Seri Pendidikan Aktuaris Indonesia

Siswadi
Windiani Erliana

Statistika Matematik Untuk Aktuaris

STATISTIKA MATEMATIK
UNTUK
AKTUARIS

Siswadi dan Windiani Erliana
e-book DOI: doi.org/10.31227/osf.io/nj3ag
ISBN : 978-XXX-XXXX-XX-X

2019

Daftar Isi

KATA SAMBUTAN PERSATUAN AKTUARIS INDONESIA

1

Persatuan Aktuaris Indonesia mengucapkan selamat atas penerbitan "Seri Catatan Kuliah Sains Aktuaria" dalam kerangka penyediaan bahan belajar bagi mereka yang menempuh pendidikan di bidang sains aktuaria. Pengajaran isi kurikulum pendidikan Persatuan Aktuaris Indonesia serta isi kurikulum pendidikan sains aktuaria di berbagai universitas di Indonesia secara umum akan terbantu dengan adanya pengadaan Seri Catatan Kuliah ini.

Sebagai organisasi profesi, Persatuan Aktuaris Indonesia akan senantiasa mendorong upaya berbagai pihak untuk menghasilkan karya bagi pembangunan di bidang sains aktuaria di Indonesia. Seri Catatan Kuliah ini dengan Editor Penerbitan yang juga sebagai Editor Utama Saudara Agah D. Garnadi merupakan salah satu kontribusi berharga dalam upaya meningkatkan kapasitas nasional di bidang sains aktuaria.

Akhir kata, Persatuan Aktuaris Indonesia menyatakan apresiasi kepada Saudara Agah D. Garnadi sebagai Editor Penerbitan atas penerbitan Seri Catatan Kuliah ini. Semoga Seri ini bisa dijadikan panduan belajar di bidang sains aktuaria serta mendapatkan tanggapan yang baik dari semua pihak.

Jakarta, 2 Mei 2016

Rianto Ahmadi Djojosugito

Ketua Persatuan Aktuaris Indonesia

Kata Pengantar

Buku ini disusun berdasarkan perkuliahan Statistika Matematik pada rentang waktu 2014/2015, 2015/2016, 2016/2017, 2017/2018, 2018/2019, yang diampu oleh penulis pertama.

Mata kuliah Statistika Matematik, merupakan salah satu bagian yang diakui untuk mata uji profesi Persatuan Aktuaris Indonesia (PAI). Sehingga dilengkapi dengan beberapa soal ujian A20 Probabilitas dan Statistika.

Pada Lampiran diberikan kumpulan soal ujian A20, terima kasih kepada Dr. Lia Yuliawati, MSi dari STKIP 11 April, Sumedang, yang telah memberikan kontribusi pada kumpulan soal tersebut. Serta saudari Grace Agustina, SMat., atas sejumlah catatannya.

Hibah PUPT-IPB, 'Pengembangan Perangkat Lunak Berbasis Finite Element Method (FEM) untuk Produk Jasa Keuangan dan Asuransi' dengan kontrak no: 079/SP2H/LT/DRPM/II/2016, secara tidak langsung memberi bantuan pendanaan kepada penulis yunior. Terima kasih kami ucapkan.

Terima kasih disampaikan kepada Dr. Ir. IG Putu Purnaba, yang memberikan koreksi atas jawaban soal ujian PAI. Dr. Ir. Sri Nurdiati, MSc., selaku peneliti utama Hibah PUPT-IPB tersebut di atas, yang memberikan kesempatan kepada penulis yunior mengembangkan ilmu aktuaria. Terakhir, kepada Drs. Agah D. Garnadi, Grad. Dip. Sci. yang berulangkali meminta agar naskah ini untuk diterbitkan dalam seri buku: 'Pendidikan Aktuaris Indonesia', sehingga bisa dimanfaatkan khususnya untuk pendidikan aktuaria di Indonesia.

Terima kasih kepada Rianto Ahmadi, PhD., ketua PAI periode 2014-2017, yang sudi memberikan kata sambutan seri buku 'Pendidikan Aktuaris Indonesia'.

Bogor, 09 Juni 2019
Siswadi
Windiani Erliana

BAB 1

Sebaran Multivariat

1.1 Vektor Peubah Acak

Definisi 1.1 *Misalkan Ω adalah ruang contoh suatu percobaan acak. Peubah acak X_i dari percobaan tersebut adalah suatu fungsi bernilai real $X_i : \Omega \to \mathbb{R}, i = 1, 2, ..., n$ dengan $(X_1, X_2, ..., X_n)$ disebut vektor peubah acak dimensi-n.*

Dalam bab ini, digunakan notasi vektor untuk menyatakan peubah acak $X_1, X_2, ..., X_n$. Sebagai contoh, $(X_1, X_2, ..., X_n)'$ merupakan vektor kolom $\mathbf{X}$ yang memiliki nilai $(x_1, x_2, ..., x_n)', \forall x_i \in \mathbb{R}$.

$$\begin{aligned} \mathbf{X} &= \mathbf{x} \\ (X_1, X_2, ..., X_n)' &= (x_1, x_2, ..., x_n)' \\ \begin{pmatrix} X_1 \\ X_2 \\ \vdots \\ X_n \end{pmatrix} &= \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \end{aligned}$$

1.2 Fungsi Sebaran Bersama

Fungsi sebaran bersama bagi vektor peubah acak tersebut didefinisikan sebagai berikut

$$F_{\mathbf{X}}(\mathbf{x}) = P\left(X_1 \le x_1, ..., X_n \le x_n\right).$$

Jika n peubah acak $X_1, X_2, ..., X_n$ adalah peubah acak diskret, maka

$$F_{\mathbf{X}}(\mathbf{x}) = \sum \underset{w_1 \leq x_1, \ldots, w_n \leq x_n}{\ldots} \sum p\left(w_1, \ldots, w_n\right),$$

sedangkan jika n peubah acak tersebut adalah peubah acak kontinu, maka

$$F_{\mathbf{X}}(\mathbf{x}) = \int \underset{w_1 \leq x_1, \ldots, w_n \leq x_n}{\ldots} \int f\left(w_1, \ldots, w_n\right) dw_1 \ldots dw_n,$$

$$\frac{\partial^n}{\partial x_1 \ldots \partial x_n} F_{\mathbf{X}}(\mathbf{x}) = f\left(\mathbf{x}\right),$$

dengan fungsi p dan f secara berturut-turut merupakan fungsi massa peluang dan fungsi kepekatan peluang bagi vektor peubah acak **x.**

Contoh 1.1 *Misalkan fungsi kepekatan peluang dari peubah acak* X_1, X_2, X_3 *adalah sebagai berikut*

$$f\left(x_1, x_2, x_3\right) = e^{-(x_1+x_2+x_3)}\ \mathbf{I}\left(0 < x_1, x_2, x_3 < \infty\right),$$

maka fungsi sebaran bersama bagi X_1, X_2, X_3 *ialah*

$$\begin{aligned} F_{X_1,X_2,X_3}\left(x_1, x_2, x_3\right) &= P\left(X_1 \leq x_1, X_2 \leq x_2, X_3 \leq x_3\right) \\ &= \int_0^{x_1}\int_0^{x_2}\int_0^{x_3} e^{-(x_1+x_2+x_3)} dx_1 dx_2 dx_3 \\ &= \left(1 - e^{-x_1}\right)\left(1 - e^{-x_2}\right)\left(1 - e^{-x_3}\right)\mathbf{I}\left(0 < x_1, x_2, x_3 < \infty\right). \end{aligned}$$

1.2.1 Fungsi Massa Peluang Marginal

Definisi 1.2 *Misalkan* $X_1, \ldots, X_n$ *adalah peubah acak diskret yang menyebar bersama dengan fungsi massa peluang bersama* $p_{X_1,\ldots,X_n}\left(x_1, \ldots, x_n\right)$. *Fungsi massa peluang marginal dari peubah acak* X_1 *ialah*

$$p_{X_1}\left(x_1\right) = \sum_{x_2} \ldots \sum_{x_n} p_{X_1,\ldots,X_n}\left(x_1, \ldots, x_n\right).$$

1.2.2 Fungsi Kepekatan Peluang Marginal

Definisi 1.3 *Misalkan* $X_1, \ldots, X_n$ *adalah peubah acak kontinu yang menyebar bersama dengan fungsi kepekatan peluang bersama* $f_{X_1,\ldots,X_n}\left(x_1, \ldots, x_n\right)$. *Fungsi kepekatan peluang marginal dari peubah acak* X_1 *ialah*

$$f_{X_1}\left(x_1\right) = \int_{-\infty}^{\infty} \ldots \int_{-\infty}^{\infty} f_{X_1,\ldots,X_n}\left(x_1, \ldots, x_n\right) dx_2 \ldots dx_n.$$

1.3 Nilai Harapan

Definisi 1.4 *Misalkan* $Y = u(X_1, X_2, ..., X_n)$ *adalah fungsi terhadap vektor peubah acak* $(X_1, X_2, ..., X_n)'$. *Jika* $X_1, X_2, ..., X_n$ *adalah peubah acak diskret yang menyebar bersama dengan fungsi massa peluang bersama* $p(x_1, x_2, ..., x_n)$, *maka nilai harapan dari* Y *didefinisikan sebagai*

$$E(Y) = \sum_{x_n} ... \sum_{x_1} u(x_1, x_2, ..., x_n)\, p(x_1, x_2, ..., x_n),$$

asalkan jumlah di atas konvergen mutlak. Jika $X_1, X_2, ..., X_n$ *adalah peubah acak kontinu yang menyebar bersama dengan fungsi kepekatan peluang bersama* $f(x_1, x_2, ..., x_n)$, *maka nilai harapan dari* Y *didefinisikan sebagai*

$$E(Y) = \int_{-\infty}^{\infty} ... \int_{-\infty}^{\infty} u(x_1, x_2, ..., x_n)\, f(x_1, x_2, ..., x_n)\, dx_1 dx_2 ... dx_n$$

asalkan integral di atas konvergen mutlak.

Teorema 1.1 *Misalkan* $X_1, X_2, ..., X_n$ *adalah peubah acak yang terdefinisi dalam ruang peluang yang sama dan* $k_1, k_2, ..., k_m$ *adalah* m *konstanta bilangan real. Jika* $Y_i = u(X_1, X_2, ..., X_n)$, $i = 1, 2, ..., m$, *adalah fungsi terhadap vektor peubah acak* $(X_1, X_2, ..., X_n)'$, *maka*

$$E\left(\sum_{i=1}^{m} k_i Y_i\right) = \sum_{i=1}^{m} k_i E(Y_i).$$

Misalkan ${}_n\mathbf{W}_m = (W_{ij})$ merupakan matriks peubah acak berukuran $n \times m$ yang berunsur peubah acak W_{ij}, maka

$$\begin{aligned} E(\mathbf{W}) &= (E(W_{ij})) \\ &= \begin{bmatrix} E(W_{11}) & E(W_{12}) & ... & E(W_{1m}) \\ E(W_{21}) & E(W_{22}) & ... & E(W_{2m}) \\ \vdots & \vdots & \ddots & \vdots \\ E(W_{n1}) & E(W_{n2}) & ... & E(W_{nm}) \end{bmatrix}. \end{aligned}$$

Berikut merupakan teorema yang menunjukkan sifat linearitas dari nilai harapan.

Teorema 1.2 *Misalkan* ${}_n\mathbf{V}_m$ *dan* ${}_n\mathbf{W}_m$ *masing-masing merupakan matriks peubah acak;* ${}_k\mathbf{A}_n$, ${}_k\mathbf{B}_n$, *dan* ${}_m\mathbf{C}_p$ *masing-masing merupakan matriks berunsur konstanta, maka*

$$\begin{aligned} E[\mathbf{AV} + \mathbf{BW}] &= \mathbf{A}E[\mathbf{V}] + \mathbf{B}E[\mathbf{W}] \\ E[\mathbf{AWC}] &= \mathbf{A}E[\mathbf{W}]\mathbf{C}. \end{aligned}$$

Bukti. Misal $\mathbf{AV} = \left(\sum_{k=1}^{n} a_{ik} V_{kj}\right)$ dan $\mathbf{BW} = \left(\sum_{k=1}^{n} b_{ik} W_{kj}\right)$, maka

$$
\begin{aligned}
E\left[\mathbf{AV} + \mathbf{BW}\right] &= \left(E\left[\sum_{k=1}^{n} a_{ik} V_{kj} + \sum_{k=1}^{n} b_{ik} W_{kj}\right]\right) \\
&= \left(\sum_{k=1}^{n} a_{ik} E\left[V_{kj}\right]\right) + \left(\sum_{k=1}^{n} b_{ik} E\left[W_{kj}\right]\right) \\
&= \mathbf{A} E\left[\mathbf{V}\right] + \mathbf{B} E\left[\mathbf{W}\right].
\end{aligned}
$$

Pembuktian $E\left[\mathbf{AWC}\right] = \mathbf{A} E\left[\mathbf{W}\right] \mathbf{C}$ digunakan sebagai latihan. ✠

1.4 Kebebasan Vektor Peubah Acak

Misalkan peubah acak $X_1, X_2, ..., X_n$ memiliki fungsi kepekatan peluang bersama $f(x_1, x_2, ..., x_n)$ dan fungsi kepekatan peluang marginal $f_1(x_1), f_2(x_2), ..., f_n(x_n)$ secara berturut-turut bagi $x_1, ..., x_n$. Peubah acak $X_1, X_2, ..., X_n$ saling bebas jika

$$f(x_1, x_2, ..., x_n) = f_1(x_1) f_2(x_2) ... f_n(x_n),$$

untuk kasus kontinu. Pada kasus diskret, $X_1, X_2, ..., X_n$ dikatakan saling bebas jika

$$p(x_1, x_2, ..., x_n) = p_1(x_1) p_2(x_2) ... p_n(x_n).$$

Misalkan $X_1, X_2, ..., X_n$ saling bebas, maka

1. $P(a_1 < X_1 < b_1, a_2 < X_2 < b_2, ..., a_n < X_n < b_n)$
   $= P(a_1 < X_1 < b_1) P(a_2 < X_2 < b_2) ... P(a_n < X_n < b_n)$
   $= \prod_{i=1}^{n} P(a_i < X_i < b_i)$
2. $E\left[\prod_{i=1}^{n} u(X_i)\right] = \prod_{i=1}^{n} E\left[u(X_i)\right].$
3. Fungsi pembangkit momen bagi peubah acak $X_1, X_2, ..., X_n$ dinotasikan dengan $M_{\mathbf{X}}(t_1, t_2, ..., t_n)$ dan didefinisikan sebagai berikut

$$
\begin{aligned}
M_{\mathbf{X}}(t_1, t_2, ..., t_n) &= E\left[\exp(t_1 X_1 + t_2 X_2 + ... + t_n X_n)\right] \\
M_{\mathbf{X}}(\mathbf{t}) &= E\left[\exp(\mathbf{t}'\mathbf{X})\right], \forall \mathbf{t} \in B \subset \mathbb{R}^n,
\end{aligned}
$$

   dengan $B = \{\mathbf{t} : -h_i < t_i < h_i, i = 1, 2, ..., n\}$. Bagi fungsi sebaran marginal dari peubah acak X_i, fungsi pembangkit momen peubah acak X_i didefinisikan sebagai $M_{\mathbf{X}}(0, 0, ..., 0, t_i, 0, ..., 0), i = 1, 2, ..., n.$

Teorema 1.3 *Misalkan $X_1, X_2, ..., X_n$ adalah peubah acak yang saling bebas dan fungsi pembangkit momen bagi X_i ialah $M_{X_i}(t), \forall i = 1, ..., n$. Misalkan pula $T = \sum_{i=1}^{n} k_i X_i$, dengan $k_1, ..., k_n$ suatu konstanta, maka fungsi pembangkit momen bagi T ialah*

$$M_T(t) = \prod_{i=1}^{n} M_{X_i}(k_i t), -\min_i \{h_i\} < t < \min_i \{h_i\}.$$

Contoh 1.2 *Misalkan X_1, X_2, X_3 adalah peubah acak saling bebas yang masing-masing memiliki fungsi kepekatan peluang sebagai berikut*

$$f(x) = 2x \, \mathbf{I}(0 < x < 1).$$

Fungsi kepekatan peluang bersama bagi peubah acak tersebut ialah

$$\begin{aligned} f(x_1, x_2, x_3) &= f(x_1) f(x_2) f(x_3) \\ &= 8x_1x_2x_3 \, \mathbf{I}(0 < x_1, x_2, x_3 < 1). \end{aligned}$$

Nilai harapan bagi $5X_1 (X_2)^3 + 3X_2 (X_3)^4$ ialah

$$\int_0^1 \int_0^1 \int_0^1 \left(5x_1 (x_2)^3 + 3x_2 (x_3)^4\right) 8x_1x_2x_3 \quad dx_1dx_2dx_3 = 2.$$

Jika peubah acak saling bebas memiliki sebaran yang sama, maka peubah acak tersebut disebut peubah acak bebas stokastik identik (bsi). Berikut ini merupakan akibat dari Teorema 1.3 untuk peubah acak bsi.

Akibat 1.1 *Misalkan $X_1, X_2, ..., X_n$ adalah peubah acak bsi dengan fungsi pembangkit peluang $M_X(t), -h < t < h$ dengan $h > 0$. Misalkan pula $T = \sum_{i=1}^{n} X_i$, maka fungsi pembangkit momen bagi T ialah*

$$M_T(t) = [M_X(t)]^n, -h < t < h.$$

1.5 Dekomposisi Spektrum

Dari aljabar linear diperoleh bahwa suatu matriks dapat didekomposisikan atau diuraikan menjadi perkalian beberapa matriks. Khususnya bagi matriks

$\mathbf{\Sigma}$ yang merupakan matriks $n \times n$, simetrik, dan semi-definit positif, maka dekomposisi spektrum dari $\mathbf{\Sigma}$ ialah sebagai berikut

$$\mathbf{\Sigma} = \mathbf{\Gamma}'\mathbf{\Lambda}\mathbf{\Gamma},$$

di mana $\mathbf{\Lambda}$ merupakan matriks diagonal,

$$\mathbf{\Lambda} = diag\left(\lambda_1, \lambda_2, ..., \lambda_n\right),$$

$\lambda_1 \geq \lambda_2 \geq ... \geq \lambda_n \geq 0$ merupakan nilai eigen dari $\mathbf{\Sigma}$, dan kolom-kolom matriks $\mathbf{\Gamma}'$ ialah $\mathbf{v}_1$, $\mathbf{v}_2$, ..., $\mathbf{v}_n$ yang merupakan vektor eigen padanannya yang ortonormal. Nilai eigen tersebut tidak perlu ditata dari yang terbesar ke yang terkecil. Untuk analisis multivariat, tataan ini kita perlukan terkait dengan tataan ragam atau pentingnya peubah baru. Matriks $\mathbf{\Gamma}$ merupakan matriks ortogonal di mana $\mathbf{\Gamma}'\mathbf{\Gamma} = \mathbf{\Gamma}\mathbf{\Gamma}' = \mathbf{I}$. Matriks $\mathbf{\Gamma}$ tidak bersifat khas, karena vektor eigen suatu matriks (walaupun dipilih yang merupakan vektor satuan) tidak bersifat khas.

Penulisan dekomposisi tersebut dapat dituliskan dalam bentuk lain, yaitu

$$\mathbf{\Sigma} = \mathbf{\Gamma}'\mathbf{\Lambda}\mathbf{\Gamma} = \sum_{i=1}^{n} \lambda_i \mathbf{v}_i \mathbf{v}_i'.$$

Secara intuisi, pendekatan matriks dengan matriks lain (misalnya yang berperingkat lebih rendah) dapat diperoleh dengan membuang beberapa matriks terakhir yang terkait dengan nilai eigen yang kecil. Karena nilai eigen dari matriks semi-definit positif selalu ≥ 0, maka matriks $\mathbf{\Sigma}$ dapat ditulis dalam bentuk lain, yaitu

$$\mathbf{\Sigma} = \mathbf{\Sigma}^{\frac{1}{2}}\mathbf{\Sigma}^{\frac{1}{2}} = \left(\mathbf{\Sigma}^{\frac{1}{2}}\right)^2,$$

di mana $\mathbf{\Sigma}^{\frac{1}{2}} = \mathbf{\Gamma}'\mathbf{\Lambda}^{\frac{1}{2}}\mathbf{\Gamma}$ dengan $\mathbf{\Lambda}^{\frac{1}{2}} = diag\left(\sqrt{\lambda_1}, \sqrt{\lambda_2}, \ldots, \sqrt{\lambda_n}\right)$. Misalkan $\mathbf{\Sigma}$ merupakan matriks yang definit positif, yang berarti nilai eigennya selalu positif, maka

$$\left(\mathbf{\Sigma}^{\frac{1}{2}}\right)^{-1} = \mathbf{\Sigma}^{-\frac{1}{2}} = \mathbf{\Gamma}'\mathbf{\Lambda}^{-\frac{1}{2}}\mathbf{\Gamma}$$

dengan

$$\mathbf{\Lambda}^{-\frac{1}{2}} = diag\left(\frac{1}{\sqrt{\lambda_1}}, \frac{1}{\sqrt{\lambda_2}}, \ldots, \frac{1}{\sqrt{\lambda_n}}\right).$$

Contoh 1.3 *Tentukan dekomposisi spektrum dari matriks berikut*

$$\mathbf{\Sigma} = \begin{bmatrix} 3 & 0 & -1 \\ 0 & 4 & 0 \\ -1 & 0 & 3 \end{bmatrix}.$$

Tentukan pula $\mathbf{\Sigma}^{\frac{1}{2}}$ *dan* $\mathbf{\Sigma}^{-\frac{1}{2}}$.

80 **Jawab.** Berikut merupakan polinom karakteristik dari matriks $\mathbf{\Sigma}$:

$$
\begin{aligned}
|\mathbf{\Sigma} - \lambda\mathbf{I}| &= \begin{vmatrix} 3-\lambda & 0 & -1 \\ 0 & 4-\lambda & 0 \\ -1 & 0 & 3-\lambda \end{vmatrix} \\
&= (-1)^{2+2}(4-\lambda)\begin{vmatrix} 3-\lambda & -1 \\ -1 & 3-\lambda \end{vmatrix} \\
&= (4-\lambda)\left[(3-\lambda)^2 - 1\right] \\
&= (4-\lambda)(3-\lambda+1)(3-\lambda-1) \\
&= (4-\lambda)^2(2-\lambda),
\end{aligned}
$$

sehingga nilai eigen dari matriks tersebut ialah

$$\lambda_1 = \lambda_2 = 4, \lambda_3 = 2,$$

dan vektor-vektor eigen yang berpadanan dengan nilai eigen tersebut ialah

$$\mathbf{v}_1 = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \ \mathbf{v}_2 = \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}, \text{ dan } \mathbf{v}_3 = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}.$$

Vektor-vektor ortonormal dari $\mathbf{v}_1, \mathbf{v}_2$, dan $\mathbf{v}_3$ tersebut secara berturut-turut ialah

$$\mathbf{w}_1 = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \ \mathbf{w}_2 = \frac{1}{\sqrt{2}}\begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}, \text{ dan } \mathbf{w}_3 = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix},$$

sehingga matriks $\mathbf{\Lambda}$ dan $\mathbf{\Gamma}'$ ialah sebagai berikut

$$\mathbf{\Lambda} = \begin{bmatrix} 4 & 0 & 0 \\ 0 & 4 & 0 \\ 0 & 0 & 2 \end{bmatrix}, \ \mathbf{\Gamma}' = \begin{bmatrix} 0 & \frac{-1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ 1 & 0 & 0 \\ 0 & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix}.$$

81 Jadi, dekomposisi spektrum dari matriks $\mathbf{\Sigma}$ ialah

$$
\begin{aligned}
\mathbf{\Sigma} &= \mathbf{\Gamma}'\mathbf{\Lambda}^{1/2}\,\mathbf{\Gamma} \\
\begin{bmatrix} 3 & 0 & -1 \\ 0 & 4 & 0 \\ -1 & 0 & 3 \end{bmatrix} &= \begin{bmatrix} 0 & \frac{-1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ 1 & 0 & 0 \\ 0 & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix} \begin{bmatrix} 4 & 0 & 0 \\ 0 & 4 & 0 \\ 0 & 0 & 2 \end{bmatrix} \begin{bmatrix} 0 & 1 & 0 \\ \frac{-1}{\sqrt{2}} & 0 & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & 0 & \frac{1}{\sqrt{2}} \end{bmatrix}.
\end{aligned}
$$

82 Selanjutnya, berikut merupakan matriks $\mathbf{\Sigma}^{1/2}$ dan $\mathbf{\Sigma}^{-1/2}$.

$$
\begin{aligned}
\mathbf{\Sigma}^{1/2} &= \mathbf{\Gamma}'\mathbf{\Lambda}^{1/2}\,\mathbf{\Gamma} \\
&= \begin{bmatrix} 0 & \frac{-1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ 1 & 0 & 0 \\ 0 & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix} \begin{bmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & \sqrt{2} \end{bmatrix} \begin{bmatrix} 0 & 1 & 0 \\ \frac{-1}{\sqrt{2}} & 0 & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & 0 & \frac{1}{\sqrt{2}} \end{bmatrix} \\
&= \begin{bmatrix} \left(\frac{1}{2}\sqrt{2}+1\right) & 0 & \left(\frac{1}{2}\sqrt{2}-1\right) \\ 0 & 2 & 0 \\ \left(\frac{1}{2}\sqrt{2}-1\right) & 0 & \left(\frac{1}{2}\sqrt{2}+1\right) \end{bmatrix}.
\end{aligned}
$$

$$
\begin{aligned}
\mathbf{\Sigma}^{-1/2} &= \mathbf{\Gamma}'\mathbf{\Lambda}^{-1/2}\,\mathbf{\Gamma} \\
&= \begin{bmatrix} 0 & \frac{-1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ 1 & 0 & 0 \\ 0 & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix} \begin{bmatrix} \frac{1}{2} & 0 & 0 \\ 0 & \frac{1}{2} & 0 \\ 0 & 0 & \frac{1}{\sqrt{2}} \end{bmatrix} \begin{bmatrix} 0 & 1 & 0 \\ \frac{-1}{\sqrt{2}} & 0 & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & 0 & \frac{1}{\sqrt{2}} \end{bmatrix} \\
&= \begin{bmatrix} \left(\frac{1}{4}\sqrt{2}+\frac{1}{4}\right) & 0 & \left(\frac{1}{4}\sqrt{2}-\frac{1}{4}\right) \\ 0 & \frac{1}{2} & 0 \\ \left(\frac{1}{4}\sqrt{2}-\frac{1}{4}\right) & 0 & \left(\frac{1}{4}\sqrt{2}+\frac{1}{4}\right) \end{bmatrix}.
\end{aligned}
$$

83 ✠

84 1.6 Matriks Koragam

Misalkan X_i dan X_j masing-masing merupakan peubah acak ke-i dan ke-j dengan nilai harapan $E[X_i] = \mu_i$ dan $E[X_j] = \mu_j$, serta koragam $\sigma_{ij} = E\left([X_i - \mu_i][X_j - \mu_j]\right)$. Bila $\mathbf{X} = (X_1, X_2, ..., X_n)'$ adalah suatu vektor dari n peubah acak, maka

$$
E\left(\mathbf{X}\right) = \begin{bmatrix} \mu_1 \\ \mu_2 \\ \vdots \\ \mu_n \end{bmatrix},
$$

dan matriks koragam (*covariance matrix*) vektor peubah acak $\mathbf{X}$ ialah

$$\begin{aligned} Cov\left(\mathbf{X}\right) &= E\left(\left[\mathbf{X}-\boldsymbol{\mu}\right]\left[\mathbf{X}-\boldsymbol{\mu}\right]'\right) \\ &= \begin{bmatrix} E\left(\left[X_1-\boldsymbol{\mu}_1\right]\left[X_1-\boldsymbol{\mu}_1\right]\right) & \ldots & E\left(\left[X_1-\boldsymbol{\mu}_1\right]\left[X_m-\boldsymbol{\mu}_m\right]\right) \\ E\left(\left[X_2-\boldsymbol{\mu}_2\right]\left[X_1-\boldsymbol{\mu}_1\right]\right) & \ldots & E\left(\left[X_2-\boldsymbol{\mu}_2\right]\left[X_m-\boldsymbol{\mu}_m\right]\right) \\ \vdots & \ddots & \vdots \\ E\left(\left[X_n-\boldsymbol{\mu}_n\right]\left[X_1-\boldsymbol{\mu}_1\right]\right) & \ldots & E\left(\left[X_n-\boldsymbol{\mu}_n\right]\left[X_m-\boldsymbol{\mu}_m\right]\right) \end{bmatrix} \\ &= \begin{bmatrix} \sigma_{11} & \sigma_{12} & \ldots & \sigma_{1m} \\ \sigma_{21} & \sigma_{22} & \ldots & \sigma_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{n1} & \sigma_{n2} & \ldots & \sigma_{nm} \end{bmatrix}. \end{aligned} \tag{1.1}$$

Misalkan $\mathbf{a}$ adalah vektor berunsur konstanta yang berukuran $n \times 1$. Misalkan pula $Y = \mathbf{a}'\mathbf{X}$ adalah peubah acak, maka Y memiliki ragam bernilai non-negatif, yaitu $0 \leq Var\left(Y\right) = Var\left(\mathbf{a}'\mathbf{X}\right) = \mathbf{a}'Cov\left(\mathbf{X}\right)\mathbf{a}$. Dengan demikian, matriks koragam pada 1.1 merupakan matriks semi-definit positif.

Secara umum, kita dapat mendefinisikan matriks koragam antara vektor p-peubah acak $\mathbf{X} = \left(X_1, X_2, ..., X_p\right)'$ dengan vektor q-peubah acak $\mathbf{Y} = \left(Y_1, Y_2, ..., Y_q\right)'$ sebagai

$$\begin{aligned} Cov\left(\mathbf{X},\mathbf{Y}\right) &= E\left(\left[\mathbf{X}-E\left(\mathbf{X}\right)\right]\left[\mathbf{Y}-E\left(\mathbf{Y}\right)\right]'\right) \\ &= \begin{bmatrix} E\left(\left[X_1-E\left(X_1\right)\right]\left[Y_1-E\left(Y_1\right)\right]\right) & \ldots & E\left(\left[X_1-E\left(X_1\right)\right]\left[Y_q-E\left(Y_q\right)\right]\right) \\ E\left(\left[X_2-E\left(X_2\right)\right]\left[Y_1-E\left(Y_1\right)\right]\right) & \ldots & E\left(\left[X_2-E\left(X_2\right)\right]\left[Y_q-E\left(Y_q\right)\right]\right) \\ \vdots & \ddots & \vdots \\ E\left(\left[X_p-E\left(X_p\right)\right]\left[Y_1-E\left(Y_1\right)\right]\right) & \ldots & E\left(\left[X_p-E\left(X_p\right)\right]\left[Y_q-E\left(Y_q\right)\right]\right) \end{bmatrix}, \end{aligned}$$

sehingga $Cov\left(\mathbf{X}\right)$ merupakan $Cov\left(\mathbf{X},\mathbf{X}\right)$. Selanjutnya, misalkan $\mathbf{X}$ dan $\mathbf{Y}$ merupakan vektor peubah acak dengan vektor nilai harapan $\boldsymbol{\mu}_{\mathbf{X}}$ dan $\boldsymbol{\mu}_{\mathbf{Y}}$, maka

$$\begin{aligned} \boldsymbol{\Sigma}_{\mathbf{X},\mathbf{Y}} &= Cov\left(\mathbf{X},\mathbf{Y}\right) \\ &= E\left(\left[\mathbf{X}-E\left(\mathbf{X}\right)\right]\left[\mathbf{Y}-E\left(\mathbf{Y}\right)\right]'\right) \\ &= E\left(\mathbf{X}\mathbf{Y}'\right) - E\left(\mathbf{X}\right)\left[E\left(\mathbf{Y}\right)\right]' \\ &= E\left(\mathbf{X}\mathbf{Y}'\right) - \boldsymbol{\mu}_{\mathbf{X}}\boldsymbol{\mu}'_{\mathbf{Y}}, \end{aligned}$$

sehingga

$$\begin{aligned} \boldsymbol{\Sigma}_{\mathbf{X}} &= \boldsymbol{\Sigma}_{\mathbf{X},\mathbf{X}} \\ &= E\left(\mathbf{X}\mathbf{X}'\right) - \boldsymbol{\mu}_{\mathbf{X}}\boldsymbol{\mu}'_{\mathbf{X}}. \end{aligned}$$

Bila $\mathbf{A}$ dan $\mathbf{B}$ merupakan matriks konstanta, maka $E(\mathbf{AX}) = \mathbf{A}\boldsymbol{\mu}_{\mathbf{X}}$ dan $E(\mathbf{BY}) = \mathbf{B}\boldsymbol{\mu}_{\mathbf{Y}}$, serta

$$\begin{aligned}
Cov(\mathbf{AX}, \mathbf{BY}) &= \boldsymbol{\Sigma}_{\mathbf{AX},\mathbf{BY}} \\
&= E\left([\mathbf{AX} - E(\mathbf{AX})]\,[\mathbf{BY} - E(\mathbf{BY})]'\right) \\
&= E\left([\mathbf{AX} - \mathbf{A}E(\mathbf{X})]\,[\mathbf{BY} - \mathbf{B}E(\mathbf{Y})]'\right) \\
&= E\left([\mathbf{AX} - \mathbf{A}\boldsymbol{\mu}_{\mathbf{X}}]\,[\mathbf{BY} - \mathbf{B}\boldsymbol{\mu}_{\mathbf{Y}}]'\right) \\
&= E\left([\mathbf{A}(\mathbf{X} - \boldsymbol{\mu}_{\mathbf{X}})]\,[\mathbf{B}(\mathbf{Y} - \boldsymbol{\mu}_{\mathbf{Y}})]'\right) \\
&= E\left(\mathbf{A}(\mathbf{X} - \boldsymbol{\mu}_{\mathbf{X}})(\mathbf{Y} - \boldsymbol{\mu}_{\mathbf{Y}})'\,\mathbf{B}'\right) \\
&= \mathbf{A}E\left((\mathbf{X} - \boldsymbol{\mu}_{\mathbf{X}})(\mathbf{Y} - \boldsymbol{\mu}_{\mathbf{Y}})'\right)\mathbf{B}' \\
&= \mathbf{A}\boldsymbol{\Sigma}_{\mathbf{X},\mathbf{Y}}\mathbf{B}',
\end{aligned}$$

sehingga

$$Cov(\mathbf{AX}) = \boldsymbol{\Sigma}_{\mathbf{AX}} = \mathbf{A}\boldsymbol{\Sigma}_{\mathbf{X}}\mathbf{A}'.$$

Misalkan pula

$$\begin{aligned}
\mathbf{V} &= \mathbf{AX} + \boldsymbol{\delta} \\
\mathbf{W} &= \mathbf{BY} + \boldsymbol{\eta}
\end{aligned}$$

dengan $\boldsymbol{\delta}$ dan $\boldsymbol{\eta}$ merupakan vektor berunsur konstanta, maka

$$\begin{aligned}
E(\mathbf{V}) &= \mathbf{A}\boldsymbol{\mu}_{\mathbf{X}} + \boldsymbol{\delta}, \\
E(\mathbf{W}) &= \mathbf{B}\boldsymbol{\mu}_{\mathbf{Y}} + \boldsymbol{\eta}, \\
Cov(\mathbf{V}, \mathbf{W}) &= \mathbf{A}\boldsymbol{\Sigma}_{\mathbf{X},\mathbf{Y}}\mathbf{B}',
\end{aligned}$$

sehingga

$$Cov(\mathbf{V}) = \mathbf{A}\boldsymbol{\Sigma}_{\mathbf{X}}\mathbf{A}'.$$

Contoh 1.4 *Misalkan* $\mathbf{X} = [X, Y, Z]'$ *merupakan vektor peubah acak dengan* $E(\mathbf{X}) = [1, 4, -6]'$ *dan matriks koragam*

$$\Sigma = \begin{bmatrix} 3 & 2 & 1 \\ 2 & 2 & 1 \\ 1 & 1 & 3 \end{bmatrix}.$$

Tentukan nilai harapan serta matriks koragam dari peubah acak T *dan* $\mathbf{W}$ *jika*

$$T = 2X - Y + 3Z + 1,\ \mathbf{W} = \begin{bmatrix} X - Y + 2Z + 2 \\ 2X + 4Y - Z + 3 \end{bmatrix}.$$

Jawab. Peubah acak T dan $\mathbf{W}$ dapat dinyatakan sebagai suatu vektor, yaitu

$$\begin{aligned} T &= 2X - Y + 3Z + 1 = [\,2, -1, 3]\,\mathbf{X} + 1, \\ \mathbf{W} &= \begin{bmatrix} X - Y + 2Z \\ 2X + 4Y - Z \end{bmatrix} = \begin{bmatrix} 1 & -1 & 2 \\ 2 & 4 & -1 \end{bmatrix} \mathbf{X} + \begin{bmatrix} 2 \\ 3 \end{bmatrix}. \end{aligned}$$

Selanjutnya, menentukan nilai harapan dan matriks koragam bagi masing-masing peubah acak.

$$\begin{aligned} E(T) &= [\,2, -1, 3]\, E(\mathbf{X}) + 1 \\ &= [\,2, -1, 3] \begin{bmatrix} 1 \\ 4 \\ -6 \end{bmatrix} + 1 \\ &= -20 + 1 \\ &= -19 \\ \boldsymbol{\Sigma}_T &= [\,2, -1, 3] \begin{bmatrix} 3 & 2 & 1 \\ 2 & 2 & 1 \\ 1 & 1 & 3 \end{bmatrix} \begin{bmatrix} 2 \\ -1 \\ 3 \end{bmatrix} \\ &= 39 \end{aligned}$$

$$\begin{aligned} E(\mathbf{W}) &= \begin{bmatrix} 1 & -1 & 2 \\ 2 & 4 & -1 \end{bmatrix} E(\mathbf{X}) + \begin{bmatrix} 2 \\ 3 \end{bmatrix} \\ &= \begin{bmatrix} 1 & -1 & 2 \\ 2 & 4 & -1 \end{bmatrix} \begin{bmatrix} 1 \\ 4 \\ -6 \end{bmatrix} + \begin{bmatrix} 2 \\ 3 \end{bmatrix} \\ &= \begin{bmatrix} 15 \\ 24 \end{bmatrix} + \begin{bmatrix} 2 \\ 3 \end{bmatrix} \\ &= \begin{bmatrix} 17 \\ 27 \end{bmatrix} \\ \boldsymbol{\Sigma}_{\mathbf{W}} &= \begin{bmatrix} 1 & -1 & 2 \\ 2 & 4 & -1 \end{bmatrix} \begin{bmatrix} 3 & 2 & 1 \\ 2 & 2 & 1 \\ 1 & 1 & 3 \end{bmatrix} \begin{bmatrix} 1 & 2 \\ -1 & 4 \\ 2 & -1 \end{bmatrix} \\ &= \begin{bmatrix} 13 & 8 \\ 8 & 67 \end{bmatrix}. \end{aligned}$$

$\maltese$

Contoh 1.5 *Misalkan* $\mathbf{X} = (X_1, X_2, X_3)'$, $\mathbf{Y} = (Y_1, Y_2, Y_3, Y_4)'$, *dan*

$$Cov\,(\mathbf{X}, \mathbf{Y}) = \begin{bmatrix} 1 & 3 & -2 & 5 \\ -1 & 2 & 4 & 6 \\ 2 & 1 & 0 & 7 \end{bmatrix}.$$

Tentukan :

1. $Cov\,(X_1 + 2X_2 - 3X_3 + 1,\ 2Y_1 + Y_2 + 4Y_3 - 5Y_4 + 2)\,.$
2. $Cov\,(\mathbf{V}, \mathbf{W})$ jika

$$\mathbf{V} = \begin{bmatrix} X_1 + 2X_2 - 3X_3 + 3 \\ 4X_1 - 2X_2 + 3X_3 + 4 \end{bmatrix} \text{ dan } \mathbf{W} = \begin{bmatrix} Y_1 - 2Y_2 + 3Y_3 - 5Y_4 + 7 \\ 2Y_1 + Y_2 - 4Y_3 - Y_4 \\ Y_1 + Y_2 - 6Y_3 + 2Y_4 \end{bmatrix}.$$

1. Misalkan $U = X_1 + 2X_2 - 3X_3 + 1$ dan $Z = 2Y_1 + Y_2 + 4Y_3 - 5Y_4 + 2$, maka $U = [1, 2, -3]\,\mathbf{X} + 1$ dan $Z = [2, 1, 4, -5]\,\mathbf{Y} + 2$.

$$Cov\,(U{,}Z) = (1, 2, -3) \begin{bmatrix} 1 & 3 & -2 & 5 \\ -1 & 2 & 4 & 6 \\ 2 & 1 & 0 & 7 \end{bmatrix} \begin{bmatrix} 2 \\ 1 \\ 4 \\ -5 \end{bmatrix} = 34.$$

2. $\mathbf{V} = \begin{bmatrix} 1 & 2 & -3 \\ 4 & -2 & 1 \end{bmatrix} \mathbf{X} + \begin{bmatrix} 3 \\ 4 \end{bmatrix}$, $\mathbf{W} = \begin{bmatrix} 1 & -2 & 3 & -5 \\ 2 & 1 & -4 & -1 \\ 1 & 1 & -6 & 2 \end{bmatrix} \mathbf{Y} + \begin{bmatrix} 7 \\ 0 \\ 0 \end{bmatrix}$.

$$\begin{aligned} Cov\,(\mathbf{V}, \mathbf{W}) &= \begin{bmatrix} 1 & 2 & -3 \\ 4 & -2 & 1 \end{bmatrix} \begin{bmatrix} 1 & 3 & -2 & 5 \\ -1 & 2 & 4 & 6 \\ 2 & 1 & 0 & 7 \end{bmatrix} \begin{bmatrix} 1 & 2 & 1 \\ -2 & 1 & 1 \\ 3 & -4 & -6 \\ -5 & -1 & 2 \end{bmatrix} \\ &= \begin{bmatrix} 23 & -30 & -47 \\ -133 & 74 & 143 \end{bmatrix}. \end{aligned}$$

1.7 Sebaran Normal Multivariat

Pada subbab ini dibahas sebaran normal multivariat untuk vektor peubah acak dimensi-n. Misalkan vektor peubah acak $\mathbf{Z} = (Z_1, ..., Z_n)'$ dengan $Z_1, ..., Z_n$

merupakan peubah acak bebas stokastik identik yang memiliki sebaran normal baku, $Z_i \sim N(0,1), i = 1, ..., n$. Fungsi kepekatan peluang bagi $\mathbf{Z}$ ialah

$$
\begin{aligned}
f_{\mathbf{Z}}(\mathbf{z}) &= \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi}} \exp\left\{-\frac{1}{2} z_i^2\right\} \\
&= \left(\frac{1}{2\pi}\right)^{\frac{n}{2}} \exp\left\{-\frac{1}{2}\sum_{i=1}^{n} z_i^2\right\} \\
&= \left(\frac{1}{2\pi}\right)^{\frac{n}{2}} \exp\left\{-\frac{1}{2}\mathbf{z}'\mathbf{z}\right\}, \mathbf{z} \in \mathbb{R}^n.
\end{aligned}
$$

Karena $E(Z_i) = 0$, $Var(Z_i) = 1$, dan $Cov(Z_i, Z_j) = 0$ untuk $\forall i \neq j$, maka

1. $E(\mathbf{Z}) = \mathbf{0}$,
2. $Cov(\mathbf{Z}) = \mathbf{\Sigma_Z} = E\left([\mathbf{Z} - E(\mathbf{Z})][\mathbf{Z} - E(\mathbf{Z})]'\right)$

$$
= \begin{bmatrix} Var(Z_1) & 0 & \cdots & 0 \\ 0 & Var(Z_1) & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & Var(Z_n) \end{bmatrix}
$$

$$
= \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{bmatrix}
$$

$$
= \mathbf{I}_n,
$$

dengan $\mathbf{I}_n$ adalah matriks identitas berukuran $n \times n$.

Selanjutnya, karena peubah acak Z_i bebas stokastik identik, maka fungsi pembangkit momen bagi $\mathbf{Z}$ ialah

$$\begin{aligned} M_{\mathbf{Z}}(\mathbf{t}) &= E\left[\exp\left\{\mathbf{t}'\mathbf{Z}\right\}\right] \\ &= E\left[\prod_{i=1}^{n} \exp\left\{t_i Z_i\right\}\right] \\ &= \prod_{i=1}^{n} E\left[\exp\left\{t_i Z_i\right\}\right] \\ &= \prod_{i=1}^{n} M_Z(t_i) \\ &= \prod_{i=1}^{n} \exp\left\{\frac{1}{2} t_i^2\right\} \\ &= \exp\left\{\frac{1}{2}\sum_{i=1}^{n} t_i^2\right\} \\ &= \exp\left\{\frac{1}{2}\mathbf{t}'\mathbf{t}\right\}, \forall \mathbf{t} \in \mathbb{R}^n. \end{aligned}$$

Dengan demikian, vektor peubah acak $\mathbf{Z}$ dikatakan memiliki sebaran normal multivariat dengan vektor nilai harapan $\mathbf{0}$ dan matriks koragam $\mathbf{I}_n$ yang dinotasikan dengan $\mathbf{Z} \sim N_n(\mathbf{0}, \mathbf{I}_n)$.

Misalkan $\mathbf{Z}$ memiliki sebaran $N_n(\mathbf{0}, \mathbf{I}_n)$. Misalkan pula $\boldsymbol{\Sigma}$ merupakan matriks simetrik dan semi-definit positif, serta $\boldsymbol{\mu}$ adalah vektor berunsur konstanta. Definisikan suatu vektor peubah acak $\mathbf{X}$ sebagai berikut

$$\mathbf{X} = \boldsymbol{\Sigma}^{\frac{1}{2}}\mathbf{Z} + \boldsymbol{\mu}, \tag{1.2}$$

maka diperoleh

$$E(\mathbf{X}) = \boldsymbol{\mu} \text{ dan } Cov(\mathbf{X}) = \boldsymbol{\Sigma}^{\frac{1}{2}}\boldsymbol{\Sigma}^{\frac{1}{2}} = \boldsymbol{\Sigma}.$$

Kemudian, fungsi pembangkit momen bagi $\mathbf{X}$ ialah

$$\begin{aligned}
M_{\mathbf{X}}(\mathbf{t}) &= E\left[\exp\left\{\mathbf{t}'\mathbf{X}\right\}\right] \\
&= E\left[\exp\left\{\mathbf{t}'\boldsymbol{\Sigma}^{\frac{1}{2}}\mathbf{Z}+\mathbf{t}'\boldsymbol{\mu}\right\}\right] \\
&= \exp\left\{\mathbf{t}'\boldsymbol{\mu}\right\} E\left[\exp\left\{\mathbf{t}'\boldsymbol{\Sigma}^{\frac{1}{2}}\mathbf{Z}\right\}\right] \\
&= \exp\left\{\mathbf{t}'\boldsymbol{\mu}\right\} \exp\left\{\frac{1}{2}\left(\boldsymbol{\Sigma}^{\frac{1}{2}}\mathbf{t}\right)'\boldsymbol{\Sigma}^{\frac{1}{2}}\mathbf{t}\right\} \\
&= \exp\left\{\mathbf{t}'\boldsymbol{\mu}\right\} \exp\left\{\frac{1}{2}\mathbf{t}'\boldsymbol{\Sigma}\mathbf{t}\right\} \\
&= \exp\left\{\mathbf{t}'\boldsymbol{\mu}+\frac{1}{2}\mathbf{t}'\boldsymbol{\Sigma}\mathbf{t}\right\}.
\end{aligned}$$

Definisi 1.5 (*Normal Multivariat*) *Vektor peubah acak* $\mathbf{X}$ *memiliki sebaran normal multivariat atau* $\mathbf{X} \sim N_n(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ *jika* $\mathbf{X}$ *memiliki fungsi pembangkit momen sebagai berikut*

$$M_{\mathbf{X}}(\mathbf{t}) = \exp\left\{\mathbf{t}'\boldsymbol{\mu}+\frac{1}{2}\mathbf{t}'\boldsymbol{\Sigma}\mathbf{t}\right\},$$

untuk setiap $\mathbf{t} \in \mathbb{R}^n$ *dan* $\boldsymbol{\Sigma}$ *merupakan matriks simetrik, semi-definit positif, serta* $\boldsymbol{\mu} \in \mathbb{R}^n$.

Jika $\boldsymbol{\Sigma}$ merupakan matriks definit positif, maka matriks $\boldsymbol{\Sigma}^{\frac{1}{2}}$ memiliki invers, sehingga transformasi satu-satu antara peubah acak $\mathbf{X}$ dan $\mathbf{Z}$ pada persamaan 1.2

$$\mathbf{Z} = \boldsymbol{\Sigma}^{-\frac{1}{2}}(\mathbf{X}-\boldsymbol{\mu})$$

dengan Jacobi $\left|\boldsymbol{\Sigma}^{-1/2}\right| = |\boldsymbol{\Sigma}|^{-1/2}$ menghasilkan fungsi kepekatan peluang bagi $\mathbf{X}$, yaitu

$$f_{\mathbf{X}}(\mathbf{x}) = \frac{1}{(2\pi)^{n/2}|\boldsymbol{\Sigma}|^{1/2}} \exp\left\{-\frac{1}{2}(\mathbf{x}-\boldsymbol{\mu})'\boldsymbol{\Sigma}^{-1}(\mathbf{x}-\boldsymbol{\mu})\right\}, \mathbf{x} \in R^n.$$

Teorema 1.4 *Misalkan* $\mathbf{X} \sim N_n(\boldsymbol{\mu}, \boldsymbol{\Sigma})$. *Misal* $\mathbf{Y} = \mathbf{A}\mathbf{X}+\mathbf{b}$ *dengan* $\mathbf{A}$ *adalah matriks berukuran* $m \times n$ *dan* $\mathbf{b} \in \mathbb{R}^m$, *maka* $\mathbf{Y} \sim N_m(\mathbf{A}\boldsymbol{\mu}+\mathbf{b}, \mathbf{A}\boldsymbol{\Sigma}\mathbf{A}')$.

Bukti. $\forall \mathbf{t} \in \mathbb{R}^m$, fungsi pembangkit momen bagi $\mathbf{Y}$ ialah

$$\begin{aligned} M_{\mathbf{Y}}(\mathbf{t}) &= E\left[\exp\left\{\mathbf{t}'\mathbf{Y}\right\}\right] \\ &= E\left[\exp\left\{\mathbf{t}'\left(\mathbf{AX}+\mathbf{b}\right)\right\}\right] \\ &= \exp\left\{\mathbf{t}'\mathbf{b}\right\} E\left[\exp\left\{\left(\mathbf{A}'\mathbf{t}\right)'\mathbf{X}\right\}\right] \\ &= \exp\left\{\mathbf{t}'\mathbf{b}\right\}\exp\left\{\left(\mathbf{A}'\mathbf{t}\right)'\boldsymbol{\mu}+\frac{1}{2}\left(\mathbf{A}'\mathbf{t}\right)'\boldsymbol{\Sigma}\left(\mathbf{A}'\mathbf{t}\right)\right\} \\ &= \exp\left\{\mathbf{t}'\left(\mathbf{A}\boldsymbol{\mu}+\mathbf{b}\right)+\frac{1}{2}\mathbf{t}'\mathbf{A}\boldsymbol{\Sigma}\mathbf{A}'\mathbf{t}\right\}, \end{aligned}$$

di mana fungsi pembangkit momen tersebut adalah fungsi pembangkit momen bagi sebaran $N_m\left(\mathbf{A}\boldsymbol{\mu}+\mathbf{b}, \mathbf{A}\boldsymbol{\Sigma}\mathbf{A}'\right)$. ✠

Misalkan $\mathbf{X}_1$ merupakan subvektor berdimensi $m < n$ dari vektor peubah acak $\mathbf{X}$, maka $\mathbf{X}$ dapat dituliskan menjadi

$$\mathbf{X} = \begin{bmatrix} \mathbf{X}_1 \\ \mathbf{X}_2 \end{bmatrix}, \tag{1.3}$$

dengan $\mathbf{X}_2$ merupakan vektor berdimensi $p = n - m$. Karena $\mathbf{X}$ dipartisi seperti pada persamaan 1.3, maka nilai harapan dan matriks koragam dari $\mathbf{X}$ juga dapat dipartisi sebagai berikut

$$\boldsymbol{\mu} = \begin{bmatrix} \boldsymbol{\mu}_1 \\ \boldsymbol{\mu}_2 \end{bmatrix} \text{ dan } \boldsymbol{\Sigma} = \begin{bmatrix} \boldsymbol{\Sigma}_{11} & \boldsymbol{\Sigma}_{12} \\ \boldsymbol{\Sigma}_{21} & \boldsymbol{\Sigma}_{22} \end{bmatrix}, \tag{1.4}$$

dengan $\boldsymbol{\Sigma}_{11}$ merupakan matriks koragam bagi $\mathbf{X}_1$ dan $\boldsymbol{\Sigma}_{12}$ merupakan matriks koragam antara $\mathbf{X}_1$ dan $\mathbf{X}_2$. Kemudian, definisikan matriks $\mathbf{A}$ sebagai berikut

$$\mathbf{A} = \begin{bmatrix} \mathbf{I}_m & \vdots & \mathbf{O}_{mp} \end{bmatrix},$$

di mana $\mathbf{O}_{mp}$ merupakan matriks nol $m \times p$, maka $\mathbf{X}_1 = \mathbf{AX}$. Dengan menggunakan Teorema 1.4 pada transformasi tersebut, diperoleh akibat sebagai berikut:

Akibat 1.2 *Misalkan* $\mathbf{X} \sim N_n\left(\boldsymbol{\mu}, \boldsymbol{\Sigma}\right)$. *Jika* $\mathbf{X}_1$ *merupakan subvektor berdimensi* $m < n$ *dari vektor peubah acak* $\mathbf{X}$ (*persamaan 1.3*), *maka* $\mathbf{X}_1 \sim N_n\left(\boldsymbol{\mu}_1, \boldsymbol{\Sigma}_{11}\right)$.

Akibat 1.2 dapat digunakan untuk mencari sebaran marginal dari peubah acak normal multivariat.

Teorema 1.5 *Misalkan* $\mathbf{X} \sim N_n(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ *dan* $\mathbf{X}$ *dipartisi seperti pada persamaan 1.3 dan 1.4.* $\mathbf{X}_1$ *dan* $\mathbf{X}_2$ *saling bebas jika dan hanya jika* $\boldsymbol{\Sigma}_{12} = \mathbf{O}$.

Bukti. Disediakan untuk latihan. $\maltese$

Teorema 1.6 *Misalkan* $\mathbf{X} \sim N_n(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ *dan* $\mathbf{X}$ *dipartisi seperti pada persamaan 1.3 dan 1.4. Asumsikan bahwa* $\boldsymbol{\Sigma}$ *adalah matriks definit positif, maka sebaran bersyarat dari* $\mathbf{X}_1|\mathbf{X}_2$ *ialah*

$$N_m\left(\boldsymbol{\mu}_1 + \boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1}(\mathbf{X}_2 - \boldsymbol{\mu}_2), \boldsymbol{\Sigma}_{11} - \boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1}\boldsymbol{\Sigma}_{21}\right).$$

Bukti. Perhatikan vektor peubah acak $\mathbf{W} = \mathbf{X}_1 - \boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1}\mathbf{X}_2$ dan $\mathbf{X}_2$, maka

$$\begin{bmatrix} \mathbf{W} \\ \mathbf{X}_2 \end{bmatrix} = \begin{bmatrix} \mathbf{I}_m & -\boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1} \\ \mathbf{O} & \mathbf{I}_p \end{bmatrix} \begin{bmatrix} \mathbf{X}_1 \\ \mathbf{X}_2 \end{bmatrix}.$$

Berdasarkan Teorema 1.4, $E(\mathbf{W}) = \boldsymbol{\mu}_1 - \boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1}\boldsymbol{\mu}_2, E(\mathbf{X}_2) = \boldsymbol{\mu}_2$, dan matriks koragam

$$\begin{aligned} & \begin{bmatrix} \mathbf{I}_m & -\boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1} \\ \mathbf{O} & \mathbf{I}_p \end{bmatrix} \begin{bmatrix} \boldsymbol{\Sigma}_{11} & \boldsymbol{\Sigma}_{12} \\ \boldsymbol{\Sigma}_{21} & \boldsymbol{\Sigma}_{22} \end{bmatrix} \begin{bmatrix} \mathbf{I}_m & \mathbf{O} \\ -\boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1} & \mathbf{I}_p \end{bmatrix} \\ = & \begin{bmatrix} \boldsymbol{\Sigma}_{11} - \boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1}\boldsymbol{\Sigma}_{21} & \mathbf{O} \\ \mathbf{O} & \boldsymbol{\Sigma}_{22} \end{bmatrix}. \end{aligned}$$

Dengan demikian, berdasarkan Teorema 1.5, vektor peubah acak $\mathbf{W}$ dan $\mathbf{X}_2$ saling bebas. Karena saling bebas, sebaran bersyarat $\mathbf{W}|\mathbf{X}_2$ sama dengan sebaran marginal dari $\mathbf{W} \sim N_m\left(\boldsymbol{\mu}_1 - \boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1}\boldsymbol{\mu}_2, \boldsymbol{\Sigma}_{11} - \boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1}\boldsymbol{\Sigma}_{21}\right)$. Akibatnya,

$$\mathbf{X}_1|\mathbf{X}_2 \sim N_m\left(\boldsymbol{\mu}_1 - \boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1}\boldsymbol{\mu}_2 + \boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1}\mathbf{X}_2, \boldsymbol{\Sigma}_{11} - \boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1}\boldsymbol{\Sigma}_{21}\right),$$

atau

$$\left(\mathbf{W} + \boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1}\mathbf{X}_2\right)|\mathbf{X}_2 \sim N_m\left(\boldsymbol{\mu}_1 - \boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1}\boldsymbol{\mu}_2 + \boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1}\mathbf{X}_2, \boldsymbol{\Sigma}_{11} - \boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1}\boldsymbol{\Sigma}_{21}\right).$$

$\maltese$

Contoh 1.6 *Misalkan dalam suatu populasi, bobot badan dan tinggi badan pria memiliki sebaran normal bivariat. Jika diketahui bahwa rata-rata dan ragam dari bobot badan pria secara berturut-turut sebesar* 60 *kg dan* 25 *kg*^{2}*, kemudian rata-rata dan ragam dari tinggi badannya secara berturut-turut sebesar* 165 *cm dan* 100 *cm*^{2}*, serta besar korelasi antara bobot badan dan tinggi badan sebesar* 0.6*, maka tentukan :*

a. $P(160 < Y < 175)$,

b. $P(160 < Y < 175 | X = 65)$,

dengan X dan Y secara berturut-turut menyatakan bobot badan dan tinggi badan pria.

Jawab.

a. $P(160 < Y < 175)$

$$\begin{aligned} P(160 < Y < 175) &= P\left(\frac{160-\mu_y}{\sigma_y} < \frac{Y-\mu_y}{\sigma_y} < \frac{175-\mu_y}{\sigma_y}\right) \\ &= P\left(\frac{160-165}{10} < Z < \frac{175-165}{10}\right) \\ &= P(-0.5 < Z < 1) \\ &= \Phi(1) - \Phi(-0.5) \\ &= \Phi(1) - [1 - \Phi(0.5)] \\ &= 0.8413 - [1 - 0.6915] \\ &= 0.5328. \end{aligned}$$

b. $P(160 < Y < 175 | X = 65)$

$$\begin{bmatrix} Y \\ X \end{bmatrix} \sim N_2\left[\begin{pmatrix} \mu_y \\ \mu_x \end{pmatrix}, \begin{pmatrix} \sigma_{11} & \sigma_{12} \\ \sigma_{21} & \sigma_{22} \end{pmatrix}\right]$$

catatan: $\sigma_y^2 = \sigma_{11}$, $\sigma_x^2 = \sigma_{22}$, dan $\sigma_{12} = \sigma_{21} = \rho\sigma_x\sigma_y$.

Berdasarkan Teorema 1.6,

$$\begin{aligned} Y|X &\sim N_2\left(\mu_y + \frac{\sigma_{12}}{\sigma_{22}}(x-\mu_x), \sigma_{11} - \frac{\sigma_{12}}{\sigma_{22}}\sigma_{21}\right) \\ &\sim N_2\left(\mu_y + \frac{\rho\sigma_x\sigma_y}{\sigma_x^2}(x-\mu_x), \sigma_y^2 - \frac{\rho\sigma_x\sigma_y}{\sigma_x^2}(\rho\sigma_x\sigma_y)\right) \\ &\sim N_2\left(\mu_y + \rho\frac{\sigma_y}{\sigma_x}(x-\mu_x), \sigma_y^2\left(1-\rho^2\right)\right) \end{aligned}$$

$$\begin{aligned} E(Y|x=65) &= 165 + 0.6\left(\frac{10}{5}\right)(65-60) = 171 \\ Var(Y|x=65) &= 100(1-0.36) = 64. \end{aligned}$$

Jadi,

$$\begin{aligned}
P\left(160 < Y < 175 | X = 65\right) &= P\left(\frac{160-171}{8} < Z < \frac{175-171}{8}\right)\\
&= \Phi\left(0.5\right) - \Phi\left(-1.375\right)\\
&= \Phi\left(0.5\right) - \left[1 - \Phi\left(1.375\right)\right]\\
&= 0.6915 - \left[1 - 0.9154\right]\\
&= 0.6069.
\end{aligned}$$

1.8 Latihan

1. Misalkan X, Y, dan Z adalah tiga peubah acak diskret yang menyebar bersama dengan fungsi massa peluang

$$p_{X,Y,Z}\left(x,y,z\right) = c\left(x+y+z\right)\ \mathbf{I}\left(x = 1,2;\ y = 1,2;\ z = 1,3\right)$$

   a. Tentukan nilai c.

   b. Tentukan fungsi massa peluang marginal $p_{X,Y}\left(x,y\right)$ dan $p_X\left(x\right)$.

   c. Tentukan $E\left(X+Y+Z\right)$.

2. Misalkan X, Y, dan Z adalah tiga peubah acak kontinu yang menyebar bersama dengan fungsi kepekatan peluang

$$f_{X,Y,Z}\left(x,y,z\right) = cxyz\ \mathbf{I}\left(0 \le x \le\ y \le z \le 2\right)$$

   a. Tentukan nilai c.

   b. Tentukan fungsi kepekatan peluang marginal $f_{X,Y}\left(x,y\right)$ dan $f_X\left(x\right)$.

   c. Tentukan $E\left(XY+Z\right)$.

3. Misalkan X dan Y memiliki sebaran normal bivariat dengan $\mu_x = 3$, $\mu_y = 1$, $\sigma_x^2 = 16$, $\sigma_y^2 = 25$, dan $\rho = \frac{3}{5}$. Tentukan:

   a. $P\left(3 < Y < 8\right)$.

   b. $P\left(3 < Y < 8 | X = 7\right)$.

   c. $P\left(-3 < X < 3\right)$.

   d. $P\left(-3 < X < 3 | X = -4\right)$.

4. Misalkan $\mathbf{X} \sim N_2(\boldsymbol{\mu}, \boldsymbol{\Sigma})$. Tentukan sebaran dari vektor peubah acak $(X_1 + X_2, X_1 - X_2)$ serta tunjukkan bahwa peubah acak $X_1 + X_2$ dan $X_1 - X_2$ saling bebas jika $Var(X_1) = Var(X_2)$.

5. Misalkan $\mathbf{X} = (X_1, X_2, X_3)'$ memiliki sebaran $N_3(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ di mana $\mu = (3, 2, 1)'$, dan

$$\boldsymbol{\Sigma} = \begin{bmatrix} 5 & -1 & 2 \\ -1 & 2 & 1 \\ 2 & 1 & 3 \end{bmatrix}$$

   a. Tentukanlah sebaran dari vektor peubah acak

$$(X_1 - X_2 + 1, 2X_2 + 3X_3 - 4)',$$

   serta berikanlah korelasi dari kedua peubah acak tersebut.

   b. Carilah $P(X_1 > X_2 - 2X_3 + 3)$.

   c. Carilah $P\left((X_1 - X_2 - X_3)^2 > 10\right)$.

6. Buktikan bahwa $\mathbf{X}_1$ dan $\mathbf{X}_2$ saling bebas jika dan hanya jika $\boldsymbol{\Sigma}_{12} = \mathbf{O}$, di mana $\mathbf{X} \sim N_n(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ dan $\mathbf{X}$ dipartisi seperti pada persamaan 1.3 dan 1.4.

BAB 2

Beberapa Sebaran Kontinu

2.1 Sebaran Gamma

Definisi 2.1 *Fungsi Gamma yang dinotasikan* $\Gamma(\alpha)$ *didefinisikan sebagai*

$$\Gamma(\alpha) = \int_0^\infty y^{\alpha-1} e^{-y} dy$$

dengan $\alpha > 0$.

Jika $\alpha = 1$, maka

$$\Gamma(1) = \int_0^\infty e^{-y} dy = 1.$$

Jika $\alpha > 1$, maka dengan menggunakan teknik integral parsial diperoleh

$$\begin{aligned} \Gamma(\alpha) &= (\alpha-1)\int_0^\infty y^{\alpha-2} e^{-y} dy \\ &= (\alpha-1)\Gamma(\alpha-1). \end{aligned}$$

Oleh karena itu, jika α adalah bilangan bulat positif yang lebih besar dari 1, maka

$$\begin{aligned} \Gamma(\alpha) &= (\alpha-1)(\alpha-2)...(3)(2)(1)\Gamma(1) \\ &= (\alpha-1)!. \end{aligned}$$

Misalkan suatu variabel baru $y = x/\beta$ di mana $\beta > 0$, maka

$$\Gamma(\alpha) = \int_0^\infty \left(\frac{x}{\beta}\right)^{\alpha-1} e^{-x/\beta} \left(\frac{1}{\beta}\right) dx$$

atau ekuivalen dengan

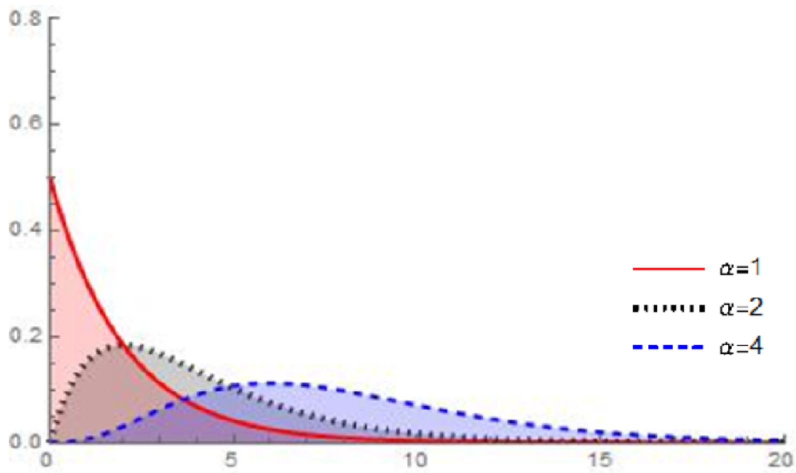
$$1 = \int_0^\infty \frac{1}{\Gamma(\alpha)\beta^\alpha} x^{\alpha-1} e^{-x/\beta} dx.$$

Karena $\alpha > 0, \beta > 0$, dan $\Gamma(\alpha) > 0$, maka

$$f(x) = \frac{1}{\Gamma(\alpha)\beta^\alpha} x^{\alpha-1} e^{-x/\beta} \mathbf{I}(x > 0)$$

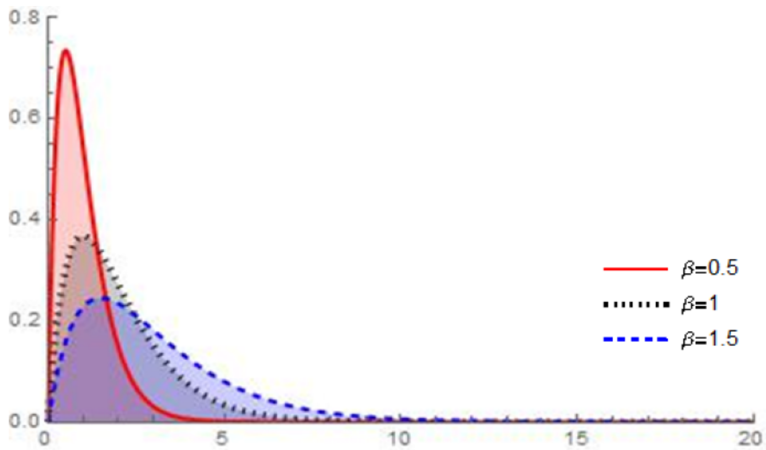
merupakan fungsi kepekatan peluang dari suatu peubah acak kontinu. Peubah acak X yang memiliki fungsi kepekatan peluang tersebut disebut memiliki sebaran gamma dengan parameter α dan β, dinotasikan $\Gamma(\alpha, \beta)$.

Berikut merupakan ilustrasi dari sebaran gamma apabila salah satu parameternya dibuat tetap. Gambar 1 merupakan sebaran gamma dengan parameter β tetap, yaitu $\beta = 2$ dan parameter α yang berbeda, yaitu $\alpha = 1, 2, 4$.



Gambar 1. Sebaran gamma dengan $\beta = 2$ dan $\alpha = 1, 2, 4$

Selanjutnya, Gambar 2 merupakan sebaran gamma dengan parameter α tetap, yaitu $\alpha = 2$ dan parameter β yang berbeda, yaitu $\beta = 0.5, 1, 1.5$.



Gambar 2. Sebaran gamma dengan $\alpha = 2$ dan $\beta = 0.5, 1, 1.5$

Fungsi pembangkit momen bagi peubah acak X yang menyebar gamma ialah

$$
\begin{aligned}
M_X(t) &= \int_0^\infty e^{tx} \frac{1}{\Gamma(\alpha)\beta^\alpha} x^{\alpha-1} e^{-x/\beta} dx \\
&= \int_0^\infty \frac{1}{\Gamma(\alpha)\beta^\alpha} x^{\alpha-1} e^{-x(1-\beta t)/\beta} dx \\
&= \frac{1}{\Gamma(\alpha)\beta^\alpha} \times \Gamma(\alpha) \times \left(\frac{\beta}{1-\beta t}\right)^\alpha \\
&= \frac{1}{(1-\beta t)^\alpha}, t < \frac{1}{\beta}.
\end{aligned}
$$

Selanjutnya,

$$M'(t) = (-\alpha)(1-\beta t)^{-\alpha-1}(-\beta)$$

dan

$$M''(t) = (-\alpha)(-\alpha-1)(1-\beta t)^{-\alpha-2}(-\beta)^2 .$$

Dengan demikian, sebaran gamma memiliki nilai harapan dan ragam sebagai berikut

$$E(X) = M'(0) = \alpha\beta$$

dan

$$
\begin{aligned}
Var(X) &= M''(0) - [E(X)]^2 \\
&= \alpha(\alpha+1)\beta^2 - \alpha^2\beta^2 \\
&= \alpha\beta^2.
\end{aligned}
$$

Teorema 2.1 *Misalkan* $X_1, X_2, ..., X_n$ *adalah peubah acak yang saling bebas di mana* $X_i \sim \Gamma(\alpha_i, \beta)$ *untuk* $i = 1, 2, ..., n$. *Misalkan pula* $Y = \sum_{i=1}^{n} X_i$, *maka* $Y \sim \Gamma\left(\sum_{i=1}^{n} \alpha_i, \beta\right)$.

Bukti. Dengan menggunakan asumsi kebebasan dan fungsi pembangkit momen dari sebaran gamma, diperoleh

$$\begin{aligned} M_Y(t) &= E\left[\exp\left\{t\sum_{i=1}^{n} X_i\right\}\right] \\ &= \prod_{i=1}^{n} E\left[\exp\left\{tX_i\right\}\right] \\ &= \prod_{i=1}^{n} (1-\beta t)^{-\alpha_i} \\ &= (1-\beta t)^{-\sum_{i=1}^{n}\alpha_i} \qquad (2.1) \end{aligned}$$

untuk $t < 1/\beta$. Persamaan 2.1 merupakan fungsi pembangkit momen dari sebaran $\Gamma\left(\sum_{i=1}^{n} \alpha_i, \beta\right)$. ✠

2.2 Sebaran Khi-Kuadrat

Sebaran khi-kuadrat merupakan kasus khusus dari sebaran gamma, yaitu jika parameter $\alpha = \frac{r}{2}$ dan $\beta = 2$ di mana r adalah bilangan bulat positif. Nilai harapan dan ragam dari sebaran khi-kuadrat ialah

$$E(X) = \alpha\beta = \left(\frac{r}{2}\right) \times 2 = r$$

dan

$$Var(X) = \alpha\beta^2 = \left(\frac{r}{2}\right) \times 2^2 = 2r.$$

Jadi, fungsi kepekatan peluang bagi peubah acak kontinu X yang memiliki sebaran khi kuadrat dengan parameter r, dinotasikan dengan $X \sim \chi^2(r)$, ialah

$$f(x) = \frac{1}{\Gamma(r/2)\, 2^{r/2}} x^{r/2-1} e^{-x/2} \mathbf{I}(x > 0),$$

di mana parameter r disebut sebagai derajat bebas dari sebaran $\chi^2(r)$. Gambar fungsi kepekatan peluang untuk $r = 2$, $r = 4$, dan $r = 8$ diberikan pada Gambar 1.

Teorema 2.2 *Misalkan* X *memiliki sebaran* $\chi^2(r)$. *Jika* $k > -r/2$, *maka nilai harapan dari* X^k *ada dan*

$$E\left(X^k\right) = \frac{2^k \Gamma\left(\frac{r}{2}+k\right)}{\Gamma\left(\frac{r}{2}\right)}.$$

Bukti.

$$\begin{aligned} E\left(X^k\right) &= \int_0^\infty \frac{1}{\Gamma(r/2)\,2^{r/2}} x^{(r/2)+k-1} e^{-x/2} dx \\ &= \frac{1}{\Gamma(r/2)\,2^{r/2}} \times \Gamma\left(\frac{r}{2}+k\right) \times 2^{r/2+k} \\ &= 2^k \frac{\Gamma\left(\frac{r}{2}+k\right)}{\Gamma(r/2)}. \end{aligned}$$

✠

Akibat 2.1 *Misalkan* $X_1, X_2, ..., X_n$ *adalah peubah acak yang saling bebas di mana* $X_i \sim \chi^2(r_i)$ *untuk* $i = 1, 2, ..., n$. *Misalkan pula* $Y = \sum_{i=1}^n X_i$, *maka* $Y \sim \chi^2\left(\sum_{i=1}^n r_i\right)$.

Bukti. Karena sebaran khi-kuadrat dengan derajat bebas r merupakan kasus khusus dari sebaran gamma yang memiliki parameter $\alpha = \frac{r}{2}$ dan $\beta = 2$, maka fungsi pembangkit momen bagi sebaran khi-kuadrat ialah

$$M(t) = (1-\beta t)^{-\alpha} = (1-2t)^{-r/2}.$$

Dengan menggunakan asumsi kebebasan dan fungsi pembangkit momen dari sebaran khi-kuadrat, diperoleh

$$\begin{aligned} M_Y(t) &= E\left[\exp\left\{t\sum_{i=1}^n X_i\right\}\right] \\ &= \prod_{i=1}^n E\left[\exp\{tX_i\}\right] \\ &= \prod_{i=1}^n (1-2t)^{-r_i/2} \\ &= (1-2t)^{-\frac{1}{2}\sum_{i=1}^n r_i} \end{aligned} \tag{2.2}$$

untuk $t < 1/2$. Persamaan 2.2 merupakan fungsi pembangkit momen dari sebaran $\chi^2\left(\sum_{i=1}^n r_i\right)$. ✠

Teorema 2.3 *Jika* $Z \sim N(0,1)$, *maka* $Z^2 \sim \chi^2(1)$.

Bukti. Misalkan $Y = Z^2$, maka fungsi sebaran dari peubah acak Y ialah

$$\begin{aligned} F(y) &= P(Y \le y) \\ &= P\left(Z^2 \le y\right) \\ &= P\left(-\sqrt{y} \le Z \le \sqrt{y}\right) \\ &= 2P\left(0 \le Z \le \sqrt{y}\right) \\ &= 2\int_0^{\sqrt{y}} \tfrac{1}{\sqrt{2\pi}} e^{-z^2/2} dz. \end{aligned}$$

Dengan menggunakan butir (8) pada Appendix untuk $\beta(y) = \sqrt{y}, \alpha(y) = 0$, dan $f(z,y) = \frac{1}{\sqrt{2\pi}} e^{-z^2/2}$, maka diperoleh fungsi kepekatan peluang peubah acak Y yaitu

$$\begin{aligned} f(y) &= F'(y) \\ &= 0 + 2 \times \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}\left(\sqrt{y}\right)^2} - \frac{1}{2} y^{-1/2} + 0 \\ &= \frac{1}{\sqrt{\pi}\sqrt{2}} y^{-1/2} e^{-y/2} \\ &= \frac{1}{\Gamma\left(\frac{1}{2}\right) 2^{1/2}} y^{1/2-1} e^{-y/2} \mathbf{I}(y > 0). \end{aligned}$$

Dengan demikian, $Y = Z^2 \sim Gamma\left(\frac{1}{2}, 2\right) = \chi^2(1)$. ✠

Akibat 2.2 *Jika peubah acak* X *menyebar* $N(\mu, \sigma^2)$, $\sigma^2 > 0$, *maka peubah acak* $V = (X - \mu)^2 / \sigma^2$ *menyebar* $\chi^2(1)$.

Teorema 2.4 *Jika* $\mathbf{X} \sim N_n(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ *dengan* $\boldsymbol{\Sigma}$ *adalah matriks definit positif, maka peubah acak* $W = (\mathbf{X} - \boldsymbol{\mu})' \boldsymbol{\Sigma}^{-1} (\mathbf{X} - \boldsymbol{\mu})$ *memiliki sebaran* $\chi^2(n)$.

Bukti. $\boldsymbol{\Sigma}$ adalah matriks definit positif, maka $\boldsymbol{\Sigma} = \boldsymbol{\Sigma}^{\frac{1}{2}} \boldsymbol{\Sigma}^{\frac{1}{2}}$ dan $\boldsymbol{\Sigma}^{\frac{1}{2}}$ memiliki invers. Vektor peubah acak $\mathbf{Z} = \boldsymbol{\Sigma}^{-\frac{1}{2}} (\mathbf{X} - \boldsymbol{\mu}) \sim N_n(\mathbf{0}, \mathbf{I}_n)$. Misalkan peubah acak $W = \mathbf{Z}'\mathbf{Z} = \sum_{i-1}^{n} Z_i^2$. Karena untuk $\forall i = 1, 2, ..., n$, Z_i adalah peubah acak normal baku bebas, $Z_i \overset{bsi}{\sim} N(0,1)$, maka $\sum_{i-1}^{n} Z_i^2$ memiliki sebaran $\chi^2(n)$. ✠

2.3 Sebaran t

Misalkan peubah acak Z menyebar $N(0,1)$ dan peubah acak U menyebar $\chi^2(r)$, kedua peubah acak tersebut saling bebas. Fungsi kepekatan peluang bersama bagi Z dan U ialah

$$\begin{aligned} f(z,u) &= f(z) \cdot f(u) \\ &= \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{z^2}{2}\right) \cdot \frac{1}{\Gamma\left(\frac{r}{2}\right) 2^{r/2}} u^{(r/2)-1} \exp\left(-\frac{u}{2}\right) \mathbf{I}(-\infty < z < \infty, u > 0) \end{aligned}$$

Definisikan suatu peubah acak

$$T = \frac{Z}{\sqrt{U/r}}.$$

Dengan transformasi peubah $t = \dfrac{z}{\sqrt{u/r}}$ dan $w = u$, maka dapat diperoleh fungsi kepekatan peluang $g_T(t)$ dari peubah acak T. Dari transformasi satu-satu,

$$t = \frac{z}{\sqrt{u/r}} \text{ dan } w = u \Leftrightarrow z = t\sqrt{w/r}, u = w,$$

diperoleh

$$\begin{aligned} J &= \begin{vmatrix} \frac{\partial z}{\partial t} & \frac{\partial z}{\partial w} \\ \frac{\partial u}{\partial t} & \frac{\partial u}{\partial w} \end{vmatrix} \\ &= \begin{vmatrix} \sqrt{w/r} & ? \\ 0 & 1 \end{vmatrix} \\ &= \frac{\sqrt{w}}{\sqrt{r}}. \end{aligned}$$

Fungsi kepekatan peluang bersama bagi peubah acak T dan $W = U$ ialah

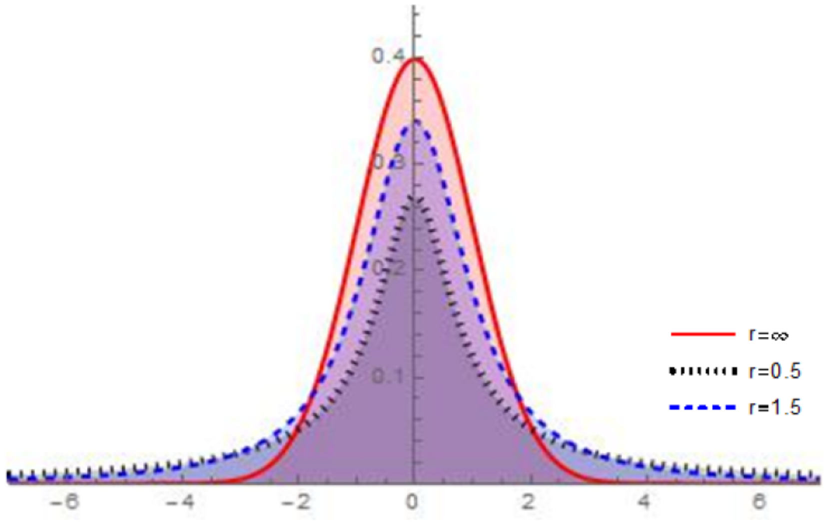
$$\begin{aligned} g(t,w) &= f\left(t\sqrt{w/r}, w\right) |J| \\ &= \frac{1}{\sqrt{2\pi r}\Gamma(r/2) 2^{r/2}} w^{\frac{r+1}{2}-1} \exp\left[-\frac{w}{2}\left(1 + \frac{t^2}{r}\right)\right] \mathbf{I}(-\infty < t < \infty, w > 0). \end{aligned}$$

Dengan demikian, fungsi kepekatan peluang marginal bagi T ialah

$$\begin{aligned}
g_T(t) &= \int_0^\infty g(t,w)\,dw \\
&= \int_0^\infty \frac{1}{\sqrt{2\pi r}\,\Gamma(r/2)\,2^{r/2}} w^{\frac{r+1}{2}-1} \exp\left[-\frac{w}{2}\left(1+\frac{t^2}{r}\right)\right] dw \\
&= \frac{1}{\sqrt{2\pi r}\,\Gamma(r/2)\,2^{r/2}} \times \Gamma\left(\frac{r+1}{2}\right) \times \left(\frac{2}{1+\frac{t^2}{r}}\right)^{(r+1)/2} \\
&= \frac{\Gamma\left(\frac{r+1}{2}\right)}{\sqrt{\pi r} \times \Gamma(r/2)} \times \left(\frac{1}{1+\frac{t^2}{r}}\right)^{(r+1)/2} \mathbf{I}(-\infty < t < \infty). \qquad (2.3)
\end{aligned}$$

Definisi 2.2 *Jika $Z \sim N(0,1)$, $U \sim \chi^2(r)$, dan dua peubah acak tersebut saling bebas, maka sebaran dari $T = \dfrac{Z}{\sqrt{U/r}}$ disebut sebaran t dengan derajat bebas r, dinotasikan $T \sim t(r)$, serta fungsi kepekatan peluang dari T seperti pada persamaan 2.3.*

Sebaran t ini dikenal juga sebagai sebaran $t-Student$ yang diperkenalkan pertama kali oleh W. S. Gosset dengan nama samaran *Student*. Berikut merupakan ilustrasi sebaran t dengan parameter $r = 0.5, 1.5$, dan ∞.



Gambar 3. Sebaran t dengan parameter $r = 0.5, 1.5, \infty$

Teorema 2.5 *Jika peubah acak T memiliki sebaran t dengan derajat bebas r, maka*

$$\begin{aligned} E(T) &= 0, && \textit{jika } r \geq 2 \\ Var(T) &= \frac{r}{r-2}, && \textit{jika } r \geq 3. \end{aligned}$$

Bukti. Karena $T \sim t(r)$, maka

$$T = \frac{Z}{\sqrt{U/r}}$$

dengan $Z \sim N(0,1)$, $U \sim \chi^2(r)$, dan dua peubah acak tersebut saling bebas.

$$\begin{aligned} E(T) &= E\left(\frac{Z}{\sqrt{U/r}}\right) \\ &= r^{1/2} E(Z) E\left(U^{-1/2}\right) \\ &= 0. \end{aligned}$$

Berdasarkan Teorema 2.2,

$$E\left(U^{-1/2}\right) = 2^{-1/2} \times \frac{\Gamma\left(\frac{r}{2} - \frac{1}{2}\right)}{\Gamma\left(\frac{r}{2}\right)},$$

di mana $E\left(U^{-1/2}\right)$ terdefinisi bila $\frac{r}{2} - \frac{1}{2} > 0 \Leftrightarrow r > 1$, sehingga $E(T) = 0$ jika $r \geq 2$.

Selanjutnya, momen ke-2 dari peubah acak T ialah

$$\begin{aligned} E\left(T^2\right) &= E\left(\frac{Z^2}{U/r}\right) \\ &= rE\left(Z^2\right) E\left(U^{-1}\right). \end{aligned} \tag{2.4}$$

Berdasarkan Teorema 2.2,

$$\begin{aligned} E\left(U^{-1}\right) &= 2^{-1} \times \frac{\Gamma\left(\frac{r}{2} - 1\right)}{\Gamma\left(\frac{r}{2}\right)} \\ &= \frac{1}{2} \times \frac{\Gamma\left(\frac{r}{2} - 1\right)}{\left(\frac{r}{2} - 1\right)\Gamma\left(\frac{r}{2} - 1\right)} \\ &= \frac{1}{r-2}, \end{aligned}$$

sehingga ruas kanan pada persamaan 2.4 menjadi

$$\begin{aligned} E\left(T^2\right) &= r \times 1 \times \frac{1}{r-2} \\ &= \frac{r}{r-2},\ r \geq 3. \end{aligned}$$

Dengan demikian,

$$\begin{aligned} Var\left(T\right) &= E\left(T^2\right) - \left(E\left(T\right)\right)^2 \\ &= \frac{r}{r-2} - 0 \\ &= \frac{r}{r-2},\ r \geq 3. \end{aligned}$$

✠

Contoh 2.1 *Misalkan X_1, X_2, X_3, X_4 adalah peubah acak bebas dari sebaran normal baku. Tentukan nilai harapan dari peubah acak*

$$W = \frac{X_1 - X_2 + X_3}{\sqrt{X_1^2 + X_2^2 + X_3^2 + X_4^2}}.$$

Jawab. Karena $X_i \sim N\left(0,1\right)$, maka

$$X_1 - X_2 + X_3 \sim N\left(0,3\right)$$

dan

$$\frac{X_1 - X_2 + X_3}{\sqrt{3}} \sim N\left(0,1\right),$$

serta $X_i^2 \sim \chi^2\left(1\right)$, sehingga

$$X_1^2 + X_2^2 + X_3^2 + X_4^2 \sim \chi^2\left(4\right).$$

Dengan demikian, peubah acak

$$\frac{\dfrac{X_1 - X_2 + X_3}{\sqrt{3}}}{\sqrt{\dfrac{X_1^2 + X_2^2 + X_3^2 + X_4^2}{4}}} = \left(\frac{2}{\sqrt{3}}\right) W \sim t\left(4\right).$$

Berdasarkan Teorema 2.5, nilai harapan dari W ialah

$$\begin{aligned} E\left(W\right) &= \frac{\sqrt{3}}{2} E\left(\frac{2}{\sqrt{3}} W\right) \\ &= \frac{\sqrt{3}}{2} \times 0 \\ &= 0. \end{aligned}$$

✠

280 2.4 Sebaran *F*

Misalkan U dan Y adalah peubah acak bebas khi-kuadrat dengan derajat bebas r_1 dan r_2. Fungsi kepekatan peluang bersama bagi peubah acak U dan Y ialah

$$h(u,y) = \frac{1}{\Gamma(r_1/2)\,\Gamma(r_2/2)\,2^{(r_1+r_2)/2}} u^{(r_1/2)-1} y^{(r_2/2)-1} e^{-(u+y)/2} \mathbf{I}(u,y>0).$$

Definisikan suatu peubah acak

$$W = \frac{U/r_1}{Y/r_2}.$$

Dengan transformasi satu-satu antara peubah $w = \dfrac{u/r_1}{y/r_2}$ dan $z = y$, maka dapat diperoleh fungsi kepekatan peluang $g_W(w)$ dari peubah acak W. Dari transformasi satu-satu,

$$\begin{aligned} & w = \frac{u/r_1}{y/r_2} \text{ dan } z = y \\ \Leftrightarrow \quad & u = \left(\tfrac{r_1}{r_2}\right) zw, \ y = z, \end{aligned}$$

281 diperoleh

$$\begin{aligned} J &= \begin{vmatrix} \frac{\partial u}{\partial w} & \frac{\partial u}{\partial z} \\ \frac{\partial y}{\partial w} & \frac{\partial y}{\partial z} \end{vmatrix} \\ &= \begin{vmatrix} \left(\tfrac{r_1}{r_2}\right) z & ? \\ 0 & 1 \end{vmatrix} \\ &= \left(\frac{r_1}{r_2}\right) z. \end{aligned}$$

282 Fungsi kepekatan peluang $g(w,z)$ dari peubah acak W dan $Z = Y$ ialah

$$\begin{aligned} g(w,z) \quad = \quad & \frac{1}{\Gamma(r_1/2)\,\Gamma(r_2/2)\,2^{(r_1+r_2)/2}} \left[\frac{r_1 z w}{r_2}\right]^{(r_1/2)-1} z^{(r_2/2)-1} \\ & \exp\left[-\frac{z}{2}\left(1+\frac{r_1}{r_2}w\right)\right] \frac{r_1 z}{r_2} \mathbf{I}(0<w,z<\infty). \end{aligned}$$

Dengan demikian, fungsi kepekatan peluang marginal bagi W ialah

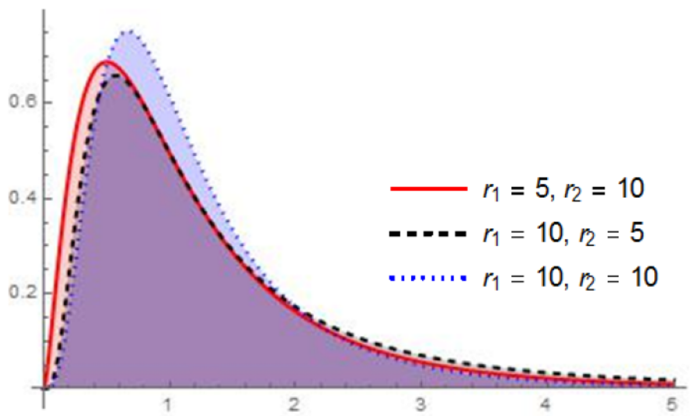
$$\begin{aligned}
g_W(w) &= \int_0^\infty g(w,z)\,dz \\
&= \int_0^\infty \frac{(r_1/r_2)^{r_1/2}\, w^{(r_1/2)-1}}{\Gamma(r_1/2)\,\Gamma(r_2/2)\,2^{(r_1+r_2)/2}} z^{(r_1+r_2)/2-1} \\
&\quad \times \exp\left[-\frac{z}{2}\left(1+\frac{r_1}{r_2}w\right)\right]\frac{r_1 z}{r_2}dz. \\
&= \frac{(r_1/r_2)^{r_1/2}\, w^{(r_1/2)-1}}{\Gamma(r_1/2)\,\Gamma(r_2/2)\,2^{(r_1+r_2)/2}} \times \Gamma\left(\frac{r_1+r_2}{2}\right) \\
&\quad \times \frac{2}{\left(1+\frac{r_1}{r_2}w\right)^{(r_1+r_2)/2}} \mathbf{I}(w>0). \\
&= \frac{\Gamma\left(\frac{r_1+r_2}{2}\right)(r_1/r_2)^{r_1/2}}{\Gamma(r_1/2)\,\Gamma(r_2/2)} \times \frac{w^{(r_1/2)-1}}{\left(1+\frac{r_1}{r_2}w\right)^{(r_1+r_2)/2}} \mathbf{I}(w>0). \quad (2.5)
\end{aligned}$$

Definisi 2.3 *Misalkan U dan Y peubah acak bebas khi-kuadrat dengan derajat bebas r_1 dan r_2. Sebaran dari*

$$W = \frac{U/r_1}{Y/r_2}$$

disebut sebaran F dengan derajat bebas r_1 dan r_2, dinotasikan $F(r_1, r_2)$, dengan fungsi kepekatan peluang seperti pada persamaan 2.5.

Berikut merupakan ilustrasi dari sebaran F, yaitu $F(5,10)$, $F(10,5)$, dan $F(10,10)$.



Gambar 4. Sebaran $F(5,10)$, $F(10,5)$, dan $F(10,10)$

Teorema 2.6 *Jika peubah acak W memiliki sebaran-F dengan derajat bebas* r_1 *dan* r_2*, maka*

$$\begin{aligned} E(X) &= \frac{r_2}{r_2-2}, && \textit{jika } r_2 \geq 3 \\ Var(X) &= \frac{2r_2^2(r_1+r_2-2)}{r_1(r_2-2)^2(r_2-4)}, && \textit{jika } r_2 \geq 5. \end{aligned}$$

Bukti. Analog dengan Teorema 2.5. ✠

Contoh 2.2 *Misalkan* X_1, X_2, X_3, X_4 *dan* Y_1, Y_2, Y_3, Y_4, Y_5 *adalah peubah acak bebas dari sebaran normal baku. Tentukan ragam dari peubah acak*

$$U = \frac{5(X_1^2+X_2^2+X_3^2+X_4^2)}{4(Y_1^2+Y_2^2+Y_3^2+Y_4^2+Y_5^2)}.$$

Jawab. Karena $X_i \sim N(0,1)$ dan $Y_i \sim N(0,1)$, maka

$$X_1^2+X_2^2+X_3^2+X_4^2 \sim \chi^2(4)$$

dan

$$Y_1^2+Y_2^2+Y_3^2+Y_4^2+Y_5^2 \sim \chi^2(5),$$

sehingga

$$U = \frac{\dfrac{(X_1^2+X_2^2+X_3^2+X_4^2)}{4}}{\dfrac{(Y_1^2+Y_2^2+Y_3^2+Y_4^2+Y_5^2)}{5}} \sim F(4,5).$$

Berdasarkan Teorema 2.6, ragam dari U ialah

$$\begin{aligned} Var(U) &= \frac{2(5)^2(4+5-2)}{4(5-2)^2(5-4)} \\ &= \frac{350}{36} \\ &= 9.72. \end{aligned}$$

✠

2.5 Teorema Student

Teorema 2.7 *Misalkan* $X_1, X_2, ..., X_n$ *adalah peubah acak bebas stokastik identik yang menyebar* $N(\mu, \sigma^2)$. *Definisikan peubah acak*

$$\bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i \textit{ dan } S^2 = \frac{1}{n-1}\sum_{i=1}^{n}\left(X_i - \bar{X}\right)^2,$$

maka

a. $\bar{X} \sim N\left(\mu, \frac{\sigma^2}{n}\right)$.

b. $\bar{X}$ *dan* S^2 *saling bebas.*

c. $(n-1)\, S^2/\sigma^2 \sim \chi^2\,(n-1)$.

d. *Peubah acak* $T = \dfrac{\bar{X}-\mu}{S/\sqrt{n}}$ *memiliki sebaran* t *dengan derajat bebas* $n-1$.

Bukti. Misalkan $\mathbf{X} = (X_1, X_2, ..., X_n)'$. Karena $X_1, X_2, ..., X_n$ adalah peubah acak bebas stokastik identik dari sebaran $N\left(\mu, \sigma^2\right)$, maka $\mathbf{X}$ memiliki sebaran normal multivariat $N\left(\mu\mathbf{1}, \sigma^2\mathbf{I}\right)$ di mana $\mathbf{1}$ merupakan vektor yang berunsur 1. Misalkan $\mathbf{v}' = (1/n, ..., 1/n) = (1/n)\,\mathbf{1}'$, maka $\bar{X} = \mathbf{v}'\mathbf{X}$. Definisikan vektor peubah acak $\mathbf{Y} = \left(X_1 - \bar{X}, ..., X_n - \bar{X}\right)'$. Perhatikan transformasi peubah berikut.

$$\mathbf{W} = \begin{bmatrix} \bar{X} \\ \mathbf{Y} \end{bmatrix} = \begin{bmatrix} \mathbf{v}' \\ \mathbf{I} - \mathbf{1}\mathbf{v}' \end{bmatrix} \mathbf{X}.$$

Karena $\mathbf{W}$ adalah transformasi linear dari vektor peubah acak normal multivariat, maka berdasarkan Teorema 1.4, $\mathbf{W}$ memiliki sebaran normal multivariat dengan nilai harapan

$$E\left(\mathbf{W}\right) = \begin{bmatrix} \mathbf{v}' \\ \mathbf{I} - \mathbf{1}\mathbf{v}' \end{bmatrix} E\left(\mathbf{X}\right) = \begin{bmatrix} \mathbf{v}' \\ \mathbf{I} - \mathbf{1}\mathbf{v}' \end{bmatrix} \mu\mathbf{1} = \begin{bmatrix} \mu \\ \mathbf{0}_n \end{bmatrix}, \tag{2.6}$$

di mana $\mathbf{0}_n$ adalah vektor yang berunsur 0 dan matriks koragam

$$\begin{aligned} \mathbf{\Sigma} &= \begin{bmatrix} \mathbf{v}' \\ \mathbf{I} - \mathbf{1}\mathbf{v}' \end{bmatrix} \sigma^2\mathbf{I} \begin{bmatrix} \mathbf{v}' \\ \mathbf{I} - \mathbf{1}\mathbf{v}' \end{bmatrix}' \qquad (2.7) \\ &= \sigma^2 \begin{bmatrix} 1/n & \mathbf{0}_n' \\ \mathbf{0}_n & \mathbf{I} - \mathbf{1}\mathbf{v}' \end{bmatrix}. \end{aligned}$$

Karena koragamnya bernilai nol, maka $\bar{X}$ saling bebas dengan $\mathbf{Y}$. Dengan demikian, $\bar{X}$ juga saling bebas dengan $S^2 = (n-1)^{-1}\,\mathbf{Y}'\mathbf{Y}$. Selanjutnya, berdasarkan persamaan 2.6 dan 2.7 diperoleh bahwa $\bar{X} \sim N\left(\mu, \frac{\sigma^2}{n}\right)$.

Misalkan suatu peubah acak

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

maka berdasarkan Teorema 2.2, $\forall i = 1, ..., n$, $\left(\frac{X_i - \mu}{\sigma}\right)^2$ menyebar $\chi^2(1)$, sehingga

$$V = \sum_{i=1}^{n} \left(\frac{X_i - \mu}{\sigma}\right)^2 \sim \chi^2(n)$$

(lihat Akibat 2.1). Peubah acak V juga dapat dituliskan menjadi

$$\begin{aligned} V &= \sum_{i=1}^{n} \left(\frac{\left(X_i - \bar{X}\right) + \left(\bar{X} - \mu\right)}{\sigma}\right)^2 \\ &= \sum_{i=1}^{n} \left(\frac{X_i - \bar{X}}{\sigma}\right)^2 + \left(\frac{\bar{X} - \mu}{\sigma/\sqrt{n}}\right)^2 \\ &= \frac{(n-1)\,S^2}{\sigma^2} + \left(\frac{\bar{X} - \mu}{\sigma/\sqrt{n}}\right)^2. \qquad (2.8) \end{aligned}$$

Karena $\bar{X}$ dan S^2 saling bebas, dua bagian pada ruas kanan persamaan 2.8, yaitu $\frac{(n-1)\,S^2}{\sigma^2}$ dan $\left(\frac{\bar{X} - \mu}{\sigma/\sqrt{n}}\right)^2$ juga saling bebas. Peubah acak $\left(\frac{\bar{X} - \mu}{\sigma/\sqrt{n}}\right)^2$ merupakan kuadrat dari peubah acak normal baku, sehingga memiliki sebaran $\chi^2(1)$. Selanjutnya, fungsi pembangkit momen dari kedua ruas pada persamaan 2.8, yaitu

$$(1-2t)^{-n/2} = E\left[\exp\left\{t\frac{(n-1)\,S^2}{\sigma^2}\right\}\right](1-2t)^{-1/2}.$$

Dengan demikian,

$$\begin{aligned} E\left[\exp\left\{t\frac{(n-1)\,S^2}{\sigma^2}\right\}\right] &= \frac{(1-2t)^{-n/2}}{(1-2t)^{-1/2}} \\ &= (1-2t)^{-(n-1)/2}, \end{aligned}$$

sehingga $\frac{(n-1)\,S^2}{\sigma^2} \sim \chi^2(n-1)$.

Selanjutnya, dapat disimpulkan bahwa peubah acak T, yaitu

$$\begin{aligned} T &= \frac{\left(\bar{X} - \mu\right)/\left(\sigma/\sqrt{n}\right)}{\sqrt{(n-1)\,S^2/\left(\sigma^2(n-1)\right)}} \\ &= \frac{\bar{X} - \mu}{S/\sqrt{n}} \end{aligned}$$

memiliki sebaran t dengan derajat bebas $(n-1)$. ✠

Contoh 2.3 *Diketahui* $X \sim N(10, 25)$ *dan* $X_1, X_2, ..., X_{501}$ *adalah contoh acak dari populasi X. Tentukan nilai harapan dari ragam contoh* S^2.

Jawab. Berdasarkan Teorema 2.7, $\dfrac{(501-1)\,S^2}{25} \sim \chi^2(500)$, sehingga

$$\begin{aligned} E\left(\frac{(501-1)\,S^2}{25}\right) &= 500 \\ \left(\frac{500}{25}\right) E\left(S^2\right) &= 500 \\ E\left(S^2\right) &= \left(\frac{25}{500}\right)(500) \\ &= 25. \end{aligned}$$

✠

2.6 Latihan

1. Misalkan X_1 dan X_2 merupakan contoh acak dari sebaran dengan fkp $f(x) = \exp(-x)\mathbf{I}(x > 0)$, maka tunjukkanlah bahwa $Z = X_1/X_2$ merupakan peubah acak bersebaran F. Tentukan pula derajat bebasnya.

2. Tunjukkan bahwa fungsi kepekatan peluang bagi peubah acak $W = \frac{U/r_1}{V/r_2}$ adalah seperti pada persamaan 2.5, serta tentukan momen ke-k dari peubah acak W.

3. Misalkan peubah acak $T = \dfrac{W}{\sqrt{V/r}}$ di mana $W \sim N(0, 1)$ dan $V \sim \chi^2(r)$ saling bebas. Tunjukkan bahwa memiliki sebaran-F dengan parameter $r_1 = 1$ dan $r_2 = r$.

4. Misalkan X_1, X_2, X_3 adalah tiga peubah acak bebas khi-kuadrat dengan derajat bebas masing-masing adalah r_1, r_2, dan r_3.

   a. Tunjukkan bahwa $Y_1 = \frac{X_1}{X_2}$ dan $Y_2 = X_1 + X_2$ saling bebas serta $Y_2 \sim \chi^2(r_1 + r_2)$.

   b. Tunjukkan bahwa $\dfrac{X_1/r_1}{X_2/r_2}$ dan $\dfrac{X_3/r_3}{(X_1 + X_2)/(r_1 + r_2)}$ merupakan peubah acak bebas yang memiliki sebaran-F.

BAB 3

Pendugaan Parameter

3.1 Penduga Titik

Dalam statistika, populasi adalah sekumpulan data yang mempunyai karakteristik yang sama dan menjadi objek inferensi. Statistika inferensia mendasarkan diri pada dua konsep dasar, yaitu populasi sebagai keseluruhan data dan sampel sebagai bagian dari populasi yang digunakan untuk melakukan pendekatan terhadap populasi. Suatu populasi biasanya dijelaskan oleh suatu peubah acak X. Jika informasi mengenai sebaran $f(x;\theta)$ dari peubah acak X diketahui, maka informasi mengenai populasi tersebut juga akan diketahui. Misalkan $X_1, X_2, ..., X_n$ adalah peubah acak yang memiliki fungsi kepekatan peluang $f(x;\theta)$ yang sama dengan populasi, maka

$$X_1 = x_1, X_2 = x_2, ..., X_n = x_n$$

merupakan data atau nilai amatan sampel yang digunakan untuk mewakili populasi.

Misalkan X adalah peubah acak yang memiliki fungsi kepekatan peluang $f(x;\theta)$ atau fungsi massa peluang $p(x;\theta)$, di mana θ merupakan suatu bilangan real atau vektor bilangan real yang tidak diketahui. Asumsikan bahwa $\theta \in \Omega$ merupakan himpunan bagian dari $\mathbb{R}^p$ untuk $p \geq 1$. Dalam tulisan ini, kadangkala notasi $f(x;\theta)$ digunakan baik sebagai fungsi kepekatan peluang atau fungsi massa peluang peubah acak X, yang hasilnya berlaku baik untuk kasus peubah acak kontinu maupun diskret. Sebagai contoh, θ merupakan vektor (μ, σ^2) jika $X \sim N(\mu, \sigma^2)$ atau θ merupakan peluang sukses p jika X memiliki sebaran binomial. Informasi mengenai parameter θ dapat diketahui berdasarkan sampel $X_1, X_2, ..., X_n$.

Definisi 3.1 *Peubah acak* $X_1, X_2, ..., X_n$ *disebut sebagai* ***sampel acak*** *dari* *peubah acak X jika peubah-peubah acak tersebut saling bebas dan masing-masing memiliki sebaran yang sama dengan sebaran X.*

Definisi 3.2 *Misalkan n peubah acak $X_1, X_2, ..., X_n$ merupakan sampel acak dari X yang memiliki sebaran $f(x;\theta)$. Fungsi dari sampel yang tidak bergantung pada parameter θ, yaitu $T = T(X_1, X_2, ..., X_n)$ disebut sebagai* ***statistik****.*

Statistik T yang didefinisikan pada definisi 3.2, digunakan untuk menduga parameter θ. Dalam hal ini, $T(X_1, X_2, ..., X_n)$ disebut sebagai **penduga (*estimator*) titik** bagi θ yang dilambangkan dengan $\hat{\theta}$ dan $T(x_1, x_2, ..., x_n)$ disebut sebagai **dugaan (*estimate*) titik** bagi θ. Sebagai contoh, andaikan $X_1, X_2, ..., X_n$ merupakan sampel acak dari suatu sebaran dengan nilai harapan μ dan ragam σ^2, maka

$$\hat{\mu} = \bar{X} = \frac{1}{n}\sum\nolimits_{i=1}^{n} X_i \text{ dan } \hat{\sigma}^2 = S^2 = \frac{1}{n-1}\sum\nolimits_{i=1}^{n}\left(X_i - \bar{X}\right)^2$$

merupakan penduga titik bagi μ dan σ^2, serta dugaan titik bagi μ dan σ^2 ialah

$$\bar{x} = \frac{1}{n}\sum\nolimits_{i=1}^{n} x_i \text{ dan } s^2 = \frac{1}{n-1}\sum\nolimits_{i=1}^{n}(x_i - \bar{x})^2.$$

Pada pendugaan titik secara parametrik, diasumsikan bahwa sebaran populasi $f(x;\theta)$ diketahui, sehingga pendugaan yang dilakukan hanya pada parameter yang tidak diketahui, yaitu parameter θ atau fungsi dari parameter θ, misalkan $k(\theta)$. Pendugaan tersebut dilakukan berdasarkan informasi yang tersedia dari sampel. Terdapat beberapa metode untuk melakukan pendugaan titik, di antaranya metode momen dan metode kemungkinan maksimum.

3.2 Metode Pendugaan Titik

3.2.1 Metode Momen

Misalkan $X_1, X_2, ..., X_n$ merupakan sampel acak dari populasi X dengan fungsi kepekatan atau fungsi massa peluang $f(x;\theta_1, \theta_2, ..., \theta_m)$ di mana θ_1, θ_2, ..., θ_m merupakan parameter-parameter yang tidak diketahui. Misalkan pula

$$E\left(X^k\right) = g_k(\theta_1, \theta_2, ..., \theta_m)$$

merupakan momen ke-k dari populasi. Selanjutnya, misalkan

$$M_k = \frac{1}{n}\sum_{i=1}^{n} X_i^k$$

merupakan momen ke-k dari sampel. Dengan menggunakan metode momen, penduga bagi m-parameter θ diperoleh dengan cara mencari solusi dari persamaan penduga m-momen pertama (jika ada) yang memberi solusi tunggal (khas) dari X, yaitu

$$\begin{aligned}
g_1\left(\hat{\theta}_1, \hat{\theta}_2, ..., \hat{\theta}_m\right) &= M_1 \\
g_2\left(\hat{\theta}_1, \hat{\theta}_2, ..., \hat{\theta}_m\right) &= M_2 \\
g_3\left(\hat{\theta}_1, \hat{\theta}_2, ..., \hat{\theta}_m\right) &= M_3 \\
&\vdots \\
g_m\left(\hat{\theta}_1, \hat{\theta}_2, ..., \hat{\theta}_m\right) &= M_m.
\end{aligned}$$

Contoh 3.1 *Misalkan $X_1, X_2, ..., X_n$ adalah sampel acak dari populasi X dengan fungsi kepekatan peluang*

$$f(x;\theta) = \frac{1}{\theta}\mathbf{I}(0 < x < \theta)$$

Tentukan penduga bagi θ dengan menggunakan metode momen.

Jawab. Berdasarkan fungsi kepekatan peluang $f(x;\theta)$, maka $X \sim$ seragam $(0,\theta)$, sehingga $E(X) = \frac{\theta}{2} = g_1(\theta)$. Dengan menggunakan metode momen diperoleh

$$\begin{aligned}
g_1\left(\hat{\theta}\right) &= M_1, \\
\frac{\hat{\theta}}{2} &= \bar{X}.
\end{aligned}$$

Jadi, penduga bagi θ ialah

$$\hat{\theta} = 2\bar{X}.$$

$\maltese$

Contoh 3.2 *Misalkan $X \sim N(\mu, \sigma^2)$ dan $X_1, X_2, ..., X_n$ adalah contoh acak dari populasi X. Tentukan penduga bagi μ dan σ^2 dengan menggunakan metode momen.*

Jawab. Karena $X \sim N\left(\mu, \sigma^2\right)$, maka

$$\begin{aligned} E\left(X\right) &= \mu \\ E\left(X^2\right) &= \sigma^2 + \mu^2, \end{aligned}$$

sehingga

$$\begin{aligned} g_1\left(\mu, \sigma^2\right) &= \mu \\ g_2\left(\mu, \sigma^2\right) &= \sigma^2 + \mu^2. \end{aligned}$$

Dengan metode momen,

$$\begin{aligned} g_1\left(\hat{\mu}, \hat{\sigma}^2\right) &= M_1 \\ g_2\left(\hat{\mu}, \hat{\sigma}^2\right) &= M_2, \end{aligned}$$

sehingga

$$\begin{aligned} \hat{\mu} &= \frac{1}{n}\sum_{i=1}^{n} X_i = \bar{X}, \\ \hat{\sigma}^2 + \hat{\mu}^2 &= \frac{1}{n}\sum_{i=1}^{n} X_i^2 \\ \hat{\sigma}^2 &= \frac{1}{n}\sum_{i=1}^{n} X_i^2 - \hat{\mu}^2 \\ &= \frac{1}{n}\sum_{i=1}^{n} X_i^2 - \bar{X}^2 \\ &= \frac{1}{n}\sum_{i=1}^{n} \left(X_i - \bar{X}\right)^2. \end{aligned}$$

Jadi, penduga bagi parameter μ dan σ^2 secara berturut-turut ialah $\hat{\mu} = \bar{X}$ dan $\hat{\sigma}^2 = \dfrac{1}{n}\sum_{i=1}^{n}\left(X_i - \bar{X}\right)^2$. $\maltese$

Contoh 3.3 *Misalkan* $X_1, X_2, ..., X_n$ *adalah sampel acak dari populasi* X *dengan fungsi kepekatan peluang*

$$f\left(x; \theta\right) = \theta x^{\theta - 1} \mathbf{I}\left(0 < x < 1\right)$$

di mana $0 < \theta < \infty$ *ialah parameter yang tidak diketahui. Dengan menggunakan metode momen, tentukan penduga bagi* θ*. Jika* $x_1 = 0.2$, $x_2 = 0.6$, $x_3 = 0.5$, $x_4 = 0.3$ *adalah nilai amatan dari empat sampel acak, maka tentukan dugaan bagi* θ*.*

Jawab. Nilai harapan bagi X dengan $f(x;\theta) = \theta x^{\theta-1}\mathbf{I}(0 < x < 1)$ ialah

$$\begin{aligned} E(X) &= \int_0^1 x f(x;\theta)\, dx \\ &= \int_0^1 x\theta x^{\theta-1} dx \\ &= \frac{\theta}{\theta+1}, \end{aligned}$$

sehingga

$$g_1(\theta) = \frac{\theta}{\theta+1}.$$

Dengan metode momen,

$$\begin{aligned} M_1 &= g_1\left(\hat{\theta}\right) \\ \bar{X} &= \frac{\hat{\theta}}{\hat{\theta}+1} \\ \hat{\theta} &= \frac{\bar{X}}{1-\bar{X}}, \end{aligned}$$

dengan $\bar{X}$ adalah rata-rata dari sampel. Dengan demikian, statistik $\dfrac{\bar{X}}{1-\bar{X}}$ adalah penduga bagi θ,

$$\hat{\theta} = \frac{\bar{X}}{1-\bar{X}}.$$

Karena $x_1 = 0.2$, $x_2 = 0.6$, $x_3 = 0.5$, $x_4 = 0.3$, diperoleh $\bar{x} = 0.4$, sehingga dugaan bagi θ ialah

$$\frac{0.4}{1-0.4} = \frac{2}{3}.$$

$\maltese$

Contoh 3.4 *Misalkan $X_1, X_2, ..., X_n$ adalah sampel acak dari sebaran seragam (α, β), $\alpha < \beta$. Dengan menggunakan metode momen, tentukan penduga bagi α dan β. Untuk ukuran sampel $n = 9$ dengan nilai-nilai amatan* 0.5, 0.6, 0.1, 0.9, 1.3, 1.6, 0.7, 0.9, *dan* 1.0*, tentukanlah dugaan bagi α dan β.*

Jawab. Karena $E(X) = \frac{\alpha+\beta}{2} = g_1(\alpha,\beta)$ dan $E(X^2) = \frac{1}{3}(\alpha^2+\alpha\beta+\beta^2) = g_2(\alpha,\beta)$, maka dengan metode momen diperoleh

$$\frac{\hat{\alpha}+\hat{\beta}}{2} = \frac{1}{n}\sum_{i=1}^{n} X_i = Q_1, \tag{3.1}$$

$$\frac{1}{3}\left(\hat{\alpha}^2+\hat{\alpha}\hat{\beta}+\hat{\beta}^2\right) = \frac{1}{n}\sum_{i=1}^{n} X_i^2 = Q_2. \tag{3.2}$$

Dari persamaan 3.1 dan 3.2, diperoleh

$$\begin{aligned} & \hat{\alpha}+\hat{\beta} = 2Q_1 \\ \Leftrightarrow\ & \hat{\alpha} = 2Q_1 - \hat{\beta} \end{aligned}$$

$$\begin{aligned} & \hat{\alpha}^2+\hat{\alpha}\hat{\beta}+\hat{\beta}^2 = 3Q_2 \\ \Leftrightarrow\ & \left(2Q_1-\hat{\beta}\right)^2 + \left(2Q_1-\hat{\beta}\right)\hat{\beta}+\hat{\beta}^2 = 3Q_2 \\ \Leftrightarrow\ & \hat{\beta}^2 - 2\hat{\beta}Q_1 + 4Q_1^2 - 3Q_2 = 0 \\ \Leftrightarrow\ & \hat{\beta}_{1,2} = \frac{2Q_1 \pm \sqrt{4Q_1^2 - 4\left(4Q_1^2-3Q_2\right)}}{2} \\ \Leftrightarrow\ & \hat{\beta}_{1,2} = Q_1 \pm \sqrt{3\left(Q_2-Q_1^2\right)} \\ \Leftrightarrow\ & \hat{\alpha}_{1,2} = Q_1 \mp \sqrt{3\left(Q_2-Q_1^2\right)}. \end{aligned}$$

Karena $\hat{\alpha} < \hat{\beta}$, maka

$$\begin{aligned} \hat{\alpha} &= Q_1 - \sqrt{3\left(Q_2-Q_1^2\right)} \\ \hat{\beta} &= Q_1 + \sqrt{3\left(Q_2-Q_1^2\right)}. \end{aligned}$$

Untuk ukuran sampel $n = 9$ dengan nilai-nilai amatan 0.5, 0.6, 0.1, 0.9, 1.3, 1.6, 0.7, 0.9, dan 1.0 akan diperoleh dugaan bagi Q_1 sebesar 0.84 dan dugaan bagi Q_2 sebesar 0.89 serta dugaan bagi α sebesar 0.096 dan dugaan bagi β sebesar 1.584. ✠

Dari Contoh 3.4, dugaan yang diperoleh dengan metode momen tidak memberikan dugaan yang layak, karena untuk menduga β yang seharusnya bernilai lebih besar sama dengan 1.6, diperoleh dugaannya sebesar 1.584.

3.2.2 Metode Kemungkinan Maksimum

Definisi 3.3 *Misalkan $X_1, X_2, ..., X_n$ adalah sampel acak dari populasi X dengan fungsi kepekatan peluang $f(x;\theta)$, di mana θ adalah parameter yang* *tidak diketahui dalam ruang parameter* Ω*. Fungsi kemungkinan atau likelihood function,* $L(\theta;\mathbf{x})$ *, didefinisikan sebagai berikut*

$$L(\theta;\mathbf{x}) = \prod_{i=1}^{n} f(x_i;\theta).$$

Catatan : $L(\theta;\mathbf{x})$ *adalah fungsi kepekatan bersama dari sampel. Karena* $L(\theta;\mathbf{x})$ *merupakan fungsi dari* θ*, maka kita juga dapat menggunakan notasi* $L(\theta)$ *.*

Definisi 3.4 *Suatu fungsi dari* $\mathbf{X}$*, katakanlah* $\hat{\theta} = \hat{\theta}(\mathbf{X})$ *disebut* ***penduga kemungkinan maksimum*** *bagi* θ *bila*

$$\hat{\theta} = Arg \sup_{\theta\in\Omega} L(\theta;\mathbf{X}).$$

Berikut merupakan langkah-langkah untuk menentukan penduga kemungkinan maksimum.

1. Definisikan fungsi kemungkinan $L(\theta)$ untuk amatan sampel x_1, x_2, ..., x_n,
$$L(\theta) = \prod_{i=1}^{n} f(x_i;\theta).$$

2. Perhatikan bahwa setiap nilai θ yang memaksimumkan $L(\theta)$ juga akan memaksimumkan fungsi log-kemungkinan $\ln L(\theta)$, sehingga untuk memudahkan perhitungan kita tentukan θ yang memaksimumkan $\ln L(\theta)$.

3. Nyatakan θ sebagai $\hat{\theta}$ untuk memperoleh penduga kemungkinan maksimum bagi θ dengan menggantikan amatan sampel dengan peubah acaknya.

Contoh 3.5 *Misalkan dan* $X_1, X_2, ..., X_n$ *adalah sampel acak dari populasi* X *dengan fungsi kepekatan peluang*

$$f(x;\theta) = \frac{1}{\theta}\mathbf{I}(0 < x < \theta).$$

Tentukan penduga bagi θ *dengan menggunakan metode kemungkinan maksimum.*

Jawab. Fungsi kemungkinan dari contoh ialah

$$\begin{aligned}
L(\theta) &= \prod_{i=1}^{n} f(x_i;\theta) \\
&= \prod_{i=1}^{n} \frac{1}{\theta},\ \theta > x_i \ \ (i = 1, 2, ..., n) \\
&= \left(\frac{1}{\theta}\right)^n,\ \theta > \max\{x_1, ..., x_n\} \\
\ln L(\theta) &= \ln \left(\frac{1}{\theta}\right)^n \\
&= -n \ln \theta.
\end{aligned}$$

Ruang parameter θ terhadap $L(\theta)$ ialah

$$\Omega = \{\theta \in \mathbb{R} \mid x_{\max} < \theta < \infty\} = (x_{\max}, \infty).$$

Selanjutnya, tentukan θ yang memaksimumkan $\ln L(\theta)$. Turunan pertama dari $\ln L(\theta)$ ialah

$$\begin{aligned}
\frac{d \ln L(\theta)}{d\theta} &= \frac{d}{d\theta}[-n \ln \theta] \\
&= -\frac{n}{\theta} < 0.
\end{aligned}$$

Karena $\frac{d \ln L(\theta)}{d\theta} < 0$, maka $\ln L(\theta)$ merupakan fungsi turun. Dengan demikian, $\theta = x_{\max}$ memaksimumkan fungsi $\ln L(\theta)$. Jadi, penduga kemungkinan maksimum bagi θ ialah

$$\hat{\theta} = X_{(n)},$$

di mana $X_{(n)}$ adalah statistik tataan ke-n dari sampel. ✠

Contoh 3.6 *Jika $X_1, X_2, ..., X_n$ adalah sampel acak dengan fungsi kepekatan peluang*

$$f(x;\theta) = \theta e^{-\theta x} \mathbf{I}(x > 0),$$

maka tentukan penduga kemungkinan maksimum bagi θ.

Jawab. Fungsi kemungkinan dari sampel ialah

$$\begin{aligned}
L(\theta) &= \prod_{i=1}^{n} f(x_i;\theta) \\
\ln L(\theta) &= \ln\left[\prod_{i=1}^{n} f(x_i;\theta)\right] \\
&= \sum_{i=1}^{n} \ln f(x_i;\theta) \\
&= \sum_{i=1}^{n} \ln\left(\theta e^{-\theta x_i}\right) \\
&= \sum_{i=1}^{n} [\ln\theta - \theta x_i] \\
&= n\ln\theta - \theta\sum_{i=1}^{n} x_i.
\end{aligned}$$

Kemudian, tentukan nilai maksimum dari $\ln L(\theta)$.

$$\begin{aligned}
\frac{d\ln L(\theta)}{d\theta} &= 0 \\
\frac{d}{d\theta}\left[n\ln\theta - \theta\sum_{i=1}^{n} x_i\right] &= 0 \\
\frac{n}{\theta} - \sum_{i=1}^{n} x_i &= 0 \\
\frac{n}{\theta} &= \sum_{i=1}^{n} x_i \\
\theta &= \frac{1}{\bar{x}}.
\end{aligned}$$

Turunan kedua bagi $\ln L(\theta)$ ialah

$$\begin{aligned}
\frac{d^2\ln L(\theta)}{d\theta^2} &= -\frac{n}{\theta^2} \\
\frac{d^2\ln L(\theta)}{d\theta^2} &= -n\bar{x}^2 < 0.
\end{aligned}$$

Dengan demikian, penduga kemungkinan maksimum bagi θ ialah

$$\hat{\theta} = \frac{1}{\bar{X}}.$$

✠

Contoh 3.7 *Jika $X_1, X_2, ..., X_n$ adalah sampel acak dengan fungsi kepekatan peluang*

$$f(x;\theta) = (1-\theta)\, x^{-\theta}\mathbf{I}(0 < x < 1),$$

maka tentukan penduga kemungkinan maksimum bagi θ.

Jawab. Fungsi kemungkinan dari sampel ialah

$$\begin{aligned} L(\theta) &= \prod_{i=1}^{n} f(x_i;\theta) \\ \ln L(\theta) &= \ln\left[\prod_{i=1}^{n} f(x_i;\theta)\right] \\ &= \sum_{i=1}^{n} \ln f(x_i;\theta) \\ &= \sum_{i=1}^{n} \ln\left[(1-\theta)\, x_i^{-\theta}\right] \\ &= n\ln(1-\theta) - \theta\sum_{i=1}^{n}\ln x_i. \end{aligned}$$

Kemudian, tentukan nilai maksimum dari $\ln L(\theta)$.

$$\begin{aligned} \frac{d\ln L(\theta)}{d\theta} &= 0 \\ \frac{d}{d\theta}\left[n\ln(1-\theta) - \theta\sum_{i=1}^{n}\ln x_i\right] &= 0 \\ -\frac{n}{1-\theta} - \sum_{i=1}^{n}\ln x_i &= 0 \\ \frac{1}{1-\theta} &= -\frac{1}{n}\sum_{i=1}^{n}\ln x_i \\ \theta &= 1 + \frac{n}{\sum_{i=1}^{n}\ln x_i}. \end{aligned}$$

Turunan kedua bagi $\ln L(\theta)$ ialah

$$\begin{aligned} \frac{d^2\ln L(\theta)}{d\theta^2} &= -\frac{n}{(1-\theta)^2} \\ &= -\frac{\left(\sum_{i=1}^{n}\ln x_i\right)^2}{n} < 0. \end{aligned}$$

Dengan demikian, $L(\theta)$ maksimum bila

$$\theta = 1 + \frac{n}{\sum_{i=1}^{n} \ln x_i},$$

sehingga penduga kemungkinan maksimum bagi θ ialah

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

$\maltese$

Contoh 3.8 *Misalkan dan $X_1, X_2, ..., X_n$ adalah sampel acak dari populasi X yang menyebar $N(\theta_1, \theta_2)$. Tentukan penduga bagi θ_1 dan θ_2 dengan menggunakan metode kemungkinan maksimum.*

Jawab. Fungsi kemungkinan dari sampel tersebut ialah

$$\begin{aligned}
L(\theta_1, \theta_2) &= \prod_{i=1}^{n} f(x_i; \theta_1, \theta_2) \\
&= \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi\theta_2}} \exp\left[-\frac{(x_i - \theta_1)^2}{2\theta_2}\right] \\
&= (2\pi\theta_2)^{-n/2} \exp\left[-\frac{\sum_{i=1}^{n} (x_i - \theta_1)^2}{2\theta_2}\right] \\
\ln L(\theta_1, \theta_2) &= -\frac{n}{2} \ln(2\pi) - \frac{n}{2} \ln \theta_2 - \frac{\sum_{i=1}^{n} (x_i - \theta_1)^2}{2\theta_2}.
\end{aligned}$$

Kemudian, tentukan nilai maksimum dari $\ln L(\theta_1, \theta_2)$.

$$\begin{aligned}
\frac{\partial \ln L(\theta_1, \theta_2)}{\partial \theta_1} &= 0 \\
\frac{2\sum_{i=1}^{n}(x_i - \theta_1)}{2\theta_2} &= 0 \\
\theta_1 &= \frac{\sum_{i=1}^{n} x_i}{n} \\
\theta_1 &= \bar{x}, \\
\hat{\theta}_1 &= \bar{X}, \qquad (3.3)
\end{aligned}$$

$$\begin{aligned}
\frac{\partial \ln L(\theta_1, \theta_2)}{\partial \theta_2} &= 0 \\
-\frac{n}{2\theta_2} + \frac{\sum_{i=1}^{n}(x_i - \theta_1)^2}{2(\theta_2)^2} &= 0 \\
\frac{\sum_{i=1}^{n}(x_i - \theta_1)^2}{2(\theta_2)^2} &= \frac{n}{2\theta_2} \\
\theta_2 &= \frac{\sum_{i=1}^{n}(x_i - \theta_1)^2}{n}. \qquad (3.4)
\end{aligned}$$

Selanjutnya, substitusikan 3.3 pada 3.4 untuk menentukan penduga kemungkinan maksimum bagi θ_2, yaitu

$$\hat{\theta}_2 = \frac{\sum_{i=1}^{n}\left(X_i - \bar{X}\right)^2}{n}.$$

✠

Contoh 3.9 *Misalkan dan $X_1, X_2, ..., X_n$ adalah sampel acak dari populasi X yang menyebar gamma dengan parameter α dan β. Tentukan penduga bagi α dan β dengan menggunakan metode kemungkinan maksimum.*

Jawab. Fungsi kemungkinan dari sampel tersebut ialah

$$
\begin{aligned}
L(\alpha,\beta) &= \prod_{i=1}^{n} f(x_i;\alpha,\beta) \\
&= \prod_{i=1}^{n} \frac{1}{\Gamma(\alpha)\beta^{\alpha}} x_i^{\alpha-1} e^{-x_i/\beta} \\
&= \frac{1}{[\Gamma(\alpha)]^n \beta^{n\alpha}} \left(\prod_{i=1}^{n} x_i\right)^{\alpha-1} \exp\left(-\frac{1}{\beta}\sum_{i=1}^{n} x_i\right) \\
\ln L(\alpha,\beta) &= -n\ln\Gamma(\alpha) - n\alpha\ln\beta + (\alpha-1)\sum_{i=1}^{n}\ln x_i - \frac{1}{\beta}\sum_{i=1}^{n} x_i.
\end{aligned}
$$

Kemudian, tentukan nilai maksimum dari $\ln L(\alpha,\beta)$.

$$
\begin{aligned}
\frac{\partial \ln L(\alpha,\beta)}{\partial \beta} &= 0 \\
-\frac{n\alpha}{\beta} + \frac{\sum_{i=1}^{n} x_i}{\beta^2} &= 0 \\
n\alpha &= \frac{\sum_{i=1}^{n} x_i}{\beta} \\
\beta &= \frac{\bar{x}}{\alpha}, \qquad (3.5)
\end{aligned}
$$

$$
\begin{aligned}
\frac{\partial \ln L(\alpha,\beta)}{\partial \alpha} &= 0 \\
-\frac{n}{\Gamma(\alpha)}\Gamma'(\alpha) - n\ln\beta + \sum_{i=1}^{n}\ln x_i &= 0 \\
-\frac{\Gamma'(\alpha)}{\Gamma(\alpha)} - \ln\beta + \frac{1}{n}\sum_{i=1}^{n}\ln x_i &= 0 \\
\frac{\Gamma'(\alpha)}{\Gamma(\alpha)} + \ln\beta &= \frac{1}{n}\sum_{i=1}^{n}\ln x_i. \qquad (3.6)
\end{aligned}
$$

Selanjutnya, substitusikan 3.5 pada 3.6 untuk menentukan penduga kemungkinan maksimum bagi α,

$$\begin{aligned}
\frac{\Gamma'(\alpha)}{\Gamma(\alpha)} + \ln\beta &= \frac{1}{n}\sum_{i=1}^{n}\ln x_i \\
\frac{\Gamma'(\alpha)}{\Gamma(\alpha)} + \ln\frac{\bar{x}}{\alpha} &= \frac{1}{n}\sum_{i=1}^{n}\ln x_i \\
\frac{\Gamma'(\alpha)}{\Gamma(\alpha)} + \ln\bar{x} - \ln\alpha &= \frac{1}{n}\sum_{i=1}^{n}\ln x_i \\
\frac{\Gamma'(\alpha)}{\Gamma(\alpha)} - \ln\alpha &= \frac{1}{n}\sum_{i=1}^{n}\ln x_i - \ln\bar{x}. \qquad (3.7)
\end{aligned}$$

Perhatikan bahwa persamaan 3.7 merupakan persamaan implisit, sehingga dibutuhkan analisis numerik untuk menentukan penduga kemungkinan maksimum bagi α, yaitu $\hat{\alpha}$ dan penduga kemungkinan maksimum bagi β ialah

$$\hat{\beta} = \frac{\bar{X}}{\hat{\alpha}}. \qquad (3.8)$$

✠

Penduga kemungkinan maksimum $\hat{\theta}$ bagi parameter θ memiliki sifat invarian. Sifat invarian tersebut menyatakan bahwa jika $\hat{\theta}$ adalah penduga kemungkinan maksimum bagi θ, maka $g\left(\hat{\theta}\right)$ adalah penduga kemungkinan maksimum bagi $g(\theta)$. Sifat tersebut dinyatakan dalam teorema berikut.

Teorema 3.1 *Misalkan $\hat{\theta}$ adalah penduga kemungkinan maksimum bagi θ, maka penduga kemungkinan maksimum bagi $g(\theta)$ adalah $g\left(\hat{\theta}\right)$.*

Berikut merupakan contoh yang menjelaskan Teorema 3.1.

Contoh 3.10 *Jika $X_1, X_2, ..., X_n$ adalah sampel acak dengan fungsi kepekatan peluang*

$$f(x;\theta) = \theta e^{-\theta x}\mathbf{I}(x > 0),$$

maka tentukan penduga kemungkinan maksimum bagi $\sqrt{\theta^2 + 5}$.

Jawab. Berdasarkan Contoh 3.6, diperoleh penduga kemungkinan maksimum dari θ ialah

$$\hat{\theta} = \frac{1}{\bar{X}},$$

maka dengan menggunakan Teorema 3.1, penduga kemungkinan maksimum bagi $\sqrt{\theta^2+5}$ ialah

$$\sqrt{\frac{1}{\bar{X}^2}+5}.$$

✠

3.3 Sifat-Sifat Penduga

Secara umum, perbedaan penggunaan metode pendugaan akan menghasilkan penduga yang berbeda. Seperti pada Contoh 3.1 dan Contoh 3.5, jika $X \sim$ seragam $(0, \theta)$ dan $X_1, X_2, ..., X_n$ adalah sampel acak dari populasi X, maka penduga yang diperoleh dengan metode momen ialah

$$\hat{\theta}_1 = 2\bar{X},$$

sedangkan penduga yang diperoleh dengan metode kemungkinan maksimum ialah

$$\hat{\theta}_2 = X_{(n)}.$$

Untuk menentukan manakah penduga yang lebih baik, perlu diperiksa apakah penduga tersebut memenuhi sifat-sifat penduga yang baik. Suatu penduga dikatakan baik apabila memenuhi sifat-sifat berikut, yaitu ketakbiasan, memiliki ragam minimum, dan konsisten.

3.3.1 Ketakbiasan

Definisi 3.5 *Misalkan X adalah peubah acak dengan fungsi kepekatan atau fungsi massa peluang $f(x;\theta)$, $\theta \in \Omega$. Misalkan pula $X_1, X_2, ..., X_n$ adalah sampel acak dari X dan T merupakan suatu statistik. T dikatakan penduga* ***takbias*** *bagi θ jika*

$$E(T) = \theta, \ \forall \theta \in \Omega.$$

Jika $E(T) \neq \theta$, maka T disebut penduga ***bias*** *bagi θ. Bias dari $\hat{\theta}$, yang dinotasikan $B\left(\hat{\theta}, \theta\right)$, didefinisikan sebagai*

$$B\left(\hat{\theta}, \theta\right) = E\left(\hat{\theta}\right) - \theta.$$

Jika

$$\lim_{n\to\infty} B\left(\hat{\theta}, \theta\right) = 0 \text{ atau } \lim_{n\to\infty} E(T) = \theta,$$

maka $\hat{\theta}$ merupakan penduga takbias asimtotik bagi θ.

Contoh 3.11 *Misalkan* $X_1, X_2, ..., X_n$ *adalah sampel acak dari sebaran normal dengan nilai harapan* μ *dan ragam* σ^2*. Apakah penduga bagi* μ *dan* σ^2 *yang telah ditunjukkan pada Contoh 3.2 merupakan penduga takbias?*

Jawab. Berdasarkan Contoh 3.2, penduga bagi μ dan σ^2 berturut-turut ialah

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

Karena $\bar{X} \sim N\left(\mu, \frac{\sigma^2}{n}\right)$, maka $E\left(\bar{X}\right) = \mu$. Jadi, $\bar{X}$ adalah penduga takbias bagi μ.

$$\begin{aligned}
E\left[\hat{\sigma}^2\right] &= E\left[\frac{1}{n}\sum_{i=1}^{n}\left(X_i - \bar{X}\right)^2\right] \\
&= E\left[\frac{n-1}{n}\frac{1}{n-1}\sum_{i=1}^{n}\left(X_i - \bar{X}\right)^2\right] \\
&= \frac{n-1}{n}E\left[\frac{1}{n-1}\sum_{i=1}^{n}\left(X_i - \bar{X}\right)^2\right] \\
&= \frac{n-1}{n}E\left[S^2\right] \\
&= \frac{\sigma^2}{n}E\left[\frac{n-1}{\sigma^2}S^2\right], \qquad (3.9)
\end{aligned}$$

Karena $\frac{n-1}{\sigma^2}S^2 \sim \chi^2\left(n-1\right)$, maka persamaan 3.9 menjadi

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

Jadi, $\hat{\sigma}^2 = \frac{1}{n}\sum_{i=1}^{n}\left(X_i - \bar{X}\right)^2$ merupakan penduga bias bagi σ^2. ✠

Pada Contoh 3.11,

$$\begin{aligned}
B\left(\hat{\sigma}^2, \sigma^2\right) &= E\left(\hat{\sigma}^2\right) - \sigma^2 \\
&= \frac{n-1}{n}\sigma^2 - \sigma^2 \\
\lim_{n\to\infty} B\left(\hat{\sigma}^2, \sigma^2\right) &= \lim_{n\to\infty}\left(\frac{n-1}{n}\sigma^2 - \sigma^2\right) \\
&= \lim_{n\to\infty} -\frac{\sigma^2}{n} \\
&= 0.
\end{aligned}$$

Karena $\lim_{n\to\infty} B\left(\hat{\sigma}^2, \sigma^2\right) = 0$, maka $\hat{\sigma}^2$ merupakan penduga takbias asimtotik bagi σ^2. Selanjutnya, berdasarkan fakta bahwa $\frac{n-1}{\sigma^2}S^2 \sim \chi^2(n-1)$, maka

$$E\left(\frac{n-1}{\sigma^2}S^2\right) = n-1,$$

yang berarti bahwa

$$E\left(S^2\right) = \frac{\sigma^2}{n-1} \times (n-1) = \sigma^2.$$

Jadi,

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

merupakan penduga takbias bagi σ^2.

Penduga takbias kadang dapat memberi dugaan yang tidak layak seperti pada contoh berikut.

Contoh 3.12 *Misalkan $X \sim Poisson(\theta)$, $\theta > 0$ dengan*

$$f(x;\theta) = \frac{e^{-\theta}\theta^x}{x!}\mathbf{I}(x = 0, 1, 2, ...).$$

Ingin diperoleh penduga takbias bagi $e^{-3\theta}$. Untuk peubah acak tersebut, berikan penduga bagi fungsi parameter $g(\theta) = e^{-3\theta}$.

Jawab. Untuk sampel acak berukuran satu, yaitu X, $T(X) = (-2)^X$, diperoleh

$$\begin{aligned} E\left(T\left(X\right)\right) &= E\left(\left(-2\right)^{X}\right) \\ &= \sum_{x=0}^{\infty}\left[\left(-2\right)^{x}\cdot\frac{e^{-\theta}\theta^{x}}{x!}\right] \\ &= e^{-\theta}\sum_{x=1}^{n}\frac{\left(-2\theta\right)^{x}}{x!} \\ &= e^{-\theta}e^{-2\theta} \\ &= e^{-3\theta}. \end{aligned}$$

Jadi, $T\left(X\right)=\left(-2\right)^{X}$ merupakan penduga takbias bagi $e^{-3\theta}$, tetapi dugaan yang diperoleh akan kurang dari nol bila x merupakan bilangan gasal, suatu dugaan yang tidak layak digunakan bagi suatu fungsi parameter yang positif. ✠

3.3.2 Efisiensi Relatif Penduga

Definisi 3.6 *Misalkan $\hat{\theta}_1$ dan $\hat{\theta}_2$ adalah dua penduga takbias bagi θ. Penduga $\hat{\theta}_1$ dikatakan lebih efisien daripada $\hat{\theta}_2$ jika $Var\left(\hat{\theta}_1\right) < Var\left(\hat{\theta}_2\right)$. Rasio η didefinisikan sebagai berikut*

$$\eta\left(\hat{\theta}_1,\hat{\theta}_2\right)=\frac{Var\left(\hat{\theta}_2\right)}{Var\left(\hat{\theta}_1\right)},$$

merupakan efisiensi relatif dari $\hat{\theta}_1$ terhadap $\hat{\theta}_2$.

Contoh 3.13 *Misalkan X_1, X_2, X_3 adalah tiga sampel acak dari populasi dengan nilai harapan μ dan ragam $\sigma^2 > 0$. Jika statistik $\bar{X}$ dan*

$$Y=\frac{X_1+2X_2+3X_3}{6}$$

adalah dua penduga takbias bagi μ, maka penduga manakah yang lebih efisien?

Jawab. Ragam dari $\bar{X}$ dan Y ialah

$$\begin{aligned} Var\left(\bar{X}\right) &= Var\left(\frac{X_1+X_2+X_3}{3}\right) \\ &= \frac{1}{9}\left[Var\left(X_1\right)+Var\left(X_2\right)+Var\left(X_3\right)\right] \\ &= \frac{1}{9}\left(3\sigma^2\right) \\ &= \frac{1}{3}\sigma^2 = \frac{12}{36}\sigma^2 \end{aligned}$$

dan

$$\begin{aligned} Var\left(Y\right) &= Var\left(\frac{X_1+2X_2+3X_3}{6}\right) \\ &= \frac{1}{36}\left[Var\left(X_1\right)+4Var\left(X_2\right)+9Var\left(X_3\right)\right] \\ &= \frac{1}{36}\left(14\sigma^2\right) \\ &= \frac{14}{36}\sigma^2. \end{aligned}$$

Jadi, $\frac{12}{36}\sigma^2 = Var\left(\bar{X}\right) < Var\left(Y\right) = \frac{14}{36}\sigma^2$, sehingga $\bar{X}$ merupakan penduga yang lebih efisien daripada Y. Efisiensi relatif dari $\bar{X}$ terhadap Y ialah

$$\eta\left(\bar{X}, Y\right) = \frac{14}{12} = \frac{7}{6}.$$

✠

3.3.3 Efisiensi dan Pertaksamaan Rao-Cramèr

Pada subbab sebelumnya dibahas efisiensi relatif dari dua penduga takbias dengan cara membandingkan ragamnya. Berdasarkan definisi efisiensi relatif penduga tersebut, timbul pertanyaan apakah terdapat penduga yang paling efisien di antara beberapa penduga takbias tersebut? Untuk mengetahui hal tersebut, kita dapat menentukan apakah ada penduga yang paling efisien berdasarkan batas bawah pertaksamaan Rao-Cramèr. Jika suatu penduga memiliki ragam yang nilainya sama dengan batas bawah pertaksamaan Rao-Cramèr, maka penduga tersebut adalah penduga yang paling efisien.

Definisi 3.7 *Misalkan X adalah peubah acak dengan fungsi kepekatan peluang $f(x;\theta)$. Andaikan $f(x;\theta)$ kontinu dan terturunkan dua kali, maka informasi Fisher, dinotasikan $I(\theta)$, didefinisikan sebagai*

$$I(\theta) = \int_{-\infty}^{\infty} \left[\frac{d \ln f(x;\theta)}{d\theta}\right]^2 f(x;\theta)\, dx.$$

Dengan kata lain, $I(\theta)$ adalah nilai harapan dari peubah acak $\left[\frac{d \ln f(X;\theta)}{d\theta}\right]^2$, yaitu

$$I(\theta) = E\left(\left[\frac{d \ln f(X;\theta)}{d\theta}\right]^2\right).$$

Lema berikut adalah formula alternatif untuk informasi Fisher.

Lema 3.1 *Informasi Fisher $I(\theta)$ juga dapat didefinisikan sebagai*

$$I(\theta) = -E\left(\frac{d^2 \ln f(X;\theta)}{d\theta^2}\right).$$

Bukti. Karena $f(x;\theta)$ adalah fungsi kepekatan peluang, maka

$$\int_{-\infty}^{\infty} f(x;\theta)\, dx = 1. \tag{3.10}$$

Turunkan persamaan 3.10 terhadap θ,

$$\frac{d}{d\theta}\int_{-\infty}^{\infty} f(x;\theta)\, dx = 0. \tag{3.11}$$

Selanjutnya, persamaan 3.11 dapat ditulis menjadi

$$\int_{-\infty}^{\infty} \frac{df(x;\theta)}{d\theta}\frac{1}{f(x;\theta)} f(x;\theta)\, dx = 0$$

$$\int_{-\infty}^{\infty} \frac{d \ln f(x;\theta)}{d\theta} f(x;\theta)\, dx = 0$$

$$\int_{-\infty}^{\infty} \left[\frac{d^2 \ln f(x;\theta)}{d\theta^2} f(x;\theta) + \frac{d \ln f(x;\theta)}{d\theta}\frac{df(x;\theta)}{d\theta}\right] dx = 0$$

$$\int_{-\infty}^{\infty} \left[\frac{d^2 \ln f(x;\theta)}{d\theta^2} f(x;\theta) + \frac{d \ln f(x;\theta)}{d\theta}\frac{df(x;\theta)}{d\theta}\frac{1}{f(x;\theta)} f(x;\theta)\right] dx = 0$$

$$\int_{-\infty}^{\infty} \left[\frac{d^2 \ln f(x;\theta)}{d\theta^2} f(x;\theta) + \left[\frac{d \ln f(x;\theta)}{d\theta}\right]^2 f(x;\theta)\right] dx = 0$$

Jadi, berdasarkan persamaan di atas diperoleh

$$\int_{-\infty}^{\infty}\left[\frac{d\ln f(x;\theta)}{d\theta}\right]^2 f(x;\theta)\,dx = -\int_{-\infty}^{\infty}\frac{d^2\ln f(x;\theta)}{d\theta^2}f(x;\theta)\,dx.$$

Dengan demikian, berdasarkan Definisi 3.7 diperoleh

$$I(\theta) = -\int_{-\infty}^{\infty}\frac{d^2\ln f(x;\theta)}{d\theta^2}f(x;\theta)\,dx.$$

✠

Lema 3.2 *Jika $X_1, X_2, ..., X_n$ adalah sampel acak dari sebaran X dengan fungsi kepekatan peluang $f(x;\theta)$, maka informasi Fisher dari sampel berukuran n ialah*

$$I_n(\theta) = nI(\theta).$$

Bukti. Dengan menggunakan Definisi 3.7,

$$\begin{aligned}
I_n(\theta) &= -E\left(\frac{d^2\ln f(X_1, X_2, ..., X_n;\theta)}{d\theta^2}\right)\\
&= -E\left(\frac{d^2}{d\theta^2}\left[\ln f(X_1;\theta) + ... + \ln f(X_n;\theta)\right]\right)\\
&= -E\left(\frac{d^2\ln f(X_1;\theta)}{d\theta^2}\right) - ... - E\left(\frac{d^2\ln f(X_n;\theta)}{d\theta^2}\right)\\
&= I(\theta) + ... + I(\theta)\\
&= nI(\theta).
\end{aligned}$$

✠

Contoh 3.14 *Misalkan X adalah peubah acak dari sebaran eksponensial dengan fungsi kepekatan peluang*

$$f(x;\theta) = \frac{1}{\theta}e^{-x/\theta}\mathbf{I}(0 \le x < \infty),\ \theta > 0.$$

Tentukan $I(\theta)$ dengan menggunakan Definisi 3.7 dan Lema 3.1.

Jawab. Karena $f(x;\theta) = \frac{1}{\theta}e^{-x/\theta}$, maka $\ln f(x;\theta) = -\ln\theta - \frac{x}{\theta}$. Turunan pertama dari $\ln f(x;\theta)$ terhadap θ ialah

$$\begin{aligned}
\frac{d\ln f(x;\theta)}{d\theta} &= -\frac{1}{\theta} + \frac{x}{\theta^2}\\
\left(\frac{d\ln f(x;\theta)}{d\theta}\right)^2 &= \frac{1}{\theta^2} - \frac{2x}{\theta^3} + \frac{x^2}{\theta^4}.
\end{aligned}$$

Karena $X \sim$ eksponensial(θ), maka $E(X) = \theta$ dan $E(X^2) = 2\theta^2$, sehingga dengan menggunakan **Definisi 3.7** diperoleh

$$
\begin{aligned}
I(\theta) &= E\left[\left(\frac{d \ln f(X;\theta)}{d\theta}\right)^2\right] \\
&= \frac{1}{\theta^2} - \frac{2\theta}{\theta^3} + \frac{2\theta^2}{\theta^4} \\
&= \frac{1}{\theta^2}.
\end{aligned}
$$

Jika menggunakan **Lema 3.1**, maka diperlukan turunan kedua dari $\ln f(x;\theta)$ terhadap θ, yaitu

$$\frac{d^2 \ln f(x;\theta)}{d\theta^2} = \frac{1}{\theta^2} - \frac{2x}{\theta^3}.$$

Dengan demikian,

$$
\begin{aligned}
I(\theta) &= -E\left[\frac{d^2 \ln f(X;\theta)}{d\theta^2}\right] \\
&= -E\left[\frac{1}{\theta^2} - \frac{2X}{\theta^3}\right] \\
&= -\left[\frac{1}{\theta^2} - \frac{2\theta}{\theta^3}\right] \\
&= \frac{1}{\theta^2}.
\end{aligned}
$$

✠

Contoh 3.15 *Misalkan $X_1, X_2, ..., X_n$ adalah sampel acak dari sebaran normal dengan parameter μ dan σ^2. Tentukan informasi Fisher $I_n(\mu)$.*

Jawab. Misalkan $X \sim N(\mu, \sigma^2)$, maka

$$
\begin{aligned}
\ln f(x;\mu) &= \ln\left(\frac{1}{\sqrt{2\pi\sigma^2}} \exp\left[-\frac{(x-\mu)^2}{2\sigma^2}\right]\right) \\
&= \ln\left(\frac{1}{\sqrt{2\pi\sigma^2}}\right) - \frac{(x-\mu)^2}{2\sigma^2},
\end{aligned}
$$

sehingga turunan pertama dan kedua dari $\ln f(x;\mu)$ ialah

$$
\begin{aligned}
\frac{d \ln f(x;\mu)}{d\mu} &= \frac{x-\mu}{\sigma^2} \\
\frac{d^2 \ln f(x;\mu)}{d\mu^2} &= -\frac{1}{\sigma^2}.
\end{aligned}
$$

Jadi,

$$\begin{aligned} I_n(\mu) &= nI(\mu) \\ &= n \times -E\left(\frac{d^2 \ln f(X;\mu)}{d\mu^2}\right) \\ &= n \times \frac{1}{\sigma^2}. \end{aligned}$$

✠

Teorema 3.2 (***Pertaksamaan Rao-Cramèr***) *Misalkan $X_1, X_2, ..., X_n$ adalah sampel dari sebaran $X \sim f(x;\theta), \theta \in \Omega$. Misalkan pula $Y = u(X_1, X_2, ..., X_n)$ adalah statistik dengan nilai harapan $E(Y) = k(\theta)$. Andaikan $L(\theta)$ adalah fungsi kemungkinan (likelihood function) yang memenuhi*

$$\begin{aligned} &\frac{d}{d\theta}\int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty} u(x_1, x_2, ..., x_n)\, L(\theta)\, dx_1 \cdots dx_n \\ = &\int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty} u(x_1, x_2, ..., x_n)\, \frac{d}{d\theta}L(\theta)\, dx_1 \cdots dx_n, \end{aligned}$$

maka

$$Var(Y) \geq \frac{(k'(\theta))^2}{nI(\theta)}.$$

Bukti. Nilai harapan $E(Y) = k(\theta)$ dapat dituliskan menjadi

$$\begin{aligned} k(\theta) &= \int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty} u(x_1, x_2, ..., x_n)\, f(x_1;\theta)\cdots f(x_n;\theta)\, dx_1 \cdots dx_n \\ &= \int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty} u(x_1, x_2, ..., x_n)\, L(\theta)\, dx_1 \cdots dx_n \end{aligned}$$

Turunkan $k(\theta)$ terhadap θ, sehingga

$$\begin{aligned} k'(\theta) &= \int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty} u(x_1, x_2, ..., x_n)\, \frac{dL(\theta)}{d\theta} \cdot \frac{1}{L(\theta)} \cdot L(\theta)\, dx_1 \cdots dx_n \\ &= \int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty} u(x_1, x_2, ..., x_n)\, \frac{d\ln L(\theta)}{d\theta} \cdot L(\theta)\, dx_1 \cdots dx_n \\ &= \int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty} u(x_1, x_2, ..., x_n) \left[\sum_{i=1}^{n} \frac{1}{f(x_i;\theta)} \cdot \frac{df(x_i;\theta)}{d\theta}\right] L(\theta)\, dx_1 \cdots dx_n \\ &= \int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty} u(x_1, x_2, ..., x_n) \left[\sum_{i=1}^{n} \frac{d\ln f(x_i;\theta)}{d\theta}\right] L(\theta)\, dx_1 \cdots dx_n. \quad (3.12) \end{aligned}$$

Definisikan peubah acak Z sebagai

$$Z = \sum_{i=1}^{n} \frac{d \ln f(X_i; \theta)}{d\theta},$$

maka $E(Z) = 0$ dan $Var(Z) = nI(\theta)$. Persamaan 3.12 dapat ditulis menjadi

$$k'(\theta) = E(YZ) = E(Y)\,E(Z) + \rho\sigma_Y\sqrt{nI(\theta)}$$

dengan ρ adalah koefisien korelasi antara Y dan Z. Karena $E(Z) = 0$ dan $\rho^2 \leq 1$, maka

$$\rho = \frac{k'(\theta)}{\sigma_Y\sqrt{nI(\theta)}}$$

dan

$$\frac{[k'(\theta)]^2}{\sigma_Y^2\; nI(\theta)} \leq 1.$$

Dengan demikian, terbukti bahwa

$$\sigma_Y^2 \geq \frac{[k'(\theta)]^2}{nI(\theta)}.$$

526 ✠

Definisi 3.8 *Misalkan Y adalah penduga takbias bagi parameter θ. Statistik Y disebut penduga efisien jika ragam dari Y sama dengan batas bawah pertaksamaan Rao-Cramèr. Efisiensi penduga takbias Y bagi parameter θ didefinisikan sebagai*

$$\frac{[k'(\theta)]^2}{nI(\theta)\,Var(Y)} \times 100\%.$$

Jika

$$\lim_{n\to\infty} \frac{[k'(\theta)]^2}{nI(\theta)\,Var(Y)} = 1,$$

527 *maka penduga takbias Y disebut sebagai penduga efisien asimtotik.*

Akibat 3.1 *Berdasarkan asumsi pada Teorema 3.2, jika $Y = u(X_1, X_2, ..., X_n)$ adalah penduga takbias bagi θ, $k(\theta) = \theta$, maka pertaksamaan Rao-Cramèr menjadi*

$$Var(Y) \geq \frac{1}{nI(\theta)}.$$

Contoh 3.16 *Misalkan $X_1, X_2, ..., X_n$ adalah sampel acak dari sebaran Poisson dengan nilai harapan $\theta > 0$. Tunjukkan bahwa $\bar{X}$ adalah penduga efisien bagi θ.*

Jawab.

$$\begin{aligned}\frac{d \ln f(x;\theta)}{d\theta} &= \frac{d}{d\theta}(x\ln\theta - \theta - \ln x!) \\ &= \frac{x}{\theta} - 1 \\ &= \frac{x-\theta}{\theta}\end{aligned}$$

$$\begin{aligned}I(\theta) &= E\left[\left(\frac{d \ln f(X;\theta)}{d\theta}\right)^2\right] \\ &= \frac{E(X-\theta)^2}{\theta^2} \\ &= \frac{\sigma^2}{\theta^2} \\ &= \frac{\theta}{\theta^2} \\ &= \frac{1}{\theta}.\end{aligned}$$

Karena $X \sim \text{Poisson}(\theta)$, maka

$$Var\left(\bar{X}\right) = \frac{Var(X)}{n} = \frac{\theta}{n}.$$

Batas bawah Rao-Cramèr:

$$\frac{1}{nI(\theta)} = \frac{\theta}{n} = Var\left(\bar{X}\right),$$

sehingga $\bar{X}$ merupakan penduga efisien bagi θ. Efisiensi penduga $\bar{X}$ bagi θ sebesar

$$\frac{1/[nI(\theta)]}{Var\left(\bar{X}\right)} \times 100\% = \frac{\theta/n}{\theta/n} \times 100\% = 100\%.$$

✠

Contoh 3.17 *Misalkan $X_1, X_2, ..., X_n$ adalah sampel acak dari sebaran normal dengan nilai harapan μ dan ragam θ. Tentukan efisiensi bagi penduga takbias θ, yaitu*

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

Jawab.

$$\begin{aligned}
\frac{d \ln f(x;\theta)}{d\theta} &= \frac{d}{d\theta}\left(-\frac{1}{2}\ln 2\pi - \frac{1}{2}\ln\theta - \frac{1}{2\theta}(x-\mu)^2\right) \\
&= -\frac{1}{2\theta} + \frac{1}{2\theta^2}(x-\mu)^2 \\
\frac{d^2 \ln f(x;\theta)}{d\theta^2} &= \frac{1}{2\theta^2} - \frac{1}{\theta^3}(x-\mu)^2 \\
I(\theta) &= -E\left[\frac{d^2 \ln f(X;\theta)}{d\theta^2}\right] \\
&= \frac{1}{2\theta^2} - \frac{1}{\theta^3}E(X-\mu)^2 \\
&= \frac{1}{2\theta^2} - \frac{1}{\theta^3}\cdot\theta \\
&= \frac{1}{2\theta^2}.
\end{aligned}$$

Batas bawah pertaksamaan Rao-Cramèr:

$$\frac{1}{nI(\theta)} = \frac{1}{n\frac{1}{2\theta^2}} = \frac{2\theta^2}{n}.$$

Karena $Y = \dfrac{\sum_{i=1}^{n}\left(X_i - \bar{X}\right)^2}{\theta} \sim \chi^2(n-1)$, maka $Var(Y) = 2(n-1)$ dan

$$\begin{aligned}
S^2 &= \frac{\theta}{(n-1)}Y, \\
Var\left(S^2\right) &= \frac{\theta^2}{(n-1)^2}Var(Y) \\
&= \frac{\theta^2}{(n-1)^2}\cdot 2(n-1) \\
&= \frac{2\theta^2}{(n-1)}.
\end{aligned}$$

Dengan demikian,

$$Var\left(S^2\right) = \frac{2\theta^2}{(n-1)} > \frac{1}{nI(\theta)} = \frac{2\theta^2}{n}.$$

Efisiensi penduga S^2 bagi θ sebesar

$$\frac{1/[nI(\theta)]}{Var(S^2)} \times 100\% = \frac{n-1}{n} \times 100\%$$

yang akan sama dengan 100% untuk $n \to \infty$, sehingga S^2 merupakan penduga efisien asimtotik bagi θ. ✠

3.3.4 Kekonsistenan

Misalkan $X_1, X_2, ..., X_n$ adalah contoh acak dari populasi X dengan fungsi kepekatan peluang $f(x;\theta)$. Misalkan $\hat{\theta}$ adalah penduga bagi sampel berukuran n. Jelas bahwa penduga tersebut akan bergantung pada ukuran contoh n. Oleh karena itu, untuk menunjukkan bahwa $\hat{\theta}$ bergantung pada n, maka notasikan $\hat{\theta}$ sebagai $\hat{\theta}_n$.

Definisi 3.9 *Misalkan $X_1, X_2, ..., X_n$ adalah sampel acak dari populasi X dengan fungsi kepekatan peluang $f(x;\theta)$. Barisan penduga $\left\{\hat{\theta}_n\right\}$ disebut konsisten jika barisan $\left\{\hat{\theta}_n\right\}$ konvergen dalam peluang ke θ, yaitu*

$$\forall \varepsilon > 0, \ \lim_{n\to\infty} P\left(\left|\hat{\theta}_n - \theta\right| \geq \varepsilon\right) = 0.$$

Teorema 3.3 *Misalkan $X_1, X_2, ..., X_n$ adalah sampel acak dari sebaran X dengan fungsi kepekatan peluang $f(x;\theta)$ dan $\left\{\hat{\theta}_n\right\}$ merupakan barisan penduga bagi θ berdasarkan sampel. Jika ragam dari $\hat{\theta}_n$ ada untuk setiap n dan berhingga, serta*

$$\lim_{n\to\infty} E\left(\left(\hat{\theta}_n - \theta\right)^2\right) = 0,$$

maka

$$\forall \varepsilon > 0, \ \lim_{n\to\infty} P\left(\left|\hat{\theta}_n - \theta\right| \geq \varepsilon\right) = 0.$$

Bukti. Berdasarkan pertaksamaan Markov diperoleh

$$P\left(\left(\hat{\theta}_n - \theta\right)^2 \geq \varepsilon^2\right) \leq \frac{E\left(\left(\hat{\theta}_n - \theta\right)^2\right)}{\varepsilon^2}$$

untuk setiap $\varepsilon > 0$. Karena $\left(\hat{\theta}_n - \theta\right)^2 \geq \varepsilon^2$ equivalen dengan $\left|\hat{\theta}_n - \theta\right| \geq \varepsilon$, maka

$$P\left(\left(\hat{\theta}_n - \theta\right)^2 \geq \varepsilon^2\right) = P\left(\left|\hat{\theta}_n - \theta\right| \geq \varepsilon\right) \leq \frac{E\left(\left(\hat{\theta}_n - \theta\right)^2\right)}{\varepsilon^2}$$

untuk setiap $n \in \mathbb{N}$. Perhatikan bahwa jika

$$\lim_{n\to\infty} E\left(\left(\hat{\theta}_n - \theta\right)^2\right) = 0$$

maka

$$\lim_{n\to\infty} P\left(\left|\hat{\theta}_n - \theta\right| \geq \varepsilon\right) = 0.$$

$\maltese$

Bias bagi penduga θ ialah $B\left(\hat{\theta}, \theta\right) = E\left(\hat{\theta}_n\right) - \theta$. Jika penduga yang diperoleh merupakan penduga takbias, maka $B\left(\hat{\theta}, \theta\right) = 0$. Selanjutnya, akan ditunjukkan bahwa

$$E\left(\left(\hat{\theta} - \theta\right)^2\right) = Var\left(\hat{\theta}\right) + \left[B\left(\hat{\theta}, \theta\right)\right]^2. \tag{3.13}$$

Perhatikan bahwa

$$\begin{aligned}
E\left(\left(\hat{\theta} - \theta\right)^2\right) &= E\left(\hat{\theta}^2 - 2\hat{\theta}\theta + \theta^2\right) \\
&= E\left(\hat{\theta}^2\right) - 2\theta E\left(\hat{\theta}\right) + \theta^2 \\
&= E\left(\hat{\theta}^2\right) - \left[E\left(\hat{\theta}\right)\right]^2 + \left[E\left(\hat{\theta}\right)\right]^2 - 2\theta E\left(\hat{\theta}\right) + \theta^2 \\
&= Var\left(\hat{\theta}\right) + \left[E\left(\hat{\theta}\right)\right]^2 - 2\theta E\left(\hat{\theta}\right) + \theta^2 \\
&= Var\left(\hat{\theta}\right) + \left[E\left(\hat{\theta}_n\right) - \theta\right]^2 \\
&= Var\left(\hat{\theta}\right) + \left[B\left(\hat{\theta}, \theta\right)\right]^2.
\end{aligned}$$

Jika

$$\lim_{n\to\infty} Var\left(\hat{\theta}\right) = 0 \text{ dan } \lim_{n\to\infty} B\left(\hat{\theta}, \theta\right) = 0, \tag{3.14}$$

maka berdasarkan persamaan 3.13,

$$\lim_{n\to\infty} E\left(\left(\hat{\theta} - \theta\right)^2\right) = 0.$$

Dengan kata lain, untuk menunjukkan bahwa suatu penduga konsisten, kita cukup menunjukkan persamaan 3.14.

Contoh 3.18 *Misalkan $X_1, X_2, \ldots, X_n$ adalah sampel acak dari populasi normal dengan nilai harapan μ dan ragam $\sigma^2 > 0$. Tentukan apakah penduga bagi σ^2, yaitu*

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

merupakan penduga konsisten bagi σ^2?

Jawab. Karena $\hat{\sigma}^2$ bergantung terhadap n, maka notasikan $\hat{\sigma}^2$ sebagai $\hat{\sigma}_n^2$, sehingga

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

Ragam bagi $\hat{\sigma}_n^2$ ialah

$$\begin{aligned}
Var\left(\hat{\sigma}_n^2\right) &= Var\left(\frac{1}{n}\sum_{i=1}^{n}\left(X_i - \bar{X}\right)^2\right) \\
&= \frac{1}{n^2}Var\left(\sigma^2\frac{(n-1)\,S^2}{\sigma^2}\right) \\
&= \frac{\sigma^4}{n^2}Var\left(\frac{(n-1)\,S^2}{\sigma^2}\right); \quad \frac{(n-1)\,S^2}{\sigma^2} \sim \chi^2\,(n-1) \\
&= \frac{2\,(n-1)\,\sigma^4}{n^2} \\
&= \left[\frac{1}{n} - \frac{1}{n^2}\right]2\sigma^4, \\
\lim_{n\to\infty} Var\left(\hat{\sigma}_n^2\right) &= \lim_{n\to\infty}\left[\frac{1}{n} - \frac{1}{n^2}\right]2\sigma^4 = 0.
\end{aligned}$$

Bias $B\left(\hat{\sigma}_n^2, \sigma^2\right)$ ialah

$$\begin{aligned}
B\left(\hat{\sigma}_n^2, \sigma^2\right) &= E\left(\hat{\sigma}_n^2\right) - \sigma^2 \\
&= E\left(\frac{1}{n}\sum_{i=1}^{n}\left(X_i - \bar{X}\right)^2\right) - \sigma^2 \\
&= \frac{1}{n}E\left(\sigma^2\frac{(n-1)\,S^2}{\sigma^2}\right) - \sigma^2 \\
&= \frac{\sigma^2}{n}E\left(\frac{(n-1)\,S^2}{\sigma^2}\right) - \sigma^2 \\
&= \frac{(n-1)\,\sigma^2}{n} - \sigma^2 \\
&= -\frac{\sigma^2}{n}, \\
\lim_{n\to\infty} B\left(\hat{\sigma}_n^2, \sigma^2\right) &= -\lim_{n\to\infty}\frac{\sigma^2}{n} = 0.
\end{aligned}$$

Jadi, $\frac{1}{n}\sum_{i=1}^{n}\left(X_i - \bar{X}\right)^2$ adalah penduga konsisten bagi σ^2. ✠

3.4 Penduga Selang

Definisi 3.10 *Misalkan* $X_1, X_2, ..., X_n$ *adalah sampel acak dari populasi* X *dengan fungsi kepekatan peluang* $f(x, \theta)$*, di mana* θ *adalah parameter yang tidak diketahui. Penduga selang bagi* θ *dengan peluang* $(1-\alpha)$ *adalah selang statistik* $[L(X_1, X_2, ..., X_n), U(X_1, X_2, ..., X_n)]$ *dengan* $L \leq U$ *sedemikian sehingga*

$$P\left(\left[L(X_1, X_2, ..., X_n) \leq \theta \leq U(X_1, X_2, ..., X_n)\right]\right) = 1 - \alpha \tag{3.15}$$

dengan $0 < \alpha < 1$.

Pada persamaan 3.15, $P(L \leq \theta \leq U)$ menyatakan bahwa peluang parameter θ berada di antara selang acak tersebut adalah sebesar $(1-\alpha)$, di mana $(1-\alpha)$ disebut sebagai **koefisien kepercayaan**. Setelah diperoleh titik atau amatan sampel, maka selang $[L(x_1, x_2, ..., x_n), U(x_1, x_2, ..., x_n)]$ disebut sebagai **selang kepercayaan** $100(1-\alpha)\,\%$ bagi θ. Selang

$$[L(x_1, x_2, ..., x_n), U(x_1, x_2, ..., x_n)]$$

tentunya dapat mencakup atau tidak mencakup nilai θ yang sebenarnya, tetapi bila sampel ini diperoleh sebanyak m kali, maka harapan kita $(1-\alpha)\times m$ kali selang ini akan mencakup nilai θ.

Contoh 3.19 *Misalkan* $X_1, X_2, ..., X_n$ *adalah sampel acak dari sebaran* $N(\mu, \sigma^2)$, *dengan asumsi* $\sigma^2 > 0$ *diketahui. Misalkan pula* $\bar{X}$ *merupakan nilai tengah dari sampel, maka peubah acak* $\left(\bar{X} - \mu\right) / \left(\sigma/\sqrt{n}\right)$ *memiliki sebaran normal baku. Dengan kata lain,*

$$\frac{\left(\bar{X} - \mu\right)}{\sigma/\sqrt{n}} \sim N(0, 1).$$

Tentukan selang kepercayaan $100(1-\alpha)\%$ *bagi* μ.

Jawab. Untuk $0 < \alpha < 1$, definisikan bahwa

$$P\left(\frac{\left(\bar{X} - \mu\right)}{\sigma/\sqrt{n}} > z_{\alpha/2}\right) = \frac{\alpha}{2}$$

sedemikian sehingga

$$\begin{aligned} 1 - \alpha &= P\left(-z_{\alpha/2} < \frac{\left(\bar{X} - \mu\right)}{\sigma/\sqrt{n}} < z_{\alpha/2}\right) \\ &= P\left(-\frac{\sigma}{\sqrt{n}} z_{\alpha/2} < \bar{X} - \mu < \frac{\sigma}{\sqrt{n}} z_{\alpha/2}\right) \\ &= P\left(\bar{X} - \frac{\sigma}{\sqrt{n}} z_{\alpha/2} < \mu < \bar{X} + \frac{\sigma}{\sqrt{n}} z_{\alpha/2}\right). \end{aligned}$$

Jadi, selang kepercayaan $100(1-\alpha)\%$ bagi μ ialah

$$\left(\bar{x} - \frac{\sigma}{\sqrt{n}} z_{\alpha/2},\ \bar{x} + \frac{\sigma}{\sqrt{n}} z_{\alpha/2}\right),$$

di mana

$$\bar{x} = \frac{1}{n} \sum\nolimits_{i=1}^{n} x_i \ .$$

✠

Contoh 3.20 *Misalkan* $X_1, X_2, ..., X_n$ *adalah sampel acak dari sebaran* $N(\mu, \sigma^2)$, *dengan asumsi* σ^2 *tidak diketahui. Misalkan* $\bar{X}$ *dan* S^2 *merupakan nilai tengah dan ragam dari sampel. Tentukan selang kepercayaan* $100(1-\alpha)\%$ *bagi* μ.

Jawab. Karena $X_1, X_2, ..., X_n \sim N(\mu, \sigma^2)$, maka

$$\bar{X} \sim N\left(\mu, \frac{\sigma^2}{n}\right) \text{ dan } \frac{n-1}{\sigma^2} S^2 \sim \chi^2(n-1),$$

dengan $\bar{X}$ dan $\frac{n-1}{n}S^2$ saling bebas. Oleh karena itu, peubah acak

$$T = \frac{\left(\bar{X}-\mu\right)}{S/\sqrt{n}} \sim t\left(n-1\right).$$

Untuk $0 < \alpha < 1$, definisikan bahwa $P\left(\frac{\left(\bar{X}-\mu\right)}{S/\sqrt{n}} > t_{\alpha/2}\right) = \frac{\alpha}{2}$ sedemikian sehingga

$$\begin{aligned} 1-\alpha &= P\left(-t_{\alpha/2} < \frac{\left(\bar{X}-\mu\right)}{S/\sqrt{n}} < t_{\alpha/2}\right) \\ &= P\left(-\frac{S}{\sqrt{n}}t_{\alpha/2} < \bar{X}-\mu < \frac{S}{\sqrt{n}}t_{\alpha/2}\right) \\ &= P\left(\bar{X}-\frac{S}{\sqrt{n}}t_{\alpha/2} < \mu < \bar{X}+\frac{S}{\sqrt{n}}t_{\alpha/2}\right). \end{aligned}$$

Jadi, selang kepercayaan $100\left(1-\alpha\right)\%$ bagi μ ialah

$$\left(\bar{x}-\frac{s}{\sqrt{n}}t_{\alpha/2},\ \bar{x}+\frac{s}{\sqrt{n}}t_{\alpha/2}\right),$$

di mana

$$\bar{x} = \frac{1}{n}\sum\nolimits_{i=1}^{n} x_i \text{ dan } s^2 = \frac{1}{n-1}\sum\nolimits_{i=1}^{n}\left(x_i-\bar{x}\right)^2.$$

✠

Contoh 3.21 *Misalkan $X_1, X_2, ..., X_n$ adalah sampel acak dari sebaran $N\left(\mu, \sigma^2\right)$, dengan asumsi σ^2 tidak diketahui. Misalkan $\bar{X}$ dan S^2 merupakan nilai tengah dan ragam dari sampel. Tentukan selang kepercayaan $100\left(1-\alpha\right)\%$ bagi σ^2.*

Jawab. Karena $X_1, X_2, ..., X_n \sim N\left(\mu, \sigma^2\right)$, maka

$$\frac{n-1}{\sigma^2}S^2 \sim \chi^2_{\alpha/2}\left(n-1\right).$$

Untuk $0 < \alpha < 1$, definisikan bahwa

$$P\left(\frac{n-1}{\sigma^2}S^2 < \chi^2_{\alpha/2}\left(n-1\right)\right) = \frac{\alpha}{2}$$

sedemikian sehingga

$$\begin{aligned}
1-\alpha &= P\left(\chi^2_{\alpha/2}(n-1) < \frac{n-1}{\sigma^2}S^2 < \chi^2_{1-\alpha/2}(n-1)\right) \\
&= P\left(\frac{\chi^2_{\alpha/2}(n-1)}{(n-1)S^2} < \frac{1}{\sigma^2} < \frac{\chi^2_{1-\alpha/2}(n-1)}{(n-1)S^2}\right) \\
&= P\left(\frac{(n-1)S^2}{\chi^2_{1-\alpha/2}(n-1)} < \sigma^2 < \frac{(n-1)S^2}{\chi^2_{\alpha/2}(n-1)}\right).
\end{aligned}$$

Jadi, selang kepercayaan $100(1-\alpha)\%$ bagi σ^2 ialah

$$\left(\frac{(n-1)s^2}{\chi^2_{1-\alpha/2}(n-1)}, \frac{(n-1)s^2}{\chi^2_{\alpha/2}(n-1)}\right),$$

di mana

$$s^2 = \frac{1}{n-1}\sum\nolimits_{i=1}^{n}(x_i-\bar{x})^2.$$

✠

Contoh 3.22 *Misalkan X adalah peubah acak Bernoulli dengan peluang sukses sebesar p dan $X_1, X_2, ..., X_n$ adalah sampel acak dari sebaran X. Misalkan $\hat{p} = \bar{X}$ adalah proporsi sampel yang menyatakan sukses, dengan $\hat{p} = \frac{1}{n}\sum_{i=1}^{n} X_i$ merupakan nilai tengah dan $Var(\hat{p}) = p(1-p)/n$. Tentukan selang kepercayaan bagi p.*

Jawab. Berdasarkan Teorema Limit Pusat, sebaran dari peubah acak

$$Z = \frac{\hat{p}-p}{\sqrt{p(1-p)/n}}$$

mendekati sebaran $N(0,1)$. Bila $p(1-p)$ diganti dengan penduganya, yaitu $\hat{p}(1-\hat{p})$, maka

$$\begin{aligned}
1-\alpha &= P\left(-z_{\alpha/2} < \frac{\hat{p}-p}{\sqrt{\hat{p}(1-\hat{p})/n}} < z_{\alpha/2}\right) \\
&= P\left(-z_{\alpha/2}\sqrt{\hat{p}(1-\hat{p})/n} < \hat{p}-p < z_{\alpha/2}\sqrt{\hat{p}(1-\hat{p})/n}\right) \\
&= P\left(\hat{p}-z_{\alpha/2}\sqrt{\hat{p}(1-\hat{p})/n} < p < \hat{p}+z_{\alpha/2}\sqrt{\hat{p}(1-\hat{p})/n}\right)
\end{aligned}$$

Jadi, selang kepercayaan $100(1-\alpha)\%$ bagi μ ialah

$$\left(\bar{p}-z_{\alpha/2}\sqrt{\bar{p}(1-\bar{p})/n}, \bar{p}+z_{\alpha/2}\sqrt{\bar{p}(1-\bar{p})/n}\right),$$

dengan $\bar{p}$ menyatakan dugaan dari p. ✠

3.4.1 Selang Kepercayaan Beda Dua Nilai Tengah

Misalkan $X_1, X_2, ..., X_m$ adalah sampel acak dari sebaran X dan $Y_1, Y_2, ..., Y_n$ adalah sampel acak dari sebaran Y. Asumsikan bahwa masing-masing sampel saling bebas dan ragam dari X dan Y berhingga, dinotasikan dengan $\sigma_1^2 = Var(X)$ dan $\sigma_2^2 = Var(Y)$. Misalkan $\bar{X} = \frac{1}{m}\sum_{i=1}^{m} X_i$ dan $\bar{Y} = \frac{1}{n}\sum_{i=1}^{n} Y_i$ adalah nilai tengah sampel dari sebaran X dan Y. Dengan demikian, $\bar{X} - \bar{Y}$ memiliki sebaran dengan

$$E\left(\bar{X} - \bar{Y}\right) = \mu_x - \mu_y \text{ dan } Var\left(\bar{X} - \bar{Y}\right) = \frac{\sigma_1^2}{m} + \frac{\sigma_2^2}{n}$$

Berdasarkan asumsi kebebasan sampel dan Teorema Limit Pusat, peubah acak

$$Z = \frac{\left(\bar{X} - \bar{Y}\right) - (\mu_x - \mu_y)}{\sqrt{\frac{S_1^2}{m} + \frac{S_2^2}{n}}} \dot{\sim} N(0,1).$$

Jadi, selang kepercayaan $100(1-\alpha)\%$ bagi $\mu_x - \mu_y$ ialah

$$\left((\bar{x} - \bar{y}) - z_{\alpha/2}\sqrt{\frac{s_1^2}{m} + \frac{s_2^2}{n}}, (\bar{x} - \bar{y}) + z_{\alpha/2}\sqrt{\frac{s_1^2}{m} + \frac{s_2^2}{n}}\right) \tag{3.16}$$

dengan

$$\begin{aligned} s_1^2 &= \frac{1}{m-1}\sum_{i=1}^{m}(x_i - \bar{x})^2, \\ s_2^2 &= \frac{1}{n-1}\sum_{i=1}^{n}(y_i - \bar{y})^2. \end{aligned}$$

Selang kepercayaan bagi beda dua nilai tengah yang dihasilkan pada persamaan 3.16 merupakan suatu pendekatan. Berikut ini akan dijelaskan hasil eksak untuk selang kepercayaan bagi beda dua nilai tengah jika diasumsikan bahwa sebaran X dan Y adalah normal dengan ragam yang sama, yaitu $\sigma_1^2 = \sigma_2^2$. Asumsikan $X \sim N(\mu_x, \sigma^2)$ dan $Y \sim N(\mu_y, \sigma^2)$ saling bebas. Misalkan $n = n_1 + n_2$ adalah total dari ukuran sampel. Misalkan pula $X_1, X_2, ..., X_{n_1}$ adalah sampel acak dari sebaran X dan $Y_1, Y_2, ..., Y_{n_2}$ adalah sampel acak dari sebaran Y.

Penduga bagi $\mu_x - \mu_y$ adalah $\bar{X} - \bar{Y}$ dengan $E\left(\bar{X} - \bar{Y}\right) = \mu_x - \mu_y$ dan $Var\left(\bar{X} - \bar{Y}\right) = \frac{\sigma^2}{n_1} + \frac{\sigma^2}{n_2}$. Karena $\bar{X}$ dan $\bar{Y}$ saling bebas, maka

$$\frac{\left(\bar{X} - \bar{Y}\right) - (\mu_x - \mu_y)}{\sigma\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} \sim N(0,1).$$

Misalkan

$$S_p^2 = \frac{(n_1 - 1)\, S_1^2 + (n_2 - 1)\, S_2^2}{n_1 + n_2 - 2}.$$

Perhatikan bahwa S_p^2 merupakan penduga tak bias bagi σ^2. Karena

$$(n_1 - 1)\, S_1^2/\sigma^2 \sim \chi^2\,(n_1 - 1) \text{ dan } (n_2 - 1)\, S_2^2/\sigma^2 \sim \chi^2\,(n_2 - 1)\,,$$

serta S_1^2 dan S_2^2 saling bebas, maka $(n-2)\, S_p^2/\sigma^2 \sim \chi^2\,(n-2)$. Dengan demikian,

$$\begin{aligned} T &= \frac{\left[\left(\bar{X} - \bar{Y}\right) - (\mu_x - \mu_y)\right] / \left[\sigma\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}\right]}{\sqrt{(n-2)\, S_p^2/\,(n-2)\,\sigma^2}} \\ &= \frac{\left[\left(\bar{X} - \bar{Y}\right) - (\mu_x - \mu_y)\right]}{\left[S_p\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}\right]} \end{aligned}$$

memiliki sebaran-t dengan derajat bebas $(n-2)$, sehingga dapat disimpulkan bahwa selang kepercayaan $100\,(1-\alpha)\,\%$ bagi $\mu_1 - \mu_2$ ialah

$$\left((\bar{x} - \bar{y}) - t_{(\alpha/2, n-2)} s_p \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}, (\bar{x} - \bar{y}) + t_{(\alpha/2, n-2)} s_p \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}\right)$$

dengan

$$\begin{aligned} s_p^2 &= \frac{(n_1 - 1)\, s_1^2 + (n_2 - 1)\, s_2^2}{n_1 + n_2 - 2}, \\ s_1^2 &= \frac{1}{n_1 - 1} \sum_{i=1}^{n_1} (x_i - \bar{x})^2\,, \\ s_2^2 &= \frac{1}{n_2 - 1} \sum_{i=1}^{n_2} (y_i - \bar{y})^2\,. \end{aligned}$$

3.4.2 Selang Kepercayaan Nisbah Dua Ragam

Misalkan $X_1, X_2, ..., X_n$ adalah sampel acak dari sebaran $X \sim N\,(\mu_x, \sigma_x^2)\,,$ maka

$$V = \frac{\sum_{i=1}^{n} \left(X_i - \bar{X}\right)^2}{\sigma_x^2} \sim \chi^2\,(n-1)\,.$$

Misalkan pula $Y_1, Y_2, ..., Y_m$ adalah sampel acak dari sebaran $Y \sim N\left(\mu_y, \sigma_y^2\right),$ maka

$$W = \frac{\sum_{i=1}^{m} \left(Y_i - \bar{Y}\right)^2}{\sigma_y^2} \sim \chi^2\,(m-1)\,.$$

Karena peubah acak V dan W saling bebas, maka

$$\frac{W/(m-1)}{V/(n-1)} \sim F(m-1, n-1),$$

sehingga untuk $0 < \alpha < 1$,

$$\begin{aligned}
1-\alpha &= P\left(F_{\alpha/2}(m-1,n-1) \leq \frac{\sum_{i=1}^{m}\left(Y_i-\bar{Y}\right)/\sigma_y^2(m-1)}{\sum_{i=1}^{n}\left(X_i-\bar{X}\right)/\sigma_x^2(n-1)} \leq F_{1-\alpha/2}(m-1,n-1)\right) \\
&= P\left(F_{\alpha/2}(m-1,n-1) \leq \frac{S_y^2/\sigma_y^2}{S_x^2/\sigma_x^2} \leq F_{1-\alpha/2}(m-1,n-1)\right) \\
&= P\left(F_{\alpha/2}(m-1,n-1)\frac{S_x^2}{S_y^2} \leq \frac{\sigma_x^2}{\sigma_y^2} \leq F_{1-\alpha/2}(m-1,n-1)\frac{S_x^2}{S_y^2}\right).
\end{aligned}$$

Jadi, selang kepercayaan $100(1-\alpha)\%$ bagi $\frac{\sigma_x^2}{\sigma_y^2}$ ialah

$$\left(F_{\alpha/2}(m-1,n-1)\frac{s_x^2}{s_y^2},\ F_{1-\alpha/2}(m-1,n-1)\frac{s_x^2}{s_y^2}\right)$$

dengan

$$\begin{aligned}
s_x^2 &= \frac{1}{n-1}\sum_{i=1}^{n}(x_i-\bar{x})^2, \\
s_y^2 &= \frac{1}{m-1}\sum_{i=1}^{m}(y_i-\bar{y})^2.
\end{aligned}$$

Catatan: Misalkan

$$Z = \frac{W/(m-1)}{V/(n-1)} \sim F(m-1,n-1)$$

atau

$$\frac{1}{Z} = \frac{V/(n-1)}{W/(m-1)} \sim F(n-1,m-1),$$

maka

$$F_{\alpha/2}(m-1,n-1) = \frac{1}{F_{1-\alpha/2}(n-1,m-1)}.$$

Bukti. Karena $Z \sim F(m-1,n-1)$, maka

$$\begin{aligned}
& P\left(Z \leq F_{\alpha/2}(m-1,n-1)\right) = \tfrac{\alpha}{2} \\
\Leftrightarrow \quad & P\left(\tfrac{1}{Z} \geq \tfrac{1}{F_{\alpha/2}(m-1,n-1)}\right) = \tfrac{\alpha}{2}.
\end{aligned}$$

Karena $\frac{1}{Z} \sim F\left(n-1, m-1\right)$, maka

$$\begin{aligned} F_{1-\alpha/2}\left(n-1, m-1\right) &= \frac{1}{F_{\alpha/2}\left(m-1, n-1\right)} \\ \Leftrightarrow \quad F_{\alpha/2}\left(m-1, n-1\right) &= \frac{1}{F_{1-\alpha/2}\left(n-1, m-1\right)}. \end{aligned}$$

✠

3.5 Latihan

1. Misalkan $X_1, X_2, ..., X_n$ adalah sampel acak dari sebaran dengan fungsi kepekatan peluang

$$f\left(x;\theta\right) = \frac{1}{\beta-\alpha}\mathbf{I}\left(\alpha < x < \beta\right)$$

   Tentukan penduga dari α dan β dengan menggunakan metode momen.

2. Misalkan $X_1, X_2, ..., X_n$ adalah sampel acak dari sebaran dengan fungsi kepekatan peluang

$$f\left(x;\theta\right) = \sqrt{\frac{2}{\pi}}\exp\left[-\frac{1}{2}\left(x-\theta\right)^2\right]\mathbf{I}\left(x \geq \theta\right)$$

   Tentukan penduga kemungkinan maksimum bagi θ.

3. Andaikan $X_1, X_2, ..., X_n$ merupakan sampel acak dari sebaran dengan fungsi kepekatan peluang

$$f\left(x;\theta\right) = \mathbf{I}\left(\theta - 0.5 \leq x \leq \theta + 0.5\right)$$

   dan $Y_1 < Y_2 < ... < Y_n$ merupakan statistik-statistik tataan (*order statistics*)-nya. Carilah penduga bagi θ dengan menggunakan:

   (a) metode momen,

   (b) metode kemungkinan maksimum. Jelaskan bahwa $(Y_1+Y_n)/2$ dan $(4Y_1+2Y_n+1)/6$ masing-masing merupakan penduga kemungkinan maksimum bagi θ.

4. Misalkan $X_1, X_2, ..., X_n$ adalah sampel acak dari sebaran $\Gamma\left(2,\theta\right)$, di mana θ merupakan parameter yang tidak diketahui.

(a) Carilah sebaran dari Y jika $Y = \sum_{i=1}^{n} X_i$ dan tentukan konstanta c sedemikian sehingga cY adalah penduga tak bias bagi θ.

(b) Jika $n = 5$, tunjukkan bahwa $P\left(9.59 < \frac{2Y}{\theta} < 34.2\right) = 0.95$.

5. Misalkan $X_1, X_2, ..., X_n$ adalah sampel acak dari sebaran dengan fungsi kepekatan peluang

$$f(x;\theta) = e^{-(x-\theta)}\mathbf{I}(x > \theta)$$

di mana $-\infty < \theta < \infty$. Periksa apakah penduga $X_{(1)}$ dan $\bar{X} - 1$ merupakan penduga tak bias bagi θ? Di antara penduga tersebut, manakah penduga yang lebih efisien?

6. Misalkan $X_1, X_2, ..., X_n$ adalah sampel acak dari sebaran $X \sim Poisson(\lambda)$ dengan $\lambda > 0$. Tentukan penduga bagi λ dan periksa apakah penduga tersebut merupakan penduga konsisten bagi λ.

7. Misalkan $X_1, X_2, ..., X_n$ adalah sampel acak dari sebaran dengan fungsi kepekatan peluang

$$f(x;\theta) = \theta e^{-\theta x}\mathbf{I}(x > 0)$$

di mana $\theta > 0$.

(a) Tentukanlah $I(\theta)$, informasi Fisher bagi θ.

(b) Berikanlah penduga takbias bagi θ.

(c) Berikanlah batas bawah pertaksamaan Rao-Cramèr untuk penduga yang diperoleh dari butir b.

(d) Carilah efisiensi penduga yang diperoleh dari butir b.

8. Andaikan fungsi kepekatan peluang $f(x;\theta) = \theta \exp(-\theta x)\mathbf{I}(x > 0)$ dan $Y = (X_1 + X_2 + ... + X_n)/n$, maka :

(a) Tunjukkanlah bahwa $W = 2nY/\theta$ memiliki sebaran Khi-kuadrat dengan derajat bebas $2n$.

(b) Berikanlah selang kepercayaan $(1 - \alpha)$ bagi θ.

9. Misalkan $X_1, X_2, ..., X_{11}$ adalah sampel acak dari sebaran normal dengan nilai tengah μ (tidak diketahui) dan ragam $\sigma^2 = 9.9$. Jika $\sum_{i=1}^{11} x_i = 132$, tentukanlah konstanta k sedemikian sehingga 90% selang kepercayaan bagi μ ialah $\left[12 - k\sqrt{0.9}, 12 + k\sqrt{0.9}\right]$.

10. Misalkan $\bar{X}$ dan $\bar{Y}$ adalah nilai tengah dari dua sampel acak yang saling bebas dari sebaran $N(\mu_x, \sigma^2)$ dan $N(\mu_y, \sigma^2)$ dengan asumsi ragam tidak diketahui. Tentukan n sedemikian sehingga

$$P\left(\bar{X} - \bar{Y} - \sigma/5 < \mu_x - \mu_y < \bar{X} - \bar{Y} + \sigma/5\right) = 0.90.$$

BAB 4

Kecukupan

4.1 Statistik Cukup

Definisi 4.1 *Misalkan* $X_1, X_2, ..., X_n$ *adalah sampel acak dari populasi* X *dengan fungsi kepekatan peluang atau fungsi massa peluang* $f(x;\theta), \theta \in \Omega$. *Andaikan* $Y = u(X_1, X_2, ..., X_n)$ *merupakan statistik dengan fungsi kepekatan peluang atau fungsi massa peluang* $g_Y(y;\theta), \theta \in \Omega$, *maka* Y *adalah statistik cukup bagi* θ *jika*

$$\frac{f(x_1, ..., x_n;\theta)}{g_Y(u(x_1, ..., x_n);\theta)} = H(x_1, ..., x_n),$$

dengan $H(x_1, ..., x_n)$ *tidak bergantung pada* $\theta \in \Omega$.

Contoh 4.1 *Misalkan* $X_1, X_2, ..., X_n$ *adalah sampel acak dari sebaran dengan* $f(x;\theta) = \theta^x (1-\theta)^{1-x} \mathbf{I}(x = 0, 1)$ *di mana* $0 < \theta < 1$. *Tunjukkan bahwa* $Y = \sum_{i=1}^n X_i$ *adalah statistik cukup bagi* θ.

Jawab.

$$\begin{aligned} f(x_1, ..., x_n;\theta) &= \prod_{i=1}^n \theta^{x_i} (1-\theta)^{1-x_i} \mathbf{I}(x_i = 0, 1) \\ &= \theta^{\Sigma_{i=1}^n x_i} (1-\theta)^{n-\Sigma_{i=1}^n x_i} \prod_{i=1}^n \mathbf{I}(x_i = 0, 1). \end{aligned}$$

Karena $X_i \sim Bernoulli(\theta)$, maka $Y = \sum_{i=1}^n X_i \sim binomial(n, \theta)$ sehingga

$$g_Y(y;\theta) = \binom{n}{y} \theta^y (1-\theta)^{n-y} \mathbf{I}(y = 0, 1, ..., n).$$

Sebaran bersyarat dari $X_1, X_2, ..., X_n$ untuk $Y = \sum_{i=1}^{n} X_i = y$ ialah

$$\begin{aligned}
\frac{f(x_1, ..., x_n; \theta)}{g_Y(y; \theta)} &= \frac{\theta^{\Sigma_{i=1}^{n} x_i} (1-\theta)^{n-\Sigma_{i=1}^{n} x_i}}{\binom{n}{\Sigma_{i=1}^{n} x_i} \theta^{\Sigma_{i=1}^{n} x_i} (1-\theta)^{n-\Sigma_{i=1}^{n} x_i}} \\
&= \frac{1}{\binom{n}{\Sigma_{i=1}^{n} x_i}} \\
&= \frac{1}{\binom{n}{y}}
\end{aligned}$$

(tidak bergantung pada parameter θ). Jadi, $Y = \sum_{i=1}^{n} X_i$ merupakan statistik cukup bagi θ. ✠

Contoh 4.2 *Misalkan $X_1, X_2, ..., X_n$ adalah sampel acak dari sebaran dengan $f(x; \theta) = e^{-(x-\theta)} \mathbf{I}(x > \theta)$. Tunjukkan bahwa $Y = X_{(1)}$ adalah statistik cukup bagi θ.*

Jawab. Fungsi kepekatan peluang bagi $Y = X_{(1)}$ ialah

$$\begin{aligned}
g_Y(y; \theta) &= n f(y; \theta) [1 - F(y)]^{n-1} \mathbf{I}(y > \theta) \\
&= n e^{-(y-\theta)} \left[1 - \left(\int_{\theta}^{y} e^{-(x-\theta)} dx\right)\right]^{n-1} \\
&= n e^{-(y-\theta)} \left[1 - \left(1 - e^{-(y-\theta)}\right)\right]^{n-1} \\
&= n e^{-(y-\theta)} \left[e^{-(y-\theta)}\right]^{n-1} \\
&= n e^{-n(y-\theta)} \mathbf{I}(y > \theta).
\end{aligned}$$

Sebaran bersyarat dari $X_1, X_2, ..., X_n$ untuk $Y = X_{(1)} = \min(X_i)$ ialah

$$\begin{aligned}
\frac{f(x_1, ..., x_n; \theta)}{g_Y(y; \theta)} &= \frac{\prod_{i=1}^{n} e^{-(x_i-\theta)}}{n e^{-n(\min(x_i)-\theta)}} \\
&= \frac{\exp\left(n\theta - \sum_{i=1}^{n} x_i\right)}{n \exp\left(-n\left(\min(x_i) - \theta\right)\right)} \\
&= \frac{\exp\left(-\sum_{i=1}^{n} x_i\right)}{n \exp\left(-n \min(x_i)\right)}
\end{aligned}$$

(tidak bergantung pada parameter θ). Jadi, $Y = X_{(1)}$ merupakan statistik cukup bagi θ. ✠

Teorema 4.1 (***Faktorisasi Neyman***) *Misalkan $X_1, X_2, ..., X_n$ adalah sampel acak dari sebaran dengan fungsi kepekatan peluang atau fungsi massa* *peluang* $f(x;\theta), \theta \in \Omega$. *Statistik* $Y_1 = u_1(X_1, X_2, ..., X_n)$ *merupakan statistik cukup bagi* θ *jika dan hanya jika terdapat dua fungsi nonnegatif* k_1 *dan* k_2 *sehingga*

$$\prod_{i=1}^{n} f(x_i;\theta) = k_1(u_1(x_1, x_2, ..., x_n);\theta) \cdot k_2(x_1, x_2, ..., x_n)$$

dengan $k_2(x_1, x_2, ..., x_n)$ *tidak bergantung pada* θ.

Bukti. ($\Rightarrow$) Jika Y_1 adalah statistik cukup, maka

$$\begin{aligned} f(x_1, x_2, ..., x_n;\theta) &= g_{Y_1}(u_1(x_1, x_2, ..., x_n);\theta) \cdot H(x_1, x_2, ..., x_n) \\ \prod_{i=1}^{n} f(x_i;\theta) &= k_1(u(x_1, x_2, ..., x_n);\theta) \cdot k_2(x_1, x_2, ..., x_n). \end{aligned}$$

($\Leftarrow$) Misalkan X merupakan peubah acak kontinu. Dalam pembuktian ini, peubah $y_1 = u_1(x_1, x_2, ..., x_n)$, $y_2 = u_2(x_1, x_2, ..., x_n), ..., y_n = u_n(x_1, x_2, ..., x_n)$ ditransformasikan satu-satu. Peubah $y_1, y_2, ...$ dan y_n memiliki invers $x_1 = w_1(y_1, y_2, ..., y_n), x_2 = w_2(y_1, y_2, ..., y_n), ..., x_n = w_n(y_1, y_2, ..., y_n)$ dan determinan matriks Jacobi J. Fungsi kepekatan peluang bagi Y_1, $Y_2, ...$ dan Y_n ialah

$$g(y_1, y_2, ..., y_n;\theta) = k_1(y_1;\theta) \cdot k_2(w_1, w_2, ..., w_n) \cdot |J|$$

di mana $w_i = w_i(y_1, y_2, ..., y_n), i = 1, 2, ..., n$. Fungsi kepekatan peluang bagi Y_1 ialah

$$\begin{aligned} g_{Y_1}(y_1;\theta) &= \int_{-\infty}^{\infty} ... \int_{-\infty}^{\infty} g(y_1, y_2, ..., y_n;\theta)\, dy_2...dy_n \\ &= k_1(y_1;\theta) \int_{-\infty}^{\infty} ... \int_{-\infty}^{\infty} k_2(w_1, w_2, ..., w_n) \cdot |J|\, dy_2...dy_n. \end{aligned}$$

Misalkan $m(y_1) = \int_{-\infty}^{\infty} ... \int_{-\infty}^{\infty} k_2(w_1, w_2, ..., w_n) \cdot |J|\, dy_2...dy_n$, maka

$$g_{Y_1}(y_1;\theta) = k_1(y_1;\theta)\, m(y_1).$$

Jika $m(y_1) = 0$, maka $g_{Y_1}(y_1;\theta) = 0$. Jika $m(y_1) > 0$, maka fungsi $k_1(y_1;\theta)$ dapat dituliskan menjadi

$$k_1(u_1(x_1, x_2, ..., x_n);\theta) = \frac{g_{Y_1}(u_1(x_1, x_2, ..., x_n);\theta)}{m(u_1(x_1, x_2, ..., x_n))}$$

dan faktorisasinya menjadi

$$\prod_{i=1}^{n} f(x_i;\theta) = g_{Y_1}(u_1(x_1, x_2, ..., x_n);\theta) \frac{k_2(x_1, x_2, ..., x_n)}{m(u_1(x_1, x_2, ..., x_n))},$$

di mana fungsi k_2 dan m tidak bergantung pada θ. Dengan demikian, berdasarkan definisinya dengan

$$H(x_1, ..., x_n) = \frac{k_2(x_1, x_2, ..., x_n)}{m(u_1(x_1, x_2, ..., x_n))},$$

maka Y_1 merupakan statistik cukup bagi θ. ✠

Contoh 4.3 *Misalkan $X_1, X_2, ..., X_n$ adalah sampel acak dari sebaran $N(\theta, \sigma^2)$, $-\infty < \theta < \infty$, di mana $\sigma^2 > 0$ diketahui. Dengan menggunakan Teorema 4.1 (Faktorisasi Neyman), tunjukkan bahwa $\bar{X}$ adalah statistik cukup bagi θ.*

Jawab. Jika $\bar{x} = \frac{\sum_{i=1}^n x_i}{n}$, maka

$$\begin{aligned}
\sum_{i=1}^n (x_i - \theta)^2 &= \sum_{i=1}^n (x_i - \bar{x} + \bar{x} - \theta)^2 \\
&= \sum_{i=1}^n \left[(x_i - \bar{x})^2 + 2(x_i - \bar{x})(\bar{x} - \theta) + (\bar{x} - \theta)^2\right] \\
&= \sum_{i=1}^n (x_i - \bar{x})^2 + n(\bar{x} - \theta)^2,
\end{aligned}$$

karena $(\bar{x} - \theta)\sum_{i=1}^n (x_i - \bar{x}) = 0$. Fungsi kepekatan peluang bersama bagi $X_1, X_2, ..., X_n$ ialah

$$\begin{aligned}
f(x_1, x_2, ..., x_n; \theta) &= \prod_{i=1}^n f(x_i; \theta) \\
&= \left(\frac{1}{\sigma\sqrt{2\pi}}\right)^n \exp\left[-\frac{\sum_{i=1}^n (x_i - \theta)^2}{2\sigma^2}\right] \\
&= \left\{\exp\left[-\frac{n(\bar{x} - \theta)^2}{2\sigma^2}\right]\right\} \times \left\{\frac{\exp\left[-\sum_{i=1}^n (x_i - \bar{x})^2 / 2\sigma^2\right]}{\left(\sigma\sqrt{2\pi}\right)^n}\right\} \\
&= k_1(\bar{x}; \theta) \times k_2(x_1, x_2, ..., x_n).
\end{aligned}$$

Karena $\prod_{i=1}^n f(x_i; \theta)$ dapat difaktorkan menjadi k_1 dan k_2 di mana k_2 tidak bergantung pada parameter θ, maka berdasarkan Teorema 4.1, $\bar{X}$ merupakan statistik cukup bagi θ. ✠

Akibat 4.1 *Jika Y adalah statistik cukup bagi θ, maka penduga kemungkinan maksimum bagi θ adalah fungsi dari Y.*

Bukti. Berdasarkan Teorema 4.1, fungsi kemungkinan $L(\theta)$ dapat difaktorisasi menjadi

$$\begin{aligned} L(\theta) &= \prod_{i=1}^{n} f(x_i, \theta) \\ &= k_1(y;\theta) \times k_2(\mathbf{x}), \end{aligned}$$

sehingga untuk memaksimumkan $L(\theta)$, hanya perlu memaksimumkan $k_1(y;\theta)$. Dengan demikian, penduga kemungkinan maksimum bagi θ merupakan fungsi dari statistik cukup Y. ✠

Teorema 4.2 (*Rao-Blackwell*) *Misalkan $X_1, X_2, ..., X_n$ adalah sampel acak dari sebaran yang memiliki fungsi kepekatan peluang atau fungsi massa peluang $f(x;\theta), \theta \in \Omega$. Misalkan $Y_1 = u_1(X_1, X_2, ..., X_n)$ merupakan statistik bagi θ dan $Y_2 = u_2(X_1, X_2, ..., X_n)$ penduga takbias bagi θ. Jika $E(Y_2|y_1) = \varphi(y_1)$, maka $E(\varphi(Y_1)) = \theta$ dan $Var(\varphi(Y_1)) \leq Var(Y_2)$.*

Bukti. Misalkan X merupakan peubah acak kontinu. Karena $E(Y_2|y_1) = \varphi(y_1)$, maka

$$\begin{aligned} E(\varphi(Y_1)) &= \int_{-\infty}^{\infty} \varphi(y_1) f_{Y_1}(y_1;\theta)\, dx \\ &= \int_{-\infty}^{\infty} \left(\int_{-\infty}^{\infty} y_2 f(y_2|y_1)\, dy_2 \right) f_{Y_1}(y_1;\theta)\, dy_1 \\ &= \int_{-\infty}^{\infty} \left(\int_{-\infty}^{\infty} y_2 \frac{f(y_1, y_2;\theta)}{f_{Y_1}(y_1;\theta)} dy_2 \right) f_{Y_1}(y_1;\theta)\, dy_1 \\ &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} y_2 f(y_1, y_2;\theta)\, dy_2 dy_1 \\ &= E(Y_2) \\ &= \theta. \end{aligned}$$

$$\begin{aligned} E[(Y_2 - \varphi(Y_1))(\varphi(Y_1) - \theta)] &= \int\int (y_2 - \varphi(y_1))(\varphi(y_1) - \theta) f(y_1, y_2;\theta)\, dy_1 dy_2 \\ &= \int\int (y_2 - \varphi(y_1))(\varphi(y_1) - \theta) f(y_2|y_1) f_{Y_1}(y_1;\theta)\, dy_1 dy_2 \\ &= \int (\varphi(y_1) - \theta) \left(\int (y_2 - \varphi(y_1)) f(y_2|y_1)\, dy_2 \right) f_{Y_1}(y_1;\theta)\, dy_1 \\ &= \int (\varphi(y_1) - \theta)(\varphi(y_1) - \varphi(y_1)) f_{Y_1}(y_1;\theta)\, dy_1 \\ &= 0 \end{aligned}$$

$$\begin{aligned}
Var\left(Y_2\right) &= E\left(Y_2-\theta\right)^2 \\
&= E\left(Y_2-\varphi\left(Y_1\right)+\varphi\left(Y_1\right)-\theta\right)^2 \\
&= E\left(Y_2-\varphi\left(Y_1\right)\right)^2+2E\left[\left(Y_2-\varphi\left(Y_1\right)\right)\left(\varphi\left(Y_1\right)-\theta\right)\right] \\
&\quad +E\left(\varphi\left(Y_1\right)-\theta\right)^2 \\
&= E\left(Y_2-\varphi\left(Y_1\right)\right)^2+E\left(\varphi\left(Y_1\right)-\theta\right)^2 \\
&= E\left(Y_2-\varphi\left(Y_1\right)\right)^2+Var\left(\varphi\left(Y_1\right)\right). \qquad (4.1)
\end{aligned}$$

Karena $E\left(Y_2-\varphi\left(Y_1\right)\right)^2 \geq 0$, maka persamaan 4.1 menjadi

$$Var\left(Y_2\right) \geq Var\left(\varphi\left(Y_1\right)\right).$$

✠

Berdasarkan Teorema 4.2, jika Y_1 adalah statistik cukup untuk suatu parameter θ, maka untuk mendapatkan penduga takbias dengan ragam minimum atau *Minimum Variance Unbiased Estimator* (MVUE), cukup dengan mencari fungsi dari statistik Y_1, $\varphi\left(Y_1\right)$, sedemikian sehingga $E\left(\varphi\left(Y_1\right)\right)=\theta$.

Contoh 4.4 *Misalkan $X_1, X_2, ..., X_n$ adalah sampel acak dengan $f\left(x;\theta\right)=\frac{1}{\theta}e^{-x/\theta}\mathbf{I}\left(x>0\right)$. Tentukan MVUE bagi θ.*

Jawab.

$$\begin{aligned}
L\left(\theta\right) &= \prod\nolimits_{i=1}^{n} f\left(x_i,\theta\right) \\
L\left(\theta\right) &= \theta^{-n}\exp\left[-\frac{\sum_{i=1}^{n}x_i}{\theta}\right]. \qquad (4.2)
\end{aligned}$$

Berdasarkan Teorema 4.1, statistik $Y_1=\sum_{i=1}^{n}X_i$ merupakan statistik cukup. Fungsi *log-likelihood* dari persamaan 4.2 adalah

$$\ln L\left(\theta\right)=-n\ln\theta-\frac{1}{\theta}\sum_{i=1}^{n}x_i,$$

sehingga diperoleh penduga kemungkinan maksimum bagi θ adalah $\bar{X}=\frac{Y_1}{n}$. Karena $X_i\sim\Gamma\left(1,\theta\right)$, maka $Y_1=\sum_{i=1}^{n}x_i\sim\Gamma\left(n,\theta\right)$.

$$E\left(\bar{X}\right)=E\left(\frac{Y_1}{n}\right)=\frac{1}{n}E\left(Y_1\right)=\frac{1}{n}\cdot n\theta=\theta,$$

sehingga $\bar{X}$ merupakan MVUE bagi θ. ✠

Misalkan $\varphi(Y_1) = E(Y_2|Y_1)$ adalah penduga takbias yang memiliki ragam lebih kecil daripada penduga takbias bagi θ, yaitu Y_2. Misalkan pula suatu fungsi dari suatu statistik Y_3, dinotasikan dengan $\Upsilon(Y_3)$ dan amatannya didefinisikan sebagai $\Upsilon(y_3) = E(\varphi(Y_1)|Y_3 = y_3)$, di mana Y_3 bukan statistik cukup. Berdasarkan Teorema 4.2 (Rao-Blackwell), $E(\Upsilon(Y_3)) = \theta$ dan $\Upsilon(Y_3)$ memiliki ragam lebih kecil daripada $\varphi(Y_1)$. Oleh karena itu, seharusnya $\Upsilon(Y_3)$ merupakan penduga takbias yang lebih baik jika dibandingkan dengan $\varphi(Y_1)$. Namun, hal ini tidak benar karena Y_3 bukan statistik cukup, sehingga sebaran bersyarat dari Y_1 jika diketahui $Y_3 = y_3$ dan nilai harapan bersyarat dari Y_3 masih mengandung θ. Jadi, meskipun $E(\Upsilon(Y_3)) = \theta$, $\Upsilon(Y_3)$ bukanlah suatu statistik karena masih bergantung pada paramater θ yang tidak diketahui. Sebagai contoh, misalkan X_1, X_2, X_3 adalah sampel acak dari sebaran eksponensial dengan nilai harapan $\theta > 0$, maka fungsi kepekatan peluang bersama dari sampel acak tersebut ialah

$$f(x_1, x_2, x_3; \theta) = \left(\frac{1}{\theta}\right)^3 \exp\left(-\frac{x_1 + x_2 + x_3}{\theta}\right) \prod_{i=1}^{3} \mathbf{I}(0 < x_i < \infty).$$

Berdasarkan Teorema 4.1, $Y_1 = X_1 + X_2 + X_3$ merupakan statistik cukup bagi θ. Nilai harapan dari Y_1 ialah

$$E(Y_1) = E(X_1 + X_2 + X_3) = 3\theta,$$

dan $Y_1/3 = \bar{X}$ ialah fungsi dari statistik cukup yang merupakan penduga takbias bagi θ.

Misalkan $Y_2 = X_2 + X_3$ dan $Y_3 = X_3$. Transformasi satu-satu $x_1 = y_1 - y_2$, $x_2 = y_2 - y_3$, $x_3 = y_3$ memiliki determinan matriks Jacobi sama dengan

$$\begin{aligned} J &= \det\left(\frac{\partial \mathbf{x}}{\partial \mathbf{y}'}\right) \\ &= \begin{vmatrix} 1 & -1 & 0 \\ 0 & 1 & -1 \\ 0 & 0 & 1 \end{vmatrix} \\ &= 1 \end{aligned}$$

dan fungsi kepekatan peluang bersama bagi Y_1, Y_2, dan Y_3 ialah

$$g(y_1, y_2, y_3; \theta) = \left(\frac{1}{\theta}\right)^3 \exp\left(-\frac{y_1}{\theta}\right) \mathbf{I}(0 < y_3 < y_2 < y_1 < \infty).$$

Fungsi kepekatan peluang marginal bagi Y_1 dan Y_3 ialah

$$\begin{aligned} g_{13}(y_1, y_3; \theta) &= \int_{y_3}^{y_1} \left(\frac{1}{\theta}\right)^3 \exp\left(-\frac{y_1}{\theta}\right) dy_2 \\ &= \left(\frac{1}{\theta}\right)^3 (y_1 - y_3) \exp\left(-\frac{y_1}{\theta}\right) \mathbf{I}(0 < y_3 < y_1 < \infty). \end{aligned}$$

Karena $Y_3 = X_3$, maka fungsi kepekatan peluang bagi Y_3 ialah

$$g_3(y_3; \theta) = \frac{1}{\theta} \exp\left(-\frac{y_3}{\theta}\right) \mathbf{I}(0 < y_3 < \infty),$$

sehingga fungsi kepekatan peluang bersyarat bagi Y_1 jika diketahui $Y_3 = y_3$ ialah

$$\begin{aligned} g_{1|3}(y_1|y_3) &= \frac{g_{13}(y_1, y_3; \theta)}{g_3(y_3; \theta)} \\ &= \left(\frac{1}{\theta}\right)^2 (y_1 - y_3) \exp\left(-\frac{1}{\theta}(y_1 - y_3)\right) \mathbf{I}(0 < y_3 < y_1 < \infty). \end{aligned}$$

$$\begin{aligned} E\left(\frac{Y_1}{3}\middle| y_3\right) &= E\left(\frac{Y_1 - Y_3}{3}\middle| y_3\right) + E\left(\frac{Y_3}{3}\middle| y_3\right) \\ &= \frac{1}{3}\int_{y_3}^{\infty} \left(\frac{1}{\theta}\right)^2 (y_1 - y_3)^2 \exp\left(-\frac{1}{\theta}(y_1 - y_3)\right) dy_1 + \frac{y_3}{3} \\ &= \frac{1}{3}\frac{\Gamma(3)\,\theta^3}{\theta^2} + \frac{y_3}{3} \\ &= \frac{2\theta}{3} + \frac{y_3}{3} \\ &= \Upsilon(y_3). \end{aligned}$$

Jelas bahwa $E(\Upsilon(Y_3)) = \theta$ dan $Var(\Upsilon(Y_3)) \leq Var(Y_1/3)$, tetapi $\Upsilon(Y_3)$ bukan statistik karena bergantung pada parameter θ dan tidak dapat digunakan sebagai penduga bagi θ.

4.2 Kelengkapan dan Kekhasan

Definisi 4.2 *Misalkan Z adalah peubah acak dengan fungsi kepekatan atau fungsi massa peluang $h(z;\theta), \theta \in \Omega$. Jika $E(u(Z)) = 0, \forall\theta \in \Omega$ memerlukan $u(Z) = 0$ kecuali pada himpunan berpeluang nol untuk setiap $h(z;\theta), \theta \in \Omega$, maka famili $\{h(z;\theta) : \theta \in \Omega\}$ disebut famili lengkap dari fungsi kepekatan atau fungsi massa peluang.*

Berikut merupakan beberapa contoh famili lengkap dan famili tak lengkap.

Contoh 4.5 *Misalkan $X_1, X_2, ..., X_n$ adalah sampel acak dari sebaran Poisson dengan parameter $\theta > 0$. Tunjukkan bahwa $Y_1 = \sum_{i=1}^{n} X_i$ adalah statistik cukup bagi θ dan himpunan fungsi massa peluang bagi Y_1 merupakan famili lengkap.*

Jawab. Karena $X_i \sim Poisson\,(\theta)$, maka

$$\begin{aligned}\prod\nolimits_{i=1}^{n} f(x_i;\theta) &= \prod\nolimits_{i=1}^{n} \frac{\theta^{x_i} e^{-\theta}}{x_i!} \mathbf{I}(x_i = 0, 1, 2, \ldots) \\ &= \frac{\theta^{\Sigma x_i} e^{-n\theta}}{\prod_{i=1}^{n}(x_i!)} \prod\nolimits_{i=1}^{n} \mathbf{I}(x_i = 0, 1, 2, \ldots) \\ &= \theta^{\Sigma x_i} e^{-n\theta} \times \frac{1}{\prod_{i=1}^{n}(x_i!)} \\ &= k_1(y_1;\theta) \times k_2(x_1, x_2, ..., x_n).\end{aligned}$$

Berdasarkan Teorema 4.1, $Y_1 = \sum_{i=1}^{n} X_i$ merupakan statistik cukup bagi θ.

Kemudian, fungsi massa peluang bagi Y_1 ialah

$$g(y_1;\theta) = \frac{(n\theta)^{y_1} e^{-n\theta}}{y_1!} \mathbf{I}(y_1 = 0, 1, 2, \ldots).$$

Misalkan $\{g(y_1;\theta) : \theta > 0\}$ merupakan famili dari fungsi massa peluang. Misalkan pula $u(Y_1)$ adalah fungsi dari Y_1 sedemikian sehingga $E(u(Y_1)) = 0$. Akan ditunjukkan bahwa bila $E(u(Y_1)) = 0$ memerlukan $u(Y_1) = 0$ kecuali pada himpunan berpeluang nol untuk setiap $g(y_1;\theta), \theta > 0$.

$$E(u(Y_1)) = 0$$
$$\sum_{y_1=0}^{\infty} u(y_1) \frac{(n\theta)^{y_1} e^{-n\theta}}{y_1!} = 0$$
$$e^{-n\theta}\left[u(0) + u(1)\frac{n\theta}{1!} + u(2)\frac{(n\theta)^2}{2!} + \cdots\right] = 0.$$

Karena $e^{-n\theta} \neq 0$, maka

$$u(0) + [nu(1)]\theta + \left[\frac{n^2 u(2)}{2}\right]\theta^2 + \cdots = 0.$$

Catatan: Jika suatu deret (pangkat) takhingga konvergen ke nol untuk setiap $\theta > 0$, maka setiap koefisiennya sama dengan nol.

Dengan demikian,

$$u(0) = 0, nu(1) = 0, \frac{n^2 u(2)}{2} = 0, \cdots .$$

Artinya, $u(0) = u(1) = u(2) = \cdots = 0$. Jadi, $\{g(y_1; \theta) : \theta > 0\}$ merupakan famili lengkap. ✠

Contoh 4.6 *Misalkan Z adalah peubah acak dari famili $\{h(z;\theta) : \theta \in \Omega\}$ di mana*

$$h(z;\theta) = \frac{1}{\theta} e^{-z/\theta} \mathbf{I}(0 < z < \infty).$$

Tunjukkan bahwa famili $\{h(z;\theta) : \theta \in \Omega\}$ merupakan famili lengkap.

Jawab. Misalkan pula $u(Z)$ adalah fungsi dari peubah acak Z sedemikian sehingga $E(u(Z)) = 0$, yaitu

$$\frac{1}{\theta} \int_0^\infty u(z) e^{-z/\theta} dz = 0, \ \theta > 0. \tag{4.3}$$

Catatan: Misalkan £ merupakan operator transformasi Laplace. Bila £f = £g, maka $f = g$ pada titik-titik kekontinuannya. Dengan kata lain,

$$\int_0^\infty f(t) e^{-st} dt = \int_0^\infty g(t) e^{-st} dt \Rightarrow f(t) = g(t).$$

Karena ruas kiri pada persamaan 4.3 merupakan transformasi Laplace bagi fungsi $u(z)$, maka $u(z) = 0$ pada titik-titik kekontinuannya. Jadi,

$$\{h(z;\theta) : \theta \in \Omega\}$$

merupakan famili lengkap. ✠

Contoh 4.7 *Tunjukkan bahwa famili yang anggotanya sebagai berikut bukan merupakan famili lengkap.*

a. $f(x;\theta) = \frac{1}{2\theta} \mathbf{I}(|x| = 1, 2, ..., \theta), \theta \in \mathbb{N}.$

b. $f(x;\theta) = \frac{1}{2\theta} \mathbf{I}(|x| < \theta).$

Jawab.

a. Karena peubah acak X memiliki fungsi massa peluang

$$f(x;\theta) = \frac{1}{2\theta}\mathbf{I}(|x| = 1, 2, ..., \theta), \theta \in \mathbb{N},$$

maka

$$\begin{aligned} E(X) &= \sum_{x=-\theta}^{\theta} x f(x;\theta) \\ &= \sum_{x=-\theta}^{\theta} x \left(\frac{1}{2\theta}\right) \\ &= 0, \end{aligned}$$

tetapi tidak mengharuskan $X = 0$ sehingga $f(x;\theta)$ bukan anggota famili lengkap.

b. Karena peubah acak X memiliki fungsi kepekatan peluang

$$f(x;\theta) = \frac{1}{2\theta}\mathbf{I}(|x| < \theta)$$

maka

$$\begin{aligned} E(X) &= \int_{-\theta}^{\theta} x f(x;\theta)\, dx \\ &= \int_{-\theta}^{\theta} x \left(\frac{1}{2\theta}\right) dx \\ &= 0, \end{aligned}$$

tetapi tidak mengharuskan $X = 0$ sehingga $f(x;\theta)$ bukan anggota famili lengkap.

✠

Misalkan $X_1, X_2, \ldots, X_n$ adalah sampel acak dengan fungsi kepekatan atau fungsi massa peluang $f(x;\theta)$ dan $Y_1 = u(X_1, X_2, \ldots, X_n)$ adalah statistik cukup bagi θ. Misalkan pula $\varphi(Y_1)$ dan $\psi(Y_1)$ masing-masing adalah fungsi dari statistik cukup Y_1 (tidak bergantung pada θ) sedemikian sehingga $E(\varphi(Y_1)) = \theta$ dan $E(\psi(Y_1)) = \theta$, $\forall\theta \in \Omega$. Nilai harapan dari selisih kedua fungsi tersebut ialah

$$E(\varphi(Y_1) - \psi(Y_1)) = 0, \theta \in \Omega.$$

Jika famili $\{f_{Y_1}(y_1;\theta) : \theta \in \Omega\}$ merupakan famili lengkap, maka $\varphi(y_1) - \psi(y_1) = 0$ kecuali pada himpunan yang berpeluang nol. Dengan kata lain, untuk setiap penduga takbias $\psi(Y_1)$ bagi θ,

$$\varphi(y_1) = \psi(y_1),$$

kecuali pada himpunan yang berpeluang nol. Dengan demikian, $\varphi(Y_1)$ merupakan fungsi dari statistik cukup Y_1 yang khas dan memiliki ragam minimum jika dibandingkan dengan penduga takbias lainnya (Teorema 4.2). Hal tersebut dinyatakan dalam teorema berikut.

Teorema 4.3 (*Lehmann dan Scheffé*) *Misalkan $X_1, X_2, \ldots, X_n$ adalah sampel acak dengan fungsi kepekatan atau fungsi massa peluang $f(x;\theta)$. Misalkan pula $Y_1 = u(X_1, X_2, \ldots, X_n)$ adalah statistik cukup bagi θ dan famili $\{f_{Y_1}(y_1;\theta) : \theta \in \Omega\}$ merupakan famili lengkap. Jika terdapat fungsi Y_1 yang merupakan penduga takbias bagi θ, maka fungsi Y_1 merupakan penduga takbias dengan ragam minimum yang khas atau Unique Minimum Variance Unbiased Estimator ($UMVUE$).*

4.3 Kelas Eksponen

Definisi 4.3 *Suatu fungsi kepekatan atau fungsi massa peluang*

$$f(x;\theta) = \exp\left[p(\theta)K(x) + S(x) + q(\theta)\right] \mathbf{I}(x \in W), \gamma < \theta < \delta \qquad (4.4)$$

disebut sebagai anggota kelas eksponen biasa fungsi kepekatan peluang kontinu atau fungsi massa peluang diskret jika

1. *W tak bergantung pada parameter θ.*
2. *$p(\theta)$ fungsi kontinu nontrivial dari θ.*
3. *a. Jika X peubah acak **kontinu**, maka setiap $K'(x) \neq 0$ dan $S(x)$ fungsi kontinu dari x, $x \in W$.*
   *b. Jika X peubah acak **diskret**, maka $K(x)$ adalah fungsi nontrivial dari $x \in W$.*

Contoh 4.8 *Misalkan $X \sim N(0,\theta)$. Apakah anggota dari famili*

$$\{f(x;\theta) :\ 0 < \theta < \infty\}$$

merupakan anggota kelas eksponen biasa fungsi kepekatan peluang kontinu?

Jawab.

$$\begin{aligned} f(x;\theta) &= \frac{1}{\sqrt{2\pi\theta}} e^{-x^2/2\theta} \\ &= \exp\left[\ln f(x;\theta)\right] \\ &= \exp\left[-\frac{1}{2\theta}x^2 - \ln\sqrt{2\pi\theta}\right]\mathbf{I}(-\infty < x < \infty). \end{aligned} \tag{4.5}$$

Berdasarkan persamaan 4.5, $W = (-\infty, \infty)$ tidak bergantung pada parameter θ,

$$\begin{aligned} f(x;\theta) &= \exp\left[p(\theta)K(x) + S(x) + q(\theta)\right] \\ &= \exp\left[-\frac{1}{2\theta}x^2 - \ln\sqrt{2\pi\theta}\right] \end{aligned}$$

$$p(\theta) = -\frac{1}{2\theta},\ K(x) = x^2,\ S(x) = 0,\ q(\theta) = -\ln\sqrt{2\pi\theta},$$

$K'(x) = 2x \neq 0$ dan $S(x)$ merupakan fungsi kontinu pada $x \in (-\infty, \infty)$. Jadi, $N(0, \theta)$ anggota kelas eksponen biasa fungsi kepekatan peluang kontinu. ✠

Contoh 4.9 *Misalkan $X \sim Poisson(\theta)$. Apakah anggota dari famili*

$$\{f(x;\theta) : 0 < \theta < \infty\}$$

merupakan anggota kelas eksponen biasa fungsi massa peluang diskret?

Jawab.

$$\begin{aligned} f(x;\theta) &= e^{-\theta}\frac{\theta^x}{x!}\mathbf{I}(x = 0, 1, 2, \ldots) \\ &= \exp\left[\ln f(x;\theta)\right]\mathbf{I}(x = 0, 1, 2, \ldots) \\ &= \exp\left[(\ln\theta)x + \ln\left(\frac{1}{x!}\right) - \theta\right]\mathbf{I}(x = 0, 1, 2, \ldots) \end{aligned} \tag{4.6}$$

Berdasarkan persamaan 4.6, $W = \{0, 1, 2, \ldots\}$ tidak bergantung pada parameter θ,

$$\begin{aligned} f(x;\theta) &= \exp\left[p(\theta)K(x) + S(x) + q(\theta)\right] \\ &= \exp\left[(\ln\theta)x + \ln\left(\frac{1}{x!}\right) - \theta\right]\mathbf{I}(x = 0, 1, 2, \ldots) \end{aligned}$$

$$p(\theta) = \ln\theta,\ K(x) = x,\ S(x) = \ln\left(\frac{1}{x!}\right),\ q(\theta) = -\theta,$$

dan $K(x) = x$ merupakan fungsi nontrivial dari $x \in W$. Jadi,

$$f(x;\theta) = e^{-\theta}\frac{\theta^x}{x!}\mathbf{I}(x = 0, 1, 2, \ldots)$$

anggota kelas eksponen biasa fungsi massa peluang diskret. ✠

Misalkan $X \sim$ *seragam* $(0, \theta)$, maka $f(x;\theta) = \frac{1}{\theta}\mathbf{I}(0 < x < \theta)$ juga dapat dituliskan seperti persamaan 4.6. Namun, perhatikan bahwa interval $(0, \theta)$ bergantung pada parameter θ sehingga fungsi kepekatan peluang sebaran *seragam* $(0, \theta)$ bukan anggota kelas eksponen biasa fungsi kepekatan peluang kontinu.

Misalkan $X_1, X_2, \ldots, X_n$ adalah sampel acak dari sebaran yang merupakan anggota kelas eksponen biasa. Fungsi kepekatan atau fungsi massa peluang bersama dari $X_1, X_2, \ldots, X_n$ ialah

$$\begin{aligned} \prod_{i=1}^{n} f(x_i;\theta) &= \prod_{i=1}^{n} \exp\left[\ln f(x_i;\theta)\right] \\ &= \exp\left[p(\theta)\sum_{i=1}^{n} K(x_i) + \sum_{i=1}^{n} S(x_i) + nq(\theta)\right] \\ &= \exp\left[p(\theta)\sum_{i=1}^{n} K(x_i) + nq(\theta)\right]\exp\left[\sum_{i=1}^{n} S(x_i)\right]\prod_{i=1}^{n}\mathbf{I}(x_i \in W). \end{aligned}$$

Berdasarkan Teorema 4.1, $Y_1 = \sum_{i=1}^{n} K(x_i)$ merupakan statistik cukup bagi parameter θ. Selanjutnya, teorema berikut ini menyatakan kelengkapan dari statistik cukup Y_1 tersebut.

Teorema 4.4 *Misalkan $f(x;\theta)$, $\gamma < \theta < \delta$, adalah fungsi kepekatan atau fungsi massa peluang dari peubah acak X yang merupakan anggota kelas eksponen biasa. Jika $X_1, X_2, \ldots, X_n$ adalah contoh acak dari sebaran X, maka*

$$Y_1 = \sum_{i=1}^{n} K(x_i)$$

merupakan statistik cukup bagi θ dan famili $\{f_{Y_1}(y_1;\theta) : \gamma < \theta < \delta\}$ dari fungsi kepekatan atau fungsi massa peluang Y_1 adalah famili lengkap. Dengan kata lain, Y_1 adalah statistik cukup lengkap.

Contoh 4.10 *Misalkan $X_1, X_2, \ldots, X_n$ adalah sampel acak dari sebaran normal $N(\theta, \sigma^2)$, $-\infty < \theta < \infty$, $\sigma^2 > 0$. Tunjukkan bahwa $Y_1 = \sum_{i=1}^{n} X_i$ merupakan statistik cukup lengkap dan Y_1/n merupakan UMVUE bagi θ.*

Jawab. Karena $X \sim N\left(\theta, \sigma^2\right)$, maka fungsi kepekatan peluangnya ialah

$$f\left(x;\theta\right) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left[-\frac{\left(x-\theta\right)^2}{2\sigma^2}\right] \mathbf{I}\left(-\infty < x < \infty\right)$$

atau

$$f\left(x;\theta\right) = \exp\left[\frac{\theta}{\sigma^2}x - \frac{x^2}{2\sigma^2} - \ln\sqrt{2\pi\sigma^2} - \frac{\theta^2}{2\sigma^2}\right] \mathbf{I}\left(-\infty < x < \infty\right),$$

sehingga

$$p\left(\theta\right) = \frac{\theta}{\sigma^2},\ K\left(x\right) = x,\ S\left(x\right) = -\frac{x^2}{2\sigma^2} - \ln\sqrt{2\pi\sigma^2},\ q\left(\theta\right) = -\frac{\theta^2}{2\sigma^2}.$$

Berdasarkan Teorema 4.4, $Y_1 = \sum_{i=1}^{n} X_i$ merupakan statistik cukup lengkap. Karena $E\left(Y_1\right) = n\theta$, maka $\varphi\left(Y_1\right) = Y_1/n = \bar{X}$ adalah penduga takbias khas bagi θ dan fungsi dari statistik cukup Y_1 yang memiliki ragam minimum. Dengan demikian, $\bar{X}$ merupakan UMVUE bagi θ. $\maltese$

Contoh 4.11 *Misalkan* $X_1, X_2, \ldots, X_n$ *adalah sampel acak dari sebaran binomial dengan parameter* $\left(1, \theta\right)$, $0 < \theta < 1$. *Tentukan UMVUE bagi* θ *dan UMVUE bagi* $Var\left(\bar{X}\right)$.

Jawab.

$$\begin{aligned} f\left(x;\theta\right) &= \theta^x\left(1-\theta\right)^{1-x} \mathbf{I}\left(x = 0, 1\right) \\ &= \left(\frac{\theta}{1-\theta}\right)^x \left(1-\theta\right) \mathbf{I}\left(x = 0, 1\right) \\ &= \exp\left[x \ln\frac{\theta}{1-\theta} + \ln\left(1-\theta\right)\right] \mathbf{I}\left(x = 0, 1\right) \end{aligned}$$

$$p\left(\theta\right) = \ln\frac{\theta}{1-\theta},\ K\left(x\right) = x,\ S\left(x\right) = 0,\ q\left(\theta\right) = \ln\left(1-\theta\right).$$

Karena $K\left(x\right) = x$ adalah fungsi nontrivial dari $x \in \{0, 1\}$, maka $b\left(1, \theta\right)$ anggota kelas eksponen biasa fungsi massa peluang diskret. Dengan demikian, $Y = \sum_{i=1}^{n} X_i$ merupakan statistik cukup lengkap bagi θ dan $Y \sim b\left(n, \theta\right)$, sehingga nilai harapan dan ragam dari Y ialah

$$\begin{aligned} E\left(Y\right) &= n\theta \text{ dan } Var\left(Y\right) = n\theta\left(1-\theta\right), \\ E\left(\frac{Y}{n}\right) &= \theta \text{ dan } Var\left(\frac{Y}{n}\right) = \frac{\theta}{n}\left(1-\theta\right). \end{aligned}$$

Jadi, Y/n adalah UMVUE bagi θ. Selanjutnya ingin ditentukan UMVUE bagi $Var\left(\frac{Y}{n}\right) = \frac{\theta}{n}(1-\theta)$, di mana $\frac{\theta}{n}(1-\theta)$ merupakan fungsi dari parameter θ.

Misalkan $\delta = \dfrac{\theta}{n}(1-\theta)$, maka dugaan kemungkinan maksimum bagi δ diberikan sebagai fungsi dari statistik cukup, yaitu

$$\begin{aligned}
\bar{\delta} &= \frac{1}{n}\left[\frac{Y}{n}\left(1-\frac{Y}{n}\right)\right]. \\
E\left(\bar{\delta}\right) &= E\left[\frac{1}{n}\left\{\frac{Y}{n}\left(1-\frac{Y}{n}\right)\right\}\right] \\
&= \frac{1}{n}\left[E\left(\frac{Y}{n}\right) - E\left(\frac{Y^2}{n^2}\right)\right].
\end{aligned}$$

Karena $E\left(\dfrac{Y}{n}\right) = \theta$ dan $E\left(\dfrac{Y^2}{n^2}\right) = \dfrac{\theta}{n}(1-\theta) + \theta^2$, maka

$$E\left[\frac{1}{n}\left\{\frac{Y}{n}\left(1-\frac{Y}{n}\right)\right\}\right] = \frac{(n-1)}{n} \times \frac{\theta}{n}(1-\theta),$$

sehingga

$$\begin{aligned}
E\left[\frac{n}{n-1} \times \frac{1}{n}\left\{\frac{Y}{n}\left(1-\frac{Y}{n}\right)\right\}\right] &= \frac{n}{n-1} \times E\left[\frac{1}{n}\left\{\frac{Y}{n}\left(1-\frac{Y}{n}\right)\right\}\right] \\
&= \frac{\theta}{n}(1-\theta).
\end{aligned}$$

Dengan kata lain, statistik

$$\hat{\delta} = \frac{n}{n-1} \times \frac{1}{n}\left[\frac{Y}{n}\left(1-\frac{Y}{n}\right)\right] = \frac{n}{n-1}\bar{\delta}$$

adalah UMVUE bagi δ (ragam dari Y/n) . ✠

Contoh 4.12 *Misalkan $X_1, X_2, \ldots, X_n$ adalah sampel acak dari sebaran normal $N(\theta, 1)$. Tentukan UMVUE dari fungsi θ yang didefinisikan sebagai berikut*

$$P(X \le c) = \int_{-\infty}^{c} \frac{1}{\sqrt{2\pi}} e^{-(x-\theta)^2/2} dx = \Phi(c-\theta),$$

di mana c adalah suatu konstanta.

Jawab. Misalkan $u(X_1)$ adalah fungsi dari X_1. Kemudian, hitung nilai harapan bersyarat dari statistik takbias jika diketahui statistik cukup $\bar{X}$, yaitu

$$E\left(u(X_1)\,|\bar{X} = \bar{x}\right) = \varphi(\bar{x}).$$

Berdasarkan Teorema Rao-Blackwell dan Lehmann-Scheffé, $\varphi\left(\bar{X}\right)$ adalah UMVUE bagi $\Phi\left(c-\theta\right)$.

Misalkan fungsi $u\left(x_1\right) = \mathbf{I}\left(-\infty < x_1 < c\right)$, maka nilai harapan dari

$$E\left(u\left(X_1\right)\right) = 1 \cdot P\left(X_1 - \theta \leq c - \theta\right) = \Phi\left(c-\theta\right).$$

Dengan demikian, $u\left(X_1\right)$ adalah penduga takbias takbias bagi $\Phi\left(c-\theta\right)$. Selanjutnya akan ditentukan fungsi kepekatan peluang bersama bagi X_1 dan $\bar{X}$ serta fungsi kepekatan peluang bersyarat dari X_1 jika diketahui $\bar{X}$. Dari fungsi kepekatan peluang bersyarat tersebut, dapat ditentukan nilai harapan bersyarat $E\left(u\left(X_1\right)|\bar{X}=\bar{x}\right) = \varphi\left(\bar{x}\right)$. Fungsi kepekatan peluang bersama bagi X_1 dan $\bar{X}$ adalah normal bivariat dengan vektor nilai harapan (θ, θ) dan ragam $\sigma_1^2 = 1$, $\sigma_2^2 = 1/n$, serta koefisien korelasi $\rho = 1/\sqrt{n}$, sehingga fungsi kepekatan peluang bersyarat dari X_1 jika diketahui $\bar{X} = \bar{x}$ adalah normal dengan nilai harapan

$$\theta + \frac{\rho\sigma_1}{\sigma_2}\left(\bar{x}-\theta\right) = \bar{x}$$

dan ragam

$$\sigma_1^2\left(1-\rho^2\right) = \frac{n-1}{n}.$$

Jadi, nilai harapan bersyarat dari $u\left(X_1\right)$ jika diketahui statistik cukup $\bar{X} = \bar{x}$ ialah

$$\begin{aligned} \varphi\left(\bar{x}\right) &= \int_{-\infty}^{\infty} u\left(x_1\right)\sqrt{\frac{n-1}{n}}\frac{1}{\sqrt{2\pi}}\exp\left[-\frac{n\left(x_1-\bar{x}\right)}{2\left(n-1\right)}\right]dx_1 \\ &= \int_{-\infty}^{c}\sqrt{\frac{n-1}{n}}\frac{1}{\sqrt{2\pi}}\exp\left[-\frac{n\left(x_1-\bar{x}\right)}{2\left(n-1\right)}\right]dx_1. \end{aligned}$$

Dengan pemisalan peubah $z = \sqrt{n}\left(x_1-\bar{x}\right)/\sqrt{n-1}$, maka nilai harapan bersyarat tersebut menjadi

$$\varphi\left(\bar{x}\right) = \int_{-\infty}^{c'}\frac{1}{\sqrt{2\pi}}e^{-z^2/2}dz = \Phi\left(c'\right),$$

di mana $c' = \sqrt{n}\left(c-\bar{x}\right)/\sqrt{n-1}$. Dengan demikian, UMVUE bagi $\Phi\left(c-\theta\right)$ untuk setiap konstanta c ialah

$$\varphi\left(\bar{X}\right) = \Phi\left(\frac{\sqrt{n}\left(c-\bar{X}\right)}{\sqrt{n-1}}\right).$$

✠

4.4 Latihan

1. Misalkan $X_1, X_2, \ldots, X_n$ adalah sampel acak dari sebaran dengan fungsi kepekatan peluang

$$f\left(x;\theta\right) = \frac{\theta}{\left(1+x\right)^{\theta+1}}\ \mathbf{I}\left(x \geq 0\right),\ \theta > 0.$$

Tentukan statistik cukup bagi parameter θ.

2. Misalkan $X_1, X_2, \ldots, X_n$ adalah sampel acak dari sebaran dengan fungsi kepekatan peluang $f\left(x;\theta\right) = \theta e^{-\theta x}\ \mathbf{I}(x > 0)\,,\ \theta > 0$. $Y = \sum_{i=1}^{n} X_i$ merupakan statistik cukup bagi θ. Tunjukkan bahwa $\left(n-1\right)/Y$ adalah MVUE bagi θ.

3. Misalkan $\{h\left(x;\theta\right) : \theta \in \Omega\}$ merupakan famili dari fungsi kepekatan peluang, di mana $h\left(x;\theta\right) = \frac{1}{\theta}\mathbf{I}\left(x > 0\right)$. Tunjukkan bahwa famili tersebut tidak lengkap jika $\Omega = \{\theta : \theta > 1\}$.

4. Misalkan $X_1, X_2, \ldots, X_n$ adalah sampel acak dari sebaran dengan fungsi kepekatan peluang $f\left(x;\theta\right) = \theta^2 x e^{-\theta x}\ \mathbf{I}(0 < x < \infty)\,,\ \theta > 0$.

   (a) Tunjukkan bahwa $Y = \sum_{i=1}^{n} X_i$ merupakan statistik cukup lengkap bagi θ.

   (b) Hitung $E\left(1/Y\right)$ dan carilah fungsi dari Y yang merupakan UMVUE bagi θ.

5. Andaikan peubah acak X dengan fungsi kepekatan peluang $f(x;\theta) = B(\theta)h(x)\exp[Q(\theta)R(x)]\mathbf{I}\left(a < x < b\right)$ merupakan anggota kelas eksponen biasa. Tunjukkanlah bahwa

$$E[R(X)] = -B'\left(\theta\right)/[(B\left(\theta\right)Q'\left(\theta\right)].$$

BAB 5

Pendugaan Bayes

5.1 Prinsip Minimax

Misalkan $X_1, X_2, \ldots, X_n$ merupakan sampel acak dari sebaran yang memiliki fungsi kepekatan atau fungsi massa peluang $f(x;\theta)$, $\theta \in \Omega$. Misalkan pula $Y = U(X_1, X_2, \ldots, X_n)$ merupakan statistik yang digunakan untuk menduga parameter θ dan $\delta(y)$ merupakan fungsi nilai amatan dari statistik Y yang menjadi dugaan titik bagi parameter θ. Fungsi δ disebut juga sebagai **fungsi keputusan**. Suatu nilai dari fungsi keputusan, sebut saja $\delta(y)$, disebut sebagai putusan. Putusan tersebut dapat benar atau salah. Hal tersebut dapat menjadi ukuran dari perbedaan, jika ada, antara nilai sebenarnya dari θ dengan dugaan titik $\delta(y)$. Oleh karena itu, definisikan $\pounds[\theta, \delta(y)]$ sebagai **fungsi kehilangan** (*loss function*). Nilai harapan dari fungsi kehilangan disebut sebagai **fungsi risiko** (*risk function*), dinotasikan dengan $R(\theta, \delta)$. Jika Y merupakan peubah acak kontinu dengan $f_Y(y;\theta)$, $\theta \in \Omega$, merupakan fungsi kepekatan peluang dari Y, maka fungsi risiko dapat dihitung sebagai berikut

$$R(\theta, \delta) = E(\pounds[\theta, \delta(Y)]) = \int_{-\infty}^{\infty} \pounds[\theta, \delta(y)]\, f_Y(y;\theta)\, dy.$$

Contoh 5.1 *Misalkan $X_1, X_2, \ldots, X_{25}$ merupakan sampel acak dari sebaran $N(\theta, 1)$, $-\infty < \theta < \infty$. Misalkan $Y = \bar{X}$, nilai tengah dari sampel acak, dan misalkan $\pounds[\theta, \delta(y)] = [\theta - \delta(y)]^2$. Bandingkanlah dua fungsi keputusan berikut:*

- $\delta_1(y) = y$,
- $\delta_2(y) = 0$.

Jawab. Fungsi risiko untuk masing-masing $\delta_1(y)$ dan $\delta_2(y)$ ialah

$$\begin{aligned} R(\theta, \delta_1) &= E\left((\theta - Y)^2\right) = \frac{1}{25} \\ R(\theta, \delta_2) &= E\left((\theta - 0)^2\right) = \theta^2. \end{aligned}$$

Jelas bahwa jika pada kenyataannya $\theta = 0$, maka $\delta_2(y) = 0$ adalah keputusan yang tepat dan $R(0, \delta_2) = 0$. Tetapi, jika $\theta \neq 0$, maka $\delta_2(y) = 0$ adalah keputusan yang kurang tepat. Sebagai contoh, misalkan pada kenyataannya $\theta = 2$, maka $R(2, \delta_2) = 4 > R(2, \delta_1) = \frac{1}{25}$. Secara umum, $R(\theta, \delta_2) < R(\theta, \delta_1)$ jika $-\frac{1}{5} < \theta < \frac{1}{5}$ dan untuk θ selainnya $R(\theta, \delta_2) \geq R(\theta, \delta_1)$. Dengan demikian, kadang kala suatu fungsi keputusan lebih baik daripada fungsi keputusan lainnya untuk beberapa nilai θ dan fungsi keputusan lainnya lebih baik untuk nilai θ lainnya, sehingga jika kita membatasi fungsi keputusan δ, seperti $E(\delta(Y)) = \theta$ untuk setiap nilai θ, $\theta \in \Omega$, maka $\delta_2(y) = 0$ tidak termasuk ke dalam pertimbangan sebagai fungsi keputusan. Berdasarkan hal tersebut, fungsi risiko merupakan ragam dari penduga takbias $\delta(Y)$.

Misalkan kita tidak ingin membatasi sendiri fungsi keputusan δ, seperti $E(\delta(Y)) = \theta$ untuk setiap nilai θ, $\theta \in \Omega$. Misalkan kita katakan bahwa fungsi keputusan yang meminimumkan fungsi risiko yang maksimum merupakan fungsi keputusan terbaik. Karena pada contoh ini $R(\theta, \delta_2) = \theta^2$ tak terbatas, maka berdasarkan kriteria tersebut, $\delta_2(y) = 0$ bukanlah fungsi keputusan terbaik. Dengan demikian, kita peroleh

$$\max_{\theta} R(\theta, \delta_1) = \max_{\theta} \left(\frac{1}{25}\right) = \frac{1}{25}.$$

Oleh karena itu, $\delta_1(y) = y = \bar{x}$ merupakan fungsi keputusan terbaik berdasarkan kriteria minimax karena $\frac{1}{25}$ nilainya terkecil. $\maltese$

Berdasarkan Contoh 5.1, dapat disimpulkan bahwa

- tanpa pemberian batasan pada fungsi keputusan, sulit untuk mencari fungsi keputusan yang memiliki fungsi risiko yang secara umum lebih kecil daripada fungsi risiko dari putusan yang lain,

- suatu prinsip mengenai pemilihan fungsi keputusan terbaik disebut sebagai prinsip minimax. Prinsip tersebut dapat dinyatakan sebagai berikut:

  Misalkan diberikan fungsi keputusan $\delta_0(y)$, untuk setiap $\theta \in \Omega$,

$$\max_{\theta} R(\theta, \delta_0(y)) \leq \max_{\theta} R(\theta, \delta(y))$$

untuk setiap fungsi keputusan $\delta(y)$, maka $\delta(y)$ disebut sebagai fungsi keputusan minimax (*minimax decision function*).

Dengan pemberian batasan bahwa $E(\delta(Y)) = \theta$ dan fungsi kehilangan $\pounds[\theta, \delta(y)] = [\theta - \delta(y)]^2$, fungsi keputusan yang meminimumkan fungsi risiko menghasilkan penduga takbias dengan ragam minimum.

5.2 Sebaran Prior dan Posterior

Definisi 5.1 *Misalkan $X_1, X_2, \ldots, X_n$ adalah sampel acak dengan sebaran $f(x|\theta)$, di mana θ adalah parameter yang tidak diketahui. Fungsi kepekatan peluang dari peubah acak Θ disebut sebaran prior bagi Θ, dinotasikan dengan $h(\theta)$.*

Definisi 5.2 *Misalkan $X_1, X_2, \ldots, X_n$ adalah sampel acak dengan sebaran $f(x|\theta)$, di mana θ adalah parameter yang tidak diketahui. Fungsi kepekatan peluang bersyarat dari peubah acak Θ jika diberikan sampel $x_1, x_2, \ldots, x_n$ disebut sebaran posterior bagi Θ, dinotasikan dengan $k(\theta|x_1, x_2, \ldots, x_n)$.*

Catatan 5.1 *Jika $X_1, X_2, \ldots, X_n$ adalah sampel acak dari populasi yang memiliki fungsi kepekatan peluang $f(x|\theta)$, maka fungsi kepekatan peluang bersama antara sampel dan parameter ialah*

$$g(x_1, \ldots, x_n, \theta) = L(x_1, \ldots, x_n|\theta)\, h(\theta)$$

dengan $L(x_1, \ldots, x_n|\theta) = \prod_{i=1}^{n} f(x_i|\theta)$. Fungsi kepekatan peluang marginal bagi $\mathbf{X}$ ialah

$$g_1(x_1, \ldots, x_n) = \int_{-\infty}^{\infty} g(x_1, \ldots, x_n, \theta)\, d\theta.$$

Selanjutnya, dengan menggunakan formula Bayes diperoleh

$$\begin{aligned} k(\theta|x_1, x_2, \ldots, x_n) &= \frac{g(x_1, \ldots, x_n, \theta)}{g_1(x_1, \ldots, x_n)} \\ &= \frac{L(x_1, \ldots, x_n|\theta)\, h(\theta)}{g_1(x_1, \ldots, x_n)}. \end{aligned}$$

Contoh 5.2 *Misalkan $X_i|\theta \overset{bsi}{\sim} Poisson(\theta)$ dan $\Theta \sim \Gamma(\alpha, \beta)$, α dan β diketahui. Tentukan sebaran posterior bagi θ.*

Jawab. Misalkan $\mathbf{X}' = (X_1, X_2, \ldots, X_n)$ dan $\mathbf{x}' = (x_1, x_2, \ldots, x_n)$.

$$L(\mathbf{x}|\theta) = \prod_{i=1}^{n} \left(\frac{\theta^{x_i} e^{-\theta}}{x_i!} \right) \mathbf{I}(x_i = 0, 1, 2, \ldots), \ \theta > 0$$

dan sebaran prior bagi Θ ialah

$$h(\theta) = \frac{\theta^{\alpha-1} e^{-\theta/\beta}}{\Gamma(\alpha)\beta^{\alpha}} \mathbf{I}(\theta > 0).$$

879 Fungsi kepekatan peluang bersama antara $\mathbf{x}$ dan θ ialah

$$\begin{aligned} g(\mathbf{x}, \theta) &= L(\mathbf{x}|\theta) h(\theta) \\ &= \frac{\theta^{\alpha-1} e^{-\theta/\beta}}{\Gamma(\alpha)\beta^{\alpha}} \mathbf{I}(\theta > 0) \prod_{i=1}^{n} \left(\frac{\theta^{x_i} e^{-\theta}}{x_i!} \right) \mathbf{I}(x_i = 0, 1, 2, \ldots) \end{aligned}$$

880 sehingga fungsi kepekatan peluang marginal bagi $\mathbf{x}$ ialah

$$\begin{aligned} g_1(\mathbf{x}) &= \int_0^{\infty} \frac{\theta^{\Sigma x_i + \alpha - 1} \exp\left(-\frac{\theta}{\beta/(n\beta+1)} \right)}{x_1! \cdots x_n! \Gamma(\alpha) \beta^{\alpha}} d\theta \\ &= \frac{\Gamma\left(\sum_{i=1}^{n} x_i + \alpha\right)}{x_1! \cdots x_n! \Gamma(\alpha) \beta^{\alpha} (n + 1/\beta)^{\Sigma x_i + \alpha}} \prod_{i=1}^{n} \mathbf{I}(x_i = 0, 1, 2, \ldots). \end{aligned}$$

Dengan demikian, sebaran posterior bagi Θ ialah

$$k(\theta|\mathbf{x}) = \frac{L(\mathbf{x}|\theta) h(\theta)}{g_1(\mathbf{x})} = \frac{\theta^{\Sigma x_i + \alpha - 1} \exp\left(-\frac{\theta}{\beta/(n\beta+1)} \right)}{\Gamma\left(\sum_{i=1}^{n} x_i + \alpha\right) [\beta/(n\beta + 1)]^{\Sigma x_i + \alpha}} \mathbf{I}(\theta > 0).$$

881 Sebaran posterior bagi Θ tersebut merupakan sebaran $\Gamma\left(\sum_{i=1}^{n} x_i + \alpha, \beta/(n\beta + 1)\right)$.

882 ✠

Pada Contoh 5.2 terlihat bahwa tidak perlu menentukan fungsi kepekatan peluang marginal $g_1(\mathbf{x})$ untuk mencari sebaran posterior $k(\theta|\mathbf{x})$. Jika kita membagi $L(\mathbf{x}|\theta) h(\theta)$ dengan $g_1(\mathbf{x})$, maka diperoleh faktor yang bergantung pada amatan peubah, $\mathbf{x}$, tetapi tidak bergantung pada θ, misalkan disebut $c(\mathbf{x})$, dan $\theta^{\Sigma x_i + \alpha - 1} e^{-\theta[\beta/(n\beta+1)]}$, yaitu

$$k(\theta|\mathbf{x}) = c(\mathbf{x}) \theta^{\Sigma x_i + \alpha - 1} \exp\left(-\frac{\theta}{\beta/(n\beta + 1)} \right) \mathbf{I}(\theta > 0)$$

dengan $x_i = 0, 1, 2, \ldots$ dan $i = 1, 2, \ldots, n$. Faktor $c(\mathbf{x})$ tersebut harus berupa konstanta sedemikian sehingga membuat $k(\theta|\mathbf{x})$ sebagai suatu fungsi kepekatan peluang, yaitu

$$c(\mathbf{x}) = \frac{1}{\Gamma\left(\sum_{i=1}^{n} x_i + \alpha\right) [\beta/(n\beta + 1)]^{\Sigma x_i + \alpha}}.$$

Oleh karena itu, seringkali $k(\theta|\mathbf{x})$ dituliskan proporsional dengan $L(\mathbf{x}|\theta)\, h(\theta)$, yaitu

$$k(\theta|\mathbf{x}) \propto L(\mathbf{x}|\theta)\, h(\theta). \tag{5.1}$$

Perhatikan bahwa ruas kanan pada 5.1, faktor yang bergantung pada peubah x dan berupa konstanta dapat diabaikan. Sebagai ilustrasi, untuk menyelesaikan soal pada Contoh 5.2, $k(\theta|\mathbf{x})$ dapat dituliskan proporsional dengan $\theta^{\Sigma x_i+\alpha-1} e^{-\theta[\beta/(n\beta+1)]}$, yaitu

$$k(\theta|\mathbf{x}) \propto \theta^{\Sigma x_i+\alpha-1} \exp\left(-\frac{\theta}{\beta/(n\beta+1)}\right) \mathbf{I}(\theta > 0). \tag{5.2}$$

Selanjutnya, tentukan konstanta $c(\mathbf{x})$ sedemikian sehingga membuat $k(\theta|\mathbf{x})$ sebagai suatu fungsi kepekatan peluang. Jelas bahwa $k(\theta|\mathbf{x})$ harus berupa fungsi kepekatan peluang dari sebaran $\Gamma\left(\sum_{i=1}^{n} x_i + \alpha, \beta/(n\beta+1)\right)$, sehingga

$$c(\mathbf{x}) = \frac{1}{\Gamma\left(\sum_{i=1}^{n} x_i + \alpha\right)[\beta/(n\beta+1)]^{\Sigma x_i+\alpha}}.$$

Misalkan terdapat statistik cukup $Y = u(\mathbf{X})$ untuk parameter θ sehingga

$$L(\mathbf{x}|\theta) = g[u(\mathbf{x})|\theta]\, H(\mathbf{x}),$$

di mana $g(y|\theta)$ adalah fungsi kepekatan peluang dari Y, jika diketahui $\Theta = \theta$. Karena faktor $H(\mathbf{x})$ tidak bergantung pada θ, maka dapat kita abaikan, sehingga

$$k(\theta|\mathbf{x}) \propto g[u(\mathbf{x})|\theta]\, h(\theta).$$

Dengan demikian, jika statistik Y untuk parameter θ ada, maka kita dapat memulai dengan fungsi kepekatan peluang Y dan menuliskan

$$k(\theta|y) \propto g[y|\theta]\, h(\theta), \tag{5.3}$$

di mana $k(\theta|y)$ adalah fungsi kepekatan peluang bersyarat dari $\Theta|Y = y$. Dalam kasus statistik cukup Y, kita juga akan menggunakan $g_1(y)$, fungsi kepekatan peluang marginal dari Y, yaitu

$$g_1(y) = \int_{-\infty}^{\infty} g[y|\theta]\, h(\theta)\, d\theta.$$

5.3 Metode Pendugaan Bayes

Dugaan Bayes adalah fungsi keputusan δ yang meminimumkan

$$E\{£[\Theta, \delta(\mathbf{x})]|\mathbf{X} = \mathbf{x}\} = \int_{-\infty}^{\infty} £[\theta, \delta(\mathbf{x})]\, k(\theta|\mathbf{x})\, d\theta,$$

jika Θ adalah peubah acak kontinu, yaitu

$$\delta(\mathbf{x}) = \operatorname{Arg\,min} \int_{-\infty}^{\infty} \pounds\left[\theta, \delta(\mathbf{x})\right] k(\theta|\mathbf{x})\, d\theta,$$

di mana $\pounds\left[\theta, \delta(\mathbf{x})\right]$ adalah fungsi kehilangan. Peubah acak terkait $\delta(\mathbf{X})$ disebut penduga Bayes bagi θ. Jika fungsi kehilangan yang diberikan ialah

$$\pounds\left[\theta, \delta(\mathbf{x})\right] = \left[\theta - \delta(\mathbf{x})\right]^2,$$

maka dugaan Bayes ialah $\delta(\mathbf{x}) = E\left[\Theta|\mathbf{x}\right]$, nilai harapan dari sebaran bersyarat $\Theta|\mathbf{X} = \mathbf{x}$. Hal tersebut berdasarkan fakta bahwa $E\left[(W-b)^2\right]$ (jika ada) adalah minimum ketika $b = E(W)$. Jika fungsi kehilangan yang diberikan ialah

$$\pounds\left[\theta, \delta(\mathbf{x})\right] = \left|\theta - \delta(\mathbf{x})\right|,$$

maka dugaan Bayes ialah median dari sebaran bersyarat $\Theta|\mathbf{X} = \mathbf{x}$. Hal tersebut berdasarkan fakta bahwa $E\left(|W-b|\right)$ (jika ada) adalah minimum ketika $b = median(W)$.

Secara umum, misalkan $\xi(\theta)$ adalah suatu fungsi dari θ. Untuk fungsi kehilangan $\pounds\left[\xi(\theta), \delta(\mathbf{x})\right]$, dugaan Bayes dari $\xi(\theta)$ adalah fungsi keputusan δ yang meminimumkan

$$E\left\{\pounds\left[\Theta, \delta(\mathbf{x})\right] | \mathbf{X} = \mathbf{x}\right\} = \int_{-\infty}^{\infty} \pounds\left[\xi(\theta), \delta(\mathbf{x})\right] k(\theta|\mathbf{x})\, d\theta.$$

Peubah acak $\delta(\mathbf{X})$ disebut penduga Bayes bagi $\xi(\theta)$.

Nilai harapan bersyarat dari $\pounds\left[\Theta, \delta(\mathbf{x})\right] | \mathbf{X} = \mathbf{x}$ mendefinisikan peubah acak yang merupakan fungsi dari contoh $\mathbf{X}$. Nilai harapan fungsi dari $\mathbf{X}$ (dalam kasus kontinu) diberikan sebagai berikut

$$\begin{aligned} & \int_{-\infty}^{\infty} \left\{ \int_{-\infty}^{\infty} \pounds\left[\theta, \delta(\mathbf{x})\right] k(\theta|\mathbf{x})\, d\theta \right\} g_1(\mathbf{x})\, d\mathbf{x} \\ = & \int_{-\infty}^{\infty} \left\{ \int_{-\infty}^{\infty} \pounds\left[\theta, \delta(\mathbf{x})\right] L(\mathbf{x}|\theta)\, d\mathbf{x} \right\} h(\theta)\, d\theta, \end{aligned} \tag{5.4}$$

di mana $\int_{-\infty}^{\infty} \pounds\left[\theta, \delta(\mathbf{x})\right] L(\mathbf{x}|\theta)\, d\mathbf{x}$ adalah fungsi risiko yang dinotasikan dengan $R(\theta, \delta)$. Oleh karena itu, persamaan 5.4 adalah nilai harapan dari risiko. Karena dugaan Bayes $\delta(\mathbf{x})$ meminimumkan

$$\int_{-\infty}^{\infty} \pounds\left[\theta, \delta(\mathbf{x})\right] k(\theta|\mathbf{x})\, d\theta$$

untuk setiap $\mathbf{x}$ di mana $g(\mathbf{x}) > 0$, maka $\delta(\mathbf{x})$ juga meminimumkan nilai harapan dari risiko.

Contoh 5.3 *Diketahui sebaran*

$$X_i|\theta \overset{bsi}{\sim} binomial\,(1,\theta)$$
$$\Theta \sim beta\,(\alpha,\beta)\,.$$

Tentukan dugaan Bayes bagi θ *jika* $\pounds\,[\theta,\delta\,(\mathbf{x})] = [\theta-\delta\,(\mathbf{x})]^2\,.$

Jawab. Sebaran prior bagi Θ ialah

$$h\,(\theta) = \frac{\Gamma\,(\alpha+\beta)}{\Gamma\,(\alpha)\,\Gamma\,(\beta)}\theta^{\alpha-1}\,(1-\theta)^{\beta-1}\,\mathbf{I}\,(0<\theta<1)\,,$$

di mana α dan β adalah konstanta positif. Statistik cukup $Y = \sum_{i=1}^{n} x_i$ menyebar binomial(n,θ) , sehingga fungsi kepekatan bersyarat $Y|\Theta=\theta$ ialah

$$g\,(y|\theta) = \binom{n}{y}\theta^{y}\,(1-\theta)^{n-y}\,\mathbf{I}\,(y=0,1,...,n)\,.$$

Dengan menggunakan 5.3,

$$k\,(\theta|y) \propto \theta^{y}\,(1-\theta)^{n-y}\,\theta^{\alpha-1}\,(1-\theta)^{\beta-1}\,\mathbf{I}\,(0<\theta<1)\,,$$

sehingga sebaran posterior $\Theta|Y=y$ menyebar beta dengan parameter $(\alpha+y,\beta+n-y)$ dan fungsi kepekatan peluangnya ialah

$$k\,(\theta|y) = \frac{\Gamma\,(n+\alpha+\beta)}{\Gamma\,(\alpha+y)\,\Gamma\,(n+\beta-y)}\theta^{\alpha+y-1}\,(1-\theta)^{\beta+n-y-1}\,\mathbf{I}\,(0<\theta<1,y=0,1,...,n)\,.$$

Untuk $\pounds\,[\theta,\delta\,(\mathbf{x})] = [\theta-\delta\,(\mathbf{x})]^2$, maka $\delta\,(y)$ merupakan nilai harapan dari sebaran beta dengan parameter $(\alpha+y,\beta+n-y)$, yaitu

$$\begin{aligned}\delta\,(y) &= \frac{\alpha+y}{\alpha+\beta+n}\\ &= \frac{\alpha+\beta}{\alpha+\beta+n}\left(\frac{\alpha}{\alpha+\beta}\right)+\frac{n}{\alpha+\beta+n}\left(\frac{y}{n}\right).\end{aligned}$$

Dugaan Bayes yang diperoleh ini merupakan rata-rata terbobot antara dugaan yang diperoleh dari sebaran prior, yaitu $\dfrac{\alpha}{\alpha+\beta}$ dengan dugaan yang diperoleh dari sebaran bersyarat $Y|\theta$, yaitu $\dfrac{y}{n}$. ✠

Contoh 5.4 *Diketahui sebaran*

$$X_i|\theta \overset{bsi}{\sim} N\left(\theta,\sigma^2\right),\ \textit{di mana } \sigma^2 \textit{ diketahui}$$
$$\Theta \sim N\left(\theta_0,\sigma_0^2\right),\ \textit{di mana } \theta_0 \textit{ dan } \sigma_0^2 \textit{ diketahui.}$$

Tentukan dugaan Bayes bagi θ *jika* $\pounds\,[\theta,\delta\,(\mathbf{x})] = [\theta-\delta\,(\mathbf{x})]^2\,.$

Jawab. Karena $Y = \bar{X}$ adalah statistik cukup bagi θ, maka

$$\begin{aligned} Y|\theta &\overset{bsi}{\sim} N\left(\theta, \sigma^2/n\right), \text{ di mana } \sigma^2 \text{ diketahui} \\ \Theta &\sim N\left(\theta_0, \sigma_0^2\right), \text{ di mana } \theta_0 \text{ dan } \sigma_0^2 \text{ diketahui.} \end{aligned}$$

Sebaran prior bagi Θ ialah

$$h(\theta) = \frac{1}{\sqrt{2\pi\sigma_0^2}} \exp\left\{-\frac{1}{2\sigma_0^2}(\theta - \theta_0)^2\right\} \mathbf{I}(-\infty < \theta < \infty).$$

Dengan menggunakan 5.3,

$$\begin{aligned} k(\theta|y) &\propto \frac{1}{\sqrt{2\pi}\sigma/\sqrt{n}} \frac{1}{\sqrt{2\pi}\sigma_0} \exp\left[-\frac{(y-\theta)^2}{2(\sigma^2/n)} - \frac{(\theta-\theta_0)^2}{2\sigma_0^2}\right] \\ &\propto \exp\left[-\frac{\left(\sigma_0^2 + \sigma^2/n\right)\theta^2 - 2\left(y\sigma_0^2 + \theta_0\sigma^2/n\right)\theta}{2(\sigma^2/n)\sigma_0^2}\right] \\ &\propto \exp\left[-\frac{\left(\theta - \dfrac{y\sigma_0^2 + \theta_0\sigma^2/n}{\sigma_0^2 + \sigma^2/n}\right)^2}{\dfrac{2(\sigma^2/n)\sigma_0^2}{\sigma_0^2 + \sigma^2/n}}\right], \end{aligned}$$

sehingga sebaran posterior $\Theta|Y = y$ menyebar normal dengan parameter

$$\left(\frac{y\sigma_0^2 + \theta_0\sigma^2/n}{\sigma_0^2 + \sigma^2/n}, \frac{(\sigma^2/n)\sigma_0^2}{\sigma_0^2 + \sigma^2/n}\right). \tag{5.5}$$

Untuk $\pounds[\theta, \delta(\mathbf{x})] = [\theta - \delta(\mathbf{x})]^2$, maka $\delta(y)$ merupakan nilai harapan dari sebaran normal dengan parameter 5.5, yaitu

$$\begin{aligned} \delta(y) &= \frac{y\sigma_0^2 + \theta_0\sigma^2/n}{\sigma_0^2 + \sigma^2/n} \\ &= \left(\frac{\sigma_0^2}{\sigma_0^2 + \sigma^2/n}\right) y + \left(\frac{\sigma^2/n}{\sigma_0^2 + \sigma^2/n}\right)\theta_0. \end{aligned}$$

Dugaan Bayes yang diperoleh ini merupakan rata-rata terbobot antara dugaan yang diperoleh dari sebaran prior, yaitu θ_0 dengan dugaan yang diperoleh dari sebaran bersyarat $Y|\theta$, yaitu y.

$\maltese$

Berikut ini merupakan contoh yang membandingkan penggunaan prinsip minimax dan metode pendugaan Bayes. Misalkan diketahui sebaran prior bagi peubah acak Θ ialah sebagai berikut:

$$\begin{aligned} P(\Theta = \theta_1) &= 0.6 \\ P(\Theta = \theta_2) &= 0.4 \end{aligned}$$

Misalkan θ_1 menyatakan Bogor hujan di sore hari dan θ_2 menyatakan Bogor tidak hujan di sore hari. Kemudian, terdapat tambahan informasi Y mengenai kondisi cuaca di Gunung Salak, yaitu

$$\begin{aligned} y_1 &: \text{ Gunung Salak cerah} \\ y_2 &: \text{ Gunung Salak berawan} \\ y_3 &: \text{ Gunung Salak gelap} \end{aligned}$$

dengan peluang bersyarat $Y|\theta$ ialah sebagai berikut:

$P(Y\|\theta)$	y_1	y_2	y_3
θ_1	0.05	0.25	0.7
θ_2	0.425	0.375	0.2

Selanjutnya, misalkan fungsi kehilangan $\pounds[\theta, \delta(y)]$ diberikan sebagai berikut:

Putusan (δ)	$\Theta \backslash Y$	y_1	y_2	y_3
δ_1	θ_1	6	3	1
(membawa payung)	θ_2	12	5	2
δ_1	θ_1	10	15	20
(tidak membawa payung)	θ_2	0	2	4

- Prinsip minimax

  $R(\theta_j, \delta_j) = \sum_{i=1}^{3} \pounds[\theta_j, \delta_j(y_i)] P(Y = y_i|\theta_j)$, $j = 1, 2$.

Putusan (δ)		y_1	y_2	y_1	$R(\theta, \delta)$
δ_1	θ_1	0.05	0.25	0.7	1.75
	$\pounds[\theta_1, \delta_1(y)]$	6	3	1	
Membawa	θ_2	0.425	0.375	0.2	7.375
payung	$\pounds[\theta_2, \delta_1(y)]$	12	5	2	
δ_2	θ_1	0.05	0.25	0.7	18.25
	$\pounds[\theta_1, \delta_2(y)]$	10	15	20	
Tidak membawa	θ_2	0.425	0.375	0.2	1.033
payung	$\pounds[\theta_2, \delta_2(y)]$	0	2	4	

  Putusan minimax: $\delta = \arg\min\max R(\theta, \delta) = \delta_1$.

Sebaran bersama dan sebaran marginal bagi peubah acak Θ dan Y diperoleh sebagai berikut:

$P(\Theta, Y)$	y_1	y_2	y_3	$P(\theta_j)$
θ_1	0.03	0.15	0.42	0.6
θ_2	0.17	0.15	0.08	0.4
$P(y_i)$	0.2	0.3	0.5	1

di mana

$$\begin{aligned} P(\Theta = \theta, Y = y) &= P(Y = y|\Theta = \theta) P(\Theta = \theta) \\ P(\theta_j) &= \sum_{i=1}^{3} P(\Theta = \theta_j, Y = y_i), \ j = 1, 2 \\ P(y_i) &= \sum_{j=1}^{2} P(\Theta = \theta_j, Y = y_i), \ i = 1, 2, 3. \end{aligned}$$

Sebaran posterior $\Theta|Y$ ialah

$P(\Theta\|Y)$	y_1	y_2	y_3
θ_1	0.15	0.5	0.84
θ_2	0.85	0.5	0.16

- Pendugaan Bayes

Amatan	Putusan	$E_{\Theta\|y}(\pounds[\Theta, \delta(y)])$	
y_1	δ_1	$(6 \times 0.15) + (12 \times 0.85)$	$= 11.1$
	δ_2	$(10 \times 0.15) + (0 \times 0.85)$	$= 1.5$
y_2	δ_1	$(3 \times 0.5) + (5 \times 0.5)$	$= 4$
	δ_2	$(15 \times 0.5) + (2 \times 0.5)$	$= 8.5$
y_3	δ_1	$(1 \times 0.84) + (2 \times 0.16)$	$= 1.16$
	δ_2	$(20 \times 0.84) + (4 \times 0.16)$	$= 17.44$

Jadi, putusan Bayes bergantung pada amatan Y yang diperoleh. Jika Y yang diperoleh y_1, maka putusan Bayes ialah δ_2, sedangkan jika Y yang diperoleh y_2 atau y_3, maka putusan Bayes ialah δ_1.

5.4 Latihan

1. Misalkan X adalah peubah acak sedemikian sehingga $E\left[(X-b)^2\right]$ ada untuk setiap bilangan Real b. Tunjukkan bahwa $E\left[(X-b)^2\right]$ minimum ketika $b = E(X)$.

2. Misalkan X adalah suatu peubah acak kontinu. Tunjukkan bahwa

$$E\left(|X-b|\right) = E\left(|X-m|\right) + 2\int_{m}^{b} (b-x)\, f(x)\, dx$$

di mana m adalah median dari sebaran X.

3. Andaikan Y_n merupakan statistik tataan ke-n dari suatu sampel acak berukuran n dari sebaran dengan fungsi kepekatan peluang $f(x;\theta) = \theta^{-1}\mathbf{I}(0 < x < \theta)$. Andai pula θ merupakan amatan dari peubah acak Θ dengan fungsi kepekatan peluang $h(\theta) = \theta^{-\beta-1}\beta\alpha^{\beta}\mathbf{I}(\theta > \alpha), \alpha > 0, \beta > 0$. Bila fungsi kehilangan $\pounds[\theta, \delta(y_n)] = [\theta - \delta(y_n)]^2$, maka tentukanlah solusi Bayes $\delta(y_n)$ yang merupakan dugaan titik bagi θ.

4. Misalkan θ merupakan nilai amatan dari peubah acak Θ dengan fungsi kepekatan peluang

$$h(\theta) = \left(\Gamma(\alpha)\,\beta^{\alpha}\right)^{-1} \theta^{\alpha-1} e^{-\theta/\beta}\mathbf{I}(\theta > 0)$$

dan $\alpha > 0, \beta > 0$ merupakan bilangan yang diketahui. Misalkan pula bahwa $X_1, X_2, \ldots, X_n$ merupakan sampel acak dari sebaran Poisson dengan nilai tengah θ. Bila $Y = \Sigma_{i=1}^{n} Xi$ dan fungsi kehilangannya ialah $\pounds[\theta, \delta(y)] = [\theta - \delta(y)]^2$, maka tentukan solusi Bayes $\delta(y)$ sebagai dugaan titik bagi θ.

5. Misalkan θ merupakan nilai amatan dari peubah acak Θ bersebaran $N(a, b)$, a dan b masing-masing merupakan konstanta yang diketahui, $a \in R$ dan $b > 0$. Misalkan pula Y merupakan rata-rata dari sampel acak $X_1, X_2, \ldots, X_n$ yang bersebaran $N(\theta, 25)$, $\theta \in R$. Bila fungsi kerugian $\pounds[\theta, \delta(y)] = |\theta - \delta(y)|$ maka berikan dugaan Bayes $\delta(y)$ sebagai dugaan titik bagi θ.

BAB 6

Pengujian Hipotesis

Misalkan X peubah acak untuk suatu populasi dengan parameter θ yang tidak diketahui, di mana $\theta \in \Omega$. Misalkan pula $X_1, X_2, \ldots, X_n$ ialah sampel acak dari sebaran peubah acak X yang memiliki fungsi kepekatan atau fungsi massa peluang $f(x, \theta)$. Suatu hipotesis statistika adalah pernyataan tentang sebaran dari populasi X. Pernyataan tersebut biasanya berkaitan dengan parameter θ jika kita sedang berhadapan dengan statistika parametrik, selain itu pernyataan tersebut ialah berkaitan dengan bentuk sebaran X. Hipotesis statistika disebut sebagai hipotesis sederhana jika hipotesis tersebut secara keseluruhan menunjukkan sebaran suatu populasi, sebaliknya disebut sebagai hipotesis majemuk.

Definisikan suatu hipotesis sebagai berikut

$$\begin{aligned} H_0 &: \theta \in \omega_0 \\ H_1 &: \theta \in \omega_1 \end{aligned}$$

di mana ω_0 dan ω_1 ialah himpunan bagian dari Ω dan $\omega_0 \cup \omega_1 = \Omega$. Hipotesis H_0 disebut sebagai hipotesis nul dan H_1 disebut sebagai hipotesis alternatif atau hipotesis tandingan.

Misalkan D adalah ruang dari sampel, yaitu $D = \text{ruang}\{(X_1, X_2, \ldots, X_n)\}$. Pengujian H_0 lawan H_1 tersebut berdasarkan himpunan bagian C dari D. Himpunan bagian C disebut sebagai daerah kritis dan himpunan tersebut berpadanan dengan aturan sebagai berikut:

$$\begin{aligned} \text{Tolak } H_0 \text{ (Terima } H_1), \text{ jika } (X_1, X_2, \ldots, X_n) &\in C \\ \text{Tidak Tolak } H_0 \text{ (Tolak } H_1), \text{ jika } (X_1, X_2, \ldots, X_n) &\in C^c. \end{aligned}$$

Penerimaan suatu hipotesis tidak berarti bahwa hipotesis tersebut benar, tetapi karena tidak ada bukti yang cukup untuk menolak atau menafikannya.

Dalam pengujian hipotesis, ada dua kemungkinan kesalahan yang dapat kita lakukan, yaitu:

1. Kesalahan untuk menolak H_0, sedangkan kenyataanya H_0 benar.
2. Kesalahan untuk tidak menolak H_0, sedangkan kenyataanya H_0 salah.

Kesalahan menolak H_0 yang benar disebut sebagai kesalahan jenis I dan kesalahan tidak menolak H_0 yang salah disebut kesalahan jenis II. Peluang terjadinya salah jenis I dinotasikan dengan α, sedangkan peluang terjadinya salah jenis II dinotasikan dengan β.

	H_0 **benar**	H_0 **salah**
Tolak H_0	Salah Jenis I	Putusan benar
Tidak Tolak H_0	Putusan benar	Salah Jenis II

Ukuran atau taraf nyata dari suatu uji hipotesis adalah peluang dari salah jenis I, yaitu

$$\alpha = \max_{\theta \in \omega_0} P_\theta \left[(X_1, X_2, \ldots, X_n) \in C \right].$$

$P_\theta \left[(X_1, X_2, \ldots, X_n) \in C \right]$ merupakan peluang bahwa $(X_1, X_2, \ldots, X_n) \in C$ ketika θ adalah parameter yang benar. Berdasarkan semua kemungkinan daerah kritis C berukuran α, kita menginginkan daerah kritis dengan peluang salah jenis II yang minimum atau daerah kritis dengan peluang dari komplemen salah jenis II maksimum. Komplemen salah jenis II ialah menolak H_0 ketika H_1 benar, di mana hal tersebut adalah putusan yang benar. Dengan demikian, untuk $\theta \in \omega_1$, kita ingin memaksimumkan

$$1 - P_\theta \left[\text{Salah jenis II} \right] = P_\theta \left[(X_1, X_2, \ldots, X_n) \in C \right]. \qquad (6.1)$$

Peluang pada ruas kanan persamaan 6.1 disebut sebagai kuasa dari uji pada θ. Jadi, meminimumkan peluang salah jenis II sama dengan memaksimumkan kuasa. Selanjutnya, definisikan fungsi kuasa dari daerah kritis C sebagai

$$\gamma_C(\theta) = P_\theta \left[(X_1, X_2, \ldots, X_n) \in C \right]; \quad \theta \in \Omega.$$

Untuk H_0 sederhana lawan H_1 sederhana, yaitu

$$\begin{aligned} H_0 &: \theta = \theta' \\ H_1 &: \theta = \theta'' \end{aligned}$$

fungsi kuasanya ialah $\gamma_C(\theta') = \alpha$ dan $\gamma_C(\theta'') = 1 - \beta$. Misalkan terdapat dua daerah kritis C_1 dan C_2 yang berukuran α. C_1 dikatakan lebih baik dari C_2 jika $\gamma_{C_1}(\theta) \geq \gamma_{C_2}(\theta)$ untuk setiap $\theta \in \omega_1$.

Contoh 6.1 *Misalkan* $X \sim N(\mu, \sigma^2)$. *Tentukan daerah kritis* C *untuk hipotesis berikut*

$$\begin{aligned} H_0 &: \mu = \mu_0 \\ H_1 &: \mu > \mu_0 \end{aligned}$$

di mana μ_0 *diketahui. Asumsikan bahwa ukuran dari uji hipotesis ini ialah* α, *untuk* $0 < \alpha < 1$.

Jawab. Misalkan $X_1, X_2, \ldots, X_n$ adalah sampel acak dari sebaran $N(\mu, \sigma^2)$. Misalkan $\bar{X}$ dan S^2 adalah nilai harapan dan ragam sampel. Secara intuisi, aturan penolakan H_0 ialah jika $\bar{X} > \mu_0$. Kemudian, kita dapat mengetahui sebaran dari statistik $\bar{X}$. Berdasarkan hipotesis H_0, statistik $T = \left(\bar{X} - \mu_0\right) / (S/\sqrt{n})$ memiliki sebaran t-Student dengan derajat bebas $(n-1)$. Dengan menggunakan sebaran T, daerah penolakan berukuran α ialah

$$T = \left(\bar{X} - \mu_0\right) / \left(S/\sqrt{n}\right) \geq t_\alpha (n-1),$$

di mana $t_\alpha (n-1)$ adalah batas atas titik kritis α dari sebaran t–Student dengan derajat bebas $(n-1)$, yaitu $\alpha = P(T > t_\alpha (n-1))$. Biasanya hal tersebut disebut *t-test* dari $H_0 : \mu = \mu_0$. Dengan demikian, daerah kritis C berukuran α ialah

$$\begin{aligned} C &= \left\{\mathbf{x} : t = (\bar{x} - \mu_0) / \left(s/\sqrt{n}\right) \geq t_\alpha (n-1)\right\} \\ &= \left\{\mathbf{x} : \bar{x} \geq \mu_0 + t_\alpha (n-1) \times \left(s/\sqrt{n}\right)\right\}. \end{aligned}$$

✠

Contoh 6.2 *Misalkan* $X_1, X_2, \ldots, X_{10}$ *adalah sampel acak dari sebaran Poisson dengan nilai harapan* θ. *Daerah kritis untuk pengujian* $H_0 : \theta = 0.1$ *lawan* $H_1 : \theta > 0.1$ *diberikan sebagai berikut*

$$Y = \sum_{i=1}^{10} X_i \geq 3.$$

Statistisk Y *memiliki sebaran dengan nilai harapan* 10θ, *sehingga bila* $\theta = 0.1$, *nilai harapan* Y *ialah* 1. *Taraf nyata dari uji ini ialah*

$$\alpha = P(Y \geq 3) = 1 - P(Y \leq 2) = 1 - 0.920 = 0.080.$$

Jika daerah kritis didefinisikan sebagai $\sum_{i=1}^{10} X_i \geq 4$, *maka taraf nyatanya ialah*

$$\alpha = P(Y \geq 4) = 1 - P(Y \leq 3) = 1 - 0.981 = 0.019.$$

Sebagai contoh, jika taraf nyata yang diinginkan ialah $\alpha = 0.05$, maka hal tersebut dapat disesuaikan dengan cara sebagai berikut. Misalkan W merupakan peubah acak yang menyebar Bernoulli dengan peluang sukses sebesar

$$P(W=1) = \frac{0.050 - 0.019}{0.080 - 0.019} = \frac{31}{61}.$$

Asumsikan bahwa W dipilih secara bebas dari sampel. Perhatikan aturan penolakan berikut :

$$\textit{Tolak } H_0 \textit{ jika } \sum_{i=1}^{10} x_i \geq 4 \textit{ atau jika } \sum_{i=1}^{10} x_i = 3 \textit{ dan } W = 1.$$

Berdasarkan aturan penolakan tersebut, diperoleh taraf nyata sebesar

$$\begin{aligned} P_{H_0}(Y \geq 4) + P_{H_0}(\{Y=3\} \cap \{W=1\}) &= P_{H_0}(Y \geq 4) + P_{H_0}(Y=3)\,P(W=1) \\ &= 0.019 + 0.061 \times \frac{31}{61} \\ &= 0.05. \end{aligned}$$

Proses melakukan percobaan tambahan untuk memutuskan apakah akan menolak H_0 atau tidak ketika $Y = 3$ disebut sebagai uji acak (randomized test).

6.1 Uji Paling Kuasa

Misalkan $f(x;\theta)$ adalah fungsi kepekatan atau fungsi massa peluang dari suatu peubah acak X di mana $\theta \in \Omega = \{\theta', \theta''\}$. Misalkan pula $\omega_0 = \{\theta'\}$ dan $\omega_1 = \{\theta''\}$, serta $\mathbf{X}' = (X_1, X_2, \ldots, X_n)$ adalah sampel acak dari sebaran peubah acak X. Selanjutnya, definisikan daerah kritis terbaik untuk uji hipotesis sederhana antara hipotesis nul H_0 dan hipotesis alternatif H_1.

Definisi 6.1 *Misalkan C adalah himpunan bagian dari ruang sampel. C disebut sebagai daerah kritis terbaik berukuran α untuk uji hipotesis sederhana $H_0 : \theta = \theta'$ dan $H_1 : \theta = \theta''$ jika $P_{\theta'}[\mathbf{X} \in C] = \alpha$ dan untuk setiap himpunan bagian A dari ruang sampel,*

$$P_{\theta'}[\mathbf{X} \in A] = \alpha \Rightarrow P_{\theta''}[\mathbf{X} \in C] \geq P_{\theta''}[\mathbf{X} \in A].$$

Teorema 6.1 *$(Neyman - Pearson)$ Misalkan $X_1, X_2, \ldots, X_n$ adalah sampel acak dari sebaran yang memiliki fungsi massa atau kepekatan peluang $f(x;\theta)$, maka fungsi kemungkinan dari $X_1, X_2, \ldots, X_n$ ialah*

$$L(\theta;\mathbf{x}) = \prod_{i=1}^{n} f(x_i;\theta), \ \textit{ untuk } \mathbf{x}' = (x_1, x_2, \ldots, x_n).$$

Misalkan θ' dan θ'' adalah nilai dari θ sedemikian sehingga $\Omega = \{\theta : \theta = \theta', \theta''\}$ dan k adalah bilangan positif. Misalkan pula C adalah himpunan bagian dari ruang sampel sedemikian sehingga: (a) $\frac{L(\theta';\mathbf{x})}{L(\theta'';\mathbf{x})} \leq k$ untuk setiap $\mathbf{x} \in C$, (b) $\frac{L(\theta';\mathbf{x})}{L(\theta'';\mathbf{x})} \geq k$ untuk setiap $\mathbf{x} \in C^c$, dan (c) $\alpha = P_{H_0}[\mathbf{X} \in C]$, maka C adalah daerah kritis terbaik berukuran α untuk pengujian hipotesis sederhana $H_0 : \theta = \theta'$ dan $H_1 : \theta = \theta''$.

Bukti. Misalkan X adalah peubah acak kontinu. Jika C adalah satu-satunya daerah kritis berukuran α, maka teorema tersebut terbukti. Untuk penyederhanaan, misalkan $\int \underset{R}{\cdots} \int L(\theta; x_1, x_2, \ldots, x_n)\, dx_1 dx_2 \ldots dx_n$ menjadi $\int_R L(\theta)$. Jika terdapat daerah kritis berukuran α yang lain, misalkan daerah tersebut adalah A, maka

$$\int_A L(\theta') = \int_C L(\theta') = \alpha.$$

Kemudian yang ingin dibuktikan ialah

$$\int_C L(\theta'') - \int_A L(\theta'') \geq 0.$$

Karena C adalah gabungan dari himpunan saling lepas $C \cap A$ dan $C \cap A^c$, maka

$$\begin{aligned}
& \int_C L(\theta'') - \int_A L(\theta'') \\
= & \int_{C\cap A} L(\theta'') + \int_{C\cap A^c} L(\theta'') - \int_{A\cap C} L(\theta'') - \int_{A\cap C^c} L(\theta'') \\
= & \int_{C\cap A^c} L(\theta'') - \int_{A\cap C^c} L(\theta''). \qquad (6.2)
\end{aligned}$$

Berdasarkan hipotesis pada teorema tersebut, $L(\theta'') \geq (1/k) L(\theta')$ untuk setiap $\mathbf{x} \in C$, juga untuk setiap $\mathbf{x} \in C \cap A^c$, sehingga

$$\int_{C\cap A^c} L(\theta'') \geq \frac{1}{k} \int_{C\cap A^c} L(\theta')$$

dan $L(\theta'') \leq (1/k) L(\theta')$ untuk setiap $\mathbf{x} \in C^c$, juga untuk setiap $\mathbf{x} \in A \cap C^c$, sehingga

$$\int_{A\cap C^c} L(\theta'') \leq \frac{1}{k} \int_{A\cap C^c} L(\theta').$$

Berdasarkan pertaksamaan tersebut diperoleh

$$\int_{C\cap A^c} L(\theta'') - \int_{A\cap C^c} L(\theta'') \geq \frac{1}{k} \int_{C\cap A^c} L(\theta') - \frac{1}{k} \int_{A\cap C^c} L(\theta'); \qquad (6.3)$$

dan dari persamaan 6.2 diperoleh

$$\int_C L(\theta'') - \int_A L(\theta'') \geq \frac{1}{k}\left[\int_{C\cap A^c} L(\theta') - \int_{A\cap C^c} L(\theta')\right].$$

Dengan demikian,

$$\begin{aligned}
&\int_{C\cap A^c} L(\theta'') - \int_{A\cap C^c} L(\theta'') \\
=\ & \int_{C\cap A} L(\theta'') + \int_{C\cap A^c} L(\theta'') - \int_{A\cap C} L(\theta'') - \int_{A\cap C^c} L(\theta'') \\
=\ & \int_C L(\theta'') - \int_A L(\theta'') \\
=\ & \alpha - \alpha = 0.
\end{aligned}$$

Jika hasil tersebut disubstitusi ke pertaksamaan 6.3, maka diperoleh

$$\int_C L(\theta'') - \int_A L(\theta'') \geq 0.$$

Jika peubah acak X adalah peubah acak diskret, maka pembuktian teorema tersebut sama seperti kasus kontinu, hanya saja notasi integral diubah menjadi notasi sigma (penjumlahan). ✠

Akibat 6.1 *Misalkan C adalah daerah kritis terbaik berukuran α untuk pengujian hipotesis sederhana $H_0 : \theta = \theta'$ dan $H_1 : \theta = \theta''$. Misalkan taraf nyata dari pengujian tersebut ialah α. Misalkan pula $\gamma_C(\theta'') = P_{\theta''}[\mathbf{X} \in C]$ menyatakan kuasa dari pengujian tersebut, maka $\alpha \leq \gamma_C(\theta'')$.*

Contoh 6.3 *Misalkan $\mathbf{X}' = (X_1, X_2, \ldots, X_n)$ adalah sampel acak dari sebaran yang memiliki fungsi kepekatan peluang*

$$f(x;\theta) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{(x-\theta)^2}{2}\right), \quad -\infty < x < \infty.$$

Tentukan daerah kritis terbaik C berukuran α untuk hipotesis sederhana berikut

$$\begin{aligned}
H_0 &: \theta = \theta' = 0 \\
H_1 &: \theta = \theta'' = 1.
\end{aligned}$$

Jawab.

$$\begin{aligned}
\frac{L(\theta';\mathbf{x})}{L(\theta'';\mathbf{x})} &= \frac{\left(1/\sqrt{2\pi}\right)^n \exp\left(-\sum_{i=1}^n x_i^2/2\right)}{\left(1/\sqrt{2\pi}\right)^n \exp\left(-\sum_{i=1}^n (x_i-1)^2/2\right)} \\
&= \exp\left(-\sum_{i=1}^n x_i + \frac{n}{2}\right).
\end{aligned}$$

Jika $k > 0$, maka himpunan dari semua titik $(x_1, x_2, \ldots, x_n)$ sedemikian sehingga

$$\begin{aligned} & \exp\left(-\sum_{i=1}^{n} x_i + \tfrac{n}{2}\right) \leq k \\ \Leftrightarrow \quad & -\sum_{i=1}^{n} x_i + \tfrac{n}{2} \leq \log k \\ \Leftrightarrow \quad & \sum_{i=1}^{n} x_i \geq \tfrac{n}{2} - \log k = c \end{aligned}$$

adalah daerah kritis terbaik, yaitu $C = \left\{(x_1, x_2, \ldots, x_n) : \sum_{i=1}^{n} x_i \geq c\right\}$, di mana c adalah suatu konstanta yang dapat ditentukan sedemikian sehingga daerah kritis tersebut berukuran α. Kejadian $\sum_{i=1}^{n} X_i \geq c$ ekuivalen dengan kejadian $\bar{X} \geq c/n = c_1$. Sebagai contoh, misalkan uji hipotesis tersebut berdasarkan statistik $\bar{X}$. Jika H_0 benar, $\theta = \theta' = 0$, maka $\bar{X} \sim N(0, 1/n)$. Berdasarkan ukuran contoh n dan taraf nyata α, konstanta c_1 dapat ditentukan dengan melihat tabel sebaran normal sedemikian sehingga

$$P_{H_0}\left(\bar{X} \geq c_1\right) = \alpha.$$

Dengan demikian, jika nilai dari $X_1, X_2, \ldots, X_n$ berturut-turut adalah $x_1, x_2, \ldots, x_n$, maka $\bar{x} = \sum_{i=1}^{n} x_i/n$. Jika $\bar{x} \geq c_1$, maka hipotesis H_0 ditolak pada taraf nyata α; jika $\bar{x} \leq c_1$, maka hipotesis H_0 tidak ditolak. Peluang menolak H_0 padahal H_0 benar ialah α; peluang menolak H_0 ketika H_0 salah adalah nilai dari kuasa uji hipotesis pada $\theta = \theta'' = 1$, yaitu

$$P_{H_1}\left(\bar{X} \geq c_1\right) = \int_{c_1}^{\infty} \frac{1}{\sqrt{2\pi}\sqrt{1/n}} \exp\left(-\frac{(\bar{x}-1)^2}{2(1/n)}\right) d\bar{x}.$$

Misalkan $n = 25$ dan $\alpha = 0.05$, maka berdasarkan tabel sebaran normal $c_1 = 1.645/\sqrt{25} = 0.329$ dan kuasa dari uji terbaik H_0 lawan H_1 ialah 0.05 ketika H_0 benar dan

$$\begin{aligned} \int_{0.329}^{\infty} \frac{1}{\sqrt{2\pi}\sqrt{1/25}} \exp\left(-\frac{(\bar{x}-1)^2}{2(1/25)}\right) d\bar{x} &= \int_{-3.355}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-w^2/2} dw \\ &= P(Z \geq -3.355) \\ &= 0.9996 \end{aligned}$$

ketika H_1 benar. ✠

6.2 Uji Selalu Paling Kuasa

Uji selalu paling kuasa atau *uniformly most powerful test* ini digunakan untuk menguji hipotesis nul sederhana dengan hipotesis alternatif majemuk. Berikut merupakan contoh dari pengujian tersebut.

Contoh 6.4 *Misalkan* X_1, X_2 *adalah sampel acak dari sebaran yang memiliki fungsi kepekatan peluang*

$$f(x;\theta) = \frac{1}{\theta} e^{-x/\theta} \mathbf{I}(x > 0).$$

1035 *Tentukan daerah kritis terbaik* C *berukuran* α *untuk hipotesis berikut*

$$\begin{aligned} H_0 &: \theta = 2 \\ H_1 &: \theta = \theta^*;\ \theta^* > 2. \end{aligned}$$

1036 **Jawab.**

$$\begin{aligned} \frac{L(\theta';\mathbf{x})}{L(\theta'';\mathbf{x})} &= \frac{L(2)}{L(\theta^*)} \\ &= \frac{\frac{1}{2^2}\exp\left[-(x_1+x_2)/2\right]}{\frac{1}{(\theta^*)^2}\exp\left[-(x_1+x_2)/\theta^*\right]}. \end{aligned}$$

Jika $k > 0$, maka himpunan dari semua titik (x_1, x_2) sedemikian sehingga

$$\frac{\frac{1}{2^2}\exp\left[-(x_1+x_2)/2\right]}{\frac{1}{(\theta^*)^2}\exp\left[-(x_1+x_2)/\theta^*\right]} \le k$$

1037 adalah daerah kritis terbaik berkuran α.

$$\begin{aligned} & \frac{\frac{1}{2^2}\exp\left[-(x_1+x_2)/2\right]}{\frac{1}{(\theta^*)^2}\exp\left[-(x_1+x_2)/\theta^*\right]} \le k \\ \Leftrightarrow\ & -\log 4 - (x_1+x_2)/2 + \log(\theta^*)^2 + (x_1+x_2)/\theta^* \le \log k \\ \Leftrightarrow\ & (x_1+x_2)\left(\tfrac{1}{\theta^*} - \tfrac{1}{2}\right) \le k_1 \\ \Leftrightarrow\ & (x_1+x_2)\left(\tfrac{1}{2} - \tfrac{1}{\theta^*}\right) \ge k_2\ , \end{aligned} \tag{6.4}$$

karena $\forall\theta^* > 2$, $\left(\frac{1}{2} - \frac{1}{\theta^*}\right) > 0$, maka pertaksamaan 6.4 dapat ditulis menjadi

$$(x_1 + x_2) \ge k_3,$$

1038 sehingga daerah kritis $C = \{(x_1, x_2) : (x_1 + x_2) \ge k_3\}$.

$$\begin{aligned} \alpha &= P(\mathbf{X} \in C | H_0) \\ 0.05 &= P_{\theta=2}(X_1 + X_2 \ge k_2) \\ 0.05 &= 1 - P(X_1 + X_2 < k_2) \\ 0.05 &= 1 - \int_0^{k_3}\int_0^{k_3 - x_1} \frac{1}{4}\exp\left[-(x_1+x_2)/2\right] dx_2 dx_1 \\ 0.05 &= e^{-k_3/2}\left(1 + \frac{k_3}{2}\right) \\ k_3 &= 9.5. \end{aligned}$$

1039 ✠

Definisi 6.2 *Daerah kritis C adalah* ***daerah kritis selalu paling kuasa*** *(uniformly most powerful critical region) berukuran α untuk pengujian hipotesis sederhana H_0 dan hipotesis majemuk H_1 jika himpunan C adalah daerah kritis terbaik berukuran α untuk pengujian hipotesis H_0 dengan setiap hipotesis sederhana H_1. Pengujian yang didefinisikan berdasarkan daerah kritis C tersebut disebut* ***uji selalu paling kuasa*** *(uniformly most powerful test) dengan taraf nyata α untuk pengujian hipotesis sederhana H_0 dan hipotesis majemuk H_1.*

Contoh 6.5 *Misalkan $X_1, X_2, \ldots, X_n$ adalah sampel acak dari sebaran normal $N(\theta, 1)$. Tentukan daerah kritis selalu paling kuasa berukuran α (jika ada) untuk hipotesis berikut*

$$\begin{aligned} H_0 &: \theta = 5 \\ H_1 &: \theta = \theta^*;\ \theta^* \neq 5. \end{aligned}$$

Jawab.

$$\begin{aligned} \frac{L(\theta = 5)}{L(\theta = \theta^*)} &= \frac{\prod_{i=1}^{n} \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2}(x-5)^2\right)}{\prod_{i=1}^{n} \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2}(x-\theta^*)^2\right)} \\ &= -\frac{1}{2}\sum_{i=1}^{n}(x_i - 5)^2 + \frac{1}{2}\sum_{i=1}^{n}(x_i - \theta^*)^2 \\ &= (5 - \theta^*)\sum_{i=1}^{n} x_i + n(\theta^*)^2 - \frac{25n}{2} \leq k_1. \end{aligned}$$

$$(5 - \theta^*)\sum_{i=1}^{n} x_i \leq k_1 - n(\theta^*)^2 + \frac{25n}{2}.$$

Jika $\theta^* < 5$, maka $C = \{(x_1, x_2, \ldots, x_n) : \sum_{i=1}^{n} x_i \leq k_2\}$, tetapi untuk $\theta^* > 5$, maka $C = \{(x_1, x_2, \ldots, x_n) : \sum_{i=1}^{n} x_i \geq k_2\}$. Dengan demikian, tidak ada daerah kritis selalu paling kuasa. ✠

6.3 Uji Rasio Kemungkinan

Misalkan peubah acak X memiliki fungsi kepekatan atau massa peluang $f(x; \boldsymbol{\theta})$, di mana $\boldsymbol{\theta}$ adalah vektor dari parameter pada Ω. Misalkan $\omega \subset \Omega$ dan

$$H_0 : \theta \in \omega \text{ lawan } H_1 : \theta \in \Omega \cap \omega^c.$$

Prinsip dari rasio kemungkinan untuk menolak H_0 adalah jika dan hanya jika $\Lambda \leq \lambda_0 < 1$, di mana

$$\Lambda(x) = \frac{\sup\limits_{\boldsymbol{\theta}\in\omega} L(\omega)}{\sup\limits_{\boldsymbol{\theta}\in\Omega} L(\Omega)} = \frac{L(\hat{\omega})}{L\left(\hat{\Omega}\right)},$$

dengan taraf nyata dari uji tersebut ialah $\alpha = \max\limits_{\theta\in\omega} P_{H_0}\left[\Lambda(X) \leq \lambda_0\right].$

Contoh 6.6 *Misalkan $X_1, X_2, \ldots, X_n$ merupakan sampel acak dari sebaran $N(\theta_1, \theta_2)$ di mana θ_1 adalah nilai tengah dan θ_2 adalah ragam yang tidak diketahui, sehingga $\Omega = \{(\theta_1, \theta_2) \in \mathbb{R} \times \mathbb{R}^+\}$. Hipotesis yang diuji ialah*

$$\begin{aligned} H_0 &: \theta_1 = 0, \theta_2 > 0 \\ H_1 &: \theta_1 \neq 0, \theta_2 > 0, \end{aligned}$$

sehingga $\omega = \{\theta_1 = 0, \theta_2 \in \mathbb{R}^+\}$. Tentukan statistik dan sebaran yang terkait hipotesis tersebut, serta berikan daerah kritis pada taraf nyata α.

Jawab. Dalam hal ini sampel acak $X_1, X_2, \ldots, X_n$ saling bebas, sehingga fungsi kemungkinannya ialah

$$L(\Omega) = (2\pi\theta_2)^{-n/2} \exp\left[-\frac{1}{2\theta_2}\Sigma_{i=1}^{n}(x_i - \theta_1)^2\right]$$

dan

$$L(\omega) = (2\pi\theta_2)^{-n/2} \exp\left[-\frac{1}{2\theta_2}\Sigma_{i=1}^{n} x_i^2\right].$$

Jika $\partial \ln L(\omega)/\partial\theta_2 = 0$, maka

$$\begin{aligned} \frac{-n}{2\theta_2} + \frac{1}{\theta_2} &= 0 \\ \theta_2 &= \frac{\Sigma_{i=1}^{n} x_i^2}{n}, \end{aligned}$$

di mana θ_2 memaksimumkan $L(\omega)$. Jadi, penduga kemungkinan maksimum bagi θ_2 ialah

$$\hat{\theta}_2 = \frac{\Sigma_{i=1}^{n} X_i^2}{n}.$$

Jika $\partial \ln L(\Omega)/\partial\theta_1$ dan $\partial \ln L(\Omega)/\partial\theta_2$ dibuat sama dengan nol, maka penduga kemungkinan maksimum bagi θ_1 dan θ_2 ialah

$$\begin{aligned} \hat{\theta}_1 &= \bar{X} \\ \hat{\theta}_2 &= \frac{\Sigma_{i=1}^{n}\left(X_i - \bar{X}\right)^2}{n}. \end{aligned}$$

Dengan demikian,

$$
\begin{aligned}
L\left(\hat{\Omega}\right) &= \left[2\pi \times \frac{\Sigma_{i=1}^{n}\left(x_i-\bar{x}\right)^2}{n}\right]^{-n/2} e^{-n/2},\\
L\left(\hat{\omega}\right) &= \left[2\pi \times \frac{\Sigma_{i=1}^{n}x_i^2}{n}\right]^{-n/2} e^{-n/2},
\end{aligned}
$$

sehingga

$$
\Lambda\left(x\right) = \frac{L\left(\hat{\omega}\right)}{L\left(\hat{\Omega}\right)} = \left[\frac{\Sigma_{i=1}^{n}\left(x_i-\bar{x}\right)^2}{\Sigma_{i=1}^{n}x_i^2}\right]^{n/2}.
$$

Karena $\Sigma_{i=1}^{n}\left(x_i-\bar{x}\right)^2 = \Sigma_{i=1}^{n}x_i^2 - n\bar{x}^2$, maka

$$
\begin{aligned}
& \Lambda\left(x\right) = \left[\frac{\Sigma_{i=1}^{n}\left(x_i-\bar{x}\right)^2}{\Sigma_{i=1}^{n}\left(x_i-\bar{x}\right)^2 + n\bar{x}^2}\right]^{n/2} \leq \lambda_0\\
\Leftrightarrow\ & \frac{1}{1+\dfrac{n\bar{x}^2}{\Sigma_{i=1}^{n}\left(x_i-\bar{x}\right)^2}} \leq \lambda_0^{2/n}\\
\Leftrightarrow\ & 1+\frac{n\bar{x}^2}{\Sigma_{i=1}^{n}\left(x_i-\bar{x}\right)^2} \geq \lambda_0^{-2/n}\\
\Leftrightarrow\ & \frac{n\bar{x}^2}{\Sigma_{i=1}^{n}\left(x_i-\bar{x}\right)^2} \geq \lambda_0^{-2/n} - 1 = \lambda_0^*\\
\Leftrightarrow\ & \frac{n\bar{x}^2/\theta_2}{\Sigma_{i=1}^{n}\left(x_i-\bar{x}\right)^2/\theta_2 \times \frac{1}{n-1}} \geq \lambda_0^*\left(n-1\right)\\
\Leftrightarrow\ & \frac{\sqrt{n}\left|\bar{x}\right|/\sqrt{\theta_2}}{\sqrt{\Sigma_{i=1}^{n}\left(x_i-\bar{x}\right)^2/\theta_2 \times \frac{1}{n-1}}} \geq \sqrt{\lambda_0^*\left(n-1\right)} = \lambda_0^{**}.
\end{aligned}
$$

Karena $\bar{X} \sim N\left(\theta_1, \theta_2\right)$ dan $\Sigma_{i=1}^{n}\left(X_i-\bar{X}\right)^2/\theta_2 \sim \chi^2\left(n-1\right)$, maka

$$
T = \frac{\sqrt{n}\left(\bar{X}-\theta_1\right)}{\sqrt{\Sigma_{i=1}^{n}\left(X_i-\bar{X}\right)^2/\left(n-1\right)}} \sim \text{t} - \text{Student}\left(n-1\right)
$$

dan bila H_0 benar,

$$
T = \frac{\sqrt{n}\bar{X}}{\sqrt{\Sigma_{i=1}^{n}\left(X_i-\bar{X}\right)^2/\left(n-1\right)}} \sim \text{t} - \text{Student}\left(n-1\right).
$$

Dengan demikian, daerah kritis C ialah

$$\begin{aligned} C &= \left\{ \mathbf{x}; \frac{\sqrt{n}\,|\bar{x}|}{\sqrt{\Sigma_{i=1}^{n}\left(x_i-\bar{x}\right)^2/\left(n-1\right)}} \geq t_{\alpha/2}\left(n-1\right) \right\} \\ &= \left\{ \mathbf{x}; |\bar{x}| \geq \frac{S}{\sqrt{n}} t_{\alpha/2}\left(n-1\right) \right\}. \end{aligned}$$

✠

Contoh 6.7 *Misalkan* $X \sim N\left(\theta_1, \theta_3\right)$ *dan* $Y \sim N\left(\theta_2, \theta_3\right)$, *di mana* θ_1, θ_2 *adalah nilai tengah dari* X *dan* Y, *serta* θ_3 *adalah ragam yang tidak diketahui, sehingga* $\Omega = \left\{\left(\theta_1, \theta_2, \theta_3\right) : \theta_1, \theta_2 \in \mathbb{R}, \theta_3 \in \mathbb{R}^+\right\}$. *Misalkan* $X_1, X_2, \ldots, X_n$ *dan* $Y_1, Y_2, \ldots, Y_m$ *merupakan sampel acak dari sebaran tersebut. Hipotesis yang diuji ialah*

$$\begin{aligned} H_0 &: \theta_1 = \theta_2, \theta_3 > 0 \\ H_1 &: \theta_1 \neq \theta_2, \theta_3 > 0, \end{aligned}$$

sehingga $\omega = \left\{\left(\theta_1, \theta_2, \theta_3\right) : \theta_1 = \theta_2 \in \mathbb{R}, \theta_3 \in \mathbb{R}^+\right\}$. *Tentukan statistik dan sebaran yang terkait hipotesis tersebut, serta berikan daerah kritis pada taraf nyata* α.

Jawab. Dalam hal ini $X_1, X_2, \ldots, X_n, Y_1, Y_2, \ldots, Y_m$ adalah $n+m > 2$ sampel acak saling bebas, sehingga fungsi kemungkinannya ialah

$$L\left(\Omega\right) = \left(2\pi\theta_3\right)^{-(n+m)/2} \exp\left[-\frac{1}{2\theta_3}\left\{\Sigma_{i=1}^{n}\left(x_i-\theta_1\right)^2 + \Sigma_{i=1}^{m}\left(y_i-\theta_2\right)^2\right\}\right]$$

dan

$$L\left(\omega\right) = \left(2\pi\theta_3\right)^{-(n+m)/2} \exp\left[-\frac{1}{2\theta_3}\left\{\Sigma_{i=1}^{n}\left(x_i-\theta_1\right)^2 + \Sigma_{i=1}^{m}\left(y_i-\theta_1\right)^2\right\}\right].$$

Jika $\partial \ln L\left(\omega\right)/\partial\theta_1$ dan $\partial \ln L\left(\omega\right)/\partial\theta_3$ dibuat sama dengan nol, maka

$$\begin{aligned} \Sigma_{i=1}^{n}\left(x_i-\theta_1\right) + \Sigma_{i=1}^{m}\left(y_i-\theta_1\right) &= 0 \\ \frac{1}{\theta_3}\left[\Sigma_{i=1}^{n}\left(x_i-\theta_1\right)^2 + \Sigma_{i=1}^{m}\left(y_i-\theta_1\right)^2\right] &= n+m. \end{aligned}$$

Solusi dari persamaan di atas untuk θ_1 dan θ_3 berturut-turut ialah

$$\begin{aligned} \theta_{1\omega} &= \left(n+m\right)^{-1}\left\{\Sigma_{i=1}^{n} x_i + \Sigma_{i=1}^{m} y_i\right\} \\ \theta_{3\omega} &= \left(n+m\right)^{-1}\left\{\Sigma_{i=1}^{n}\left(x_i-\theta_{1\omega}\right)^2 + \Sigma_{i=1}^{m}\left(y_i-\theta_{1\omega}\right)^2\right\}, \end{aligned}$$

di mana $\theta_{1\omega}$ dan $\theta_{3\omega}$ memaksimumkan $L(\omega)$. Selanjutnya, jika $\partial \ln L(\Omega)/\partial\theta_1$, $\partial \ln L(\Omega)/\partial\theta_2$, dan $\partial \ln L(\Omega)/\partial\theta_3$ dibuat sama dengan nol, maka

$$\begin{aligned} \Sigma_{i=1}^{n}(x_i-\theta_1) &= 0 \\ \Sigma_{i=1}^{m}(y_i-\theta_2) &= 0 \\ -(n+m)+\frac{1}{\theta_3}\left[\Sigma_{i=1}^{n}(x_i-\theta_1)^2+\Sigma_{i=1}^{m}(y_i-\theta_1)^2\right] &= 0. \end{aligned}$$

Solusi dari persamaan di atas untuk θ_1, θ_2 dan θ_3 berturut-turut ialah

$$\begin{aligned} \theta_1 &= \bar{x} \\ \theta_2 &= \bar{y} \\ \theta_3 &= (n+m)^{-1}\left\{\Sigma_{i=1}^{n}(x_i-\theta_1)^2+\Sigma_{i=1}^{m}(y_i-\theta_2)^2\right\} \end{aligned}$$

di mana θ_1, θ_2 dan θ_3 memaksimumkan $L(\Omega)$. Dengan demikian,

$$\begin{aligned} L\left(\hat{\Omega}\right) &= \left[\frac{e^{-1}}{2\pi\theta_3}\right]^{(n+m)/2} \\ L(\hat{\omega}) &= \left[\frac{e^{-1}}{2\pi\theta_{3\omega}}\right]^{(n+m)/2}, \end{aligned}$$

sehingga

$$\begin{aligned} \Lambda(x) &= \frac{L(\hat{\omega})}{L\left(\hat{\Omega}\right)} = \left[\frac{\theta_3}{\theta_{3\omega}}\right]^{(n+m)/2} \\ \Lambda^{2/(n+m)}(X) &= \frac{\hat{\theta}_3}{\hat{\theta}_{3\omega}} \\ &= \frac{\Sigma_{i=1}^{n}\left(X_i-\bar{X}\right)^2+\Sigma_{i=1}^{m}\left(Y_i-\bar{Y}\right)^2}{\sum_{i=1}^{n}\left(X_i-\frac{n\bar{X}+m\bar{Y}}{n+m}\right)^2+\sum_{i=1}^{m}\left(Y_i-\frac{n\bar{X}+m\bar{Y}}{n+m}\right)^2}. \end{aligned}$$

Karena

$$\begin{aligned} \sum_{i=1}^{n}\left(X_i-\frac{n\bar{X}+m\bar{Y}}{n+m}\right)^2 &= \sum_{i=1}^{n}\left(\left[X_i-\bar{X}\right]+\left[\bar{X}-\frac{n\bar{X}+m\bar{Y}}{n+m}\right]\right)^2 \\ &= \sum_{i=1}^{n}\left(X_i-\bar{X}\right)^2+n\left(\bar{X}-\frac{n\bar{X}+m\bar{Y}}{n+m}\right)^2 \\ &= \sum_{i=1}^{n}\left(X_i-\bar{X}\right)^2+\frac{m^2n}{(n+m)^2}\left(\bar{X}-\bar{Y}\right)^2 \end{aligned}$$

dan

$$\begin{aligned}\sum_{i=1}^{m}\left(Y_i-\frac{n\bar{X}+m\bar{Y}}{n+m}\right)^2 &= \sum_{i=1}^{m}\left(\left[Y_i-\bar{Y}\right]+\left[\bar{Y}-\frac{n\bar{X}+m\bar{Y}}{n+m}\right]\right)^2\\ &= \sum_{i=1}^{m}\left(Y_i-\bar{Y}\right)^2+m\left(\bar{Y}-\frac{n\bar{X}+m\bar{Y}}{n+m}\right)^2\\ &= \sum_{i=1}^{m}\left(Y_i-\bar{Y}\right)^2+\frac{n^2m}{(n+m)^2}\left(\bar{X}-\bar{Y}\right)^2,\end{aligned}$$

maka

$$\begin{aligned}\Lambda^{2/(n+m)}(X) &= \frac{\Sigma_{i=1}^{n}\left(X_i-\bar{X}\right)^2+\Sigma_{i=1}^{m}\left(Y_i-\bar{Y}\right)^2}{\sum_{i=1}^{n}\left(X_i-\bar{X}\right)^2+\sum_{i=1}^{m}\left(Y_i-\bar{Y}\right)^2+\left[nm/(n+m)\right]\left(\bar{X}-\bar{Y}\right)^2}\\ &= \frac{1}{1+\dfrac{\left[nm/(n+m)\right]\left(\bar{X}-\bar{Y}\right)^2}{\sum_{i=1}^{n}\left(X_i-\bar{X}\right)^2+\sum_{i=1}^{m}\left(Y_i-\bar{Y}\right)^2}}.\end{aligned}$$

Jika hipotesis $H_0:\theta_1=\theta_2$ benar, maka peubah acak

$$T=\sqrt{\frac{nm}{n+m}}\left(\bar{X}-\bar{Y}\right)\left\{(n+m-2)^{-1}\left[\sum_{i=1}^{n}\left(X_i-\bar{X}\right)^2+\sum_{i=1}^{m}\left(Y_i-\bar{Y}\right)^2\right]\right\}^{-1/2}$$

memiliki sebaran t$-$Student dengan derajat bebas $n+m-2$. Dengan demikian, peubah acak $\Lambda^{2/(n+m)}(X)$ dapat dituliskan menjadi

$$\frac{n+m-2}{(n+m-2)+T^2},$$

sehingga daerah kritis C ialah $C=\left\{|t|\geq t_{\alpha/2}(n+m-2)\right\}$, di mana t adalah nilai amatan dari peubah acak T. ✠

Definisi 6.3 *Misalkan peubah acak* $Z\sim N(\delta,1)$ *dan peubah acak* $U\sim\chi^2(r)$, *serta dua peubah acak tersebut saling bebas, maka peubah acak*

$$T=\frac{Z}{\sqrt{U/r}}$$

dikatakan memiliki sebaran t *takterpusat dengan derajat bebas* r *dan parameter takterpusat* δ. *Jika* $\delta=0$, *maka* T *disebut memiliki sebaran terpusat* t.

Pada dua contoh di atas, Contoh 6.6 dan Contoh 6.7, penentuan uji rasio kemungkinan tersebut dilakukan berdasarkan statistik di mana jika H_0 benar, maka memiliki sebaran t. Kemudian, kita ingin menguji statistik pada contoh-contoh tersebut dengan menggunakan Definisi 6.3. Pada Contoh 6.6, diperoleh

$$T = \frac{\sqrt{n}\bar{X}}{\sqrt{\Sigma_{i=1}^{n}\left(X_i - \bar{X}\right)^2 / (n-1)}}$$

di mana

$$\begin{aligned} Z_1 &= \frac{\sqrt{n}\bar{X}}{\sigma} \sim N\left(\frac{\sqrt{n}\theta_1}{\sigma}, 1\right) \\ U_1 &= \frac{\Sigma_{i=1}^{n}\left(X_i - \bar{X}\right)^2}{\sigma^2} \sim \chi^2\left(n-1\right), \end{aligned}$$

peubah acak Z_1 dan U_1 saling bebas. Dengan demikian, jika $\theta_1 \neq 0$, maka T memiliki sebaran t takterpusat dengan derajat bebas $n-1$ dan parameter terpusatnya $\delta_1 = \dfrac{\sqrt{n}\theta_1}{\sigma}$.

Dari Contoh 6.7 diperoleh

$$T = \frac{Z_2}{\sqrt{U_2 / \left(n+m-2\right)}}$$

di mana

$$Z_2 = \sqrt{\frac{nm}{n+m}}\left(\bar{X} - \bar{Y}\right)/\sigma$$

dan

$$U_2 = \frac{1}{\sigma^2}\left[\sum_{i=1}^{n}\left(X_i - \bar{X}\right)^2 + \sum_{i=1}^{m}\left(Y_i - \bar{Y}\right)^2\right].$$

Kedua peubah acak tersebut saling bebas dan

$$\begin{aligned} Z_2 &\sim N\left(\sqrt{nm/\left(n+m\right)}\left(\theta_1 - \theta_2\right)/\sigma, 1\right) \\ U_2 &\sim \chi^2\left(n+m-2\right). \end{aligned}$$

Dengan demikian, jika $\theta_1 \neq \theta_2$, maka T memiliki sebaran t takterpusat dengan derajat bebas $n+m-2$ dan parameter takterpusatnya

$$\delta_2 = \sqrt{nm/\left(n+m\right)}\left(\theta_1 - \theta_2\right)/\sigma.$$

Misalkan suatu statistik $-2\ln\Lambda(X)$ dengan

$$\Lambda(x) = \frac{L(\hat{\omega})}{L\left(\hat{\Omega}\right)}.$$

Berdasarkan prinsip dari uji rasio kemungkinan, H_0 ditolak jika dan hanya jika $\Lambda \leq \lambda_0$, sehingga untuk statistik $-2\ln\Lambda$, H_0 ditolak jika $-2\ln\Lambda > \lambda_1$ untuk suatu konstanta λ_1. Dengan demikian, uji rasio kemungkinan ekuivalen dengan uji yang menggunakan statistik $-2\ln\Lambda$. Berikut ini merupakan teorema yang menyatakan sebaran asimtotik bagi $-2\ln\Lambda$.

Teorema 6.2 *Misalkan $X_1, X_2, ..., X_n$ adalah sampel acak dengan fungsi kepekatan peluang $f(\mathbf{x};\theta)$, $\theta \in \Omega$, di mana Ω merupakan himpunan bagian berdimensi r dari $\mathbb{R}^r$ dan misalkan ω merupakan himpunan bagian berdimensi m dari Ω. Misalkan pula himpunan positif dari fungsi kepekatan peluang tidak bergantung pada θ, maka sebaran asimtotik dari $-2\ln\Lambda$ ialah $\chi^2(r-m)$, asalkan $\theta \in \omega$, yaitu saat $n \to \infty$,*

$$P_\theta(-2\ln\Lambda \leq x) \to G(x),\ x \geq 0,\ \forall\theta \in \omega,$$

dengan G adalah fungsi sebaran dari sebaran $\chi^2(r-m)$.

6.4 Uji Rasio Peluang Bersekuens

Pada Teorema 6.1 diberikan suatu metode untuk menentukan daerah kritis terbaik untuk pengujian hipotesis sederhana lawan hipotesis alternatif sederhana. Misalkan $X_1, X_2, \ldots, X_n$ adalah sampel acak dengan ukuran contoh n yang tetap, yang berasal dari sebaran dengan fungsi massa atau kepekatan peluang $f(x;\theta)$, di mana $\theta = \{\theta : \theta = \theta', \theta''\}$ dan θ', θ'' diketahui nilainya. Fungsi kemungkinan dari $X_1, X_2, \ldots, X_n$ ialah

$$L(\theta;n) = f(x_1;\theta)\, f(x_2;\theta)\cdots f(x_n;\theta).$$

Jika $H_0 : \theta = \theta'$ ditolak dan $H_1 : \theta = \theta''$ diterima ketika

$$\frac{L(\theta';n)}{L(\theta'';n)} \leq k,$$

di mana $k > 0$, maka berdasarkan Teorema 6.1, hal tersebut merupakan uji terbaik dari H_0 lawan H_1.

Selanjutnya, misalkan ukuran sampel n merupakan suatu peubah acak. Misalkan N adalah peubah acak yang menyatakan ukuran sampel dengan ruang sampel $\{1, 2, 3, ...\}$. Prosedur untuk pengujian hipotesis sederhana $H_0 : \theta = \theta'$ lawan $H_1 : \theta = \theta''$ ialah sebagai berikut: misalkan k_0 dan k_1 merupakan konstanta positif dengan $k_0 < k_1$. Amati nilai dari $X_1, X_2, \ldots, X_n$ pada suatu sekuens, sebagai contoh $x_1, x_2, x_3 \ldots$ dan hitung

$$\frac{L(\theta'; 1)}{L(\theta''; 1)}, \frac{L(\theta'; 2)}{L(\theta''; 2)}, \frac{L(\theta'; 3)}{L(\theta''; 3)}, \ldots .$$

Hipotesis H_0 ditolak jika dan hanya jika terdapat bilangan bulat positif n sedemikian sehingga $\mathbf{x}_n = (x_1, x_2, \ldots, x_n)$ merupakan anggota dari himpunan

$$C_n = \left\{ \mathbf{x}_n : k_0 < \frac{L(\theta'; j)}{L(\theta''; j)} < k_1,\ j = 1, 2, ..., n-1 \text{ dan } \frac{L(\theta'; n)}{L(\theta''; n)} \leq k_0 \right\}.$$

Di samping itu, hipotesis H_0 diterima jika dan hanya jika terdapat bilangan bulat positif n sedemikian sehingga $\mathbf{x}_n = (x_1, x_2, \ldots, x_n)$ merupakan anggota dari himpunan

$$A_n = \left\{ \mathbf{x}_n : k_0 < \frac{L(\theta'; j)}{L(\theta''; j)} < k_1,\ j = 1, 2, ..., n-1 \text{ dan } \frac{L(\theta'; n)}{L(\theta''; n)} \geq k_1 \right\},$$

Dengan kata lain, pengamatan contoh dilanjutkan selama

$$k_0 < \frac{L(\theta'; n)}{L(\theta''; n)} < k_1 \tag{6.5}$$

dan berhenti jika:

1. Dengan penolakan $H_0 : \theta = \theta'$,
$$\frac{L(\theta'; n)}{L(\theta''; n)} \leq k_0.$$
2. Dengan penerimaan $H_0 : \theta = \theta'$,
$$\frac{L(\theta'; n)}{L(\theta''; n)} \geq k_1.$$

Pengujian hipotesis tersebut disebut uji rasio peluang bersekuens Wald (*Wald's sequential probability ratio test*). Seringkali, pertaksamaan 6.5 dituliskan menjadi

$$c_0(n) < u(x_1, x_2, \ldots, x_n) < c_1(n)$$

di mana $u(X_1, X_2, \ldots, X_n)$ adalah suatu statistik dan $c_0(n), c_1(n)$ adalah konstanta yang bergantung terhadap $k_0, k_1, \theta', \theta''$, dan n, sehingga pengamatan dihentikan serta putusan ditentukan ketika

$$u(x_1, x_2, \ldots, x_n) \leq c_0(n) \text{ atau } u(x_1, x_2, \ldots, x_n) \geq c_1(n).$$

Contoh 6.8 *Misalkan X memiliki fungsi massa peluang*

$$f(x;\theta) = \theta^x (1-\theta)^{1-x} \mathbf{I}(x = 0, 1).$$

Tentukan daerah kritis bagi hipotesis

$$\begin{aligned} H_0 &: \theta = \frac{1}{3} \\ H_1 &: \theta = \frac{2}{3} \end{aligned}$$

dengan menggunakan uji rasio peluang bersekuens Wald.

Jawab.

$$\frac{L\left(\frac{1}{3}; n\right)}{L\left(\frac{2}{3}; n\right)} = \frac{\left(\frac{1}{3}\right)^{\Sigma x_i} \left(\frac{2}{3}\right)^{n-\Sigma x_i}}{\left(\frac{2}{3}\right)^{\Sigma x_i} \left(\frac{1}{3}\right)^{n-\Sigma x_i}} = 2^{n-2\Sigma x_i},$$

di mana $\Sigma x_i = \Sigma_{i=1}^n x_i$.

$$\begin{aligned} k_0 < &\frac{L\left(\frac{1}{3}; n\right)}{L\left(\frac{2}{3}; n\right)} < k_1 \\ \log_2 k_0 < &\, n - 2\Sigma_{i=1}^n x_i < \log_2 k_1 \\ c_0(n) = \frac{n}{2} - \frac{1}{2}\log_2 k_1 < &\, \Sigma_{i=1}^n x_i < \frac{n}{2} - \frac{1}{2}\log_2 k_0 = c_1(n). \end{aligned}$$

Perhatikan bahwa $L\left(\frac{1}{3}; n\right)/L\left(\frac{2}{3}; n\right) \le k_0$ jika dan hanya jika $c_1(n) \le \Sigma_{i=1}^n x_i$ dan $L\left(\frac{1}{3}; n\right)/L\left(\frac{2}{3}; n\right) \ge k_1$ jika dan hanya jika $c_0(n) \ge \Sigma_{i=1}^n x_i$. Dengan demikian, pengamatan contoh dilanjutkan selama

$$c_0(n) < \Sigma_{i=1}^n x_i < c_1(n).$$

Daerah penolakan $H_0 : \theta = \frac{1}{3}$ ialah $\Sigma_{i=1}^n x_i \ge c_1(n)$ dan daerah penerimaan $H_0 : \theta = \frac{1}{3}$ ialah $\Sigma_{i=1}^n x_i \le c_0(n)$. ✠

Pada subbab ini, kuasa dari uji ketika H_0 benar dinotasikan α dan kuasa dari uji ketika H_1 benar dinotasikan $1-\beta$. Dengan demikian, α adalah peluang terjadinya kesalahan jenis I (penolakan H_0 padahal H_0 benar) dan β adalah peluang terjadinya kesalahan jenis II (penerimaan H_0 padahal H_0 salah), sehingga (dalam kasus kontinu)

$$\alpha = \sum_{n=1}^{\infty} \int_{C_n} L(\theta'; n), \quad (1-\beta) = \sum_{n=1}^{\infty} \int_{C_n} L(\theta''; n).$$

Karena peluang bernilai satu ketika prosedur berakhir, maka

$$1-\alpha=\sum_{n=1}^{\infty}\int_{A_n}L\left(\theta';n\right),\quad \beta=\sum_{n=1}^{\infty}\int_{A_n}L\left(\theta'';n\right).$$

Jika $(x_1,x_2,\ldots,x_n)\in C_n$, maka $L\left(\theta';n\right)\le k_0L\left(\theta'';n\right)$, sehingga jelas bahwa

$$\alpha=\sum_{n=1}^{\infty}\int_{C_n}L\left(\theta';n\right)\le\sum_{n=1}^{\infty}\int_{C_n}k_0L\left(\theta'';n\right)=k_0\left(1-\beta\right).$$

Karena $L\left(\theta';n\right)\ge k_1L\left(\theta'';n\right)$ pada setiap titik dari himpunan A_n, maka

$$1-\alpha=\sum_{n=1}^{\infty}\int_{A_n}L\left(\theta';n\right)\ge\sum_{n=1}^{\infty}\int_{A_n}k_1L\left(\theta'';n\right)=k_1\beta.$$

1113 Oleh karena itu,

$$\frac{\alpha}{1-\beta}\le k_0,\quad k_1\le\frac{1-\alpha}{\beta}\tag{6.6}$$

1114 asalkan β tidak bernilai 0 atau 1.

Misalkan α_0 dan β_0 adalah nilai yang ditetapkan sebelumnya. Jika kita mengambil

$$k_0=\frac{\alpha_0}{1-\beta_0}\text{ dan }k_1=\frac{1-\alpha_0}{\beta_0},$$

maka pertaksamaan 6.6 menjadi

$$\frac{\alpha}{1-\beta}\le\frac{\alpha_0}{1-\beta_0},\quad \frac{1-\alpha_0}{\beta_0}\le\frac{1-\alpha}{\beta}$$

1115 atau ekuivalen dengan

$$\alpha\left(1-\beta_0\right)\le\alpha_0\left(1-\beta\right),\quad \beta\left(1-\alpha_0\right)\le\beta_0\left(1-\alpha\right).\tag{6.7}$$

Apabila kedua pertaksamaan 6.7 dijumlahkan, maka diperoleh

$$\alpha+\beta\le\alpha_0+\beta_0.$$

Dengan demikian, diperoleh batas atas dari α dan β ialah

$$\alpha\le\frac{\alpha_0}{1-\beta_0},\ \beta\le\frac{\beta_0}{1-\alpha_0},$$

1116 serta secara empiris

$$\begin{aligned}\alpha&\approx\alpha_0,\\ \beta&\approx\beta_0.\end{aligned}$$

Pada Contoh 6.8, bila $\alpha = 0.01$ dan $\beta = 0.05$, maka

$$\begin{aligned} k_0 &= \frac{0.01}{1-0.05} = \frac{1}{95} \\ k_1 &= \frac{1-0.01}{0.05} = \frac{99}{5}, \end{aligned}$$

sehingga tolak H_0 bila

$$\Sigma_{i=1}^{n} x_i \geq \frac{n}{2} - \frac{1}{2}\log_2 \frac{1}{95} = c_1(n)$$

dan terima H_1 bila

$$\Sigma_{i=1}^{n} x_i \leq \frac{n}{2} - \frac{1}{2}\log_2 \frac{99}{5} = c_0(n).$$

Catatan 6.1 *Peluang bahwa prosedur dari pengujian tersebut terus berlanjut sampai $n \to \infty$ ialah nol. Jika $\theta = \theta'$ atau $\theta = \theta''$, maka*

$$\begin{aligned} E(N|H_0) &= \frac{1}{E(Z|H_0)}[\alpha \ln k_0 + (1-\alpha)\ln k_1], \\ E(N|H_1) &= \frac{1}{E(Z|H_1)}[(1-\beta)\ln k_0 + \beta \ln k_1], \end{aligned}$$

di mana

$$Z = \ln \frac{f(X;\theta')}{f(X;\theta'')}$$

dan N merupakan peubah acak yang menyatakan ukuran sampel.

Contoh 6.9 *Misalkan X memiliki sebaran $N \sim (\mu, 1)$. Gunakan uji rasio peluang bersekuens Wald untuk menentukan rata-rata ukuran contoh yang diharapkan untuk sampai pada suatu putusan bila diketahui hipotesis sebagai berikut :*

$$\begin{aligned} H_0 &: \mu = 0 \\ H_1 &: \mu = 1 \end{aligned}$$

dengan $\alpha = \beta = 0.01$.

1126 **Jawab.**

$$\begin{aligned} k_0 &= \frac{0.01}{1-0.01} = \frac{1}{99} \\ k_1 &= \frac{1-0.01}{0.01} = 99 \end{aligned}$$

$$\begin{aligned} Z &= \ln \frac{f(X;\theta')}{f(X;\theta'')} \\ &= \ln \left[\frac{\exp\left(-\frac{1}{2}X^2\right)}{\exp\left(-\frac{1}{2}[X-1]^2\right)} \right] \\ &= -\frac{1}{2}\left[X^2 - (X-1)^2\right] \\ &= \frac{1}{2} - X \end{aligned}$$

1127

$$\begin{aligned} E(Z|H_0) &= E\left(\frac{1}{2} - X \mid H_0\right) \\ &= \frac{1}{2} - E(X \mid \mu = 0) \\ &= \frac{1}{2} \end{aligned}$$

$$\begin{aligned} E(Z|H_1) &= E\left(\frac{1}{2} - X \mid H_1\right) \\ &= \frac{1}{2} - E(X \mid \mu = 1) \\ &= \frac{1}{2} - 1 \\ &= -\frac{1}{2} \end{aligned}$$

1128

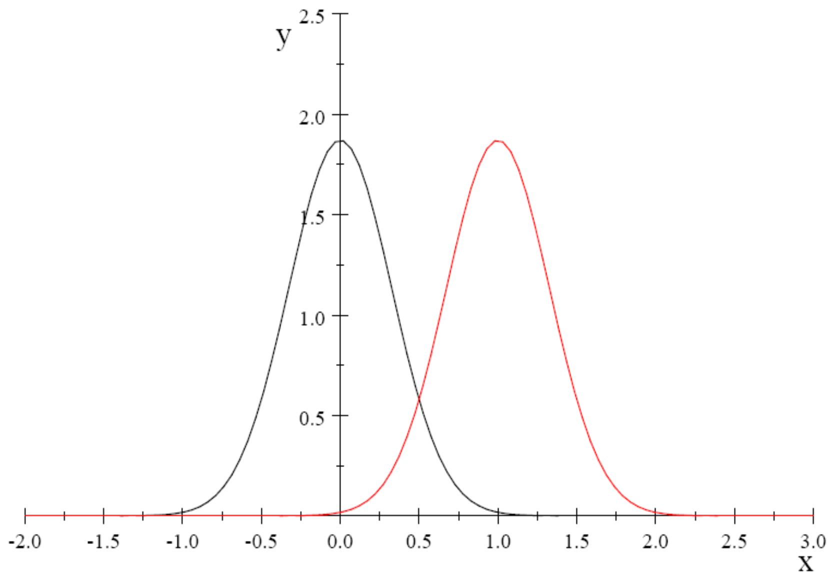
$$\begin{aligned} E(N|H_0) &= \frac{1}{1/2}\left[0.01 \ln\left(\frac{1}{99}\right) + 0.99 \ln(99)\right] = 9 \\ E(N|H_1) &= \frac{1}{-1/2}\left[0.99 \ln\left(\frac{1}{99}\right) + 0.01 \ln(99)\right] = 9 \end{aligned}$$

Sebagai pembanding, gunakan uji terbaiknya.

$$
\begin{aligned}
\alpha &= P\left(\bar{X} > k \mid H_0\right) \\
0.01 &= P\left(\frac{\bar{X}}{\sqrt{\frac{1}{n}}} > \frac{k}{\sqrt{\frac{1}{n}}}\right) \\
0.01 &= P\left(Z > \frac{k}{\sqrt{\frac{1}{n}}}\right) \\
0.01 &= P\left(Z > 2.325\right),
\end{aligned}
$$

sehingga

$$
\begin{aligned}
\frac{k}{\sqrt{\frac{1}{n}}} &= 2.325 \\
\frac{\frac{1}{2}}{\sqrt{\frac{1}{n}}} &= 2.325 \\
n &= 22.
\end{aligned}
$$



✠

6.5 Latihan

1. Misalkan $X_1, X_2, \ldots, X_n$ merupakan sampel acak dari sebaran $N(\theta, 25)$.

(a) Tunjukkanlah bahwa $C = \{(x_1, x_2, \ldots, x_n) : \frac{\Sigma_{i=1}^{n} x_i}{n} \geq c\}$ merupakan daerah kritis terbaik untuk pengujian $H_0 : \theta = 10$ lawan $H_1 : \theta = 12$.

(b) Tentukan pula n dan c sehingga $\alpha = 0.05$ dan $1 - \beta = 0.90$.

Catatan:

Bila $Z\tilde{}N(0, 1)$, maka $P(Z \geq 1.645) = 0.05$ dan $P(Z \geq 1.285) = 0.10$.

2. Misalkan $X_1, X_2, \ldots, X_{10}$ merupakan sampel acak dari sebaran $N(0, \theta)$ untuk pengujian

$$H_0 : \theta = 5 \text{ lawan } H_1 : \theta > 5.$$

(a) Tunjukkanlah bahwa ada Uji Selalu Paling Kuasa (UMPT, *Uniformly Most Powerful Test*) dalam pengujian ini.

(b) Bila $\alpha = 0.05$ maka berikanlah daerah kritis dan fungsi kuasanya.

3. Misalkan $X_1, X_2, \ldots, X_n$ merupakan sampel acak dari sebaran $N(5, \theta)$, $\theta > 0$, untuk pengujian

$$H_0 : \theta = \theta_0 \text{ lawan } H_1 : \theta \neq \theta_0$$

(a) Apakah dapat diperoleh uji selalu paling kuasa (*Uniformly Most Powerful Test*)? Jelaskan.

(b) Tunjukkanlah bahwa uji rasio kemungkinan (LRT, *Likelihood Ratio Test*) $H_0 : \theta = \theta_0$ lawan $H_1 : \theta \neq \theta_0$ dapat didasarkan pada statistik $W = \Sigma_{i=1}^{n}(X_i - 5)^2/\theta_0$. Tentukan sebaran dari W bila H_0 benar dan berikan daerah kritisnya untuk taraf nyata (*level of significance*) α.

4. Berikut fungsi massa peluang peubah acak Z dengan $0 < \theta < 1$:

Amatan (z)	$f(z;\theta_0)$	$f(z;\theta)$	$f(z;\theta_0)/\sup f(z;\theta)$
z_1	$1/12$	$\theta/3$	$\cdots$
z_2	$1/12$	$(1/3)(1-\theta)$	$\cdots$
z_3	$1/6$	$1/2$	$\cdots$
z_4	$2/3$	$1/6$	$\cdots$

Isilah nilai $f(z;\theta_0)/\sup f(z;\theta)$ pada tabel di atas dan berikanlah uji rasio kemungkinan (LRT, *Likelihood Ratio Test*) untuk pengujian $H_0 : \theta = \theta_0$ lawan $H_1 : \theta \neq \theta_0$ dengan amatan tunggal untuk $\alpha = 1/6$. Berikan pula kuasa uji-nya $(1 - \beta)$.

5. Misalkan peubah acak X bersebaran $N(0, \theta)$. Berikanlah uji rasio peluang bersekuens (SPRT, *Sequential Probability Ratio Test*) bagi $H_0 : \theta = 4$ lawan $H_1 : \theta = 9$ dengan $\alpha = 0.05$ dan $\beta = 0.10$. Berikan pula rata-rata ukuran contoh yang diharapkan untuk sampai pada suatu putusan.

LAMPIRAN

1. Terdapat 7 karyawan di departemen Aktuaria dan 6 karyawan di departemen *Market Management*. Akan dibentuk suatu tim khusus untuk melakukan analisa terhadap suatu produk asuransi yang akan diluncurkan. Hitunglah banyaknya cara untuk membentuk tim khusus dengan memilih 3 karyawan dari departemen Aktuaria dan 2 karyawan dari departemen *Market Management*.

   Jawab:

   Banyaknya cara untuk membentuk tim khusus dengan memilih 3 karyawan dari departemen Aktuaria dan 2 karyawan dari departemen *Market Management* adalah

$$\begin{pmatrix} 7 \\ 3 \end{pmatrix} \times \begin{pmatrix} 6 \\ 2 \end{pmatrix} = 35 \times 15 = 525.$$

2. Dari soal nomor 1, hitunglah probabilitas bahwa tim khusus dari 5 karyawan yang dipilih secara acak tersebut, terdiri dari 3 karyawan dari departemen Aktuaria dan 2 karyawan dari departemen *Market Management*.

   Jawab:

$$\begin{aligned}
\binom{7}{0} \times \binom{6}{5} &= 1 \times 6 = 6. \\
\binom{7}{1} \times \binom{6}{4} &= 7 \times 15 = 105. \\
\binom{7}{2} \times \binom{6}{3} &= 21 \times 20 = 420. \\
\binom{7}{3} \times \binom{6}{2} &= 35 \times 15 = 525. \\
\binom{7}{4} \times \binom{6}{1} &= 35 \times 6 = 210. \\
\binom{7}{5} \times \binom{6}{0} &= 21 \times 1 = 21.
\end{aligned}$$

   Jadi probabilitas bahwa tim khusus dari 5 karyawan yang dipilih secara acak tersebut, terdiri dari 3 karyawan dari departemen Aktuaria dan 2 karyawan dari departemen *Market Management* adalah

$$\frac{525}{6 + 105 + 420 + 525 + 210 + 21} = \frac{525}{1287} = 0,41.$$

3. Di sebuah perusahaan asuransi terdapat 3 sub bagian di bagian *product development* yang akan diisi oleh 7 karyawan baru yang telah memenuhi seleksi. Sub bagian pertama tersedia 2 posisi, sub bagian kedua tersedia 2 posisi, dan sub bagian ketiga tersedia 3 posisi. Hitunglah banyaknya cara yang mungkin untuk menempatkan karyawan tersebut ke dalam posisi yang tersedia.

   Jawab:

   Banyaknya cara yang mungkin untuk menempatkan karyawan tersebut ke dalam posisi yang tersedia adalah

$$\begin{pmatrix} & 7 & \\ 2 & 2 & 3 \end{pmatrix} = \frac{7!}{2!2!3!} = 210.$$

4. Misalkan N adalah jumlah klaim per bulan, dengan

$$\Pr(N = n) = \frac{1}{(n+1)(n+2)}$$

dimana $0 \leq n$. Tentukanlah probabilitas dari sedikitnya satu klaim selama bulan tertentu, diberikan paling banyak empat klaim selama bulan tersebut.

Jawab:

Probabilitas dari sedikitnya satu klaim selama bulan tertentu adalah

$$\begin{aligned}
\Pr(N \geq 1 | N \leq 4) &= \frac{\Pr(1 \leq N \leq 4)}{\Pr(N \leq 4)} \\
&= \frac{\Sigma_1^4 \frac{1}{(n+1)(n+2)}}{\Sigma_0^4 \frac{1}{(n+1)(n+2)}} \\
&= \frac{2}{5} \\
&= 0,4.
\end{aligned}$$

5. Misalkan kejadian A dan B saling bebas dengan $\Pr(A \cap B^c) = 0,2$ dan $\Pr(A^c \cap B) = 0,3$.
   Hitunglah $\Pr(A \cup B)$.

   Jawab:

   Karena kejadian A dan B saling bebas maka diperoleh

$$\begin{aligned}
\Pr(A \cap B) &= \Pr(A) \cdot \Pr(B) \\
\Pr(A \cap B) &= (\Pr(A \cap B^c) + \Pr(A \cap B))(\Pr(A^c \cap B) + \Pr(A \cap B)) \\
x &= (0,2 + x)(0,3 + x) \\
x &= 0,06 + 0,5x + x^2 \\
0 &= x^2 - 0,5x + 0,06.
\end{aligned}$$

   Nilai $\Pr(A \cap B) = x$ yang memenuhi persamaan di atas adalah $0,2$ atau $0,3$. Karena $\Pr(A \cup B) = \Pr(A \cap B^c) + \Pr(A^c \cap B) + \Pr(A \cap B)$ maka nilai $\Pr(A \cup B)$ yang mungkin adalah 0,7 atau 0,8.

6. Sebuah perusahaan asuransi menentukan bahwa N, jumlah klaim yang diterima dalam satu minggu, memenuhi $\Pr[N = n] = \frac{1}{2^{n+1}}$, dimana $n \geq 0$. Perusahaan juga menentukan bahwa jumlah klaim yang diterima dalam satu minggu adalah saling bebas dari jumlah klaim yang diterima dalam minggu yang lain. Tentukanlah probabilitas bahwa tepatnya tujuh klaim akan diterima selama periode dua minggu.

   Jawab:

   Probabilitas bahwa tepatnya tujuh klaim akan diterima selama periode dua minggu adalah

$$
\begin{aligned}
&2(\Pr(N=0)\Pr(N=7)+\Pr(N=1)\Pr(N=6)+\Pr(N=2)\Pr(N=5)\\
&+\Pr(N=3)\Pr(N=4))\\
&=2(\tfrac{1}{2^1}\tfrac{1}{2^8}+\tfrac{1}{2^2}\tfrac{1}{2^7}+\tfrac{1}{2^3}\tfrac{1}{2^6}+\tfrac{1}{2^4}\tfrac{1}{2^5})\\
&=2\times 4\times\tfrac{1}{2^9}\\
&=\tfrac{1}{64}.
\end{aligned}
$$

7. Sebuah polis asuransi memberikan manfaat individu sebesar 100 per hari untuk biaya perawatan sampai 3 hari dan 25 per hari untuk biaya perawatan setelah 3 hari seterusnya. Jumlah hari perawatan, X, adalah variabel acak diskrit dengan fungsi probabilitas

$$P(X = k) = \begin{cases} \frac{6-k}{15} & \text{untuk } k = 1, 2, 3, 4, 5 \\ 0 & \text{lainnya} \end{cases}$$

Hitunglah nilai variansi pembayaran biaya perawatan dari polis asuransi tersebut.

Jawab:

Misal B merupakan variabel acak yang menyatakan biaya perawatan, maka diperoleh:

$$B(x) = \begin{cases} 100x & ; \quad x = 1, 2, 3 \\ 25x & ; \quad x = 4, 5 \end{cases}$$

Jadi,

$$\begin{aligned}
E[B(x)] \quad & = \Sigma_{k=1}^{k=5} B(k)P(X = k) \\
& = 100 \cdot (1 \cdot \tfrac{6-1}{15} + 2 \cdot \tfrac{6-2}{15} + 3 \cdot \tfrac{6-3}{15}) + 25 \cdot (4 \cdot \tfrac{6-4}{15} + 5 \cdot \tfrac{6-5}{15}) \\
& = 168,33. \\
E[B^2(x)] \quad & = \Sigma_{k=1}^{k=5} B^2(k)P(X = k) \\
& = 100^2 \cdot (1^2 \cdot \tfrac{6-1}{15} + 2^2 \cdot \tfrac{6-2}{15} + 3^2 \cdot \tfrac{6-3}{15}) + 25^2 \cdot (4^2 \cdot \tfrac{6-4}{15} + 5^2 \cdot \tfrac{6-5}{15}) \\
& = 34375. \\
Var[B(x)] \quad & = E[B^2(x)] - (E[B(x)])^2 \\
& = 34375 - (168,33)^2 \\
& = 6038,9.
\end{aligned}$$

8. Misalkan X adalah variabel acak diskrit dengan distribusi probabilitas sebagai berikut:

x	2	4	6	7	12
$\Pr(X = x)$	3/7	1/7	1/7	1/7	1/7

Hitunglah nilai persentil ke-60 dari X.

Jawab:

Karena $60 = \frac{100i}{8}$ sehingga $i = 4,8$. Berdasarkan data di atas, dengan mengasumsikan bahwa data di atas sebanyak 7 data, kita ketahui bahwa data ke-empat,$x_4 = 4$, dan data ke-lima, $x_5 = 6$. Oleh karena itu persentil ke-60 dari data di atas adalah 80% dari 4 ke 6 atau

$$4 + (6 - 4) * (0,8) = 5,6.$$

9. Distribusi probabilitas dari ukuran klaim untuk sebuah polis asuransi diberikan dalam tabel berikut.

Ukuran klaim	20	30	40	50	60	70	80
Probabilitas	0,15	0,10	0,05	0,20	0,10	0,10	0,30

Tentukanlah persentase dari klaim yang terletak dalam nilai satu standar deviasi dari rata-rata ukuran klaim.

Jawab:

Rata-rata ukuran klaim

$$\begin{aligned} &= 20 \cdot 0,15 + 30 \cdot 0,10 + 40 \cdot 0,05 + 50 \cdot 0,20 + 60 \cdot 0,10 + 70 \cdot 0,10 + 80 \cdot 0,30 \\ &= 3 + 3 + 2 + 10 + 6 + 7 + 24 \\ &= 55 \end{aligned}$$

Rata-rata ukuran klaim kuadrat

$$\begin{aligned} &= 20^2 \cdot 0,15 + 30^2 \cdot 0,10 + 40^2 \cdot 0,05 + 50^2 \cdot 0,20 + 60^2 \cdot 0,10 + 70^2 \cdot 0,10 + 80^2 \cdot 0,30 \\ &= 60 + 90 + 80 + 500 + 360 + 490 + 1920 \\ &= 3500 \end{aligned}$$

Standar deviasi ukuran klaim $= \sqrt{3500 - 3025} = \sqrt{475} = 21,8$ Jadi persentase dari klaim yang terletak dalam nilai satu standar deviasi dari rata-rata ukuran klaim (ukuran klaim yang terletak pada interval [55-21,8;22+21,8]=[33,2;76,8]) adalah 45%.

10. Misalkan dipilih sebuah angka W yang merupakan variabel acak dari 50, 55, 60, 65, 70. Hitunglah $\Pr[W \geq 55]$.
    Petunjuk: $W = 5X + 45$, dimana X mempunyai distribusi uniform diskrit pada 1, 2, 3, 4, 5.

    Jawab:

    Karena X mempunyai distribusi uniform diskrit pada 1, 2, 3, 4, 5, maka
$$\Pr[X = k] = \frac{1}{5} \text{ untuk } k = 1, 2, 3, 4, 5$$
$$\begin{aligned} \Pr[W \geq 55] &= 1 - \Pr[W < 55] \\ &= 1 - \Pr[W = 50] \\ &= 1 - \Pr[5X + 45 = 50] \\ &= 1 - \Pr[X = 1] \\ &= 1 - \tfrac{1}{5} \\ &= \tfrac{4}{5}. \end{aligned}$$

11. Misalkan X adalah variabel acak yang mempunyai distribusi binomial dengan $n = 10$ dan $Var[X] = 0,25E[X]$.
    Hitunglah probabilitas untuk $X = 7$.

    Jawab:

    Karena X adalah variabel acak yang mempunyai distribusi binomial dengan $n = 10$ maka $E[X] = 10p$ dan $Var[X] = 10p(1-p)$ dengan p merupakan peluang sukses.
    Karena $Var[X] = 0,25E[X]$ maka diperoleh $10p(1-p) = 0,25 \cdot 10p$.
    Nilai p yang memenuhi adalah $p = 0,75$.
    Jadi
    $$\Pr[X = 7] = \binom{10}{7}(0,75)^7(0,25)^3 = 0,25.$$

12. Misalkan $Z \sim Poisson(\lambda)$ dan $E[Z^2] = 12$.
Hitunglah $\Pr(Z = 1)$.

Jawab:

Pertama akan ditentukan parameter λ.
Karena $Z \sim Poisson(\lambda)$ maka $E[Z] = Var[Z] = \lambda$.
Karena $Var[Z] = E[Z^2] - E[Z]^2$ maka diperoleh $\lambda^2 + \lambda - 12 = 0$ atau $\lambda = 3$.
Jadi,

$$\Pr[Z = 1] = \frac{3^1 e^{-3}}{1!} = 0,149.$$

13. Masa beroperasi dari sebuah bagian mesin mempunyai distribusi kontinu pada interval $[0, 40]$ dengan fungsi probabilitas densitas, $f(x)$, yang proporsional dengan $(10 + x)^{-2}$. Hitunglah probabilitas untuk masa beroperasi dari sebuah bagian mesin yang lebih kecil dari 5.

    Jawab:

    Misal $f(x) = k(10 + x)^{-2}$ untuk x pada interval $[0, 40]$.
    Karena $\int_0^{40} k(10 + x)^{-2} dx = 1$ maka diperoleh $k = 12, 5$.
    Jadi

$$\begin{aligned}
\Pr[X < 5] &= \int_0^5 12,5(10 + x)^{-2} dx \\
&= -\frac{12,5}{10 + x}\bigg|_0^5 \\
&= -0,83 + 1,25 \\
&= 0,42.
\end{aligned}$$

14. Misalkan X adalah variabel acak yang mempunyai fungsi distribusi kumulatif

$$F(x) = \begin{cases} 0 & , x < 1 \\ \frac{x^2-2x+2}{2} & , 1 \leq x < 2 \\ 1 & , x \geq 2 \end{cases}$$

Hitunglah variansi dari X.

Jawab:

Berdasarkan fungsi distribusi kumulatif di atas, maka diperoleh fkp sebagai berikut:

$$f(x) = x - 1, \quad 1 \leq x < 2$$

$$\begin{aligned} E[X] &= \int_1^2 x(x-1)dx \\ &= \int_1^2 (x^2 - x)dx \\ &= \left(\frac{1}{3}x^3 - \frac{1}{2}x^2\right)\Bigg|_1^2 \\ &= \frac{7}{3} - \frac{3}{2} \\ &= \frac{5}{6} \end{aligned}$$

$$\begin{aligned} E[X^2] &= \int_1^2 x^2(x-1)dx \\ &= \int_1^2 (x^3 - x^2)dx \\ &= \left(\frac{1}{4}x^4 - \frac{1}{3}x^3\right)\Bigg|_1^2 \\ &= \frac{15}{4} - \frac{7}{3} \\ &= \frac{17}{12} \end{aligned}$$

$$\begin{aligned} Var[X] &= E[X^2] - (E[X])^2 \\ &= \frac{17}{12} - \frac{25}{36} \\ &= \frac{13}{18}. \end{aligned}$$

15. Misalkan X_1, X_2, X_3 adalah sampel acak dari distribusi diskrit dengan fungsi probabilitas

$$p(x) = \begin{cases} \frac{1}{3} & , x = 0 \\ \frac{2}{3} & , x = 1 \\ 0 & , lainnya \end{cases}$$

Tentukanlah fungsi pembangkit momen, $M_Y(t)$, dari perkalian $Y = X_1 \cdot X_2 \cdot X_3$.

Jawab:

X_1	X_2	X_3	e^{tY}	$p(y)$	$e^{tY}p(y)$
0	0	0	1	$\frac{1}{27}$	$\frac{1}{27}$
0	0	1	1	$\frac{2}{27}$	$\frac{2}{27}$
0	1	0	1	$\frac{2}{27}$	$\frac{2}{27}$
0	1	1	1	$\frac{4}{27}$	$\frac{4}{27}$
1	0	0	1	$\frac{2}{27}$	$\frac{2}{27}$
1	0	1	1	$\frac{4}{27}$	$\frac{4}{27}$
1	1	0	1	$\frac{4}{27}$	$\frac{4}{27}$
1	1	1	e^t	$\frac{8}{27}$	$e^t \frac{8}{27}$

Jadi diperoleh

$$E[e^{tY}] = \Sigma e^{ty} p(y) = \frac{19}{27} + \frac{8}{27} e^t.$$

16. Pemilik kendaraan mengasuransikan kendaraannya dari kerusakan dengan membeli sebuah polis asuransi dengan pemotongan 250 dari total klaim yang dibayar. Dalam kejadian bahwa kendaraan mengalami kerusakan, biaya perbaikan dapat dimodelkan dengan variabel acak uniform pada interval [0,1500]. Hitunglah standar deviasi dari manfaat polis asuransi untuk kejadian bahwa kendaraan tersebut mengalami kerusakan.

    Jawab:

    Misal X adalah variabel acak yang menyatakan biaya perbaikan, $X \sim uniform[0, 1500]$ sehingga $f(x) = \frac{1}{1500},\ x \in [0, 1500]$.

$$\begin{aligned} E[X] &= \int_0^{1500} \frac{1}{1500} x dx \\ &= \frac{1}{3000} x^2 |_0^{1500} \\ &= \frac{1}{3000} 1500^2 \\ &= 750. \end{aligned}$$

$$\begin{aligned} E[X^2] &= \int_0^{1500} \frac{1}{1500} x^2 dx \\ &= \frac{1}{4500} x^3 |_0^{1500} \\ &= \frac{1}{4500} 1500^3 \\ &= 750000. \end{aligned}$$

$$\begin{aligned} \sigma(X) &= \sqrt{E[X^2] - (E[X])^2} \\ &= \sqrt{750000 - 750^2} \\ &= 433. \end{aligned}$$

17. Misalkan X adalah variabel acak yang berdistribusi eksponensial dengan variansi 9. Hitunglah $E(X^4)$.

    Jawab:

    Karena X merupakan peubah acak yang berdistribusi eksponensial dengan variansi 9 maka diperoleh parameter λ sebagai berikut

$$Var[X] = \frac{1}{\lambda^2} = 9$$

    sehingga diperoleh $\lambda = \frac{1}{3}$ dan fkp peubah acak X:

$$f(x) = \frac{1}{3}e^{-\frac{1}{3}x}$$

$$\begin{aligned} E[X^4] &= \int_0^\infty x^4 \frac{1}{3} e^{-\frac{1}{3}x} dx \\ &= -\frac{1}{3}\left(3x^4 + 36x^3 + 324x^2 + 1944x + 5832\right) e^{-\frac{1}{3}x}\Big|_0^\infty \\ &= \frac{1}{3} \times 5832 \\ &= 1944. \end{aligned}$$

18. Misalkan X adalah variabel acak dengan mean μ_x dan variansi σ_x^2. X disebut sebagai variabel acak populasi. Misalkan $X_1, X_2, ...X_n$ adalah pengamatan-pengamatan yang berdistribusi identik dan saling bebas dari X.
    $X_1, X_2, ...X_n$ adalah sampel acak dari X. Misalkan $S = \Sigma_{i=1}^n X_i$ dan $\bar{X} = \frac{1}{n}\Sigma_{i=1}^n X_i$. $\bar{X}$ adalah mean dari sampel.
    Manakah dari pernyataan berikut yang benar.

    (i) $E(S) = \mu_X$

    (ii) $Var(S) = n \cdot \sigma_X^2$

    (iii) $E(X) = n \cdot \mu_X$

    (iv) $Var(\bar{X}) = \frac{\sigma_X^2}{n}$

    Jawab:

    Pernyataan yang benar adalah (ii) dan (iv).

19. Sebuah koin yang kedua sisinya seimbang dilempar sebanyak dua kali. Misalkan X adalah banyaknya sisi pertama yang terlihat pada lemparan pertama. Misalkan Y adalah banyaknya sisi pertama yang terlihat pada dua lemparan pertama. Hitunglah $Var(X+Y)$.

    Jawab:

Luaran	X	Y	$Z = X + Y$	$\Pr(Z = z)$
S_1S_1	1	2	3	1/4
S_1S_2	1	1	2	1/4
S_2S_1	0	1	1	1/4
S_2S_2	0	0	0	1/4

    $E(Z) = \frac{1}{4}(0+1+2+3) = 3/2$
    $E(Z^2) = \frac{1}{4}(0+1+4+9) = 7/2$
    $Var(X+Y) = Var(Z) = E(Z^2) - (E(Z))^2 = \frac{7}{2} - (\frac{3}{2})^2 = \frac{5}{4} = 1,25.$

20. Dari soal nomor 19, hitunglah $E(Y|X = 1)$.

    Jawab:

$x \downarrow, y \longrightarrow$	0	1	2	$p_X(x)$
0	1/4	1/4	0	1/2
1	0	1/4	1/4	1/2
$p_Y(y)$	1/4	1/2	1/4	1

$$
\begin{aligned}
E(Y|X=1) &= \Sigma y \cdot p(y|x=1) \\
&= 0 \cdot p(0|x=1) + 1 \cdot p(1|x=1) + 2 \cdot p(2|x=1) \\
&= 0 \cdot \tfrac{p(1,0)}{p_X(1)} + 1 \cdot \tfrac{p(1,1)}{p_X(1)} + 2 \cdot \tfrac{p(1,2)}{p_X(1)} \\
&= 0 \cdot \tfrac{0}{1/2} + 1 \cdot \tfrac{1/4}{1/2} + 2 \cdot \tfrac{1/4}{1/2} \\
&= 0 + \tfrac{1}{2} + 1 \\
&= \tfrac{3}{2} = 1,50.
\end{aligned}
$$

21. Misalkan N_1 dan N_2 mewakili banyaknya klaim yang diajukan ke sebuah perusahaan asuransi di bulan April dan Mei. Fungsi probabilitas bersama dari N_1 dan N_2 adalah

$$p(n_1, n_2) = \begin{cases} \frac{1}{3}(\frac{1}{4})^{n_1-1}e^{-n_1}(1-e^{-n_1})^{n_2-1} & , n_1 = 1,2,3,...\ n_2 = 1,2,3,... \\ 0 & , lainnya \end{cases}$$

Hitunglah nilai ekspektasi banyaknya klaim yang akan diajukan ke perusahaan asuransi di bulan Mei jika tepatnya 2 klaim diajukan di bulan April.

Jawab:

$$\begin{aligned} E[N_2|N_1=2] &= \Sigma_{n_2=1}^{\infty} n_2\ p(n_2|N_1=2) \\ &= \Sigma_{n_2=1}^{\infty} n_2\ \frac{p(2,n_2)}{p(N_1=2)} \end{aligned}$$

$$\begin{aligned} p(N_1=2) &= \Sigma_{n_2=1}^{\infty} \left(\frac{1}{3}\frac{1}{4}e^{-2}(1-e^{-2})^{n_2-1}\right) \\ &= \frac{1}{12e^2}\Sigma_{n_2=0}^{\infty}(1-e^{-2})^{n_2} \\ &= \frac{1}{12e^2}\frac{1}{1-(1-e^{-2})} \\ &= \frac{1}{12} \end{aligned}$$

Oleh karena itu diperoleh

$$\begin{aligned} E[N_2|N_1=2] &= \Sigma_{n_2=1}^{\infty} n_2\ \frac{\frac{1}{3}\frac{1}{4}e^{-2}(1-e^{-2})^{n_2-1}}{\frac{1}{12}} \\ &= \Sigma_{n_2=1}^{\infty} n_2\ e^{-2}(1-e^{-2})^{n_2-1} \\ &= \Sigma_{n_2=0}^{\infty} n_2\ e^{-2}(1-e^{-2})^{n_2} \end{aligned}$$

Perhatikan bahwa nilai ekspektasi di atas adalah nilai ekspektasi untuk peubah acak geometri N_2. Maka diperoleh

$$\begin{aligned} E[N_2|N_1=1] &= \frac{1-e^{-2}}{e^{-2}} \\ &= e^2 - 1. \end{aligned}$$

22. Sebuah model aktuaria untuk lama penggunaan peralatan menggunakan variabel acak $Y = 10 \cdot X^{0,8}$, dimana X adalah variabel acak yang berdistribusi eksponensial dengan mean 1 tahun.
    Tentukanlah fungsi probabilitas densitas $f(y)$, untuk $y > 0$, dari variabel acak Y.

    Jawab:

    Karena $X \sim \exp(\lambda = 1)$ maka diperoleh $f_X(x) = e^{-x}$.
    Karena $Y = 10 \cdot X^{0,8}$ adalah fungsi naik, maka

$$
\begin{aligned}
f(y) = f_Y(y) &= f_X(x)\tfrac{dx}{dy} \\
&= e^{-x}\tfrac{dx}{dy} \\
&= e^{-(0,1y)^{1,25}}(0,125(0,1y)^{0,25}) \\
&= 0,125(0,1y)^{0,25}exp(-(0,1y)^{1,25}).
\end{aligned}
$$

23. Misalkan variabel acak X berdistribusi uniform pada interval [2,6]. Misalkan variabel acak $Y = \frac{1}{X}$. Tentukanlah fungsi probabilitas densitas untuk variabel acak Y pada interval [1/6,1/2].

Jawab:

Karena $X \sim uniform[2,6]$ maka diperoleh $f_X(x) = \frac{1}{4},\ x \in [2,6]$.
Karena $Y = \frac{1}{X}$ adalah fungsi turun maka

$$\begin{aligned} f(y) = f_Y(y) &= -f_X(x)\frac{dx}{dy} \\ &= -\frac{1}{4}\frac{dx}{dy} \\ &= -\frac{1}{4}\left(-\frac{1}{y^2}\right) \\ &= \frac{1}{4y^2}. \end{aligned}$$

24. Misalkan S adalah variabel acak yang berdistribusi eksponensial dengan mean 1 dan T adalah variabel acak yang berdistribusi eksponensial dengan mean 2. Tentukanlah fungsi probabilitas densitas untuk $f_X(x)$, dimana $X = S + T$.

    Jawab:

    Perhatikan bahwa

    $$S \sim f_S(s) = e^{-s}$$

    dan

    $$T \sim f_T(t) = \frac{1}{2}e^{-\frac{1}{2}t}.$$

    Sehingga diperoleh

    $$\begin{aligned}
    &f_X(x) \\
    &= \int_0^x f_S(s) f_T(x-s)ds \\
    &= \int_0^x e^{-s}\tfrac{1}{2}e^{-\frac{1}{2}(x-s)}ds \\
    &= \tfrac{1}{2}e^{-\frac{1}{2}x}\int_0^x e^{-\frac{1}{2}s}ds \\
    &= e^{-\frac{1}{2}x}(1-e^{-\frac{1}{2}x}) \\
    &= e^{-\frac{1}{2}x} - e^{-x}.
    \end{aligned}$$

25. Suatu populasi mempunyai mean 100 dan standar deviasi 16. Hitunglah probabilitas bahwa mean sampel berada dalam interval ± 3 dari mean populasi jika ukuran sampel $n = 100$.

    Jawab:

$$
\begin{aligned}
&\Pr(|\bar{X} - 100| \leq 3) \\
&= \Pr(-3 \leq \bar{X} - 100 \leq 3) \\
&= \Pr(\tfrac{-3}{16/\sqrt{100}} \leq \tfrac{\bar{X}-100}{16/\sqrt{100}} \leq \tfrac{3}{16/\sqrt{100}}) \\
&= \Pr(\tfrac{-3}{16/\sqrt{100}} \leq Z \leq \tfrac{3}{16/\sqrt{100}}) \\
&= \Pr(|Z| \leq \tfrac{3\sqrt{100}}{16}) \\
&= \Pr(|Z| \leq 1,875) \\
&= 2\ Pr(0 \leq Z \leq 1,875) \\
&= 0,9392.
\end{aligned}
$$

26. Variabel acak Y diketahui mempunyai distribusi normal. Sampel acak berukuran $n = 14$ menghasilkan nilai $\bar{Y} = -43,2$ dan $S_Y = 17,9$. Hitunglah interval kepercayaan 98% untuk mean populasi.

    Jawab:

$$\varepsilon = t_{\alpha/2}(n-1)\frac{S_Y}{\sqrt{n}} \longrightarrow \varepsilon = t_{0,01}(13)\frac{17,9}{\sqrt{14}} = 12,7.$$

    Jadi interval kepercayaan 98% untuk mean populasi adalah

$$(-43,2 - 12,7; -43,2 + 12,7) = (-55,9; -30,5).$$

27. Sampel acak dari 9 pengamatan diambil dari populasi normal menghasilkan statistik

$$\bar{X} = 5 \quad \text{dan} \quad \frac{1}{8}\Sigma_{i=1}^{9}(X_i - \bar{X})^2 = 36.$$

Tentukanlah batas kepercayaan 95% simetris untuk mean populasi.

Jawab:

$$\varepsilon = t_{\alpha/2}(n-1)\frac{S_X}{\sqrt{n}} \longrightarrow \varepsilon = t_{0,025}(8)\frac{6}{\sqrt{9}} = 4,6.$$

Jadi batas kepercayaan 95% simetris untuk mean populasi adalah $5 \pm 4,6$.

28. Diberikan informasi sampel acak sebagai berikut:

    - $Y = X_1 + ... + X_n$ dimana ukuran sampel n sama dengan 25 dan variabel acak saling bebas dan berdistribusi identik.
    - X_i berdistribusi Poisson dengan parameter λ
    - $H_0 : \lambda = 0,1$
    - $H_0 : \lambda < 0,1$
    - Daerah kritis untuk menolak H_0 adalah $Y \leq 2$

    Hitunglah tingkat signifikansi dari uji tersebut.

    Jawab:

    Tingkat signifikansi dari uji di atas adalah peluang menolak H_0 jika H_0 benar, yaitu diberikan sebagai berikut:

    $$P(Y \leq 2|\lambda = 0,1)$$

    Karena Y adalah jumlah dari peubah acak Poisson yang saling bebas dengan rata-rata 0,1 maka rata-rata Y adalah $25 \times (0,1) = 2,5$. Oleh karena itu diperoleh:

    $$P(Y \leq 2) = P(Y = 0; 1; 2) = e^{-2,5} + e^{-2,5}\frac{2.5}{1!} + e^{-2,5}\frac{2.5^2}{2!} = 0.5438.$$

    Jadi tingkat signifikansi uji di atas sedikitnya 0,50 dan lebih kecil dari 0,6.

29. Sampel acak dari 21 pengamatan dari distribusi normal memberikan hasil sebagai berikut:

$$\bar{X} = 3,5 \qquad \frac{1}{20}\Sigma_{i=1}^{21}(X_i - \bar{X})^2 = 0,445$$

Digunakan uji dua arah

$$\begin{aligned} H_0 &: \mu = 3 \\ H_1 &: \mu \neq 3 \end{aligned}$$

Hitunglah nilai p (probabilitas uji statistik) dari uji tersebut.

Jawab:

Nilai p (probabilitas uji statistik) dari uji di atas adalah:

$$\begin{aligned} P\left(t(n-1) \geq \frac{|\bar{X}-\mu|}{\frac{S}{\sqrt{n}}}\right) &= P\left(t(n-1) \geq \frac{0,5}{\frac{\sqrt{0,445}}{\sqrt{21}}}\right) \\ &= P(t(n-1) \geq 3,43) \\ &= 1 - P(t(n-1) \leq 3,43) \\ &= 1 - 0,997423 \\ &= 0,002577. \end{aligned}$$

30. Misalkan $X_1, ..., X_{12}$ adalah variabel-variabel acak yang masing-masing berdistribusi normal dengan mean 0 dan variansi $\sigma^2 > 0$. Dengan nilai p = 0,025 untuk uji hipotesis:

$$H_0 : \sigma^2 = 10$$
$$H_1 : \sigma^2 < 10$$

Tentukanlah daerah penolakan untuk uji hipotesis tersebut.

Jawab:

Perhatikan bahwa

$$\chi^2(n-1) = \frac{(n-1)S^2}{\sigma^2}.$$

Diperoleh

$$\chi^2(11) = \frac{\Sigma_{i=1}^{12}(X_i - 0)^2}{10} = \frac{\Sigma_{i=1}^{12}X_i^2}{10}.$$

Daerah penolakan untuk uji hipotesis di atas adalah

$$\chi^2(11) < \chi^2_{0,975}(11) = 3,816$$

atau

$$\Sigma_{i=1}^{12}X_i^2 < 38,16.$$

31. Kode peserta ujian dari suatu ujian profesi aktuaris terdiri dari satu *digit* angka, tiga huruf, dan 3 *digit* angka (contohnya 1SAM123), kecuali huruf I, O, dan Q tidak dimasukkan pada posisi *alphabet* pertama dan ketiga. Hitunglah banyaknya kode peserta ujian yang mungkin dari ujian profesi aktuaris tersebut.

    Jawab:

    Misalkan $K_1K_2K_3K_4K_5K_6K_7$ menyatakan kode yang mungkin dari ujian profesi aktuaris tersebut.
    Karena K_1, K_5, K_6 dan K_7 merupakan digit-digit angka maka 10 kemungkinan untuk masing-masing digit K_1, K_5, K_6 dan K_7.
    Sedangkan K_2 dan K_4 merupakan digit-digit huruf A,B,...,Z kecuali I,O dan Q, sehingga ada 23 kemungkinan untuk masing-masing digit K_2 dan K_4.
    Untuk K_3 mempunyai 26 kemungkinan.
    Dengan menggunakan prinsip perkalian maka diperoleh banyaknya kode yang mungkin dari ujian profesi aktuaris tersebut adalah:

    $$10 \times 23 \times 26 \times 23 \times 10 \times 10 \times 10 = 137.540.000$$

32. Banyaknya kombinasi dari r obyek yang dipilih dari kumpulan n obyek yang berbeda diberikan oleh $\binom{n}{r}$. Tentukanlah persamaan yang tepat dari $\binom{n}{r}$.

    Jawab:

$$
\begin{aligned}
\binom{n}{r} &= \frac{n!}{r!(n-r)!} \\
\binom{n}{r} &= \frac{(n-1)!n}{r!(n-r-1)!(n-r)} \\
\binom{n}{r} &= \frac{(n-1)!}{r!((n-1)-r)!}\frac{n}{n-r} \\
\binom{n}{r} &= \frac{(n-1)!}{r!((n-1)-r)!}\frac{(n-r)+r}{n-r} \\
\binom{n}{r} &= \frac{(n-1)!}{r!((n-1)-r)!}(1+\frac{r}{n-r}) \\
\binom{n}{r} &= \frac{(n-1)!}{r!((n-1)-r)!} + \frac{(n-1)!r}{r!((n-1)-r)!(n-r)} \\
\binom{n}{r} &= \frac{(n-1)!}{r!((n-1)-r)!} + \frac{(n-1)!r}{(r-1)!((n-r)-1)!r(n-r)} \\
\binom{n}{r} &= \frac{(n-1)!}{r!((n-1)-r)!} + \frac{(n-1)!}{(r-1)!(n-r)!} \\
\binom{n}{r} &= \frac{(n-1)!}{r!((n-1)-r)!} + \frac{(n-1)!}{(r-1)!((n-1)-(r-1))!} \\
\binom{n}{r} &= \binom{n-1}{r} + \binom{n-1}{r-1}
\end{aligned}
$$

33. Tentukanlah ekspansi dari $(x - 2y)^3$.

Jawab:

Berdasarkan Teorema Multinomial, diperoleh:

$$(ax+by)^n = \begin{pmatrix} n \\ 0 \end{pmatrix} a^n b^0 x^n y^0 + \begin{pmatrix} n \\ 1 \end{pmatrix} a^{n-1} b^1 x^{n-1} y^1 + ... + \begin{pmatrix} n \\ n \end{pmatrix} a^0 b^n x^0 y^n.$$

Jadi ekspansi dari $(x - 2y)^3$ adalah

$$\begin{aligned} (x-2y)^3 &= \begin{pmatrix} 3 \\ 0 \end{pmatrix} 1^3(-2)^0 x^3 y^0 + \begin{pmatrix} 3 \\ 1 \end{pmatrix} 1^2(-2)^1 x^2 y^1 + \begin{pmatrix} 3 \\ 2 \end{pmatrix} 1^1(-2)^2 x^1 y^2 \\ &\quad + \begin{pmatrix} 3 \\ 3 \end{pmatrix} 1^0(-2)^0 x^2 y^3 \\ &= x^3 - 6x^2y + 12xy^2 - 8y^3. \end{aligned}$$

34. Misalkan himpunan *universal* adalah $U = \{1, 2, \pi, Jamaal, water, gum\}$. Misalkan $A = \{1, 2, \pi\}$ dan $B = \{1, 2, Jamaal, gum\}$. Tentukanlah $(A \cup B)^c$.

    Jawab:

    Karena $A = \{1, 2, \pi\}$ dan $B = \{1, 2, Jamaal, gum\}$ maka

    $$A \cup B = \{1, 2, \pi, Jamaal, gum\}$$

    Oleh karena itu diperoleh

    $$\begin{aligned} (A \cup B)^c &= U - (A \cup B) \\ &= \{1, 2, \pi, Jamaal, water, gum\} - \{1, 2, \pi, Jamaal, gum\} \\ &= \{water\} \end{aligned}$$

35. Misalkan $\Pr(A \cup B) = 0,7$, $\Pr(A^c) = 0,6$, dan $\Pr(A \cap B^c) = 0,3$. Hitunglah $\Pr(B)$.

    Jawab:

    Kita ketahui bahwa

    $$\Pr(A) = 1 - \Pr(A^c) = 1 - 0,6 = 0,4.$$

    Karena $\Pr(A) = \Pr(A \cap B) + \Pr(A \cap B^c)$ maka diperoleh

    $$\Pr(A \cap B) = \Pr(A) - \Pr(A \cap B^c) = 0,4 - 0,3 = 0,1.$$

    Berdasarkan aturan peluang, diketahui:

    $$\begin{aligned} \Pr(A \cup B) &= \Pr(A) + \Pr(B) - \Pr(A \cap B) \\ 0,7 &= 0,4 + Pr(B) - 0,1 \end{aligned}$$

    sehingga diperoleh $\Pr(B) = 0,4$.

36. Dari soal nomor 35, hitunglah $\Pr(A \cap B)$.

    Jawab:

    Kita ketahui bahwa

$$\Pr(A) = 1 - \Pr(A^c) = 1 - 0,6 = 0,4.$$

Karena $\Pr(A) = \Pr(A \cap B) + \Pr(A \cap B^c)$ maka diperoleh

$$\Pr(A \cap B) = \Pr(A) - \Pr(A \cap B^c) = 0,4 - 0,3 = 0,1.$$

37. Sebuah polis asuransi memberikan manfaat individu sebesar 100 per hari untuk biaya perawatan sampai 3 hari dan 50 per hari untuk biaya perawatan setelah 3 hari seterusnya. Jumlah hari perawatan, X, adalah variabel acak diskrit dengan fungsi probabilitas

$$P(X=k)=\begin{cases} \frac{6-k}{15} & \text{untuk } k=1,2,3,4,5 \\ 0 & \text{lainnya} \end{cases}$$

Hitunglah nilai variansi pembayaran biaya perawatan dari polis asuransi tersebut.

Jawab:

Misal $B(x)$ menyatakan biaya perawatan ,

$$B(X=k)=\begin{cases} 100k & ,k=1,2,3 \\ 300+50(k-3) & ,k=4,5 \end{cases}$$

$$\begin{aligned} E[B(x)] &= \Sigma_{k=1}^{k=5} B(k)P(X=k) \\ &= 100\cdot\tfrac{6-1}{15}+200\cdot\tfrac{6-2}{15}+300\cdot\tfrac{6-3}{15}+350\cdot\tfrac{6-4}{15}+400\cdot\tfrac{6-5}{15} \\ &= 220. \end{aligned}$$

$$\begin{aligned} E[(B(x))^2] &= \Sigma_{k=1}^{k=5} (B(k))^2 P(X=k) \\ &= 100^2\cdot\tfrac{6-1}{15}+200^2\cdot\tfrac{6-2}{15}+300^2\cdot\tfrac{6-3}{15}+350^2\cdot\tfrac{6-4}{15}+400^2\cdot\tfrac{6-5}{15} \\ &= 59000. \end{aligned}$$

$$\begin{aligned} Var[B(x)] &= E[(B(x))^2]-(E[B(x)])^2 \\ &= 59000-220^2 \\ &= 10600. \end{aligned}$$

38. Misalkan X adalah variabel acak diskrit dengan distribusi probabilitas sebagai berikut:

x	2	4	6	7	12
$\Pr(X = x)$	6/11	2/11	1/11	1/11	1/11

Hitunglah nilai kuartil ketiga dari X.

Jawab:

Misalkan Q_3 menyatakan kuartil ketiga dari data di atas. Dengan mengasumsikan bahwa data di atas sebanyak 11 data, maka diperoleh:

$$Q_3 = x_{\frac{3}{4}\cdot 12} = x_9 = 6.$$

39. Distribusi probabilitas dari ukuran klaim untuk sebuah polis asuransi diberikan dalam tabel berikut.

Ukuran klaim	30	40	50	60	70	80	90
Probabilitas	0,15	0,10	0,05	0,20	0,10	0,1	0,30

Tentukanlah persentase dari klaim yang terletak dalam nilai satu standar deviasi dari *median* ukuran klaim.

Jawab:

Median ukuran klaim $= \frac{60+70}{2} = 65$. Rata-rata ukuran klaim

$$\begin{aligned} &= 30 \cdot 0,15 + 40 \cdot 0,10 + 50 \cdot 0,05 + 60 \cdot 0,20 + 70 \cdot 0,10 + 80 \cdot 0,10 + 90 \cdot 0,30 \\ &= 4,5 + 4 + 2,5 + 12 + 7 + 8 + 27 \\ &= 65 \end{aligned}$$

Rata-rata ukuran klaim kuadrat

$$\begin{aligned} &= 30^2 \cdot 0,15 + 40^2 \cdot 0,10 + 50^2 \cdot 0,05 + 60^2 \cdot 0,20 + 70^2 \cdot 0,10 + 80^2 \cdot 0,10 + 90^2 \cdot 0,30 \\ &= 135 + 160 + 125 + 720 + 490 + 640 + 2430 \\ &= 4700 \end{aligned}$$

Standar deviasi ukuran klaim $= \sqrt{4700 - 4225} = \sqrt{475} = 21,8$ Jadi persentase dari klaim yang terletak dalam nilai satu standar deviasi dari median ukuran klaim (ukuran klaim yang terletak pada interval [65-21,8;65+21,8]=[43,2;86,8]) adalah 45%.

40. Misalkan X adalah variabel acak yang berdistribusi *uniform* dari bilangan asli yang dipilih acak dari 1 sampai 100. Tentukanlah nilai modus dari X.

    Jawab:

    Karena $\Pr(X = x) = \frac{1}{100}$ maka peluang untuk setiap variabel acak adalah sama. Hal ini berarti modus dari X tidak ada.

41. Misalkan seorang peserta pertandingan memanah mempunyai kemampuan tepat mengenai sasaran adalah 65% dan mengambil $n = 5$ percobaan memanah sasaran. Misalkan X menunjukkan banyaknya percobaan memanah sasaran dimana X mempunyai distribusi *binomial*. Hitunglah $\Pr(X > E[X])$.

    Jawab:

    Misalkan p menyatakan peluang perserta tersebut sukses memanah sasaran, maka diperoleh $p = 65\% = 0,65$. Karena $E[X] = n \cdot p$ maka untuk $n = 5$ diperoleh $E[X] = 5 \cdot (0,65) = 3,25$.
    Jadi

$$\begin{aligned}
\Pr(X > E[X]) &= \Pr(X > 3,25) \\
&= \Pr(X = 4) + \Pr(X = 5) \\
&= \begin{pmatrix} 5 \\ 4 \end{pmatrix} (0,65)^4(1-0,65)^1 + \begin{pmatrix} 5 \\ 5 \end{pmatrix} (0,65)^5(1-0,65)^0 \\
&= 0,4284.
\end{aligned}$$

42. Misalkan X adalah variabel acak yang mempunyai distribusi geometrik dan standar deviasi $\sigma_X = \frac{\sqrt{15}}{2}$. Hitunglah $\Pr(X < E[X])$.

Jawab:

Misalkan p menyatakan peluang sukses. Karena $\sigma_X = \sqrt{\frac{1-p}{p^2}}$ dan $\sigma_X = \frac{\sqrt{15}}{2}$ maka diperoleh $15p^2+4p-4=0$. Sedangkan $0 \leq p \leq 1$ maka nilai p yang memenuhi adalah $p = \frac{2}{5} = 0,4$.
Karena $E[X] = \frac{1-p}{p}$ maka diperoleh $E[X] = 1,5$. Jadi

$$\begin{aligned}
\Pr(X < E[X]) &= \Pr(X < 1,5) \\
&= \Pr(X = 0) + \Pr(X = 1) \\
&= p(1-p)^0 + p(1-p)^1 \\
&= 0,4(1-0,4)^0 + 0,4(1-0,4)^1 \\
&= 0,4 + 0,24 \\
&= 0,64.
\end{aligned}$$

43. Fungsi probabilitas densitas untuk variabel acak X diberikan oleh

$$f(x) = k(10+x)^{-2}; \quad 0 \leq x < \infty.$$

Hitunglah $\Pr(X \leq 15)$.

Jawab:

Karena $\int_0^\infty f(x)dx = 1$ maka diperoleh

$$\begin{aligned} \int_0^\infty k(10+x)^{-2}dx &= 1 \\ \frac{k}{10} &= 1 \\ k &= 10 \end{aligned}$$

Jadi

$$\begin{aligned} \Pr(X \leq 15) &= \int_0^{15} 10k(10+x)^{-2}dx \\ &= -10(10+x)^{-1}\Big|_{x=0}^{x=15} \\ &= -\frac{10}{25} + 1 \\ &= \frac{15}{25} \\ &= 0,6. \end{aligned}$$

44. Kerugian yang disebabkan oleh kebakaran pada sebuah bangunan dimodelkan dengan variabel acak X yang mempunyai fungsi probabilitas densitas

$$f(x) = \begin{cases} 0,005(20-x) & ,0 < x < 20 \\ 0 & ,lainnya \end{cases}$$

Diberikan bahwa kerugian kebakaran melebihi 8, hitunglah probabilitas bahwa kerugian kebakaran melebihi 16.

Jawab:

$$\begin{aligned} P(X > 8) &= P(8 < X < 20) \\ &= \int_8^{20} 0,005(20-x)dx \\ &= 0,005(20x - \tfrac{1}{2}x^2)|_{x=8}^{x=20} \\ &= 0,36. \end{aligned}$$

$$\begin{aligned} P(X > 16) &= P(16 < X < 20) \\ &= \int_{16}^{20} 0,005(20-x)dx \\ &= 0,005(20x - \tfrac{1}{2}x^2)|_{x=16}^{x=20} \\ &= 0,04 \end{aligned}$$

$$\begin{aligned} P(X > 16|X > 8) &= \frac{P(X>16 \cap X>8)}{P(X>8)} \\ &= \frac{P(X>16)}{P(X>8)} \\ &= \frac{0,04}{0,36} \\ &= \frac{1}{9}. \end{aligned}$$

45. Misalkan X berdistribusi *uniform* pada interval [-2,6] dengan fungsi probabilitas densitas

$$f(x) = \begin{cases} \frac{1}{8} & ,-2 \leq x \leq 6 \\ 0 & ,lainnya \end{cases}$$

Hitunglah nilai ekspektasi $g(X) = X^3 + \sqrt{X+2}$.

Jawab:

$$\begin{aligned} E[g(X)] &= \int_{-2}^{6} g(x) \cdot f(x) dx \\ &= \int_{-2}^{6} (x^3 + \sqrt{x+2}) \cdot \frac{1}{8} \\ &= \frac{1}{8} \left[(\frac{1}{4} x^4 + \frac{2}{3} (x+2)^{3/2} \right]_{-2}^{6} \\ &= 41,89. \end{aligned}$$

46. Misalkan variabel acak Y berdistribusi *uniform* kontinu pada interval [-2,5]. Hitunglah $\Pr(Y = 1)$.

    Jawab:

    Karena Y berdistribusi *uniform* kontinu pada interval [-2,5] maka

$$\Pr(Y = 1) = \int_{1}^{1} \frac{1}{5 - (-2)} dy = 0.$$

47. Usia dari sebuah mesin cetak dengan harga 200 mempunyai distribusi eksponensial dengan mean 2 tahun. Pabrik yang memproduksi mesin cetak akan membayar *full refund* kepada pembeli jika mesin cetak rusak selama tahun pertama dan *one-half refund* jika mesin cetak rusak selama tahun kedua. Jika pabrik tersebut menjual 100 mesin cetak, hitunglah nilai harapan dari *refund* yang akan dibayar.

    Jawab:

    Misal X adalah variabel acak yang menyatakan usia sebuah mesin cetak. Karena $X \sim eksponensial(1/2)$ maka diperoleh $f(x) = \frac{1}{2}e^{-\frac{1}{2}x}$. Misal $R(x)$ menyatakan *refund* yang akan dibayar untuk satu mesin cetak, maka diperoleh

$$R(X = x) = \begin{cases} 200 & , 0 \leq x \leq 1 \\ 100 & , 1 < x \leq 2 \\ 0 & , selainnya \end{cases}$$

$$\begin{aligned} E[R(X)] &= \int_0^1 200\frac{1}{2}e^{-\frac{1}{2}x} + \int_1^2 100\frac{1}{2}e^{-\frac{1}{2}x} \\ &= 200[-e^{-\frac{1}{2}x}]_0^1 + 100[-e^{-\frac{1}{2}x}]_1^2 \\ &= 200(1 - e^{-\frac{1}{2}}) + 100(e^{-\frac{1}{2}} - e^{-1}) \\ &= 102,56. \end{aligned}$$

    Jadi jika pabrik tersebut menjual 100 mesin cetak, maka nilai harapan dari *refund* yang akan dibayar adalah

$$E[100R(X)] = 100 \cdot E[R(X)] = 10256.$$

48. Misalkan Z adalah variabel acak yang berdistribusi normal standar. Hitunglah $\Pr(|Z| > 2,05)$.

    Jawab:

$$
\begin{aligned}
\Pr(|Z| > 2,05) &= 1 - \Pr(|Z| \leq 2,05) \\
&= 1 - 2\Pr(0 \leq Z \leq 2,05) \\
&= 1 - 2(0,4798) \\
&= 0,0404.
\end{aligned}
$$

49. Sebuah uji diagnostik mengenai ada atau tidaknya suatu penyakit mempunyai dua hasil yang mungkin: 1 untuk ada penyakit dan 0 untuk tidak ada penyakit. Misalkan X menunjukkan adanya atau tidaknya penyakit berdasarkan pernyataan pasien, dan Y menunjukkan hasil dari uji diagnostik. Fungsi probabilitas bersama dari X dan Y diberikan oleh

$$\begin{array}{ll} \Pr(X=0,Y=0)=0,800 & \Pr(X=1,Y=0)=0,050 \\ \Pr(X=0,Y=1)=0,025 & \Pr(X=1,Y=1)=0,125 \end{array}$$

Hitunglah $Var(X+Y)$.

Jawab:

		Y		
		0	1	$\Pr(X=x)$
X	0	0,800	0,025	0,825
	1	0,050	0,125	0,175
$\Pr(Y=y)$		0,850	0,150	1

X	Y	$Z=X+Y$	$\Pr(Z=z)$
0	0	0	0,800
0	1	1	0,025
1	0	1	0,050
1	1	2	0,125

$$E(Z)=0\cdot 0,800+1\cdot(0,025+0,050)+2\cdot 0,125=0,325$$
$$E(Z^2)=0^2\cdot 0,800+1^2\cdot(0,025+0,050)+2^2\cdot 0,125=0,575$$
$$Var(X+Y)=Var(Z)=E(Z^2)-(E(Z))^2=0,575-(0,325)^2=0,4694.$$

50. Dari soal nomor 49, hitunglah $Var(Y|X=1)$.

    Jawab:

$$\begin{aligned}
&\Pr(Y=0|X=1) = \tfrac{\Pr(X=1,Y=0)}{\Pr(X=1)} = \tfrac{0{,}050}{0{,}175} = 0,286 \\
&\Pr(Y=1|X=1) = \tfrac{\Pr(X=1,Y=1)}{\Pr(X=1)} = \tfrac{0{,}125}{0{,}175} = 0,714 \\
&E(Y|X=1) = \Pr(Y=1|X=1) = 0,714 \\
&E(Y^2|X=1) = \Pr(Y=1|X=1) = 0,714 \\
&Var(Y|X=1) = E(Y^2|X=1) - (E(Y|X=1))^2 = 0,714 - (0,714)^2 = 0,20.
\end{aligned}$$

51. Misalkan X dan Y adalah variabel-variabel acak. Diberikan persamaan sebagai berikut:

    (i) $Cov[X,Y] = E[X \cdot Y] - E[X]E[Y]$

    (ii) $Cov[X,Y] = a \cdot Var[X]$ jika $Y = a \cdot X + b$

    (iii) $Cov[X,Y] = Var[Y]$

    Manakah yang benar dari tiga persamaan diatas?

    Jawab:

    Yang benar adalah persamaan (i) dan (ii).

52. Sebuah investasi memberikan tingkat bunga tahunan R yang mengikuti distribusi uniform pada interval (0,04 ; 0,08). Untuk investasi awal sebesar 10.000 akan memberikan hasil setelah satu tahun mengikuti variabel acak $V = 10.000e^R$. Tentukanlah fungsi distribusi kumulatif, $F(v)$, untuk nilai v yang memenuhi $0 < F(v) < 1$.

Jawab:

Karena R mengikuti distribusi uniform pada interval (0,04 ; 0,08) maka diperoleh $f(r) = \frac{1}{0{,}04}$.

$$\begin{aligned}
\Pr(R \le r) &= \int_{0{,}04}^{r} f(s)ds \\
&= \int_{0{,}04}^{r} \frac{1}{0,04} ds \\
&= \frac{1}{0,04} r - 1.
\end{aligned}$$

Dengan demikian, diperoleh

$$\begin{aligned}
F(v) &= \Pr(V \le v) \\
&= \Pr(10.000e^R \le v) \\
&= \Pr(\ln 10.000 + R \le \ln v) \\
&= \Pr(R \le \ln v - \ln 10.000) \\
&= \Pr\left(R \le \ln\left(\frac{v}{10000}\right)\right) \\
&= \frac{1}{0,04} \ln\left(\frac{v}{10000}\right) - 1 \\
&= 25\left[\ln\left(\frac{v}{10000}\right) - 0,04\right].
\end{aligned}$$

53. Misalkan X_1 dan X_2 adalah variabel-variabel acak dari distribusi normal standar. Tentukanlah fungsi densitas probabilitas dari $Y_1 = \frac{X_1}{X_2}$.

Jawab:

$$X_1 \sim f_{X_1}(x_1) = \frac{1}{\sqrt{2\pi}} e^{-x_1^2/2}, \quad -\infty < x_1 < \infty$$

$$X_2 \sim f_{X_2}(x_2) = \frac{1}{\sqrt{2\pi}} e^{-x_2^2/2}, \quad -\infty < x_2 < \infty$$

$$\begin{aligned}
g_1(y_1) &= 2\int_0^\infty f_{X_2}(x_2) f_{X_1}(x_2 y_1) x_2 dx_2 \\
&= 2\int_0^\infty \frac{1}{\sqrt{2\pi}} e^{-\frac{x_2^2}{2}} \frac{1}{\sqrt{2\pi}} e^{-\frac{x_2^2 y_1^2}{2}} x_2 dx_2 \\
&= \frac{1}{\pi}\int_0^\infty e^{-\frac{x_2^2(1+y_1^2)}{2}} x_2 dx_2 \\
&= \frac{1}{\pi}\frac{1}{(1+y_1^2)} e^{-\frac{x_2^2(1+y_1^2)}{2}} \Big|_{-\infty}^{0} \\
&= \frac{1}{\pi(1+y_1^2)}.
\end{aligned}$$

54. Dari soal nomor 53, tentukanlah mean dari variabel acak Y_1.

    Jawab:

    $$E[Y_1] \;=\; \int_{-\infty}^{\infty} y_1 \frac{1}{\pi(1+y_1^2)} dy_1$$

    Karena $f(y_1) = \frac{y_1}{\pi(1+y_1^2)}$ adalah fungsi ganjil, maka

    $$E[Y_1] \;=\; 0.$$

55. Suatu populasi mempunyai mean 100 dan standar deviasi 16. Hitunglah probabilitas bahwa mean sampel berada dalam interval ± 3 dari mean populasi jika ukuran sampel $n = 200$.

    Jawab:

$$\begin{aligned}
&Pr(|\bar{X} - 100| \leq 3) \\
&= Pr(|Z| \leq \tfrac{3\sqrt{200}}{16}) \\
&= Pr(|Z| \leq 2,652) \\
&= 2\ Pr(0 \leq Z \leq 2,652) \\
&= 0,992.
\end{aligned}$$

56. Misalkan X adalah variabel acak populasi dengan mean μ_X dan variansi σ_X^2 , dan misalkan $(X_i)_{i=1}^n$ adalah sampel acak dari X. Tentukanlah bentuk lain dari $\Sigma_{i=1}^n (X_i - \bar{X})^2$.

    Jawab:

$$
\begin{aligned}
\Sigma_{i=1}^n (X_i - \bar{X})^2 &= \Sigma_{i=1}^n ((X_i - \mu_X) - (\bar{X} - \mu_X))^2 \\
&= \Sigma_{i=1}^n ((X_i - \mu_X)^2 - 2(X_i - \mu_X)(\bar{X} - \mu_X) + (\bar{X} - \mu_X)^2) \\
&= \Sigma_{i=1}^n (X_i - \mu_X)^2 - 2(\bar{X} - \mu_X)\Sigma_{i=1}^n (X_i - \mu_X) + \Sigma_{i=1}^n (\bar{X} - \mu_X)^2 \\
&= \Sigma_{i=1}^n (X_i - \mu_X)^2 - 2(\bar{X} - \mu_X)(n\bar{X} - n\mu_X) + n(\bar{X} - \mu_X)^2 \\
&= \Sigma_{i=1}^n (X_i - \mu_X)^2 - 2n(\bar{X} - \mu_X)^2 + n(\bar{X} - \mu_X)^2 \\
&= \Sigma_{i=1}^n (X_i - \mu_X)^2 - n(\bar{X} - \mu_X)^2.
\end{aligned}
$$

57. Variabel acak Y diketahui mempunyai distribusi normal. Sampel acak berukuran $n = 14$ menghasilkan $\Sigma_{i=1}^{n} Y_i = -602$ dan $S_Y = 18$. Tentukanlah interval kepercayaan 95% untuk *mean* populasi.

    Jawab:

    $$\varepsilon = t_{\alpha/2}(n-1)\frac{S_Y}{\sqrt{n}} \longrightarrow \varepsilon = t_{0,025}(13)\frac{18}{\sqrt{14}} = 10,4.$$

    Jadi Hitunglah interval kepercayaan 95% untuk mean populasi adalah

    $$(-43 - 10,4; -43 + 10,4) = (-53,4; -32,6).$$

58. Diberikan sampel acak dari distribusi *loss* normal:

    - Mean sampel adalah 42.000
    - Standar deviasi sampel adalah 8.000
    - Terdapat 25 *loss* pengamatan dalam sampel

    Dengan menggunakan uji dua arah dengan

    $$H_0 : \mu = 45.000 \quad H_1 : \mu \neq 45.000.$$

    Tentukanlah nilai α yang akan menolak H_0.

    Jawab:

    $$t(n-1) = \frac{\bar{X} - \mu_X}{\frac{S}{\sqrt{n}}} \longrightarrow t(24) = \frac{42 - 45}{\frac{8}{\sqrt{25}}} = -1,875.$$

    $\alpha = 0,01 : t_{0,005}(24) = 2,797.$
    $\alpha = 0,02 : t_{0,01}(24) = 2,492.$
    $\alpha = 0,05 : t_{0,025}(24) = 2,064.$
    $\alpha = 0,1 : t_{0,05}(24) = 1,711.$
    Karena daerah penolakan untuk α adalah $|t(24)| > t_{\alpha/2}(24)$ maka kesimpulannya, menolak H_0 pada $\alpha = 0,1$ dan tidak menolak H_0 pada $\alpha = 0,01$; $\alpha = 0,02$; dan $\alpha = 0,05$.

59. Laporan survei menyebutkan biaya rata-rata dari kamar hotel di suatu kota lebih kecil dari 49,21. Untuk menguji ini, peneliti memilih sampel dari 24 kamar hotel dan memperoleh biaya rata-rata sampel adalah 47,37 dengan standar deviasi adalah 3,42. Pada $\alpha = 0,01$, berilah kesimpulan dari uji tersebut.

    Jawab:

    Dengan menggunakan uji satu arah:

$$\begin{aligned} H_0 &: \mu = 49,21 \\ H_1 &: \mu < 49,21. \end{aligned}$$

    Perhatikan bahwa

$$t(23) = \frac{47,37 - 49,21}{\frac{3,42}{\sqrt{24}}} = -2,636$$

    dan daerah penolakan untuk $\alpha = 0,01$ adalah $t(24) < t_{0,005}(24) = 2,807$ maka kesimpulannya menolak H_0 pada $\alpha = 0,01$. Jadi, biaya kamar di kota tersebut lebih kecil dari $49,21$ pada $\alpha = 0,01$.

60. Misalkan X_1, X_2, X_3, X_4 adalah sampel acak dari distribusi normal dengan mean μ yang tidak diketahui dan variansi σ^2 yang tidak diketahui. *Significance level* $\alpha = 0,05$ menggunakan statistik *Student-t* untuk uji hipotesis:

$$H_0 : \mu = 10$$
$$H_1 : \mu \neq 10$$

Diketahui $\bar{X} = 15,84$ dan $S^2 = \frac{\Sigma_{i=1}^4 (X_i - \bar{X})^2}{3} = 16$. Tentukanlah *critical t-value* dan keputusan yang diperoleh.

Jawab:

Perhatikan bahwa

$$t(3) = \frac{15,84 - 10}{\frac{4}{\sqrt{4}}} = 2,92$$

dan daerah penolakan untuk $\alpha = 0,05$ adalah $|t(4)| > t_{0,025}(4) = 2,78$. Dengan *critical t-value* $t = 2,78$ maka kesimpulannya menolak H_0 pada $\alpha = 0,05$.

ERRATA

1. Halaman : 7
   Nomor baris : 17
   Tertulis : $\Sigma = \Gamma'\Lambda^{1/2}\Gamma$
   Seharusnya : $\Sigma = \Gamma'\Lambda\Gamma$
2. Halaman : 26
   Nomor baris : 12
   Tertulis : $= 0 + 2 \times \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}\left(\sqrt{y}\right)^2} - \frac{1}{2} y^{-\frac{1}{2}} + 0$
   Seharusnya : $= 2 \times \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}\left(\sqrt{y}\right)^2} \left(\frac{1}{2} y^{-\frac{1}{2}}\right)$
3. Halaman : 81
   Nomor baris : 15
   Tertulis : dx
   Seharusnya : dy_1
4. Halaman : 90
   Nomor baris : 26
   Tertulis : $K\left(x_i\right)$
   Seharusnya : $K\left(X_i\right)$
5. Halaman : 102
   Nomor baris : 7
   Tertulis : $Y = \Sigma_{i=1}^{n} x_i$
   Seharusnya : $Y = \Sigma_{i=1}^{n} X_i$

6. Halaman : 104
   Nomor baris : 20
   Tertulis : δ_1
   Seharusnya : δ_2
7. Halaman : 114
   Nomor baris : 21 dan 22
   Tertulis : k_2
   Seharusnya : k_3
8. Halaman : 115
   Nomor baris : 17 dan 18
   Tertulis : $= (5 - \theta^*)\, \Sigma_{i=1}^{n} x_i + n\,(\theta^*)^2 - \dfrac{25n}{2} \leq k_1$
   Seharusnya : $= (5 - \theta^*)\, \Sigma_{i=1}^{n} x_i + \dfrac{n\,(\theta^*)^2}{2} - \dfrac{25n}{2} \leq k_1$
9. Halaman : 115
   Nomor baris : 18
   Tertulis : $= (5 - \theta^*)\, \Sigma_{i=1}^{n} x_i \leq n\,(\theta^*)^2 - \dfrac{25n}{2} + k_1$
   Seharusnya : $= (5 - \theta^*)\, \Sigma_{i=1}^{n} x_i \leq \dfrac{n\,(\theta^*)^2}{2} - \dfrac{25n}{2} + k_1$
10. Halaman : 118
    Nomor baris : 3
    Tertulis : S
    Seharusnya : s

Daftar Pustaka

[1] Hogg RV, McKean J, Craig AT. 2005. *Introduction to Mathematical Statistics*. Sixth Edition. New Jersey: Pearson Education, Inc.

[2] Hogg RV, McKean J, Craig AT. 2014. *Introduction to Mathematical Statistics*. Seventh Edition. New Jersey: Pearson Education, Inc.

[3] Lindgren BW. 1976. *Statistical Theory*. Third Edition. New York: Macmillan Publishing Co, Inc.

[4] Rice JA. 2007. *Mathematical Statistics and Data Analysis*. Duxbury: Thomson Brooks/Cole.

[5] Roussas GG. 1973. *A First Course in Mathematical Statistics*. Boston: Addison-Wesley Publishing Company.

INDEX

	Istilah di buku	Muncul pertama di Bab
Alternative hypothesis, 244, 255-256, 373, 439, 446, 448, 476, 563	Hipotesis alternatif	
Bernoulli trials, 141-142, 219 sequence of, 141 Bias, 210, 218, 299, 339, 357, 606-607 sample, 210, 218, 299, 606-607 Biased estimator, 210, 299, 387 Binomial distribution, 142-143, 145, 147-151, 157, 201-203, 207, 209, 228, 238-239, 246, 248, 250-251, 309-310, 316-317, 333, 354, 409, 440, 541, 544	Bernoulli trials Bias Bias Estimator Binomial (Distribusi)	
Central limit theorem, 170, 220, 225, 248, 313, 317, 323-324, 340, 342, 354, 370, 448, 456, 463, 541, 552, 570, 605, 616	Teorema Limit Pusat	
Chi-square distribution, 161, 167-168, 192, 259,	Ch-Squared (Distribusi)	
Composite null hypothesis, 457, 544 Conditional probability, 21-24, 28, 30-32, 96-97, 121, 150, 169, 178, 186, 189, 389-390 Conditional probability density functions, 121	Hipotesis Null-Komposit Peluang Bersyarat Fungsi Kepekatan Peluang Bersyarat	
Confidence intervals, 207, 218-219, 221-223, 227, 238, 244, 248, 254, 267, 269, 277, 281, 283-284, 286, 291, 315, 342, 347, 383, 410, 501-502, 513, 515-516, 537, 542, 561, 565	Selang Kepercayaan	
Continuous random variable, 33-34, 44-47, 52, 63, 97, 169-171, 211, 213, 238-239, 272, 301, 405, 417, 538, 540, 570, 607	Peubah acak kontinu	
Convergence, 54, 295-296, 298, 300-304, 306-307, 321-323, 343, 401	Konvergensi	
Correlation, 104-106, 108-111, 116, 118, 128, 138-139, 147, 157, 184, 186, 189, 240, 337, 413, 415, 421, 433, 401, 518-519, 521, 577, 594-600	Korelasi	
Covariance, 104-106, 111, 116, 125-126, 136, 150, 181, 183-185, 188, 190, 196, 241, 322-324, 358-359, 421-422, 482, 513, 517-518, 522-523, 526-527, 529-530, 534 Critical point, 195, 219, 249, 493, 541, 554, 561, 569 Critical region, 245-248, 250, 252, 255-258, 373, 439-451, 453, 455-457, 463, 469, 476-478, 502	Kovarians Titik Kritis Daerah Kritis	
Data, 28-29, 176, 212-213, 215-218, 223-224, 235-240, 243-244, 249-250, 256-257, 261, 264-265, 279, 282-284, 286-288, 373-374, 377, 379-380, 387, 389, 423, 433, 444, 466, 468, 496, 510, 514-516, 518, 541-543, 550, 552, 556-557, 562, 565, 570-571, 578-579, 582-584, 587-588, 590-591, 596-601, 606-607, 612-613, 615, 629	Data	
Density function, 33, 44, 77, 95, 173, 244, 341, 406 Dependent random variables, 130, 147, 244	Fungsi Kepekatan Peubah acak (Dependent)	
Discrete random variable, 32, 34, 40-41, 45, 53, 58, 61, 176, 211, 227, 406	Peubah Acak Diskret	
Distribution, 11, 32-36, 41-52, 54, 57-67, 69-71, 76,	Distribusi	

INDEX

INDEX

INDEX

Median, 50-51, 65, 167, 216, 223, 232, 234-236, 239-240, 243, 285-286, 329, 332-333, 342, 345-346, 352, 358, 409, 425-426, 429, 537-540, 542-543, 546, 548, 550-553, 561-562, 575, 582, 584, 588, 590, 601-607, 611-613 Method of least squares, 510, 518	Median Metode kuadrat terkecil	
Multivariate distributions, 75-140, 320 conditional expectations, 98, 122 conditional probabilities, 98 joint probability density function, 77, 95 marginal probability density functions, 84, 96, 98, 112-113, 121-122, 127	Multivariat (Distribusi)	
Normal distribution, 170-174, 176, 178-184, 186-190, 195-197, 199, 201, 207, 210, 215, 220, 224, 236, 242, 249, 252-253, 258, 261, 270-271, 275, 282, 284, 292, 295, 300-301, 303, 311-313, 315-319, 339, 342, 344, 353-354, 364, 366, 372, 387, 394, 408, 420-421, 423, 433-435, 448, 455-456, 458, 462-463, 466, 468, 474, 479-482, 486-490, 494, 512-513, 518-521, 526, 529-530, 533-534, 542, 545, 548-549, 551, 553, 559-560, 566, 571, 577-578, 583, 597, 615, 630, 635	Normal (Distribusi)	
Normal random variables, 173, 187, 195, 270, 512, 518, 523	Normal (peubah acak)	
Null hypothesis, 244, 247, 251, 284-285, 348-350, 353, 355, 368, 371, 373, 439, 452-453, 457, 476, 490, 496, 541, 544-545, 553-554, 563, 570, 572, 586, 589, 594-595, 597, 599-600 composite, 247, 251, 371, 452, 457, 490, 544 simple, 247, 251, 453, 457, 476, 544, 553, 595, 599-600	Hipotesis Null Hipotesis komposit Hipotesis sederhana	
Order statistics, 231-243, 251, 278, 269, 291-292, 313, 332, 356, 365, 371, 388, 392, 396, 400,	Statistik terurut	
Sample mean, 137-138, 176, 219, 230, 248-249, 252, 271, 279, 284, 286, 295, 298, 315, 318, 322, 342, 345, 389, 423, 464, 474, 508, 523, 545, 547-548, 550, 552, 559, 601-602, 604-605, 607, 611-612, 622	Rataan Sampel	
Sample variance, 137, 139, 219, 223, 298-299, 322, 347, 360, 364, 389, 426, 429, 465, 523, 526	Varians Sampel	
Statistics, 1-2, 42, 75, 137, 141, 161, 187, 191, 207-208, 211, 218-220, 231-243, 247, 251, 254, 265, 278, 282, 289, 291-292, 295-296, 298, 313, 318, 327, 332, 350-351, 356, 365, 371, 383, 388-389, 392, 396-397, 400, 405, 416-436, 439, 455, 458, 461-463, 465, 467, 481, 485-487, 489, 493, 508, 537-616, 617, 620, 621-623, 631, 641 population, 220, 232, 238, 298, 596, 598	Statistik	
Significance level of the test, 251, 255, 259, 271, 354, 441, 444, 449, 461, 465, 478-479, 493, 584	Uji …	
Simple null hypothesis, 251, 453, 476, 544	Hipotesis Null sederhana	
Standard deviations, 180, 186, 223	Simpangan Baku	
Standard error, 219, 221, 223, 280, 383, 626-627	Galat baku	
Tests of hypotheses, 207, 244, 295, 439-483, 501	Uji Hipotesis	
Unbiased estimator, 137, 208-210, 214, 217, 221-222, 267, 269, 279, 298-300, 337-339, 346, 383-385, 387-388, 397-400, 402, 408, 410, 412, 414-415, 419, 595 Uniform distribution, 45, 52, 141, 168, 211, 241-242, 268, 270, 274, 289, 298, 304, 319, 356, 365, 388, 395, 400, 403, 422-424, 428	Estimator tak bias	

INDEX

Variables, 32-35, 40, 42, 44, 58, 60, 62, 70, 75-79, 81, 83-89, 93, 95-98, 101-105, 107-123, 125, 127-134, 136-139, 141, 144-149, 151, 155, 157, 161, 163-166, 168, 170, 173, 176, 178-181, 184, 187, 189-190, 193-199, 203, 208, 219, 221, 223, 226-227, 235, 242, 244, 258, 262, 266, 268, 270, 272, 276, 289-290, 295-296, 299-301, 305-309, 320, 327, 331, 364, 370-371, 373, 377, 383, 389, 391-392, 396, 398, 400, 409, 411, 417, 420, 426-427, 434-435, 442-443, 446, 458, 465-466, 472, 474-478, 485-486, 488, 490, 493, 495-498, 502-503, 505, 509, 512-513, 516-519, 522-528, 530-533, 537, 539-541, 547, 550-551, 554, 569, 573, 575-576, 580, 589, 593-594, 600, 604, 607, 610-612	Peubah	
Variance, 58-59, 64-67, 69-70, 97-100, 102, 107-108, 111, 118-119, 125-126, 136-139, 141, 143, 148, 150, 152, 154, 161, 165, 168, 171-172, 176-179, 181, 184, 186, 188, 190-191, 193, 195, 198-201, 203, 207, 216-220, 222-226, 236, 248-249, 252, 254, 256, 258, 261, 280, 296, 298-299, 310, 312-313, 315-324, 333,	Varians	

Figure 6.1: `Siswadi` lulus Sarjana dari IPB. Kemudian, ia melanjutkan pendidikan S2 Statistika di The University of Minnesota. Pendidikan S3 ditempuh di Universitas North Carolina, Raleigh. Terhitung mulai tahun 2013 diangkat sebagai Guru Besar di Institut Pertanian Bogor. SINTA ID : 6029265

Figure 6.2: `Windiani Erliyana` lulus program sarjana di Departemen Matematika FMIPA IPB. Kemudian, ia melanjutkan pendidikan S2 Matematika Terapan di IPB melalui Program Sinergi S1-S2.