2 mempunyai bentuk :  $H_0 : \text{Exp}(\theta_0)$  vs  $H_1 : \text{Exp}(\theta_1)$  dengan  $\theta_0 = 5$  dan  $\theta_1 = 1.5$ . Kedua contoh tersebut dihipotesiskan bahwa sampel random berasal dari distribusi yang sama tetapi parameternya berbeda, sekarang akan diberikan contoh dimana sampel random dihipotesiskan bukan dari distribusi yang sama tetapi memang dari distribusi yang berbeda dengan hipotesis :

$H_0 : f_0(x|\theta_0)$  vs  $H_1 : f_1(x|\theta_1)$  sebagai berikut :

**Contoh 2.3**

Dipunyai suatu sampel random berukuran  $n$ , akan diuji  $H_0 : X \sim \text{Unif}(0,1)$  vs  $H_1 : X \sim \text{Exp}(1)$ .

**Penyelesaian :**

$X_i \sim \text{Unif}(0,1)$  maka densitas dari  $X_i$  adalah :  $f(x_i) = 1$  sehingga densitas bersama dari  $X_1, X_2, \dots, X_n$  adalah :  $f(\tilde{x}) = f(x_1).f(x_2) \dots f(x_n) = 1$ .

$X_i \sim \text{Exp}(1)$  maka densitas dari  $X_i$  adalah :  $f(x_i) = e^{-x_i}$  sehingga densitas bersama dari  $X_1, X_2, \dots, X_n$  adalah :

$$\begin{aligned}
 f(\tilde{x}) &= f(x_1).f(x_2) \dots f(x_n). \\
 &= e^{-x_1}.e^{-x_2} \dots e^{-x_n} = \exp\left(-\sum_{i=1}^n x_i\right)
 \end{aligned}$$

$$\begin{aligned}\lambda(\underline{x}, f_0, f_1) &= \frac{f_0(\underline{x})}{f_1(\underline{x})} \\ &= \frac{1}{\exp\left(-\sum_{i=1}^n x_i\right)} = \exp\left(\sum_{i=1}^n x_i\right)\end{aligned}$$

Tolak  $H_0$  jika :  $\lambda(\underline{x}, f_0, f_1) \leq k$

$$\exp\left(\sum_{i=1}^n x_i\right) \leq k$$

$$\sum_{i=1}^n x_i \leq c \dots \text{dengan : } c = \ln k$$

Sehingga diperoleh fungsi uji :

$$\phi(\underline{X}) = \begin{cases} 1, & \sum_{i=1}^n X_i \leq c \\ 0, & \sum_{i=1}^n X_i > c \end{cases}$$

$$\text{dengan : } P\left(\sum_{i=1}^n X_i \leq c \mid \text{Unif}(0,1)\right) = \alpha$$

Untuk menentukan C digunakan teorema limit pusat sebagai berikut:

$X_i \sim \text{Unif}(0,1)$  maka  $E(X_i) = 1/2$  dan  $\text{Var}(X_i) = 1/12$

$$\frac{\sum_{i=1}^n X_i - E\left(\sum_{i=1}^n X_i\right)}{\sigma\left(\sum_{i=1}^n X_i\right)} \sim N(0,1)$$

$$\frac{\sum_{i=1}^n X_i - n \cdot \frac{1}{2}}{\sqrt{n \cdot \frac{1}{12}}} \sim N(0,1)$$

$$\sqrt{12 \cdot n} \left( \bar{X} - \frac{1}{2} \right) \sim N(0,1)$$

$$P\left(\sum_{i=1}^n X_i \leq c \mid \text{Unif}(0,1)\right) = \alpha$$

$$P\left(\sqrt{12 \cdot n} \left( \bar{X} - \frac{1}{2} \right) \leq \sqrt{12 \cdot n} \left( \frac{c}{n} - \frac{1}{2} \right)\right) = \alpha$$

$$P\left(Z \leq \sqrt{12 \cdot n} \left( \frac{c}{n} - \frac{1}{2} \right)\right) = \alpha$$

$$\text{sehingga: } \sqrt{12.n} \left( \frac{c}{n} - \frac{1}{2} \right) = -Z_{\alpha/2}$$

Dari persamaan terakhir akan diperoleh harga C, selanjutnya bisa diperoleh uji paling kuasa untuk hipotesis :  $H_0 : X \sim \text{Unif}(0,1)$  vs  $H_1 : X \sim \text{Exp}(1)$ .

### 3 Uji Paling Kuasa Seragam (Uniformly Most Powerful Test / UMP Test)

Konsep dari uji paling kuasa diperluas untuk kasus hipotesis alternative berbentuk composite. Jika suatu uji paling kuasa dengan hipotesis alternative berbentuk composite satu sisi maka uji tersebut menjadi paling kuasa seragam.

#### Definisi 3.1

Misalkan  $X_1, X_2, \dots, X_n$  mempunyai densitas bersama  $f(x|\theta)$  untuk

$\theta \in \Omega$ , pandang uji hipotesis dengan bentuk :

$H_0 : \theta \in \Omega_0$  vs  $H_1 : \theta \in \Omega_0^c$ .  $R^*$  merupakan daerah penolakan yang sesuai uji tersebut disebut uji paling kuasa seragam ukuran  $\alpha$  jika :

$$\max_{\theta \in \Omega_0} \pi_{R^*}(\theta) = \alpha \quad ; \quad \pi_{R^*}(\theta) \geq \pi_R(\theta); \forall \theta \in \Omega_0^c \quad \text{dan daerah}$$

penolakan R ukuran  $\alpha$ .

Uji paling kuasa seragam ada untuk kasus hipotesis alternatif berbentuk composite satu sisi. Cara untuk menentukan uji paling kuasa seragam ini pertama-tama menentukan rasio seperti pada Lemma Neyman-Pearson untuk hipotesis alternatif dengan harga sebarang dan kemudian ditunjukkan bahwa uji tidak tergantung pada harga hipotesis alternatif yang spesifik. Untuk lebih jelasnya perhatikan contoh berikut :

#### Contoh 3.3.1

Misalkan  $X_1, X_2, \dots, X_n$  sampel random dari populasi berdistribusi normal dengan rata-rata  $\theta$  dan variansi 9. Diambil hipotesis  $H_0 : \theta = \theta_0$  vs  $H_1 : \theta > \theta_0$ .

Tentukan uji paling kuasa seragam ukuran  $\alpha$  untuk hipotesis tersebut.

**Penyelesaian :**

Hipotesis pada masalah ini dapat dinyatakan sebagai  $H_0 : \theta = \theta_0$  vs  $H_1 : \theta = \theta_1$  dengan  $\theta_1 > \theta_0$ . Berdasarkan contoh 3.2.1 dengan mengambil sebarang harga  $\theta_1 > \theta_0$  diperoleh uji paling kuasa :

$$\phi(\underline{X}) = \begin{cases} 1, & \bar{X} \geq c \\ 0, & \bar{X} < c \end{cases}$$

Dalam hal ini uji tersebut berlaku untuk semua  $\theta_1 > \theta_0$  tetapi bukan harga  $\theta_1$  tertentu, maka uji tersebut menjadi uji paling kuasa seragam untuk hipotesis  $H_0 : \theta = \theta_0$  vs  $H_1 : \theta > \theta_0$ .

Berikut ini akan dibicarakan sifat momotone likelihood ratio (MLR) yang sangat berguna untuk menentukan uji paling kuasa seragam.

### Definisi 3.2

Densitas bersama  $f(\underline{x} | \theta)$  dikatakan mempunyai sifat momotone likelihood ratio (MLR) dalam statistik  $T(\underline{X})$  jika dua harga parameter  $\theta_1 < \theta_2$  rasio  $f(\underline{x} | \theta_2) / f(\underline{x} | \theta_1)$  tergantung  $\underline{X}$  hanya melalui  $T(\underline{X})$  dan rasio tersebut merupakan fungsi tidak turun dari  $T(\underline{X})$ .

### Contoh 3.2

Misalkan  $X_1, X_2, \dots, X_n$  sampel random independen dari populasi berdistribusi Eksponensial dengan parameter  $\theta$ . Akan ditunjukkan mempunyai sifat monotone likelihood ratio (MLR).

**Penyelesaian :**

$$X_i \sim \text{Exp}(\theta) \text{ maka densitasnya adalah : } f(x_i | \theta) = \frac{1}{\theta} \exp\left(-\frac{x_i}{\theta}\right)$$

Densitas bersama dari  $X_1, X_2, \dots, X_n$  adalah :

$$\begin{aligned} f(\underline{x} | \theta) &= f(x_1 | \theta) \cdot f(x_2 | \theta) \cdot \dots \cdot f(x_n | \theta) \\ &= \frac{1}{\theta} \exp\left(-\frac{x_1}{\theta}\right) \cdot \frac{1}{\theta} \exp\left(-\frac{x_2}{\theta}\right) \cdot \dots \cdot \frac{1}{\theta} \exp\left(-\frac{x_n}{\theta}\right) \\ &= \theta^{-n} \cdot \exp\left(-\frac{1}{\theta} \sum_{i=1}^n x_i\right) \end{aligned}$$

$$\begin{aligned} \frac{f(\underline{x} | \theta_2)}{f(\underline{x} | \theta_1)} &= \frac{\theta_2^{-n} \cdot \exp\left(-\frac{1}{\theta_2} \sum_{i=1}^n x_i\right)}{\theta_1^{-n} \cdot \exp\left(-\frac{1}{\theta_1} \sum_{i=1}^n x_i\right)} \\ &= \left(\frac{\theta_1}{\theta_2}\right)^n \exp\left(-\sum_{i=1}^n x_i \left(\frac{1}{\theta_2} - \frac{1}{\theta_1}\right)\right) \end{aligned}$$

Ambil  $T(\underline{X}) = \sum_{i=1}^n X_i$ , maka ratio dari  $\frac{f(\underline{x} | \theta_2)}{f(\underline{x} | \theta_1)}$  tidak turun jika  $\theta_1 < \theta_2$ .

Sehingga  $f(\underline{x} | \theta)$  mempunyai sifat MLR dalam statistik  $T(\underline{X}) = \sum_{i=1}^n X_i$ .

### Teorema 3.1

Jika densitas bersama  $f(\underline{x} | \theta)$  mempunyai sifat MLR dalam statistik

$T(\underline{X})$  maka uji paling kuasa seragam ukuran  $\alpha$  untuk :

$H_0 : \theta \leq \theta_0$  vs  $H_1 : \theta > \theta_0$  adalah :

$$\phi(\underline{X}) = \begin{cases} 1 & T(\underline{X}) \geq k \\ 0 & \text{yanglain} \end{cases}$$

dengan  $k$  ditentukan dari  $P(T(\underline{X}) \geq k | \theta = \theta_0) = \alpha$ .

Dengan mengingat sifat baik dari keluarga eksponensial diantaranya sifat MLR ini, yaitu setiap anggota keluarga eksponensial regular dengan bentuk :

$f(x|\theta) = h(x)c(\theta). \exp\{q(\theta)t(x)\}$  mempunyai sifat MLR dalam  $T(\underline{X})$  bila  $q(\theta)$  fungsi tidak turun dari  $\theta$ . Dari hal ini diperoleh teorema sebagai berikut :

**Teorema 3.2**

Andaikan  $X_1, X_2, \dots, X_n$  mempunyai densitas bersama dengan bentuk :

$$f(x|\theta) = h(x)c(\theta). \exp\{q(\theta)t(x)\} \text{ dengan } q(\theta) \text{ fungsi naik dari } \theta.$$

1. Uji paling kuasa seragam ukuran  $\alpha$  untuk :  $H_0 : \theta \leq \theta_0$  vs  $H_1 : \theta > \theta_0$

$$\text{adalah : } \phi(\underline{X}) = \begin{cases} 1 & T(\underline{X}) \geq k \\ 0 & \text{yanglain} \end{cases}$$

$$\text{dengan } P(T(\underline{X}) \geq k | \theta = \theta_0) = \alpha$$

2. Uji paling kuasa seragam ukuran  $\alpha$  untuk :  $H_0 : \theta \geq \theta_0$  vs  $H_1 : \theta < \theta_0$

$$\text{adalah : } \phi(\underline{X}) = \begin{cases} 1 & T(\underline{X}) \leq k \\ 0 & \text{yanglain} \end{cases}$$

$$\text{dengan } P(T(\underline{X}) \leq k | \theta = \theta_0) = \alpha$$

**Contoh 3.3**

Misalkan  $X_1, X_2, \dots, X_n$  sampel random independen dari populasi berdistribusi normal dengan rata-rata 5 dan variansi  $\theta$ .

Ambil  $H_0 : \theta \leq \theta_0$  vs  $H_1 : \theta > \theta_0$

- Tentukan uji paling kuasa seragam ukuran  $\alpha$  untuk hipotesis tersebut.
- Jika  $\theta_0 = 20$ ,  $n = 10$  dan  $\alpha = 2,5\%$  tentukan uji paling kuasa seragam.
- Jika dipunyai data : 3,1095 8,9175 5,9284 9,4713 7,6403 -3,1282 3,4169 8,4759 -1,5752 12,9482 bagaimana kesimpulan anda.
- Tentukan kuasa ujinya jika  $\theta = 75$ .

**Penyelesaian :**

- a.  $X_i \sim N(5, \theta)$ , maka densitasnya adalah :

$$f(x_i | \theta) = (2\pi)^{-1/2} \theta^{-1/2} \exp\left(-\frac{1}{2\theta}(x_i - 5)^2\right)$$

Densitas bersama dari  $X_1, X_2, \dots, X_n$  adalah :

$$\begin{aligned} f(\underline{x} | \theta) &= f(x_1 | \theta) \cdot f(x_2 | \theta) \cdot \dots \cdot f(x_n | \theta) \\ &= (2\pi)^{-1/2} \theta^{-1/2} \exp\left(-\frac{1}{2\theta}(x_1-5)^2\right) \cdot \dots \cdot (2\pi)^{-1/2} \theta^{-1/2} \exp\left(-\frac{1}{2\theta}(x_n-5)^2\right) \\ &= (2\pi)^{-n/2} \theta^{-n/2} \exp\left(-\frac{1}{2\theta} \sum_{i=1}^n (x_i-5)^2\right) \end{aligned}$$

Densitas bersama dari  $X_1, X_2, \dots, X_n$  ini memenuhi syarat pada teorema 3.3.2 dengan mengambil :

$$h(\underline{x}) = (2\pi)^{-n/2}$$

$$c(\theta) = \theta^{-n/2}$$

$$t(\underline{x}) = \sum_{i=1}^n (x_i-5)^2$$

$$q(\theta) = -\frac{1}{2\theta}$$

dalam hal ini  $q(\theta)$  merupakan fungsi yang naik dari  $\theta$ .

Uji paling kuasa seragam dari hipotesis tersebut adalah :

$$\phi(\underline{X}) = \begin{cases} 1 & \sum_{i=1}^n (X_i-5)^2 \geq c \\ 0 & \text{yanglain} \end{cases}$$

dengan C ditentukan dari :

$$P\left(\sum_{i=1}^n (X_i-5)^2 \geq c \mid \theta = \theta_0\right) = \alpha$$

$$P\left(\frac{1}{\theta_0} \sum_{i=1}^n (X_i-5)^2 \geq \frac{c}{\theta_0}\right) = \alpha$$

$$P\left(\chi_n^2 \geq \frac{c}{\theta_0}\right) = \alpha$$

dari persamaan terakhir ini akan diperoleh harga C.

b. Jika  $\theta_0 = 20$ ;  $n = 10$  dan  $\alpha = 2,5\%$  maka  $C/20 = 20,48$  sehingga  $C = 409,6$

Uji paling kuasa seragamnya adalah :

$$\phi(\underline{X}) = \begin{cases} 1 & \sum_{i=1}^n (X_i - 5)^2 \geq 409,6 \\ 0 & \text{yanglain} \end{cases}$$

c. Berdasarkan data diperoleh ;  $\sum_{i=1}^{10} (X_i - 5)^2 = 233,8093$  sehingga  $H_0$  diterima.

d. Kuasa uji jika  $\theta = 75$  adalah :

$$\begin{aligned} 1 - \beta &= 1 - P\left(\sum_{i=1}^{10} (X_i - 5)^2 < 409,6 \mid \theta = 75\right) \\ &= 1 - P\left(\frac{1}{75} \sum_{i=1}^{10} (X_i - 5)^2 < \frac{409,6}{75}\right) \\ &= 1 - P(\chi_{10}^2 < 5,46) \\ &= 1 - 0,1452 = 0,8548 \end{aligned}$$

#### 4. Uji Rasio Likelihood (Generalized Likelihood Ratio Test/ GLRT)

Lemma Neyman-Pearson merupakan suatu metode untuk menentukan uji paling kuasa dari hipotesis sederhana , kadang-kadang juga untuk uji paling kuasa seragam bila hipotesis alternatifnya merupakan bentuk composite satu sisi. Teorema 3.1 dan 3.2 digunakan untuk menentukan uji paling kuasa seragam dengan satu parameter yang tidak diketahui dengan hipotesis berbentuk composite satu sisi.

Diperlukan suatu metode untuk menentukan suatu uji dimana dalam uji tersebut terdapat parameter pengganggu (nuisance) atau untuk menangani hipotesis alternatif berbentuk composite dua sisi. Metode rasio likelihood dalam uji hipotesis berhubungan dengan estimator likelihood maksimum dan uji tersebut secara luas dapat diterapkan. Generalized likelihood ratio test merupakan perluasan dari uji Neyman-Pearson.

##### Definisi 4.1

Bila  $X_1, X_2, \dots, X_n$  mempunyai densitas bersama  $f(x \mid \theta)$  untuk

$\theta \in \Omega$ , pandang uji hipotesis  $H_0 : \theta \in \Omega_0$  vs  $H_1 : \theta \in \Omega_0^c$ .

Generalized likelihood ratio (GLR) didefinisikan dengan :

$$\lambda(\underline{x}) = \frac{\max_{\theta \in \Omega_0} f(\underline{x} | \theta)}{\max_{\theta \in \Omega} f(\underline{x} | \theta)} = \frac{f(\underline{x} | \hat{\theta}_0)}{f(\underline{x} | \hat{\theta})}$$

dengan  $\hat{\theta}$  adalah estimator likelihood dari  $\theta$ , dan  $\hat{\theta}_0$  adalah estimator likelihood dari  $\theta$  dengan batasan bahwa  $H_0$  benar.

Generalized likelihood ratio test (GLRT) adalah tolak  $H_0$  jika  $\lambda(\underline{x}) \leq k$  dengan  $k$  dipilih sehingga uji mempunyai ukuran  $\alpha$ .  $\lambda(\underline{x})$  merupakan statistic uji berupa fungsi yang tidak tergantung dari parameter yang tidak diketahui. Dalam beberapa kasus, distribusi dari  $\lambda(\underline{x})$  bebas dari parameter dan harga kritis  $k$  dapat ditentukan secara exact. Pada kasus distribusi  $\lambda(\underline{x})$  dibawah  $H_0$  benar, masih tergantung dari parameter yang tidak diketahui, daerah kritis  $k$  secara exact tidak dapat ditentukan, sehingga dengan kondisi regularitas digunakan distribusi asyptotik dari  $\lambda(\underline{x})$ .

Jika  $X_1, X_2, \dots, X_n$  mempunyai densitas bersama  $f(\underline{x} | \theta_1, \theta_2, \dots, \theta_k)$ , diambil hipotesis:  $H_0 : (\theta_1, \theta_2, \dots, \theta_r) = (\theta_{10}, \theta_{20}, \dots, \theta_{r0})$  dengan  $r < k$ , untuk  $n$  besar maka  $-2 \ln(\lambda(\underline{x})) \sim \chi_r^2$  sehingga tolak  $H_0$  jika  $-2 \ln(\lambda(\underline{x})) \geq \chi_{1-\alpha; r}^2$

#### Contoh 4.1

Misalkan  $X_1, X_2, \dots, X_n$  sampel random dari populasi berdistribusi normal dengan rata-rata  $\theta$ , dan variansi 9. Ambil  $H_0 : \theta = \theta_0$  vs  $H_1 : \theta \neq \theta_0$ . Tentukan GLRT ukuran  $\alpha$ .

Penyelesaian :

Karena :  $X_i \sim N(\theta, 9)$

maka densitasnya :  $f(x_i | \theta) = \frac{1}{\sqrt{2\pi} \cdot \sqrt{9}} \exp\left(-\frac{1}{2 \cdot 9} (x - \theta)^2\right)$

$$f(x_i | \theta) = (18\pi)^{-1/2} \cdot \exp\left(-\frac{1}{18}(x-\theta)^2\right)$$

Densitas bersama dari  $X_1, X_2, \dots, X_n$  :

$$f(\underline{x} | \theta) = f(x_1 | \theta) \cdot f(x_2 | \theta) \cdot \dots \cdot f(x_n | \theta)$$

$$f(\underline{x} | \theta) = (18\pi)^{-1/2} \exp\left(-\frac{(x_1-\theta)^2}{18}\right) \cdot \dots \cdot (18\pi)^{-1/2} \exp\left(-\frac{(x_n-\theta)^2}{18}\right)$$

$$f(\underline{x} | \theta) = (18\pi)^{-n/2} \cdot \exp\left(-\sum_{i=1}^n \frac{(x_i-\theta)^2}{18}\right)$$

Telah diperoleh bahwa estimator maksimum likelihood untuk  $\theta$  adalah  $\bar{X}$ .

Untuk  $\theta \in \Omega_0$ ,  $\hat{\theta}_{\Omega_0} = \theta_0$

$$\max_{\theta \in \Omega_0} f(\underline{x} | \theta) = (18\pi)^{-n/2} \exp\left(-\frac{1}{18} \sum_{i=1}^n (x_i - \theta)^2\right)$$

Untuk  $\theta \in \Omega$ ,  $\hat{\theta}_{\Omega} = \bar{X}$

$$\max_{\theta \in \Omega_0} f(\underline{x} | \theta) = (18\pi)^{-n/2} \exp\left(-\frac{1}{18} \sum_{i=1}^n (x_i - \bar{X})^2\right)$$

$$\lambda(\underline{x}) = \frac{\max_{\theta \in \Omega_0} f(\underline{x} | \theta)}{\max_{\theta \in \Omega} f(\underline{x} | \theta)} = \frac{f(\underline{x} | \hat{\theta}_0)}{f(\underline{x} | \hat{\theta})}$$

$$= \frac{(18\pi)^{-n/2} \exp\left(-\frac{1}{18} \sum_{i=1}^n (x_i - \theta_0)^2\right)}{(18\pi)^{-n/2} \exp\left(-\frac{1}{18} \sum_{i=1}^n (x_i - \bar{X})^2\right)}$$

$$= \exp\left(-\frac{1}{18} \left\{ \sum_{i=1}^n (x_i - \theta_0)^2 - \sum_{i=1}^n (x_i - \bar{X})^2 \right\}\right)$$

$$= \exp\left(-\frac{1}{18} \left\{ \sum_{i=1}^n (x_i - \bar{X} + \bar{X} - \theta_0)^2 - \sum_{i=1}^n (x_i - \bar{X})^2 \right\}\right)$$

$$= \exp\left(-\frac{1}{18} \left\{ \sum_{i=1}^n (x_i - \bar{X})^2 - 2 \sum_{i=1}^n (x_i - \bar{X})(\bar{X} - \theta_0) + n(\bar{X} - \theta_0)^2 - \sum_{i=1}^n (x_i - \bar{X})^2 \right\}\right)$$

$$= \exp\left(-\frac{1}{2} \cdot \frac{n(\bar{X} - \theta_0)^2}{9}\right)$$

$$= \exp\left(-\frac{1}{2} \left(\frac{\bar{X} - \theta_0}{3/\sqrt{n}}\right)^2\right)$$

Tolak  $H_0$  jika  $\lambda(x) \leq k$

$$\exp\left(-\frac{1}{2} \left(\frac{\bar{X} - \theta_0}{3/\sqrt{n}}\right)^2\right) \leq k$$

$$-\frac{1}{2} \left(\frac{\bar{X} - \theta_0}{3/\sqrt{n}}\right)^2 \leq k_1 \quad \text{dengan} \quad k_1 = \log k$$

$$\left(\frac{\bar{X} - \theta_0}{3/\sqrt{n}}\right)^2 \geq c \quad \text{dengan} \quad c = -2k_1$$

sehingga diperoleh GLRT

$$\phi(\underline{X}) = \begin{cases} 1 & \left(\frac{\bar{X} - \theta_0}{3/\sqrt{n}}\right)^2 \geq c \\ 0 & \text{yanglain} \end{cases}$$

dengan  $c$  ditentukan melalui

$$P\left(\left(\frac{\bar{X} - \theta_0}{3/\sqrt{n}}\right)^2 \geq c\right) = \alpha$$

$$P(\chi_1^2 \geq c) = \alpha$$

$$\text{karena} \quad \frac{\bar{X} - \theta_0}{3/\sqrt{n}} \sim N(0,1)$$

Sehingga GLRT tersebut dapat juga dinyatakan dengan :

$$\phi(\underline{X}) = \begin{cases} 1 & \frac{\bar{X} - \theta_0}{3/\sqrt{n}} \geq Z_{\alpha/2} \text{ atau } \frac{\bar{X} - \theta_0}{3/\sqrt{n}} \leq -Z_{\alpha/2} \\ 0 & \text{yanglain} \end{cases}$$

**Contoh 4.2**

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

Ambil  $H_0 : \mu = \mu_0$  vs  $H_1 : \mu \neq \mu_0$ . Tentukan GLRT ukuran  $\alpha$ .

**Penyelesaian :**

$X_i \sim N(\mu, \sigma^2)$  , maka densitasnya adalah :

$$f(x_i | \mu, \sigma^2) = (2\pi)^{-1/2} (\sigma^2)^{-1/2} \exp\left(-\frac{1}{2\sigma^2}(x_i - \mu)^2\right)$$

Densitas bersama dari  $X_1, X_2, \dots, X_n$  :

$$f(\underline{x} | \mu, \sigma^2) = f(x_1 | \mu, \sigma^2) \cdot f(x_2 | \mu, \sigma^2) \cdot \dots \cdot f(x_n | \mu, \sigma^2)$$

$$f(\underline{x} | \mu, \sigma^2) = (2\pi)^{-1/2} (\sigma^2)^{-1/2} \exp\left(-\frac{1}{2\sigma^2}(x_1 - \mu)^2\right) \cdot \dots \cdot (2\pi)^{-1/2} (\sigma^2)^{-1/2} \exp\left(-\frac{1}{2\sigma^2}(x_n - \mu)^2\right)$$

$$f(\underline{x} | \mu, \sigma^2) = (2\pi)^{-n/2} (\sigma^2)^{-n/2} \cdot \exp\left(-\frac{1}{2\sigma^2} \sum_{i=1}^n (x_i - \mu)^2\right)$$

Telah diperoleh estimator maksimum likelihood :

$$\hat{\mu} = \bar{X} \quad \text{dan} \quad \hat{\sigma}^2 = \frac{1}{n} \sum_{i=1}^n (X_i - \hat{\mu})^2$$

Untuk  $\theta \in \Omega_0$  maka :

$$\hat{\mu}_{\Omega_0} = \mu_0 \quad \hat{\sigma}_{\Omega_0}^2 = \frac{1}{n} \sum_{i=1}^n (X_i - \mu_0)^2$$

$$\begin{aligned} \max_{\theta \in \Omega_0} f(\underline{x} | \mu, \sigma^2) &= (2\pi)^{-n/2} \left( \frac{1}{n} \sum_{i=1}^n (x_i - \mu_0)^2 \right)^{-n/2} \exp \left( - \frac{\sum_{i=1}^n (x_i - \mu_0)^2}{2 \cdot \frac{1}{n} \sum_{i=1}^n (x_i - \mu_0)^2} \right) \\ &= (2\pi)^{-n/2} \left( \frac{1}{n} \sum_{i=1}^n (x_i - \mu_0)^2 \right)^{-n/2} \exp \left( - \frac{n}{2} \right) \end{aligned}$$

Untuk  $\theta \in \Omega$ , maka :

$$\hat{\mu}_{\Omega_0} = \bar{X} \quad \hat{\sigma}_{\Omega_0}^2 = \frac{1}{n} \sum_{i=1}^n (X_i - \bar{X})^2$$

$$\begin{aligned} \max_{\theta \in \Omega_0} f(\underline{x} | \mu, \sigma^2) &= (2\pi)^{-n/2} \left( \frac{1}{n} \sum_{i=1}^n (x_i - \bar{X})^2 \right)^{-n/2} \exp \left( - \frac{\sum_{i=1}^n (x_i - \bar{X})^2}{2 \cdot \frac{1}{n} \sum_{i=1}^n (x_i - \bar{X})^2} \right) \\ &= (2\pi)^{-n/2} \left( \frac{1}{n} \sum_{i=1}^n (x_i - \bar{X})^2 \right)^{-n/2} \exp \left( - \frac{n}{2} \right) \end{aligned}$$

$$\begin{aligned} \lambda(\underline{x}) &= \frac{\max_{\theta \in \Omega_0} f(\underline{x} | \mu, \sigma^2)}{\max_{\theta \in \Omega} f(\underline{x} | \mu, \sigma^2)} \\ &= \frac{(2\pi)^{-n/2} \left( \frac{1}{n} \sum_{i=1}^n (x_i - \mu_0)^2 \right)^{-n/2} \exp \left( - \frac{n}{2} \right)}{(2\pi)^{-n/2} \left( \frac{1}{n} \sum_{i=1}^n (x_i - \bar{X})^2 \right)^{-n/2} \exp \left( - \frac{n}{2} \right)} \\ &= \left( \frac{\sum_{i=1}^n (x_i - \mu_0)^2}{\sum_{i=1}^n (x_i - \bar{X})^2} \right)^{-n/2} \\ &= \left( \frac{\sum_{i=1}^n (x_i - \bar{X})^2 + n(\bar{X} - \mu_0)^2}{\sum_{i=1}^n (x_i - \bar{X})^2} \right)^{-n/2} \end{aligned}$$

$$= \left( 1 + \frac{n(\bar{X} - \mu_0)^2}{(n-1) \cdot \frac{1}{(n-1)} \sum_{i=1}^n (x_i - \bar{X})^2} \right)^{-n/2}$$

$$= \left( 1 + \frac{1}{(n-1)} \cdot t^2 \right)^{-n/2}$$

dengan  $t = \frac{\bar{X} - \mu_0}{S/\sqrt{n}}$       $S = \sqrt{\frac{1}{(n-1)} \sum_{i=1}^n (x_i - \bar{X})^2}$

Tolak  $H_0$  jika  $\lambda(\underline{x}) \leq k$

$$\left( 1 + \frac{1}{(n-1)} \cdot t^2 \right)^{-n/2} \leq k$$

$$1 + \frac{1}{(n-1)} \cdot t^2 \geq k_1 \quad \text{dengan} \quad k_1 = k^{-2/n}$$

$$t^2 \geq c \quad \text{dengan} \quad c = (k_1 - 1)(n-1)$$

Sehingga GLRT :

$$\phi(\underline{X}) = \begin{cases} 1, & T^2 \geq c \\ 0, & \text{yanglain} \end{cases}$$

dengan C ditentukan dari :

$$P(T^2 \geq c) = \alpha$$

$$P(F_{1,(n-1)} \geq c) = \alpha$$

karena  $T = \frac{\bar{X} - \mu_0}{S/\sqrt{n}} \sim t_{n-1}$  maka GLRT dapat dinyatakan :

$$\phi(\underline{X}) = \begin{cases} 1, & \frac{\bar{X} - \mu_0}{S/\sqrt{n}} \geq t_{\alpha/2;n-1} \text{ atau } \frac{\bar{X} - \mu_0}{S/\sqrt{n}} \leq -t_{\alpha/2;n-1} \\ 0, & \text{yanglain} \end{cases}$$

### 3.4.2 Latihan

1. Jelaskan apa bedanya antara hipotesis sederhana dengan hipotesis composite
2. Jelaskan kapan digunakan uji paling kuasa.
3. Jelaskan kapan digunakan uji paling kuasa seragam
4. Jelaskan kapan digunakan GLRT.

### 3.4.3 Rangkuman

1. Dalam uji hipotesis pertama-tama yang dilihat adalah bentuk hipotesisnya untuk menentukan metode yang sesuai.
2. Bila  $H_0$  sederhana dan  $H_1$  sederhana digunakan uji paling kuasa.
3. Bila  $H_0$  sederhana dan  $H_1$  composite satu sisi atau  $H_0$  composite satu sisi dan  $H_1$  composite satu sisi digunakan uji paling kuasa seragam.
4. Dalam uji paling kuasa seragam, distribusi dari populasi harus memenuhi sifat MLR, untuk kasus distribusi dengan 2 parameter atau lebih digunakan GLRT. Dalam hal  $H_0$  sederhana dan  $H_1$  composite dua sisi digunakan GLRT.

### 3.4.4 Tes Formatif

1. Misalkan  $X_1, X_2, \dots, X_{10}$  sampel random dari populasi berdistribusi Gamma  $(2, \theta)$ . Diambil hipotesis  $H_0 : \theta = 3,0$  vs  $H_1 : \theta = 5,2$ .
  - a. Tentukan uji paling kuasa dari hipotesis tersebut dengan  $\alpha = 5\%$ .
  - b. Bila dipunyai data: 4,9 7,4 11,21 25,4 200,5 170,4 69,2 24,1 2,1 30,6 bagaimana kesimpulan anda.
  - c. Tentukan kuasa ujinya jika  $\theta = 5,2$ .
2. Misalkan  $X_1, X_2, \dots, X_8$  sampel random dari populasi berdistribusi normal dengan rata-rata 30 dan variansi  $\theta$ .  
Ambil hipotesis  $H_0 : \theta = 0,87$  vs  $H_1 : \theta = 0,12$ .
  - a. Tentukan uji paling kuasa dari hipotesis tersebut dengan  $\alpha = 2,5\%$ .
  - b. Bila dipunyai data: 29,8 29,7 30,1 30,4 29,4 30,8 30,4 30,2 bagaimana kesimpulan anda.
  - c. Tentukan kuasa ujinya jika  $\theta = 0,12$ .

3. Misalkan  $X_1, X_2, \dots, X_7$  sampel random dari populasi dengan densitas :

$$f(x | \theta) = \frac{2x}{\theta^2} \exp\left(-\frac{x^2}{\theta^2}\right).$$

Diambil hipotesis :  $H_0 : \theta \geq 120$  vs  $H_1 : \theta < 120$ .

- Tentukan uji paling kuasa seragam dari hipotesis tersebut dengan  $\alpha = 2,5\%$ .
  - Bila dipunyai data: 2,1723 9,2142 17,5948 25,6573 51,3431 70,9184 92,2471 bagaimana kesimpulan anda.
  - Tentukan kuasa ujinya jika  $\theta = 50$ .
4. Misalkan  $X_1, X_2, \dots, X_n$  sampel random dari populasi berdistribusi eksponensial dengan rata-rata  $\theta$ . Ambil  $H_0 : \theta = \theta_0$  vs  $H_1 : \theta \neq \theta_0$ . Tentukan GLRT ukuran  $\alpha$ .

### 3.4.5 Umpan Balik dan tidak Lanjut

Cocokkanlah jawaban anda dengan kunci jawaban dari modul ini. Hitunglah jawaban anda yang benar, kemudian gunakanlah rumus dibawah ini untuk mengetahui tingkat penguasaan anda terhadap materi ini.

Tingkat penguasaan = (jumlah jawaban benar)/10 x 100%

Arti tingkat penguasaan yang anda capai :

90% - 100% : baik sekali

80% - 89% : baik

70% - 79% : sedang

< 70% : kurang

Bila anda telah mencapai tingkat penguasaan 80% atau lebih, anda dapat meneruskan kegiatan belajar selanjutnya. Tetapi apabila tingkat penguasaan anda kurang dari 80% , sebaiknya anda mengulangi materi ini terutama bagian yang belum anda kuasai sebelum meneruskan kegiatan belajar selanjutnya.

### 3.5 Jawaban Tes Formatif

1. a. Tolak  $H_0$  jika  $\sum_{i=1}^{10} X_i \geq 83,64$

b. Tolak  $H_0$

c. 80%

2. a. Tolak  $H_0$  jika  $\sum_{i=1}^8 (X_i - 30)^2 \leq 1,8966$

b. Tolak  $H_0$

c. 95,37%

3. a. Tolak  $H_0$  jika  $\sum_{i=1}^7 X_i^2 \leq 40536$

b. Tolak  $H_0$

c. 95%

4. Tolak  $H_0$  jika  $-2n \left[ \ln \frac{\bar{X}}{\theta_0} - \frac{\bar{X}}{\theta_1} + 1 \right] \geq \chi_{\alpha,1}^2$

### 3.6 Referensi

1. Bain, L.J. and Engelhart, M. (1992); Introduction to Probability and Mathematical Statistics, Second Edition. Duxbury Press, Belmont, California.
2. Casella, G and Berger, R.L. (1990); Statistical Inference. Wadsworth Inc. Belmont, California.
3. Dudewicz, E.J and Mishra, S.N. (1998); Modern Mathematical Statistics. John Wiley and Sons, Singapore.
4. Roussas, G.G. (1976); A First Course in Mathematical Statistics. Mei Ya Publication Inc. Taipei, Taiwan.
5. -----, (2007) "[http://en.wikipedia.org/wiki/Statistical\\_hypothesis\\_testing](http://en.wikipedia.org/wiki/Statistical_hypothesis_testing)"

## MODUL IV ESTIMASI INTERVAL

### 4.1 Pengantar

Pada modul ini akan dipelajari bagaimana mengkonstruksikan interval konfindensi, limit konfindensi satu sisi atas dan limit konfindensi satu sisi bawah. Sebelum mempelajari modul ini diharapkan pembaca telah menguasai uji hipotesis dan transformasi peubah acak. Metode penentuan interval konfindensi yang akan dipelajari pada modul ini ada dua cara yaitu metode inversi uji hipotesis dan metode dengan menggunakan besaran pivot.

### 4.2 Tujuan Instruksional Umum :

Setelah mempelajari bagian ini mahasiswa dapat mengkonstruksikan interval konfindensi suatu parameter populasi.

### 4.3 Tujuan Instruksional Khusus :

Setelah mempelajari bagian ini mahasiswa dapat :

1. Menentukan interval konfindensi suatu parameter populasi dengan menggunakan inverse uji hipotesis.
2. Menentukan limit konfindensi satu sisi atas atau bawah dengan menggunakan inverse uji hipotesis.
3. Memilih bentuk transformasi yang sesuai sehingga akan menemukan bentuk besaran pivot.
4. Menentukan interval konfindensi suatu parameter populasi dengan menggunakan metode besaran pivot.
5. Menentukan limit konfindensi satu sisi atas atau bawah dengan menggunakan metode besaran pivot.

#### 4.4 Kegiatan Belajar

### ESTIMASI INTERVAL

#### 4.4.1 Uraian dan Contoh

##### 1. Pengertian.

Estimasi titik untuk parameter  $\theta$  inferensinya adalah pendugaan dengan harga tunggal dari  $\theta$ . Pada modul ini akan dibahas estimasi interval, atau secara lebih umum estimasi himpunan. Inferensi dalam persoalan estimasi himpunan adalah pernyataan bahwa " $\theta \in C$ " dengan  $C \subset \Omega$  dan  $C = C(\underline{x})$  adalah himpunan yang ditentukan oleh harga  $\underline{X} = \underline{x}$  terobservasi. Bila  $\theta$  berharga real, maka biasanya estimator  $C$  berupa interval.

##### Definisi 1.1

Estimasi parameter  $\theta$  adalah pasangan fungsi  $L(\underline{x})$  dan  $U(\underline{x})$  dari sampel yang memenuhi  $L(\underline{x}) \leq U(\underline{x})$  untuk semua  $\underline{x}$ . Jika  $\underline{X} = \underline{x}$  telah terobservasi, dibuat inferensi dengan bentuk  $L(\underline{x}) < \theta < U(\underline{x})$ . Interval random  $(L(\underline{X}), U(\underline{X}))$  disebut estimator interval.

Notasi yang dipergunakan  $(L(\underline{X}), U(\underline{X}))$  untuk estimator interval  $\theta$  berdasarkan sampel random  $\underline{X} = (X_1, X_2, \dots, X_n)$  dan  $(L(\underline{x}), U(\underline{x}))$  harga terealisasi dari interval. Meskipun kasus mayoritas yang dibicarakan harga - harga berhingga untuk  $L(\underline{x})$  dan  $U(\underline{x})$ , kadang-kadang kita juga tertarik pada estimator interval satu sisi. Perhatikan bila  $L(\underline{x}) = -\infty$ ,

maka dipunyai interval satu sisi  $(-\infty, U(\underline{x}))$ , demikian juga bila  $U(\underline{x}) = \infty$ , dipunyai interval satu sisi  $(U(\underline{x}), \infty)$ .

### Contoh 1.1

Misalakan  $X_1, X_2, X_3, X_4$  adalah sampel random dari populasi berdistribusi normal dengan rata-rata  $\theta$  dan variansi 1. Andaikan dipilih estimasi interval untuk  $\theta$  adalah :  $(\bar{X} - 1, \bar{X} + 1)$ , masalahnya adalah berapa peluangnya  $\theta$  akan tercakup dalam interval tersebut.

Peluang bahwa  $\theta$  akan tercakup dalam interval  $(\bar{X} - 1, \bar{X} + 1)$  dapat dihitung sebagai berikut :

$$\begin{aligned} P(\theta \in (\bar{X} - 1, \bar{X} + 1)) &= P[\bar{X} - 1 < \theta < \bar{X} + 1] \\ &= P[-1 < -\bar{X} + \theta < 1] \\ &= P[1 > \bar{X} - \theta > -1] \\ &= P\left[-\frac{1}{\sqrt{1/4}} < \frac{\bar{X} - \theta}{\sqrt{1/4}} < 1\right] \\ &= P[-2 < Z < 2] \text{ dengan } Z \sim N(0,1) \\ &= 0,9544 \end{aligned}$$

Jadi kita mempunyai lebih dari 95% peluang untuk mencakup parameter  $\theta$ .

Maksud penggunaan estimator interval sebagai pengganti estimator titik adalah untuk mendapat jaminan pencakupan parameter yang diselidiki. Kepastian jaminan ini dikuantifikasikan dalam definisi berikut :

### Definisi 1.2

Untuk estimator interval  $(L(\underline{x}), U(\underline{x}))$  dari parameter  $\theta$ , cakupan peluang dari  $(L(\underline{x}), U(\underline{x}))$  adalah peluang bahwa interval random

$(L(\underline{x}), U(\underline{x}))$  mencakup parameter  $\theta$ , dinotasikan dengan :  
 $P(\theta \in (L(\underline{X}), U(\underline{X})))$ .

## 2. Interval Konfindensi.

Misalkan  $X_1, X_2, \dots, X_n$  mempunyai densitas bersama  $f(\underline{x}|\theta)$ , untuk  $\theta \in \Omega$  dengan  $\Omega$  suatu interval,  $L(\underline{X})$  dan  $U(\underline{X})$  suatu statistik, berdasarkan data sampel  $x_1, x_2, \dots, x_n$  dapat ditentukan harga  $L(\underline{x})$  dan  $U(\underline{x})$ .

### Definisi 2.1

Suatu interval  $(L(\underline{x}), U(\underline{x}))$  disebut interval konfindensi  $(1-\alpha).100\%$  untuk  $\theta$  jika :  $P(L(\underline{X}) < \theta < U(\underline{X})) = 1 - \alpha$  dengan  $0 < \alpha < 1$ . Harga terobservasi  $L(\underline{x})$  dan  $U(\underline{x})$  disebut batas atas dan batas bawah limit konfindensi.

Dalam definisi 2.1 menganggap bahwa  $L(\underline{x})$  dan  $U(\underline{x})$  berhingga.

Pada kasus salah satu dari  $L(\underline{x})$  atau  $U(\underline{x})$  tidak berhingga, maka diperoleh interval satu sisi.

**Definisi 2.2**

1. Jika  $P(L(\underline{X}) < \theta) = 1 - \alpha$ , maka  $L(\underline{X})$  disebut limit konfidensi satu sisi bawah  $(1-\alpha).100\%$  untuk  $\theta$ .
2. Jika  $P(\theta < U(\underline{X})) = 1 - \alpha$ , maka  $U(\underline{X})$  disebut limit konfidensi satu sisi bawah  $(1-\alpha).100\%$  untuk  $\theta$ .

Masalah utama dalam estimasi interval adalah bagaimana menentukan statistik  $L(\underline{x})$  dan  $U(\underline{x})$ . Pada buku ajar ini akan dibahas metode mendapatkan estimator interval yaitu dengan metode inversi uji hipotesis dan dengan metode basaran pivot.

**3. Metode Inversi Uji Hipotesis.**

Terdapat hubungan yang kuat antara uji hipotesis dan interval konfidensi. Secara umum dapat dikatakan bahwa setiap interval konfidensi berkorespondensi dengan uji hipotesis dan sebaliknya. Untuk lebih jelasnya perhatikan contoh berikut :

**Contoh 3.1**

Misalkan  $\alpha$  sampel random dari populasi normal dengan rata-rata  $\theta$  dan variansi 9. Pandang uji hipotesis  $H_0 : \theta = \theta_0$  vs  $H_1 : \theta \neq \theta_0$ . Untuk taraf  $\alpha$  tertentu dan berdasar contoh 4.1, uji tersebut mempunyai daerah penolakan :

$$\left\{ \underline{X} ; \left| \bar{X} - \theta_0 \right| \geq Z_{\alpha/2} \cdot \frac{3}{\sqrt{n}} \right\}$$

Sehingga  $H_0$  diterima untuk titik sampel dengan :

$$\begin{aligned}
 |\bar{x} - \theta_0| &< Z_{\alpha/2} \cdot \frac{3}{\sqrt{n}} \quad \text{atau ekuivalen} \\
 -Z_{\alpha/2} \cdot \frac{3}{\sqrt{n}} &< \bar{x} - \theta_0 < Z_{\alpha/2} \cdot \frac{3}{\sqrt{n}} \\
 -\bar{x} - Z_{\alpha/2} \cdot \frac{3}{\sqrt{n}} &< -\theta_0 < -\bar{x} + Z_{\alpha/2} \cdot \frac{3}{\sqrt{n}} \\
 \bar{x} + Z_{\alpha/2} \cdot \frac{3}{\sqrt{n}} &> \theta_0 > \bar{x} - Z_{\alpha/2} \cdot \frac{3}{\sqrt{n}} \quad \text{karena } \theta = \theta_0 \\
 \bar{x} - Z_{\alpha/2} \cdot \frac{3}{\sqrt{n}} &< \theta < \bar{x} + Z_{\alpha/2} \cdot \frac{3}{\sqrt{n}}
 \end{aligned}$$

Karena uji ini mempunyai ukuran  $\alpha$ , berarti  $P(\text{tolak } H_0 \mid \theta = \theta_0) = \alpha$ , yang berarti  $P(\text{terima } H_0 \mid \theta = \theta_0) = 1 - \alpha$ .

$$P\left(\bar{X} - Z_{\alpha/2} \cdot \frac{3}{\sqrt{n}} < \theta < \bar{X} + Z_{\alpha/2} \cdot \frac{3}{\sqrt{n}}\right) = 1 - \alpha$$

Bentuk persamaan terakhir ini sesuai dengan definisi 4.2.1, sehingga :

$$L(\underline{X}) = \bar{X} - Z_{\alpha/2} \cdot \frac{3}{\sqrt{n}} \quad \text{dan} \quad U(\underline{X}) = \bar{X} + Z_{\alpha/2} \cdot \frac{3}{\sqrt{n}}$$

$$\text{interval } \left(\bar{x} - Z_{\alpha/2} \cdot \frac{3}{\sqrt{n}}, \bar{x} + Z_{\alpha/2} \cdot \frac{3}{\sqrt{n}}\right)$$

yang didapat dengan menginversikan daerah penerimaan dari uji taraf  $\alpha$  adalah interval konfindensi  $(1-\alpha).100\%$  untuk  $\theta$ .

Dari contoh tersebut mengilustrasikan hubungan antara interval konfindensi dan uji hipotesis. Daerah penerimaan uji hipotesis, yaitu himpunan dalam ruang sampel dimana  $H_0 : \theta = \theta_0$  dari populasi normal dengan rata-rata  $\theta$  dan variansi  $\sigma^2$  (dimana  $\sigma^2$  diketahui) diterima, diberikan

$$\text{dengan: } A(\theta_0) = \left\{ \bar{x} \mid \theta_0 - Z_{\alpha/2} \cdot \frac{\sigma}{\sqrt{n}} < \bar{x} < \theta_0 + Z_{\alpha/2} \cdot \frac{\sigma}{\sqrt{n}} \right\}$$

Interval konfindensi, himpunan dalam ruang parameter dengan harga-harga  $\theta$  yang masuk akal diberikan oleh :

$$C(\underline{x}) = \left\{ \theta \mid \bar{x} - Z_{\alpha/2} \cdot \frac{\sigma}{\sqrt{n}} < \theta < \bar{x} + Z_{\alpha/2} \cdot \frac{\sigma}{\sqrt{n}} \right\}$$

Kedua himpunan tersebut dihubungkan satu dengan lainnya dengan tautologi:  $\bar{x} \in A(\theta_0) \Leftrightarrow \theta_0 \in C(\underline{x})$ .

### Teorema 3.1

Untuk setiap  $\theta_0 \in \Omega$ , misalkan  $A(\theta_0)$  adalah daerah penerimaan taraf  $\alpha$  dari uji:  $H_0: \theta = \theta_0$ , untuk setiap  $\underline{x}$  didefinisikan himpunan  $C(\underline{x})$  dalam ruang parameter dengan  $C(\underline{x}) = \{ \theta_0 \mid \underline{x} \in A(\theta_0) \}$ , maka himpunan random  $C(\underline{x})$  adalah interval konfindensi  $(1-\alpha).100\%$  untuk  $\theta$ . Sebaliknya, misalkan  $C(\underline{x})$  adalah interval konfindensi  $(1-\alpha).100\%$  untuk  $\theta$ , untuk setiap  $\theta_0 \in \Omega$  didefinisikan  $A(\theta_0) = \{ \underline{x} \mid \theta_0 \in C(\underline{x}) \}$ , maka  $A(\theta_0)$  adalah daerah penerimaan taraf  $\alpha$  dari uji:  $H_0: \theta = \theta_0$ .

Kenyataan bahwa uji hipotesis dapat diinversikan untuk mendapatkan interval konfindensi dan sebaliknya secara teoritis menarik, tetapi bagian yang benar-benar berguna dari teorema 4.3.1 adalah bagian pertama, relatif mudah untuk mengkonstruksikan daerah penerimaan uji hipotesis taraf  $\alpha$ . Semua teknik yang telah dipunyai untuk mendapatkan daerah penerimaan uji hipotesis, dengan sendirinya dapat digunakan untuk membangun interval konfindensi.

### Contoh 3.2

Misalkan  $X_1, X_2, \dots, X_n$  adalah sampel random independen dari populasi berdistribusi normal dengan rata-rata  $\mu$  dan variansi  $\sigma^2$  dengan  $\mu$  dan  $\sigma^2$  keduanya tidak diketahui. Akan dibangun  $(1-\alpha).100\%$  limit konfindensi atas

untuk  $\mu$ . Dengan perkataan lain kita menginginkan interval konfindensi dengan bentuk :  $C(\underline{x}) = \left( -\infty, U(\underline{x}) \right)$ . Untuk mendapatkan interval tersebut dengan menggunakan theorema 4.3.1, dengan menginversikan uji satu sisi dengan bentuk :  $H_0 : \mu = \mu_0$  vs  $H_1 : \mu < \mu_0$ . Dengan GLRT ukuran  $\alpha$  dari uji tersebut, tolak  $H_0$  bila :

$$\frac{\bar{X} - \mu_0}{S/\sqrt{n}} < -t_{\alpha, n-1} \quad \text{dengan} \quad S = \sqrt{\frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X})^2}$$

Jadi daerah penerimaan untuk uji ini adalah :

$$A(\mu_0) = \left\{ \bar{x} \mid \bar{x} \geq \mu_0 - t_{\alpha, n-1} \cdot \frac{S}{\sqrt{n}} \right\}$$

$$\bar{x} \in A(\mu_0) \Leftrightarrow \bar{x} + t_{\alpha, n-1} \cdot \frac{S}{\sqrt{n}} \geq \mu_0$$

dengan teorema 3.1 dapat didefinisikan :

$$\begin{aligned} C(\underline{x}) &= \left\{ \mu_0 \mid \bar{x} \in A(\mu_0) \right\} \\ &= \left\{ \mu_0 \mid \bar{x} + t_{\alpha, n-1} \cdot \frac{S}{\sqrt{n}} \geq \mu_0 \right\} \end{aligned}$$

$$\text{himpunan random : } C(\underline{x}) = \left( -\infty, \bar{x} + t_{\alpha, n-1} \cdot \frac{S}{\sqrt{n}} \right)$$

adalah interval konfindensi  $(1-\alpha).100\%$  untuk  $\mu$ .

#### 4. Metode Besaran Pivot

Barang kali metode yang paling bagus dalam mengkonstruksikan estimator interval dan menghitung cakupan promodulilitas adalah menggunakan metode besaran pivot. Penggunaan besaran pivot untuk mengkonstruksikan interval konfindensi disebut "pivotal inference".

**Definisi 4.1**

Peubah acak  $Q(\underline{X}, \theta)$  disebut besaran pivot bila distribusi dari  $Q(\underline{X}, \theta)$  tidak tergantung dari  $\theta$  atau parameter lain yang tidak diketahui. Jika  $\underline{X} \sim F(\underline{X}, \theta)$  maka  $Q(\underline{X}, \theta)$  mempunyai distribusi yang sama untuk semua  $\theta$ .

Fungsi  $Q(\underline{X}, \theta)$  biasanya secara eksplisit memuat parameter dan statistik, tetapi sebarang himpunan  $A$ ,  $P(Q(\underline{X}, \theta) \in A)$  tidak dapat tergantung  $\theta$ . Cara mengkonstruksikan himpunan konfindensi dari besaran pivot tergantung pada kemampuannya mendapatkan besaran pivot dan himpunan  $A$  sedemikian sehingga himpunan  $\{\theta \mid Q(\underline{X}, \theta) \in A\}$ .

Langkah-langkah mengkonstruksikan interval konfindensi untuk  $\theta$  dengan menggunakan besaran pivot adalah :

1. Dari sampel random  $X_1, X_2, \dots, X_n$  ditentukan estimator titik untuk  $\theta$ , notasikan estimator titik tersebut dengan  $T(\underline{X})$ .
2. Dicari bentuk transformasi yang memuat  $\theta$  dan estimatornya ( $T(\underline{X})$ ), dapat juga memuat konstan-konstan lain yang diketahui, yang distribusinya mudah (telah ditabelkan). Transformasi ini dikenal sebagai besaran pivot (pivotal quantity).
3. Interval konfindensi dapat diturunkan berdasar langkah 2).

Beberapa transformasi yang telah dikenal :

Misalkan  $X_1, X_2, \dots, X_n$  adalah sampel random independen dari populasi berdistribusi normal dengan rata-rata  $\mu$  dan variansi  $\sigma^2$  , maka :

1. Jika  $\sigma^2$  diketahui,  $\frac{\bar{X} - \mu}{\sigma / \sqrt{n}} \sim N(0,1)$

2. Jika  $\sigma^2$  tidak diketahui diduga dengan variansi sampel

$$S^2 = \frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X})^2 \quad \text{maka} \quad \frac{\bar{X} - \mu}{S/\sqrt{n}} \sim t_{n-1}$$

3. Jika  $\mu$  diketahui,  $\sum_{i=1}^n \left( \frac{X_i - \bar{X}}{\sigma} \right)^2 \sim \chi_n^2$

4. Jika  $\mu$  tidak diketahui diduga dengan rata-rata sampel

$$\bar{X} = \frac{1}{n} \sum_{i=1}^n X_i \quad \text{maka} \quad \sum_{i=1}^n \left( \frac{X_i - \bar{X}}{\sigma} \right)^2 \sim \chi_{n-1}^2$$

#### Contoh 4.1

Misalkan  $X_1, X_2, \dots, X_n$  adalah sampel random independen dari populasi berdistribusi normal dengan rata-rata  $\theta$  dan variansi 0,25.

- dengan menggunakan metode besaran pivot tentukan interval konfindensi  $(1 - \alpha).100\%$  untuk  $\theta$ .
- Bila dipunyai data 2,55 3,59 3,25 2,50 2,75 3,60 3,79 3,15 2,98 3,49 tentukan interval konfindensi 98% untuk rata-rata.

#### Penyelesaian :

a.  $X \sim N(\theta, 0,25)$

$$E(X) = \theta$$

Dengan menggunakan metode momen :  $\hat{\theta} = \frac{1}{n} \sum_{i=1}^n X_i = \bar{X}$

Ambil :  $Z = \frac{\bar{X} - \theta}{0,5/\sqrt{n}} \sim N(0,1)$

Dalam hal ini Z sebagai besaran pivot.

Interval konfindensi  $(1 - \alpha).100\%$  untuk  $\theta$  adalah :

$$\begin{aligned}
 P(-Z_{\alpha/2} < Z < Z_{\alpha/2}) &= 1 - \alpha \\
 -Z_{\alpha/2} < \frac{\bar{X} - \theta}{0,5/\sqrt{n}} < Z_{\alpha/2} \\
 -Z_{\alpha/2} \cdot \frac{0,5}{\sqrt{n}} < \bar{X} - \theta < Z_{\alpha/2} \cdot \frac{0,5}{\sqrt{n}} \\
 -\bar{X} - Z_{\alpha/2} \cdot \frac{0,5}{\sqrt{n}} < -\theta < -\bar{X} + Z_{\alpha/2} \cdot \frac{0,5}{\sqrt{n}} \\
 \bar{X} + Z_{\alpha/2} \cdot \frac{0,5}{\sqrt{n}} > \theta > \bar{X} - Z_{\alpha/2} \cdot \frac{0,5}{\sqrt{n}} \\
 \bar{X} - Z_{\alpha/2} \cdot \frac{0,5}{\sqrt{n}} < \theta < \bar{X} + Z_{\alpha/2} \cdot \frac{0,5}{\sqrt{n}}
 \end{aligned}$$

b. Dari data diperoleh :  $n=10$  ,  $\bar{X} = 3,165$  ,  $\alpha = 2\%$

Interval konfindensi 98% untuk  $\theta$  adalah :

$$\begin{aligned}
 \bar{X} - Z_{1\%} \cdot \frac{0,5}{\sqrt{n}} < \theta < \bar{X} + Z_{1\%} \cdot \frac{0,5}{\sqrt{n}} \\
 3,165 - 2,33 \cdot \frac{0,5}{\sqrt{10}} < \theta < 3,165 + 2,33 \cdot \frac{0,5}{\sqrt{10}} \\
 2,7966 < \theta < 3,5334
 \end{aligned}$$

**Contoh 4.2**

Misalkan  $X_1, X_2, \dots, X_n$  adalah sampel random independen dari populasi berdistribusi normal dengan rata-rata 3 dan variansi  $\theta$ .

- a. dengan menggunakan metode besaran pivot tentukan interval konfindensi  $(1 - \alpha) \cdot 100\%$  untuk  $\theta$ .
- b. Bila dipunyai data : 2,55 3,59 3,25 2,50 2,75 3,60 3,79 3,15 2,98 3,49 tentukan interval konfindensi 98% untuk  $\theta$ .

**Penyelesaian :**

a.  $X_i \sim N(3, \theta)$  maka densitasnya adalah :

$$f(x_i | \theta) = (2\pi)^{-1/2} \theta^{-1/2} \exp\left(-\frac{1}{2\theta}(x_i - 3)^2\right)$$

Estimator titik untuk  $\theta$  dicari dengan metode likelihood maksimum sebagai berikut :

Fungsi likelihood :

$$\begin{aligned} L(\theta | X) &= f(x_1 | \theta) \dots f(x_n | \theta) \\ &= (2\pi)^{-1/2} \theta^{-1/2} \exp\left(-\frac{1}{2\theta}(x_1-3)^2\right) \dots (2\pi)^{-1/2} \theta^{-1/2} \exp\left(-\frac{1}{2\theta}(x_n-3)^2\right) \\ &= (2\pi)^{-n/2} \theta^{-n/2} \exp\left(-\frac{1}{2\theta} \sum_{i=1}^n (x_i-3)^2\right) \end{aligned}$$

Log likelihood :

$$\begin{aligned} l &= \log\left((2\pi)^{-n/2} \theta^{-n/2} \exp\left(-\frac{1}{2\theta} \sum_{i=1}^n (x_i-3)^2\right)\right) \\ &= -\frac{n}{2} \log 2\pi - \frac{n}{2} \log \theta - \frac{1}{2\theta} \sum_{i=1}^n (x_i-3)^2 \end{aligned}$$

$$\frac{\partial l}{\partial \theta} = -\frac{n}{2\theta} + \frac{1}{\theta^2} \sum_{i=1}^n (x_i-3)^2$$

Persamaan likelihood :

$$\frac{\partial l}{\partial \theta} = 0 \Leftrightarrow -\frac{n}{2\theta} + \frac{1}{\theta^2} \sum_{i=1}^n (x_i-3)^2 = 0$$

$$\hat{\theta} = \frac{1}{n} \sum_{i=1}^n (x_i-3)^2$$

$$n \hat{\theta} = \sum_{i=1}^n (x_i-3)^2$$

Ambil  $\chi^2 = \frac{n \hat{\theta}}{\theta} = \sum_{i=1}^n \frac{(X_i-3)^2}{\theta} \sim \chi_n^2$

Interval konfidens i  $(1 - \alpha).100\%$  untuk  $\theta$  :

$$P(\chi_{\alpha/2;n}^2 < \chi^2 < \chi_{1-\alpha/2;n}^2) = 1 - \alpha$$

notasikan  $\chi_{\alpha/2;n}^2 = a$  dan  $\chi_{1-\alpha/2;n}^2 = b$

$$a < \sum_{i=1}^n \frac{(X_i-3)^2}{\theta} < b$$

$$\frac{a}{\sum_{i=1}^n (X_i-3)^2} < \frac{1}{\theta} < \frac{b}{\sum_{i=1}^n (X_i-3)^2}$$

$$\frac{\sum_{i=1}^n (X_i-3)^2}{a} > \theta > \frac{\sum_{i=1}^n (X_i-3)^2}{b}$$

$$\frac{\sum_{i=1}^n (X_i-3)^2}{b} < \theta < \frac{\sum_{i=1}^n (X_i-3)^2}{a}$$

b. Dari data diperoleh :

$$\sum_{i=1}^{10} (X_i - 3)^2 = 2,1727 \quad n = 10 \quad \alpha = 2\%$$

dari tabel  $a = \chi^2_{1\%,10} = 2,56$  dan  $b = \chi^2_{99\%,10} = 23,21$

Interval konfindensi 98% untuk  $\theta$  adalah :

$$\frac{\sum_{i=1}^{10} (X_i - 3)^2}{b} < \theta < \frac{\sum_{i=1}^{10} (X_i - 3)^2}{a}$$

$$\frac{2,1727}{23,21} < \theta < \frac{2,1727}{2,56}$$

$$0,0936 < \theta < 0,8487$$

**Contoh 4.3**

Misalkan  $X_1, X_2, \dots, X_n$  adalah sampel random independen dari populasi berdistribusi normal dengan rata-rata  $\mu$  dan variansi  $\sigma^2$ , keduanya tidak diketahui.

- a. Dengan metode besaran pivot tentukan interval konfindensi  $(1-\alpha).100\%$  untuk  $\mu$  dan  $\sigma^2$ .
- b. Bila dipunyai data : 2,55 3,59 3,25 2,50 2,75 3,60 3,79 3,15 2,98 3,49 tentukan interval konfindensi 98% untuk  $\mu$  dan  $\sigma^2$ .

**Penyelesaian:**

$$X_i \sim N(\mu, \sigma^2)$$

Telah diperoleh MLE untuk  $\mu$  dan  $\sigma^2$  sebagai berikut :

$$\hat{\mu} = \frac{1}{n} \sum_{i=1}^n X_i \quad \text{dan} \quad \hat{\sigma}^2 = \frac{1}{n} \sum_{i=1}^n (X_i - \bar{X})^2 \quad \text{dimana } \hat{\sigma}^2 \text{ merupakan}$$

estimator yang bias untuk  $\sigma^2$ . Ambil estimator untuk  $\sigma^2$  adalah :

$$S^2 = \frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X})^2 \quad \text{yang merupakan estimator takbias.}$$

Ambil  $T = \frac{\bar{X} - \mu}{S / \sqrt{n}}$  maka T akan berdistribusi T student dengan derajat

bebas  $n - 1$

Interval konfindensi  $(1-\alpha).100\%$  untuk  $\mu$  adalah :

$$P(-t_{\alpha/2;n-1} < T < t_{\alpha/2;n-1}) = 1 - \alpha$$

$$-t_{\alpha/2;n-1} < \frac{\bar{X} - \mu}{S/\sqrt{n}} < t_{\alpha/2;n-1}$$

$$-t_{\alpha/2;n-1} \cdot \frac{S}{\sqrt{n}} < \bar{X} - \mu < t_{\alpha/2;n-1} \cdot \frac{S}{\sqrt{n}}$$

$$-\bar{X} - t_{\alpha/2;n-1} \cdot \frac{S}{\sqrt{n}} < -\mu < -\bar{X} + t_{\alpha/2;n-1} \cdot \frac{S}{\sqrt{n}}$$

$$\bar{X} + t_{\alpha/2;n-1} \cdot \frac{S}{\sqrt{n}} > \mu > \bar{X} - t_{\alpha/2;n-1} \cdot \frac{S}{\sqrt{n}}$$

$$\bar{X} - t_{\alpha/2;n-1} \cdot \frac{S}{\sqrt{n}} < \mu < \bar{X} + t_{\alpha/2;n-1} \cdot \frac{S}{\sqrt{n}}$$

b. Dari data diperoleh :  $n = 10$ ,  $\bar{X} = 3,165$ ;  $\alpha = 2\%$ ,  $S = 0,4595$

Interval konfindensi 98% untuk  $\mu$  adalah :

$$\bar{X} - t_{1\%;n-1} \cdot \frac{S}{\sqrt{n}} < \mu < \bar{X} + t_{1\%;n-1} \cdot \frac{S}{\sqrt{n}}$$

$$3,165 - 2,821 \cdot \frac{0,4595}{\sqrt{10}} < \mu < 3,165 + 2,821 \cdot \frac{0,4595}{\sqrt{10}}$$

$$2,7551 < \mu < 3,5749$$

Ambil  $\chi^2 = \frac{n \cdot \hat{\sigma}^2}{\sigma^2} = \frac{1}{\sigma^2} \sum_{i=1}^n (X_i - \bar{X})^2 \sim \chi_{n-1}^2$

Interval konfindensi  $(1 - \alpha).100\%$  untuk  $\sigma^2$  adalah :

$$P(\chi_{\alpha/2;n-1}^2 < \chi^2 < \chi_{1-\alpha/2;n-1}^2) = 1 - \alpha$$

Notasikan  $\chi_{\alpha/2;n-1}^2 = a$  dan  $\chi_{1-\alpha/2;n-1}^2 = b$

maka  $a < \chi^2 < b$

$$a < \frac{1}{\sigma^2} \sum_{i=1}^n (X_i - \bar{X})^2 < b$$

$$\frac{a}{\sum_{i=1}^n (X_i - \bar{X})^2} < \frac{1}{\sigma^2} < \frac{b}{\sum_{i=1}^n (X_i - \bar{X})^2}$$

$$\frac{\sum_{i=1}^n (X_i - \bar{X})^2}{a} > \sigma^2 > \frac{\sum_{i=1}^n (X_i - \bar{X})^2}{b}$$

$$\frac{\sum_{i=1}^n (X_i - \bar{X})^2}{b} < \sigma^2 < \frac{\sum_{i=1}^n (X_i - \bar{X})^2}{a}$$

dari data diperoleh:  $\sum_{i=1}^n (X_i - \bar{X})^2 = 4,1357$

$$\chi_{1\%;9}^2 = 2,09 = a \quad \text{dan} \quad \chi_{99\%;9}^2 = 21,67 = b$$

Interval konfindensi 98% untuk  $\sigma^2$  adalah :

$$\frac{4,1357}{21,67} < \sigma^2 < \frac{4,1357}{2,09}$$

$$0,1908 < \sigma^2 < 1,9788$$

**Contoh 4.4**

Misalkan  $X_1, X_2, \dots, X_n$  adalah sampel random independen dari populasi berdistribusi Exponensial dengan rata-rata  $\theta$ .

- a. Dengan metode besaran pivot tentukan interval konfindensi  $(1-\alpha).100\%$  untuk  $\theta$ .
- b. Bila dipunyai data : 2,55 3,59 3,25 2,50 2,75 3,60 3,79 3,15 2,98 3,49 tentukan interval konfindensi 98% untuk  $\theta$ .

**Penyelesaian :**

$X_i \sim \text{Exp}(\theta)$ , maka densitasnya adalah :

$$f(x_i | \theta) = \theta^{-1} \cdot \exp\left(-\frac{x_i}{\theta}\right)$$

Fungsi likelihood :

$$\begin{aligned} L(\theta | \underline{X}) &= f(x_1 | \theta) \cdot f(x_2 | \theta) \cdot \dots \cdot f(x_n | \theta) \\ &= \theta^{-1} \cdot \exp\left(-\frac{x_1}{\theta}\right) \cdot \theta^{-1} \cdot \exp\left(-\frac{x_2}{\theta}\right) \cdot \dots \cdot \theta^{-1} \cdot \exp\left(-\frac{x_n}{\theta}\right) \\ &= \theta^{-n} \cdot \exp\left(-\frac{1}{\theta} \sum_{i=1}^n x_i\right) \end{aligned}$$

Log likelihood :

$$l = \log\left(L(\theta | \underline{x})\right)$$

$$= -n \log \theta - \frac{1}{\theta} \sum_{i=1}^n x_i$$

$$\frac{\partial l}{\partial \theta} = -\frac{n}{\theta} + \frac{1}{\theta^2} \sum_{i=1}^n x_i$$

Persamaan likelihood :

$$\frac{\partial l}{\partial \hat{\theta}} = 0 \Leftrightarrow -\frac{n}{\hat{\theta}} + \frac{1}{\hat{\theta}^2} \sum_{i=1}^n x_i = 0$$

$$n \hat{\theta} = \sum_{i=1}^n x_i$$

untuk menentukan besaran pivot, ditentukan dahulu distribusi dari

$n \hat{\theta} = \sum_{i=1}^n x_i$ , dengan menggunakan fungsi pembangkit momen berikut :

$$M_{X_1}(t) = \int_0^{\infty} \exp(t x_i) \cdot \frac{1}{\theta} \cdot \exp\left(-\frac{x_i}{\theta}\right) d x_i$$

$$= \frac{1}{1-\theta t} \int_0^{\infty} \exp\left(-\frac{x_i}{\theta}(1-\theta t)\right) d \frac{x_i}{\theta}(1-\theta t)$$

$$= (1-\theta t)^{-1} \Gamma(1)$$

$$= (1-\theta t)^{-1}$$

Karena  $X_1, X_2, \dots, X_n$  independen maka :

$$M_{\sum_{i=1}^n X_i}(t) = M_{X_1}(t) \cdot M_{X_2}(t) \cdot \dots \cdot M_{X_n}(t)$$

$$= (1-\theta t)^{-1} \cdot (1-\theta t)^{-1} \cdot \dots \cdot (1-\theta t)^{-1}$$

$$= (1-\theta t)^{-n}$$

Padahal apabila  $X \sim \chi_n^2$  maka  $M_X(t) = (1-2t)^{-n/2}$

Ambil  $Q = \frac{2 \sum_{i=1}^n X_i}{\theta}$  maka :

$$\begin{aligned}
 M_{\theta}^{2\sum_{i=1}^n X_i}(t) &= M_{\sum_{i=1}^n X_i}\left(\frac{2t}{\theta}\right) \\
 &= \left(1 - \theta \cdot \frac{2t}{\theta}\right)^{-n} \\
 &= (1 - 2t)^{-2n/2} \text{ merupakan fungsi pembangkit momen dari} \\
 &\quad \text{distribusi Khi kuadrat dengan derajat bebas } 2n
 \end{aligned}$$

Berarti:  $Q = \frac{2\sum_{i=1}^n X_i}{\theta} \sim \chi^2_{2n}$

Interval konfidensi  $(1 - \alpha).100\%$  untuk  $\theta$  adalah :

$$P(\chi^2_{\alpha/2; 2n} < Q < \chi^2_{1-\alpha/2; 2n}) = 1 - \alpha$$

Notasikan  $\chi^2_{\alpha/2; 2n} = a$  dan  $\chi^2_{1-\alpha/2; 2n} = b$

$$a < Q < b$$

$$a < \frac{2\sum_{i=1}^n X_i}{\theta} < b$$

$$\frac{a}{2\sum_{i=1}^n X_i} < \frac{1}{\theta} < \frac{b}{2\sum_{i=1}^n X_i}$$

$$\frac{2\sum_{i=1}^n X_i}{b} < \theta < \frac{2\sum_{i=1}^n X_i}{a}$$

b. Dari data diperoleh:  $\sum_{i=1}^n X_i = 31,65$ ;  $n = 10$ ;  $\alpha = 2\%$

$$\chi^2_{1\%; 20} = 8,26 = a \text{ dan } \chi^2_{99\%; 20} = 37,57 = b$$

$$\frac{2\sum_{i=1}^n X_i}{b} < \theta < \frac{2\sum_{i=1}^n X_i}{a}$$

$$\frac{2 \times 31,65}{37,57} < \theta < \frac{2 \times 31,65}{8,26}$$

$$1,6849 < \theta < 7,6634$$

### 5. Interval konfindensi Pendekatan.

Interval konfindensi pendekatan dilakukan bila ditemukan kesulitan menentukan besaran pivot, sehingga harus ditempuh cara lain, misalnya dengan theorema limit pusat.

#### Contoh 5.1

Misalkan  $X_1, X_2, \dots, X_n$  adalah sampel random independen dari populasi berdistribusi bernoulli dengan parameter  $\theta$ . Akan ditentukan interval konfindensi  $(1-\alpha).100\%$  untuk  $\theta$ .

**Penyelesaian :**

$$X_i \sim B(1, \theta), \text{ berarti } E(X) = \theta$$

Dengan menggunakan metode momen diperoleh :

$$E(X) = \frac{1}{n} \sum_{i=1}^n X_i$$

$$\hat{\theta} = \frac{1}{n} \sum_{i=1}^n X_i$$

$$n \cdot \hat{\theta} = \sum_{i=1}^n X_i$$

Berdasarkan contoh 2.2.1 diperoleh  $\sum_{i=1}^n X_i \sim B(n, \theta)$

Dengan theorema limit pusat :

$$\frac{\hat{\theta} - \theta}{\sqrt{\theta(1-\theta)/n}} \xrightarrow{d} Z \sim N(0,1)$$

Sehingga untuk n besar :

$$P\left(-Z_{\alpha/2} < \frac{\hat{\theta} - \theta}{\sqrt{\theta(1-\theta)/n}} < Z_{\alpha/2}\right) = 1 - \alpha$$

Limit pendekatan interval konfindensi  $(\theta_0, \theta_1)$  ditentukan dengan menyelesaikan

$$\frac{\hat{\theta} - \theta_0}{\sqrt{\theta_0(1-\theta_0)/n}} = -Z_{\alpha/2} \text{ dan } \frac{\hat{\theta} - \theta_1}{\sqrt{\theta_1(1-\theta_1)/n}} = Z_{\alpha/2}$$

Akhirnya diperoleh :

$$\frac{\hat{\theta} - \theta}{\sqrt{\hat{\theta}(1-\hat{\theta})/n}} \xrightarrow{d} Z \sim N(0,1)$$

Sehingga interval konfindensi  $(1 - \alpha).100\%$  untuk  $\theta$  adalah :

$$P\left(-Z_{\alpha/2} < \frac{\hat{\theta} - \theta}{\sqrt{\hat{\theta}(1-\hat{\theta})/n}} < Z_{\alpha/2}\right) = 1 - \alpha$$

$$-Z_{\alpha/2} \cdot \sqrt{\hat{\theta}(1-\hat{\theta})/n} < \hat{\theta} - \theta < Z_{\alpha/2} \cdot \sqrt{\hat{\theta}(1-\hat{\theta})/n}$$

$$-\hat{\theta} - Z_{\alpha/2} \cdot \sqrt{\hat{\theta}(1-\hat{\theta})/n} < -\theta < -\hat{\theta} + Z_{\alpha/2} \cdot \sqrt{\hat{\theta}(1-\hat{\theta})/n}$$

$$\hat{\theta} + Z_{\alpha/2} \cdot \sqrt{\hat{\theta}(1-\hat{\theta})/n} > \theta > \hat{\theta} - Z_{\alpha/2} \cdot \sqrt{\hat{\theta}(1-\hat{\theta})/n}$$

$$\hat{\theta} - Z_{\alpha/2} \cdot \sqrt{\hat{\theta}(1-\hat{\theta})/n} < \theta < \hat{\theta} + Z_{\alpha/2} \cdot \sqrt{\hat{\theta}(1-\hat{\theta})/n}$$

**Contoh 5.2**

Misalkan  $X_1, X_2, \dots, X_n$  adalah sampel random independen dari populasi berdistribusi Poisson dengan parameter  $\theta$ . Akan ditentukan interval konfindensi  $(1-\alpha).100\%$  untuk  $\theta$ .

**Penyelesaian :**

Berdasarkan contoh 1.1 pada modul 1 telah diperoleh :  $\hat{\theta} = \frac{1}{n} \sum_{i=1}^n X_i$

Dengan theorema limit pusat :

$$\frac{\hat{\theta} - \theta}{\sqrt{\theta/n}} \xrightarrow{d} Z \sim N(0,1)$$

Dengan cara yang analog pada contoh 1.5.1 diperoleh :

$$\frac{\hat{\theta} - \theta}{\sqrt{\hat{\theta}/n}} \xrightarrow{d} Z \sim N(0,1)$$

Sehingga interval konfindensi  $(1 - \alpha).100\%$  untuk  $\theta$  adalah :

$$P\left(-Z_{\alpha/2} < \frac{\hat{\theta} - \theta}{\sqrt{\hat{\theta}/n}} < Z_{\alpha/2}\right) = 1 - \alpha$$

$$\begin{aligned}
 -Z_{\alpha/2} \cdot \sqrt{\hat{\theta}/n} < \hat{\theta} - \theta < Z_{\alpha/2} \cdot \sqrt{\hat{\theta}/n} \\
 -\hat{\theta} - Z_{\alpha/2} \cdot \sqrt{\hat{\theta}/n} < -\theta < -\hat{\theta} + Z_{\alpha/2} \cdot \sqrt{\hat{\theta}/n} \\
 \hat{\theta} + Z_{\alpha/2} \cdot \sqrt{\hat{\theta}/n} > \theta > \hat{\theta} - Z_{\alpha/2} \cdot \sqrt{\hat{\theta}/n} \\
 \hat{\theta} - Z_{\alpha/2} \cdot \sqrt{\hat{\theta}/n} < \theta < \hat{\theta} + Z_{\alpha/2} \cdot \sqrt{\hat{\theta}/n}
 \end{aligned}$$

#### 4.4.2. Latihan

1. Apakah perbedaannya antara interval konfindensi dengan limit konfindensi satu sisi atas atau bawah.
2. Tuliskan bentuk-bentuk besaran pivot yang berdistribusi Chi-Kuadrat.
3. Apakah interval konfindensi yang dikonstruksikan dengan menggunakan inversi uji hipotesis selalu sama dengan yang dikonstruksikan dengan menggunakan besaran pivot?. Berikan penjelasan secukupnya.

#### 4.4.3 Rangkuman

1. Setiap daerah penerimaan uji hipotesis merupakan interval konfindensi dari parameter yang diuji dalam uji hipotesis tersebut.
2. Penentuan besaran pivot selalu mengarah pada bentuk-bentuk distribusi yang telah ditabelkan.

#### 4.4.4 Tes Formatif

1. Misalkan  $X_1, X_2, \dots, X_n$  adalah sampel random independen dari populasi berdistribusi  $\text{Gamma}(2, \theta)$ .
  - a. Tentukan interval konfindensi  $(1-\alpha).100\%$  untuk  $\theta$ .
  - b. Bila dipunyai data : 74 57 48 29 502 12 70 21 29 386 59 27 153 26 326. Tentukan interval konfindensi 95% untuk  $\theta$ .
  - c. Tentukan pula limit konfindendi satu sisi atas 95% untuk  $\theta$ .
2. Misalkan  $X_1, X_2, \dots, X_n$  adalah sampel random independen dari populasi berdistribusi  $\text{Weibull}(2, \theta)$ .

- a. Tunjukkan bahwa :  $Q = \frac{2 \sum_{i=1}^n X_i^2}{\theta^2} \sim \chi_{2n}^2$
- b. Gunakan Q untuk menentukan interval konfindensi  $(1-\alpha).100\%$  untuk  $\theta$ .
3. Misalkan  $X_1, X_2, \dots, X_n$  adalah sampel random independen dari populasi berdistribusi normal dengan rata-rata  $\mu$  dan variansi  $\sigma^2$ .
- a. Jika  $\sigma^2 = 16$  , tentukan interval konfindensi 95% untuk  $\mu$  berdasarkan rata-rata sampel sama dengan 19,3 dengan ukuran sampel  $n=25$ .
- b. Berdasarkan ketentuan dalam a) tentukan pula limit konfindendi satu sisi atas 95% untuk  $\mu$ .
- c. Andaikan  $\mu = 32,4$  tentukan interval konfindensi 95% untuk  $\sigma^2$  berdasar data sampel  $\frac{1}{n} \sum_{i=1}^n (X_i - 3,24)^2 = 3,25$  dengan  $n=25$ .
4. . Misalkan  $X_1, X_2, \dots, X_n$  adalah sampel random independen dari populasi berdistribusi Exponensial dengan rata-rata  $\theta$ .
- a. Tentukan UMP test ukuran  $\alpha$  untuk uji  $H_0 : \theta = \theta_0$  vs  $H_1 : \theta < \theta_0$ .
- b. Gunakan metode inversi uji hipotesis untuk menentukan interval konfindensi  $(1-\alpha).100\%$  untuk  $\theta$ .

#### 4.4.5 Umpati Balik dan Tindak Lanjut :

Cocokkanlah jawaban anda dengan kunci jawaban dari modul ini. Hitunglah jawaban anda yang benar, kemudian gunakanlah rumus dibawah ini untuk mengetahui tingkat penguasaan anda terhadap materi ini.

Tingkat penguasaan = (jumlah jawaban benar)/10 x 100%

Arti tingkat penguasaan yang anda capai :

90% - 100% : baik sekali

80% - 89% : baik

70% - 79% : sedang

< 70% : kurang

Bila anda telah mencapai tingkat penguasaan 80% atau lebih, anda dapat meneruskan kegiatan belajar selanjutnya. Tetapi apabila tingkat penguasaan anda kurang dari 80% , sebaiknya anda mengulangi materi ini terutama bagian yang belum anda kuasai sebelum meneruskan kegiatan belajar selanjutnya.

#### 4.5. Kunci Jawaban Tes Formatif

$$1. a \left( \frac{2 \sum_{i=1}^n X_i}{\chi^2_{\alpha/2; 4n}}, \frac{2 \sum_{i=1}^n X_i}{\chi^2_{1-\alpha/2; 4n}} \right)$$

b. (43,6735 ; 89,8715)

c. (0 ; 84,2325)

2. a. Hint : tentukan dahulu fungsi pembentuk momen dari  $X_i^2$  , kemudian tentukan fungsi pembangkit momen dari Q

$$b. \left( \frac{2 \sum_{i=1}^n X_i^2}{\chi^2_{\alpha/2; 2n}}, \frac{2 \sum_{i=1}^n X_i^2}{\chi^2_{1-\alpha/2; 2n}} \right)$$

3. a. (17,732 ; 20,868)

b.  $(-\infty ; 20,616)$

c. (2,1580 ; 5,5613)

d. (0 ; 2,3633)

4. a. Tolak  $H_0$  jika  $\sum_{i=1}^n X_i \leq \frac{\theta_0}{2} \cdot \chi^2_{1-\alpha/2; 2n}$

$$b. \left( \frac{\theta_0}{2} \cdot \chi^2_{1-\alpha/2; 2n} \right)$$

**4.5 Referensi :**

1. Bain, L.J. and Engelhart, M. (1992); Introduction to Probability and Mathematical Statistics, Second Edition. Duxbury Press, Belmont, California.
2. Casella, G and Berger, R.L. (1990); Statistical Inference. Wadsworth Inc. Belmont, California.
3. Dudewicz, E.J and Mishra, S.N. (1998); Modern Mathematical Statistics. John Wiley and Sons, Singapore.
4. Roussas, G.G. (1976); A First Course in Mathematical Statistics. Mei Ya Publication Inc. Taipei, Taiwan.
5. ....; (2007). Statistical theory . "[http://en.wikipedia.org/wiki/Binomial\\_proportion\\_confidence\\_interval](http://en.wikipedia.org/wiki/Binomial_proportion_confidence_interval)". This page was last modified 16:20, 30 July 2007

**LAMPIRAN**  
**TABEL STATISTIKA**

## I. Cumulatif Standart Normal Distribution

$$\Phi(z) = \int_{-\infty}^z \frac{1}{\sqrt{2\pi}} e^{-z^2/2} dz$$

z	$\Phi(z)$	z	$\Phi(z)$	z	$\Phi(z)$	z	$\Phi(z)$
-4.00	0.000032	-3.57	0.000178	-3.14	0.000845	-2.71	0.003364
-3.99	0.000033	-3.56	0.000185	-3.13	0.000874	-2.70	0.003467
-3.98	0.000034	-3.55	0.000193	-3.12	0.000904	-2.69	0.003573
-3.97	0.000036	-3.54	0.000200	-3.11	0.000935	-2.68	0.003681
-3.96	0.000037	-3.53	0.000208	-3.10	0.000968	-2.67	0.003793
-3.95	0.000039	-3.52	0.000216	-3.09	0.001001	-2.66	0.003907
-3.94	0.000041	-3.51	0.000224	-3.08	0.001035	-2.65	0.004025
-3.93	0.000042	-3.50	0.000233	-3.07	0.001070	-2.64	0.004145
-3.92	0.000044	-3.49	0.000242	-3.06	0.001107	-2.63	0.004269
-3.91	0.000046	-3.48	0.000251	-3.05	0.001144	-2.62	0.004396
-3.90	0.000048	-3.47	0.000260	-3.04	0.001183	-2.61	0.004527
-3.89	0.000050	-3.46	0.000270	-3.03	0.001223	-2.60	0.004661
-3.88	0.000052	-3.45	0.000280	-3.02	0.001264	-2.59	0.004799
-3.87	0.000054	-3.44	0.000291	-3.01	0.001306	-2.58	0.004940
-3.86	0.000057	-3.43	0.000302	-3.00	0.001350	-2.57	0.005085
-3.85	0.000059	-3.42	0.000313	-2.99	0.001395	-2.56	0.005234
-3.84	0.000062	-3.41	0.000325	-2.98	0.001441	-2.55	0.005386
-3.83	0.000064	-3.40	0.000337	-2.97	0.001489	-2.54	0.005543
-3.82	0.000067	-3.39	0.000349	-2.96	0.001538	-2.53	0.005703
-3.81	0.000069	-3.38	0.000362	-2.95	0.001589	-2.52	0.005868
-3.80	0.000072	-3.37	0.000376	-2.94	0.001641	-2.51	0.006037
-3.79	0.000075	-3.36	0.000390	-2.93	0.001695	-2.50	0.006210
-3.78	0.000078	-3.35	0.000404	-2.92	0.001750	-2.49	0.006387
-3.77	0.000082	-3.34	0.000419	-2.91	0.001807	-2.48	0.006569
-3.76	0.000085	-3.33	0.000434	-2.90	0.001866	-2.47	0.006756
-3.75	0.000088	-3.32	0.000450	-2.89	0.001926	-2.46	0.006947
-3.74	0.000092	-3.31	0.000466	-2.88	0.001988	-2.45	0.007143
-3.73	0.000096	-3.30	0.000483	-2.87	0.002052	-2.44	0.007344
-3.72	0.000100	-3.29	0.000501	-2.86	0.002118	-2.43	0.007549
-3.71	0.000104	-3.28	0.000519	-2.85	0.002186	-2.42	0.007760
-3.70	0.000108	-3.27	0.000538	-2.84	0.002256	-2.41	0.007976
-3.69	0.000112	-3.26	0.000557	-2.83	0.002327	-2.40	0.008198
-3.68	0.000117	-3.25	0.000577	-2.82	0.002401	-2.39	0.008424
-3.67	0.000121	-3.24	0.000598	-2.81	0.002477	-2.38	0.008656
-3.66	0.000126	-3.23	0.000619	-2.80	0.002555	-2.37	0.008894
-3.65	0.000131	-3.22	0.000641	-2.79	0.002635	-2.36	0.009137
-3.64	0.000136	-3.21	0.000664	-2.78	0.002718	-2.35	0.009387
-3.63	0.000142	-3.20	0.000687	-2.77	0.002803	-2.34	0.009642
-3.62	0.000147	-3.19	0.000711	-2.76	0.002890	-2.33	0.009903
-3.61	0.000153	-3.18	0.000736	-2.75	0.002980	-2.32	0.010170
-3.60	0.000159	-3.17	0.000762	-2.74	0.003072	-2.31	0.010444
-3.59	0.000165	-3.16	0.000789	-2.73	0.003167	-2.30	0.010724
-3.58	0.000172	-3.15	0.000816	-2.72	0.003264	-2.29	0.011011

$z$	$\Phi(z)$	$z$	$\Phi(z)$	$z$	$\Phi(z)$	$z$	$\Phi(z)$
-2.28	0.011304	-1.80	0.035930	-1.32	0.093418	-0.84	0.200454
-2.27	0.011604	-1.79	0.036727	-1.31	0.095098	-0.83	0.203269
-2.26	0.011911	-1.78	0.037538	-1.30	0.096800	-0.82	0.206108
-2.25	0.012224	-1.77	0.038364	-1.29	0.098525	-0.81	0.208970
-2.24	0.012545	-1.76	0.039204	-1.28	0.100273	-0.80	0.211855
-2.23	0.012874	-1.75	0.040059	-1.27	0.102042	-0.79	0.214764
-2.22	0.013209	-1.74	0.040930	-1.26	0.103835	-0.78	0.217695
-2.21	0.013553	-1.73	0.041815	-1.25	0.105650	-0.77	0.220650
-2.20	0.013903	-1.72	0.042716	-1.24	0.107488	-0.76	0.223627
-2.19	0.014262	-1.71	0.043633	-1.23	0.109349	-0.75	0.226627
-2.18	0.014629	-1.70	0.044565	-1.22	0.111232	-0.74	0.229650
-2.17	0.015003	-1.69	0.045514	-1.21	0.113139	-0.73	0.232695
-2.16	0.015386	-1.68	0.046479	-1.20	0.115070	-0.72	0.235763
-2.15	0.015778	-1.67	0.047460	-1.19	0.117023	-0.71	0.238852
-2.14	0.016177	-1.66	0.048457	-1.18	0.119000	-0.70	0.241964
-2.13	0.016586	-1.65	0.049471	-1.17	0.121000	-0.69	0.245097
-2.12	0.017003	-1.64	0.050503	-1.16	0.123024	-0.68	0.248252
-2.11	0.017429	-1.63	0.051551	-1.15	0.125072	-0.67	0.251429
-2.10	0.017864	-1.62	0.052616	-1.14	0.127143	-0.66	0.254627
-2.09	0.018309	-1.61	0.053699	-1.13	0.129238	-0.65	0.257846
-2.08	0.018763	-1.60	0.054799	-1.12	0.131357	-0.64	0.261086
-2.07	0.019226	-1.59	0.055917	-1.11	0.133500	-0.63	0.264347
-2.06	0.019699	-1.58	0.057053	-1.10	0.135666	-0.62	0.267629
-2.05	0.020182	-1.57	0.058208	-1.09	0.137857	-0.61	0.270931
-2.04	0.020675	-1.56	0.059380	-1.08	0.140071	-0.60	0.274253
-2.03	0.021178	-1.55	0.060571	-1.07	0.142310	-0.59	0.277595
-2.02	0.021692	-1.54	0.061780	-1.06	0.144572	-0.58	0.280957
-2.01	0.022216	-1.53	0.063008	-1.05	0.146859	-0.57	0.284339
-2.00	0.022750	-1.52	0.064255	-1.04	0.149170	-0.56	0.287740
-1.99	0.023295	-1.51	0.065522	-1.03	0.151505	-0.55	0.291160
-1.98	0.023852	-1.50	0.066807	-1.02	0.153864	-0.54	0.294599
-1.97	0.024419	-1.49	0.068112	-1.01	0.156248	-0.53	0.298056
-1.96	0.024998	-1.48	0.069437	-1.00	0.158655	-0.52	0.301532
-1.95	0.025588	-1.47	0.070781	-0.99	0.161087	-0.51	0.305026
-1.94	0.026190	-1.46	0.072145	-0.98	0.163543	-0.50	0.308538
-1.93	0.026803	-1.45	0.073529	-0.97	0.166023	-0.49	0.312067
-1.92	0.027429	-1.44	0.074934	-0.96	0.168528	-0.48	0.315614
-1.91	0.028067	-1.43	0.076359	-0.95	0.171056	-0.47	0.319178
-1.90	0.028717	-1.42	0.077804	-0.94	0.173609	-0.46	0.322758
-1.89	0.029379	-1.41	0.079270	-0.93	0.176186	-0.45	0.326355
-1.88	0.030054	-1.40	0.080757	-0.92	0.178786	-0.44	0.329969
-1.87	0.030742	-1.39	0.082264	-0.91	0.181411	-0.43	0.333598
-1.86	0.031443	-1.38	0.083793	-0.90	0.184060	-0.42	0.337243
-1.85	0.032157	-1.37	0.085343	-0.89	0.186733	-0.41	0.340903
-1.84	0.032884	-1.36	0.086915	-0.88	0.189430	-0.40	0.344578
-1.83	0.033625	-1.35	0.088508	-0.87	0.192150	-0.39	0.348268
-1.82	0.034380	-1.34	0.090123	-0.86	0.194895	-0.38	0.351973
-1.81	0.035148	-1.33	0.091759	-0.85	0.197663	-0.37	0.355691

$z$	$\Phi(z)$	$z$	$\Phi(z)$	$z$	$\Phi(z)$	$z$	$\Phi(z)$
-0.36	0.359424	0.20	0.579260	0.67	0.748571	1.14	0.872857
-0.35	0.363169	0.21	0.583166	0.68	0.751748	1.15	0.874928
-0.34	0.366928	0.22	0.587064	0.69	0.754903	1.16	0.876976
-0.33	0.370700	0.23	0.590954	0.70	0.758036	1.17	0.879000
-0.32	0.374484	0.24	0.594835	0.71	0.761148	1.18	0.881000
-0.31	0.378280	0.25	0.598706	0.72	0.764237	1.19	0.882977
-0.30	0.382089	0.26	0.602568	0.73	0.767305	1.20	0.884930
-0.29	0.385908	0.27	0.606420	0.74	0.770350	1.21	0.886861
-0.28	0.389739	0.28	0.610261	0.75	0.773373	1.22	0.888768
-0.27	0.393580	0.29	0.614092	0.76	0.776373	1.23	0.890651
-0.26	0.397432	0.30	0.617911	0.77	0.779350	1.24	0.892512
-0.25	0.401294	0.31	0.621720	0.78	0.782305	1.25	0.894350
-0.24	0.405165	0.32	0.625516	0.79	0.785236	1.26	0.896165
-0.23	0.409046	0.33	0.629300	0.80	0.788145	1.27	0.897958
-0.22	0.412936	0.34	0.633072	0.81	0.791030	1.28	0.899727
-0.21	0.416834	0.35	0.636831	0.82	0.793892	1.29	0.901475
-0.20	0.420740	0.36	0.640576	0.83	0.796731	1.30	0.903200
-0.19	0.424655	0.37	0.644309	0.84	0.799546	1.31	0.904902
-0.18	0.428576	0.38	0.648027	0.85	0.802337	1.32	0.906582
-0.17	0.432505	0.39	0.651732	0.86	0.805105	1.33	0.908241
-0.16	0.436441	0.40	0.655422	0.87	0.807850	1.34	0.909877
-0.15	0.440382	0.41	0.659097	0.88	0.810570	1.35	0.911492
-0.14	0.444330	0.42	0.662757	0.89	0.813267	1.36	0.913085
-0.13	0.448283	0.43	0.666402	0.90	0.815940	1.37	0.914657
-0.12	0.452242	0.44	0.670031	0.91	0.818589	1.38	0.916207
-0.11	0.456205	0.45	0.673645	0.92	0.821214	1.39	0.917736
-0.10	0.460172	0.46	0.677242	0.93	0.823814	1.40	0.919243
-0.09	0.464144	0.47	0.680822	0.94	0.826391	1.41	0.920730
-0.08	0.468119	0.48	0.684386	0.95	0.828944	1.42	0.922196
-0.07	0.472097	0.49	0.687933	0.96	0.831472	1.43	0.923641
-0.06	0.476078	0.50	0.691462	0.97	0.833977	1.44	0.925066
-0.05	0.480061	0.51	0.694974	0.98	0.836457	1.45	0.926471
-0.04	0.484047	0.52	0.698468	0.99	0.838913	1.46	0.927855
-0.03	0.488034	0.53	0.701944	1.00	0.841345	1.47	0.929219
-0.02	0.492022	0.54	0.705401	1.01	0.843752	1.48	0.930563
-0.01	0.496011	0.55	0.708840	1.02	0.846136	1.49	0.931888
0.00	0.500000	0.56	0.712260	1.03	0.848495	1.50	0.933193
0.01	0.503989	0.57	0.715661	1.04	0.850830	1.51	0.934478
0.02	0.507978	0.58	0.719043	1.05	0.853141	1.52	0.935745
0.03	0.511966	0.59	0.722405	1.06	0.855428	1.53	0.936992
0.04	0.515953	0.60	0.725747	1.07	0.857690	1.54	0.938220
0.05	0.519939	0.61	0.729069	1.08	0.859929	1.55	0.939429
0.06	0.523922	0.62	0.732371	1.09	0.862143	1.56	0.940620
0.07	0.527903	0.63	0.735653	1.10	0.864334	1.57	0.941792
0.08	0.531881	0.64	0.738914	1.11	0.866500	1.58	0.942947
0.09	0.535856	0.65	0.742154	1.12	0.868643	1.59	0.944083
0.10	0.539828	0.66	0.745373	1.13	0.870762	1.60	0.945201

$z$	$\Phi(z)$	$z$	$\Phi(z)$	$z$	$\Phi(z)$	$z$	$\Phi(z)$
1.61	0.946301	2.08	0.981237	2.60	0.995339	3.07	0.998930
1.62	0.947384	2.09	0.981691	2.61	0.995473	3.08	0.998965
1.63	0.948449	2.10	0.982136	2.62	0.995604	3.09	0.998999
1.64	0.949497	2.11	0.982571	2.63	0.995731	3.10	0.999032
1.65	0.950529	2.12	0.982997	2.64	0.995855	3.11	0.999065
1.66	0.951543	2.13	0.983414	2.65	0.995975	3.12	0.999096
1.67	0.952540	2.14	0.983823	2.66	0.996093	3.13	0.999126
1.68	0.953521	2.15	0.984222	2.67	0.996207	3.14	0.999155
1.69	0.954486	2.16	0.984614	2.68	0.996319	3.15	0.999184
1.70	0.955435	2.17	0.984997	2.69	0.996427	3.16	0.999211
1.71	0.956367	2.18	0.985371	2.70	0.996533	3.17	0.999238
1.72	0.957284	2.19	0.985738	2.71	0.996636	3.18	0.999264
1.73	0.958185	2.20	0.986097	2.72	0.996736	3.19	0.999289
1.74	0.959070	2.21	0.986447	2.73	0.996833	3.20	0.999313
1.75	0.959941	2.22	0.986791	2.74	0.996928	3.21	0.999336
1.76	0.960796	2.23	0.987126	2.75	0.997020	3.22	0.999359
1.77	0.961636	2.24	0.987455	2.76	0.997110	3.23	0.999381
1.78	0.962462	2.25	0.987776	2.77	0.997197	3.24	0.999402
1.79	0.963273	2.26	0.988089	2.78	0.997282	3.25	0.999423
1.80	0.964070	2.27	0.988396	2.79	0.997365	3.26	0.999443
1.81	0.964852	2.28	0.988696	2.80	0.997445	3.27	0.999462
1.82	0.965620	2.29	0.988989	2.81	0.997523	3.28	0.999481
1.83	0.966375	2.30	0.989276	2.82	0.997599	3.29	0.999499
1.84	0.967116	2.31	0.989556	2.83	0.997673	3.30	0.999517
1.85	0.967843	2.32	0.989830	2.84	0.997744	3.31	0.999534
1.86	0.968557	2.33	0.990097	2.85	0.997814	3.32	0.999550
1.87	0.969258	2.34	0.990358	2.86	0.997882	3.33	0.999566
1.88	0.969946	2.35	0.990613	2.87	0.997948	3.34	0.999581
1.89	0.970621	2.36	0.990863	2.88	0.998012	3.35	0.999596
1.90	0.971283	2.37	0.991106	2.89	0.998074	3.36	0.999610
1.91	0.971933	2.38	0.991344	2.90	0.998134	3.37	0.999624
1.92	0.972571	2.39	0.991576	2.91	0.998193	3.38	0.999638
1.93	0.973197	2.40	0.991802	2.92	0.998250	3.39	0.999651
1.94	0.973810	2.41	0.992024	2.93	0.998305	3.40	0.999663
1.95	0.974412	2.42	0.992240	2.94	0.998359	3.41	0.999675
1.96	0.975002	2.43	0.992451	2.95	0.998411	3.42	0.999687
1.97	0.975581	2.44	0.992656	2.96	0.998462	3.43	0.999698
1.98	0.976148	2.45	0.992857	2.97	0.998511	3.44	0.999709
1.99	0.976705	2.46	0.993053	2.98	0.998559	3.45	0.999720
2.00	0.977250	2.47	0.993244	2.99	0.998605	3.46	0.999730
2.01	0.977784	2.48	0.993431	3.00	0.998650	3.47	0.999740
2.02	0.978308	2.49	0.993613	3.01	0.998694	3.48	0.999749
2.03	0.978822	2.50	0.993790	3.02	0.998736	3.49	0.999758
2.04	0.979325	2.51	0.993963	3.03	0.998777	3.50	0.999767
2.05	0.979818	2.52	0.994132	3.04	0.998817	3.51	0.999776
2.06	0.980301	2.53	0.994297	3.05	0.998856	3.52	0.999784
2.07	0.980774	2.54	0.994457	3.06	0.998893	3.53	0.999792

$z$	$\Phi(z)$	$z$	$\Phi(z)$	$z$	$\Phi(z)$	$z$	$\Phi(z)$
3.54	0.999800	3.66	0.999874	3.78	0.999922	3.90	0.999952
3.55	0.999807	3.67	0.999879	3.79	0.999925	3.91	0.999954
3.56	0.999815	3.68	0.999883	3.80	0.999928	3.92	0.999956
3.57	0.999822	3.69	0.999888	3.81	0.999931	3.93	0.999958
3.58	0.999828	3.70	0.999892	3.82	0.999933	3.94	0.999959
3.59	0.999835	3.71	0.999896	3.83	0.999936	3.95	0.999961
3.60	0.999841	3.72	0.999900	3.84	0.999938	3.96	0.999963
3.61	0.999847	3.73	0.999904	3.85	0.999941	3.97	0.999964
3.62	0.999853	3.74	0.999908	3.86	0.999943	3.98	0.999966
3.63	0.999858	3.75	0.999912	3.87	0.999946	3.99	0.999967
3.64	0.999864	3.76	0.999915	3.88	0.999948	4.00	0.999968
3.65	0.999869	3.77	0.999918	3.89	0.999950		

## II. Percentage Points of the t Distribution

df	$\alpha$									
	0.40	0.25	0.10	0.05	0.025	0.01	0.005	0.0025	0.001	0.0005
1	0.325	1.000	3.078	6.314	12.706	31.821	63.657	127.321	318.309	636.619
2	0.289	0.817	1.886	2.920	4.303	6.965	9.925	14.089	22.327	31.599
3	0.277	0.765	1.638	2.353	3.182	4.541	5.841	7.453	10.215	12.924
4	0.271	0.741	1.533	2.132	2.776	3.747	4.604	5.598	7.173	8.610
5	0.267	0.727	1.476	2.015	2.571	3.365	4.032	4.773	5.893	6.869
6	0.265	0.718	1.440	1.943	2.447	3.143	3.707	4.317	5.208	5.959
7	0.263	0.711	1.415	1.895	2.365	2.998	3.500	4.029	4.785	5.408
8	0.262	0.706	1.397	1.860	2.306	2.897	3.355	3.833	4.501	5.041
9	0.261	0.703	1.383	1.833	2.262	2.821	3.250	3.690	4.297	4.781
10	0.260	0.700	1.372	1.813	2.228	2.764	3.169	3.581	4.144	4.587
11	0.260	0.697	1.363	1.796	2.201	2.718	3.106	3.497	4.025	4.437
12	0.259	0.696	1.356	1.782	2.179	2.681	3.055	3.428	3.930	4.318
13	0.259	0.694	1.350	1.771	2.160	2.650	3.012	3.373	3.852	4.221
14	0.258	0.692	1.345	1.761	2.145	2.625	2.977	3.326	3.787	4.141
15	0.258	0.691	1.341	1.753	2.131	2.603	2.947	3.286	3.733	4.073
16	0.258	0.690	1.337	1.746	2.120	2.584	2.921	3.252	3.686	4.015
17	0.257	0.689	1.333	1.740	2.110	2.567	2.898	3.222	3.646	3.965
18	0.257	0.688	1.330	1.734	2.101	2.552	2.878	3.197	3.611	3.922
19	0.257	0.688	1.328	1.729	2.093	2.540	2.861	3.174	3.579	3.883
20	0.257	0.687	1.325	1.725	2.086	2.528	2.845	3.153	3.552	3.850
21	0.257	0.686	1.323	1.721	2.080	2.518	2.831	3.135	3.527	3.819
22	0.256	0.686	1.321	1.717	2.074	2.508	2.819	3.119	3.505	3.792
23	0.256	0.685	1.320	1.714	2.069	2.500	2.807	3.104	3.485	3.768
24	0.256	0.685	1.318	1.711	2.064	2.492	2.797	3.091	3.467	3.745
25	0.256	0.684	1.316	1.708	2.060	2.485	2.787	3.078	3.450	3.725
26	0.256	0.684	1.315	1.706	2.056	2.479	2.779	3.067	3.435	3.707
27	0.256	0.684	1.314	1.703	2.052	2.473	2.771	3.057	3.421	3.690
28	0.256	0.683	1.313	1.701	2.048	2.467	2.763	3.047	3.408	3.674
29	0.256	0.683	1.311	1.699	2.045	2.462	2.756	3.038	3.396	3.659
30	0.256	0.683	1.310	1.697	2.042	2.457	2.750	3.030	3.385	3.646
40	0.255	0.681	1.303	1.684	2.021	2.423	2.705	2.971	3.307	3.551
50	0.255	0.679	1.299	1.676	2.009	2.403	2.678	2.937	3.261	3.496
60	0.255	0.679	1.296	1.671	2.000	2.390	2.660	2.915	3.232	3.460
70	0.254	0.678	1.294	1.667	1.994	2.381	2.648	2.899	3.211	3.435
80	0.254	0.678	1.292	1.664	1.990	2.374	2.639	2.887	3.195	3.416
90	0.254	0.677	1.291	1.662	1.987	2.369	2.632	2.878	3.183	3.402
100	0.254	0.677	1.290	1.660	1.984	2.364	2.626	2.871	3.174	3.391
120	0.254	0.677	1.289	1.658	1.980	2.358	2.617	2.860	3.160	3.374

III. Percentage Points of the  $\chi^2$  Distribution

df	$\alpha$								
	0.995	0.990	0.975	0.950	0.500	0.050	0.025	0.01	0.005
1	0.00	0.00	0.00	0.00	0.45	3.84	5.02	6.63	7.88
2	0.01	0.02	0.05	0.1	1.39	5.99	7.38	9.21	10.6
3	0.07	0.11	0.22	0.35	2.37	7.81	9.35	11.34	12.84
4	0.21	0.3	0.48	0.71	3.36	9.49	11.14	13.28	14.86
5	0.41	0.55	0.83	1.15	4.35	11.07	12.83	15.09	16.75
6	0.68	0.87	1.24	1.64	5.35	12.59	14.45	16.81	18.55
7	0.99	1.24	1.69	2.17	6.35	14.07	16.01	18.48	20.28
8	1.34	1.65	2.18	2.73	7.34	15.51	17.53	20.09	21.95
9	1.73	2.09	2.70	3.33	8.34	16.92	19.02	21.67	23.59
10	2.16	2.56	3.25	3.94	9.34	18.31	20.48	23.21	25.19
11	2.60	3.05	3.82	4.57	10.34	19.68	21.92	24.72	26.76
12	3.07	3.57	4.40	5.23	11.34	21.03	23.34	26.22	28.3
13	3.57	4.11	5.01	5.89	12.34	22.36	24.74	27.69	29.82
14	4.07	4.66	5.63	6.57	13.34	23.68	26.12	29.14	31.32
15	4.60	5.23	6.26	7.26	14.34	25.00	27.49	30.58	32.8
16	5.14	5.81	6.91	7.96	15.34	26.30	28.85	32.00	34.27
17	5.70	6.41	7.56	8.67	16.34	27.59	30.19	33.41	35.72
18	6.26	7.01	8.23	9.39	17.34	28.87	31.53	34.81	37.16
19	6.84	7.63	8.91	10.12	18.34	30.14	32.85	36.19	38.58
20	7.43	8.26	9.59	10.85	19.34	31.41	34.17	37.57	40.00
21	8.03	8.90	10.28	11.59	20.34	32.67	35.48	38.93	41.40
22	8.64	9.54	10.98	12.34	21.34	33.92	36.78	40.29	42.80
23	9.26	10.2	11.69	13.09	22.34	35.17	38.08	41.64	44.18
24	9.89	10.86	12.4	13.85	23.34	36.42	39.36	42.98	45.56
25	10.52	11.52	13.12	14.61	24.34	37.65	40.65	44.31	46.93
26	11.16	12.2	13.84	15.38	25.34	38.89	41.92	45.64	48.29
27	11.81	12.88	14.57	16.15	26.34	40.11	43.19	46.96	49.64
28	12.46	13.56	15.31	16.93	27.34	41.34	44.46	48.28	50.99
29	13.12	14.26	16.05	17.71	28.34	42.56	45.72	49.59	52.34
30	13.79	14.95	16.79	18.49	29.34	43.77	46.98	50.89	53.67
40	20.71	22.16	24.43	26.51	39.34	55.76	59.34	63.69	66.77
50	27.99	29.71	32.36	34.76	49.33	67.50	71.42	76.15	79.49
60	35.53	37.48	40.48	43.19	59.33	79.08	83.3	88.38	91.95
70	43.28	45.44	48.76	51.74	69.33	90.53	95.02	100.43	104.21
80	51.17	53.54	57.15	60.39	79.33	101.88	106.63	112.33	116.32
90	59.20	61.75	65.65	69.13	89.33	113.15	118.14	124.12	128.30
100	67.33	70.06	74.22	77.93	99.33	124.34	129.56	135.81	140.17
120	83.85	86.92	91.57	95.7	119.33	146.57	152.21	158.95	163.65