Buku Referensi

PENGANTAR STATISTIKA MATEMATIKA

Edisi Pertama

Sigit Nugroho, Ph.D.

UNIB Press

PENGANTAR
STATISTIKA MATEMATIKA

Pengantar Statistika Matematika

Sigit Nugroho, Ph.D.
Universitas Bengkulu

UNIB Press
Bengkulu
2008

PENGANTAR STATISTIKA MATEMATIKA
Sigit Nugroho, Ph.D.
ISBN : 978-979-9431-33-2 360hal.
Cetakan Pertama. Edisi 1. 2008.

Penyeleksi Naskah : Fachri Faisal

Editor : Jose Rizal
Desain Sampul : Ratna Astuti Nugrahaeni

Diterbitkan pertama kali oleh **UNIB Press**, Jalan WR Supratman, Bengkulu.

Kata Pengantar

Buku yang berjudul "**Pengantar Statistika Matematika**" ini diperlukan sebagai landasan, pedoman atau rujukan bagi para mahasiswa atau siapa saja yang ingin mempelajari statistika matematika atau sering juga disebut dengan teori statistika dengan baik, mudah dan benar.

Materi buku ini biasanya disajikan untuk mata kuliah Teori Statistika atau Statistika Matematika di jurusan Matematika atau jurusan Statistika selama 2 semester, yang masing-masing bobotnya 4 SKS. Materi untuk bagian pertama adalah Peluang, Peubah Acak, Sebaran Bersama, Sebaran Fungsi Peubah Acak dan Sifat Beberapa Sebaran Kontinu khususnya yang berkaitan dengan Sebaran Normal. Sedangkan materi untuk bagian kedua berisikan Limit Sebaran Peubah Acak dan Sebaran Nilai Ekstrim, yang dilanjutkan dengan Teori Pendugaan Titik, Statistik Cukup dan Lengkap, Pendugaan Interval dan Teori Pengujian Hipotesis Statistika tentang parameter parameter populasi.

Alternatif dari beberapa definisi maupun teorema juga diberikan dalam buku ini, agar pemakai memperoleh gambaran adanya variasi cara penyampaian notasi akan hal tersebut.

Ucapan terima kasih penulis sampaikan sebelum dan sesudahnya kepada mereka yang memberikan masukan baik berupa saran atau kritik yang membangun. Penulis sampaikan kepada istri, *Ir. Mucharromah, M.Sc., Ph.D.*, dan anak-anak, *Shofa Ulfiyati Nugrahaeni* dan *Ratna Astuti Nugrahaeni*, yang telah dengan penuh pengertian dan sabar memberikan motivasi serta dorongan guna penyelesaian tulisan ini. Kepada rekan-rekan sejawat yang telah memberikan masukan-masukan untuk usaha penerbitan, penulis ucapkan ribuan terima kasih. Kepada semua pengguna, juga diucapkan terima kasih dan semoga mendapatkan manfaat dari karya tulis saya ini.

Bengkulu, 20 Juli 2008

Sigit Nugroho, Ph.D.

Oentoek:

Mucharromah Nugroho, Ph.D.,
Shofa Ulfiyati Nugrahaeni, dan
Ratna Astuti Nugrahaeni

Daftar Isi

Peluang

Pendahuluan

Dalam setiap studi ilmiah mengenai fenomena fisik, kita mungkin ingin menggunakan model matematik untuk menggambarkan atau memprediksi nilai observasi beberapa karakter yang diinginkan. Sebagai contoh, dalam ruang hampa udara, kecepatan benda yang jatuh setelah beberapa saat, t. Rumus $v = gt$ dimana g = percepatan gravitasi bumi memberikan gambaran model matematik yang bermanfaat untuk kecepatan sebuah benda jatuh dalam ruang hampa udara. Ini merupakan contoh suatu **model deterministik** (*deterministic model*). Percobaan berulang dalam kondisi yang ideal akan tetap menghasilkan hal yang sama, yang sudah tentu dapat diprediksi dengan menggunakan model yang sudah ada. Namun, apabila kondisi ideal tak mungkin diperoleh, maka hasil tersebut tak dapat kita peroleh. Mungkin dapat diakibatkan oleh variabel yang tak diketahui atau tak dapat dikendalikan, seperti misalnya suhu dan kelembaban udara, yang dapat mempengaruhi hasil studi. Demikian juga dapat dikarenakan karena kesalahan pengukuran atau hal lain yang dapat mempengaruhi hasil studi ilmiah tersebut. Lebih jauh lagi, kita tidak memeliki pengetahuan untuk menurunkan model yang lebih rumit yang sudah memperhitungkan semua penyebab keragaman.

Terdapat tipe fenomena lain dimana perbedaan hasil mungkin secara alami muncul karena suatu kebetulan, dan dengan demikian model deterministik tidak akan cocok lagi. Misalnya, dalam pengamatan jumlah partikel yang dipancarkan oleh sumber radioaktif, waktu hingga suatu komponen elektronik gagal berfungsi, atau memenangkan suatu permainan. Motivasi mempelajari peluang ini adalah sebagai dasar dalam mempelajari **model probabilistik** (*probabilitic model*) atau **model stokastik** (*stochastic model*).

Terminologi **percobaan** (*experiment*) menunjuk kepada proses untuk mendapatkan hasil observasi beberapa fenomena, dan performans suatu percobaan disebut sebagai suatu **tindakan** (*trial*) dari suatu percobaan. Sedangkan hasil observasi, pada suatu tindakan percobaan, disebut dengan **keluaran** (*outcome*).

Definisi 1. 1.

Suatu percobaan acak atau percobaan statistik adalah suatu percobaan dimana (a) semua jenis keluaran percobaan diketahui terlebih dahulu (b) hasil dari percobaan tidak diketahui , dan (c) percobaan dapat diulang pada kondisi yang sama.

Dalam teori peluang, yang kita pelajari adalah ketidakmenentuan percobaan acak atau percobaan statistik ini. Percobaan semacam ini biasanya diasosiasikan dengan suatu gugus Ω, yaitu suatu gugus semua kemungkinan keluaran suatu percobaan. Lebih jauh lagi kita hubungkan Ω, suatu σ-field $\mathscr{S}$ anak gugus dari Ω. Kita ingat bahwa σ-field merupakan kelas anak gugus yang tak kosong dari Ω yang tertutup pada operasi gabungan dan komplemen terhingga dan mencakup gugus $\emptyset$. Elemen-elemen dari Ω disebut dengan titik contoh.

Definisi 1.2.

Gugus semua keluaran yang mungkin dari suatu percobaan disebut dengan ruang contoh (*sample space*), dan dinotasikan dengan *S*.

Bila dikaitkan dengan percobaan acak, maka kita punya definisi seperti berikut

Definisi 1. 3.

Ruang contoh percobaan statistik merupakan pasangan $(\Omega,\mathscr{S})$ dimana (a) Ω, yaitu suatu gugus semua kemungkinan keluaran suatu percobaan, dan (b) $\mathscr{S}$ σ-field anak gugus dari Ω.

Teladan 1. 1.

Dalam percobaan pelemparan dua koin, dan muka dari tiap koin adalah yang akan diamati, maka ruang contohnya adalah S = {GG, GA, AG, AA}. Catatan G = Gambar dan A = Angka.

Teladan 1. 2.

Seperti dalam Teladan 1.1 namun kita tidak ingin pengamatan muka dari tiap koin, tapi ingin melihat atau mencatat banyaknya G atau munculnya Gambar keluar, maka $S = \{0,1,2\}$

Teladan 1. 3.

Jika sebuah bola lampu dinyalakan dan yang ingin diamati adalah umur sampai lampu tersebut tidak berfungsi lagi atau mati. Minimal secara konsep, kita dapat tuliskan $S = \{t \mid 0 \leq t < \infty\}$.

Suatu ruang contoh S dikatakan **terhingga** (*finite*) jika memiliki keluaran yang jumlahnya terhingga, seperti $S = \{e_1, e_2, \ldots, e_N\}$ dan dikatakan **terbilang tak terhingga** (*countably infinite*) atau tak terbilang jika keluarannya dapat dituliskan sebagai fungsi satu-satu dengan bilangan bulat positif $S = \{e_1, e_2, \ldots\}$. Pengamatan banyaknya Gambar yang muncul seperti dalam Teladan 1.2. membentuk ruang contoh terhingga. Sedangkan pengamatan umur lampu seperti dalam Teladan 1.3. membentuk ruang contoh tek terhingga. Sedangkan pengamatan seperti banyaknya semut yang berada dalam suatu kawasan tertentu membentuk suatu ruang contoh terbilang tak terhingga.

Definisi 1.4.

Jika suatu ruang contoh S terhingga atau terbilang tak terhingga, maka ruang contoh ini disebut ruang contoh diskrit (*discrete sample* *space*). Ruang konseptual yang mencakup keluaran yang mengambil sembarang nilai bilangan nyata diantara dua titik, maka ruang contoh dikatakan ruang contoh kontinu (*continuous sample space*).

Suatu gugus yang terhingga atau terbilang tak terhingga disebut **terbilang** (*countable*).

Teladan 1. 4.

Sebuah lampu pijar dinyalakan dan X = banyaknya sinar yang dihasilkan (lumen) dan Y = banyak energi panas (joule) diukur atau diamati. Ruang contoh dari percobaan ini adalah $S = [0, \infty)\text{x}[0, \infty)$ atau $S = \{(x,y) \mid 0 \le x < \infty \text{ dan } 0 \le y < \infty\}$ ini merupakan ruang contoh kontinu. Namun apabila misalnya kita bicara banyaknya lampu pijar di sebuah wilayah yang mati atau tidak berfungsi (Z) akan membentuk ruang contoh S = {0, 1, 2, …}

Definisi 1.5.

Suatu **peristiwa** (*event*) adalah anak gugus dari ruang contoh. Jika *A* adalah suatu peristiwa, kita katakan bahwa *A* terjadi, apabila *A* mencakup keluaran yang terjadi. Sembarang gugus $A \in \mathscr{S}$ dikenal dengan istilah peristiwa atau kejadian.

Irisan atau Gabungan dari peristiwa terhingga atau terbilang tak terhingga dikatakan sebagai Irisan atau Gabungan terhingga. Keseluruhan ruang contoh *S* merupakan suatu peristiwa khusus yang disebut dengan **peristiwa pasti** (*sure event*), dan himpunan kosong $\varnothing$ disebut dengan **peristiwa nol** (*null event*).

Definisi 1.6.

Peristiwa yang hanya terdiri satu keluaran disebut sebagai peristiwa sederhana (*simple event*) atau peristiwa elementer (*elementary* *event*) sedangkan yang terdiri lebih dari satu keluaran disebut dengan peristiwa majemuk.

Dalam ruang contoh diskrit, sembarang anak gugus dapat dituliskan sebagai gabungan terhingga dari peristiwa sederhana, dan tak ada kesulitan dalam penggabungan tiap anak gugus dengan suatu peristiwa dalam kasus diskrit. Tidaklah mudah untuk merepresentasikan peristiwa-peristiwa kontinu.

Definisi 1.7.

Dua peristiwa A dan B disebut saling lepas (*mutually exclusive*) jika $A \cap B = \varnothing$.

Jika dua peristiwa saling lepas, maka kedua peristiwa tersebut tidak memiliki keluaran yang sama. Dengan demikian, keluaran yang muncul disalah satu peristiwa tidak mungkin muncul di peristiwa lainnya.

Teladan 1. 5.

Misalkan dalam tindakan pelemparan dua koin seperti dalam Teladan terdahulu, bila A adalah peristiwa munculnya paling sedikit satu Gambar, sedangkan B adalah peristiwa munculnya kedua sisi koin Angka, adalah contoh dua peristiwa yang saling lepas.

Definisi 1.8.

Peristiwa-peristiwa $A_1, A_2, A_3. \ldots$ dikatakan saling lepas berpasangan (*pairwise mutually exclusive*) jika $A_i \cap A_j = \varnothing$ untuk $i \neq j$.

Cara Menyatakan Peluang

Metode Klasik atau A Priori. Jika diketahui bahwa peristiwa A dapat muncul dalam m cara dan total seluruh kemungkinan peristiwa adalah n, maka peluang munculnya peristiwa A

$$P(A) = \frac{banyaknya \quad cara \quad A}{total \quad semua \quad cara} = \frac{m}{n}$$

Teladan 1. 6.

1. Tanpa percobaan melempar mata uang logam (misalkan sisi yang dapat muncul diberi nama: Gambar dan Angka), maka peluang munculnya Gambar = ½, karena m = 1 = banyaknya cara Gambar muncul, dari total munculnya semua cara = 2.
2. Dari setumpuk kartu bridge standar, bila diambil sebanyak satu kartu secara acak, maka peluang muncuknya kartu ♥ dapat dihitung sebagai berikut: karena banyaknya cara kartu ♥ muncul ada 13 yaitu dari A♥ ,K♥ ,Q♥ ,J♥ ,10♥ ,9♥ ,8♥ ,7♥ ,6♥ ,5♥ ,4♥ ,3♥ ,2♥ sedangkan total semua cara ada 52 yaitu banyaknya semua kartu dalam tumpukan, maka peluang munculnya kartu ♥ adalah 13/52 atau ¼.
3. Dari pelemparan sebuah dadu yang seimbang, maka peluang munculnya mata paling kecil 3 dapat dihitung sebagai berikut: mata paling sedikit 3 berarti mata 3, 4, 5, atau 6 sehingga banyaknya cara ini ada 4, sedangkan total semua cara ada 6 (yaitu munculnya mata 1, 2, 3, 4, 5, atau 6); sehingga peluang munculnya mata dadu sedikitnya 3 adalah 4/6 atau 2/3.
4. Apabila peserta pelatihan Statistika diikuti oleh 24 peserta dengan nama-nama berikut: Ali, Amir, Antyo, Arief, Artika, Badu, Candra, Dwi, Edo, Fahmi, Fifi, Gatot, Hamdani, Hamzah Hasan, Indra, Jarot, Kuncung, Nugroho, Retno, Soegeng, Utari, Wawan, dan Yudi. Apabila dilakukan pemilihan secara acak untuk pemberian hadiah tiket gratis pesawat Walet Airlines, maka peluang Soegeng mendapatkan tiket tersebut adalah

1/24. Sedangkan peluang seorang peserta dengan nama depan A memenangkan hadiah adalah 5/24, karena ada lima orang yang menggunakan nama berawalan A.

Metode Frekuensi *atau* ***A Posteriori****.* Jika peristiwa serupa A muncul m kali dalam total percobaan n, maka peluang pengamatan

$$P(A) = \frac{banyaknya \quad A \quad muncul}{total \quad percobaan} = \frac{m}{n}$$

Teladan 1. 7.

Metode ini dipakai dengan menggunakan data percobaan. Sebagai teladan perhatikan hal-hal berikut:

1. Bila dalam 80 kali pelemparan mata uang (yang tak harus seimbang) munculnya Angka sebanyak 45 kali (sisanya Gambar sebanyak 35 kali), maka P(munculnya Angka) = 45/80.
2. Dari 50 kali pelemparan sebuah dadu (juga tak perlu asumsi seimbang) diperoleh data munculnya mata 1 sebanyak 8 kali, mata 2 sebanyak 7 kali, mata 3 sebanyak 10 kali, mata 4 sebanyak 6 kali, mata 5 sebanyak 9, dan mata 6 sebanyak 10 kali, dengan demikian peluang munculnya mata dadu 1 atau disingkat p(1) = 8/50; dan seterusnya p(2) = 7/50; p(3) = 10/50; p(4) = 6/50; p(5) = 9/50 dan p(6) = 10/50; dan peluang munculnya mata dadu paling besar 4 adalah sama dengan jumlah peluang munculnya mata dadu satu, dua, tiga dan empat yaitu sama dengan = p(1)+p(2)+p(3)+p(4) = 8/50 + 7/50 + 10/50 + 6/50 = (8+7+10+6)/50 = 31/50.

Metode Subyektif*.* Kadang merupakan dugaan atau perkiraan terbaik dari peluang akan munculnya suatu peristiwa A; yang tentunya hanya diperlukan dan **sah**, apabila tidak terdapat cukup data numerik.

Misalkan saudara punya usaha baru dengan barang yang diproduksi belum pernah ada di pasaran; bahkan barang yang miripun bisa jadi belum pernah ada di pasaran. Pertanyaan seperti berapa peluang barang yang diproduksi tersebut akan habis dalam sebulan, merupakan pertanyaan peluang yang dapat dijawab dengan metode subyektif, karena tidak terdapat data yang cukup untuk menyatakannya dengan metode frekuensi.

Jika untuk suatu peristiwa A, limit dari $fA = m(A)/M$ apabila $m(A)$ adalah banyaknya peristiwa A terjadi dari seluruh M tindakan, untuk $M \to \infty$, maka peluang dari peristiwa A juga dapat dinyatakan sebagai

$$P(A) = \lim_{M \to \infty} f_A = \lim_{M \to \infty} {m(A)}/{M}$$

Untuk memotivasi pendefinisian aksioma peluang, perhatikan properti frekuensi relatif berikut. Jika S adalah ruang contoh suatu percobaan dan A adalah suatu peristiwa, maka jelaslah berlaku $0 \leq m(A)$ dan $m(S) = M$, karena $m(A)$ adalah jumlah munculnya keluaran peristiwa A, dan S muncul pada setiap tindakan. Lebih jauh lagi , jika A dan B merupakan peristiwa yang saling lepas, keluaran dalam A berbeda dari keluaran dalam B, dengan demikian $m(A \cup B) = m(A) + m(B)$.

Lebih umum lagi, jika $A_1, A_2, \ldots$ adalah saling lepas berpasangan, maka $m(A_1 \cup A_2 \cup \ldots) = m(A_1) + m(A_2) + \ldots$. Dengan demikian properti berikut berlaku untuk frekuensi relatif

$$0 \leq f_A$$

$$f_S = 1$$

$$f_{A_1 \cup A_2 \cup \ldots} = f_{A_1} + f_{A_2} + \ldots \quad \text{atau} \quad f_{\bigcup_i A_i} = \sum_i f_{A_i}$$

jika $A_1, A_2, A_3. \ldots$ adalah peristiwa-peristiwa yang saling lepas berpasangan.

Definisi Peluang

Bila diberikan suatu percobaan dengan ruang contoh *S*, tujuan utama pemodelan peluang adalah memberikan tiap peristiwa *A* dengan suatu bilangan nyata *P(A)*, yang disebut dengan peluang peristiwa *A*, yang merupakan ukuran kemungkinan bahwa peristiwa *A* akan muncul atau terjadi apabila percobaan dilakukan.

Secara matematis, kita dapat memandang bahwa *P(A)* sebagai suatu **fungsi gugus** (*set function*). Dengan perkataan lain, peluang adalah suatu fungsi dimana daerah fungsinya adalah kumpulan gugus (peristiwa), dan range atau jangkauannya adalah anak gugus dari bilangan nyata.

Definisi 1.9.

Untuk suatu percobaan, *S* melambangkan ruang contoh dan A_1, A_2, A_3. … melambangkan peristiwa-peristiwa yang mungkin. Suatu fungsi gugus yang berkenaan dengan bilangan nyata *P(A)* dengan tiap peristiwa *A* disebut dengan suatu fungsi gugus peluang, dan *P(A)* disebut sebagai peluang peristiwa *A*, jika properti berikut dipenuhi:

$$0 \le P(A) \quad untuk\ setiap\ A$$

$$P(S) = 1$$

$$P\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} A_i$$

jika A_1, A_2, A_3. … adalah peristiwa-peristiwa yang saling lepas berpasangan.

Definisi 1. 10.

Triplet (Ω,$\mathscr{S}$,*P*) disebut dengan ruang peluang.

Seluruh properti diatas tampaknya sesuai dengan konsep peluang secara intuitif, dan beberapa properti ini cukup untuk membentuk suatu struktur matematik. Salah satu dari properti ini adalah peristiwa nol yang memiliki peluang nol, $P(\emptyset) = 0$. Juga, jika A dan B adalah peristiwa yang saling lepas, maka

$$P(A \cup B) = P(A) + P(B)$$

Dengan cara yang sama, jika $A_1, A_2, \ldots, A_k$ adalah peristiwa-peristiwa yang saling lepas berpasangan terhingga, maka

$$P(A_1 \cup A_2 \cup \ldots \cup A_k) = P(A_1) + P(A_2) + \ldots + P(A_k)$$

atau $P\left(\bigcup_{i=1}^{k} A_i\right) = \sum_{i=1}^{k} P(A_i)$

Peluang dalam Ruang Diskrit

Misalkan untuk setiap peristiwa elementer $\{e_i\}$ diberikan suatu nilai bilangan bulat p_i, sedemikian rupa sehingga $P(\{e_i\}) = p_i$. Agar kondisi peluang terpenuhi sesuai dengan definisi peluang, maka

$p_i \geq 0$ untuk semua i.

$$\sum_i p_i = 1$$

Notasi jumlah terakhir mengandung arti bahwa ini merupakan penjumlahan biasa apabila S terhingga, dan merupakan jumlah dari suatu deret tak hingga apabila S terbilang tak terhingga. Sebagai konsekuensinya,

$$P(A) = \sum_{e_i \in A} P(\{e_i\})$$

penjumlahan peluang dilakukan untuk seluruh subskrip i dimana ei sebagai keluaran dari peristiwa A.

Teladan 1. 8.

Tindakan melempar sebuah koin akan menghasilkan ruang contoh $(\Omega, \mathscr{S})$, dimana Ω ={M, B}, dan $\mathscr{S}$ merupakan σ-field semua anak gugus Ω. Bila didefinisikan *P* pada $\mathscr{S}$ dengan *P{M}* = *p* dan *P{B}* = *1-p*, maka *P* mendefinisikan suatu peluang atau ukuran peluang (*probability measure*). $0 \le p \le 1$.

Teladan 1. 9.

Misalkan Ω = {1, 2, 3, …} merupakan gugus bilangan bulat positif, dan misalkan $\mathscr{S}$ merupakan kelas semua anak gugus Ω. Bila didefinisikan *P* pada $\mathscr{S}$ seperti berikut

$$P\{i\} = \frac{1}{2^i}, \quad i = 1, 2, 3, \ldots$$

Maka $\sum_{i=1}^{\infty} P\{i\} = 1$, dan *P* mendefinisikan sebagai peluang.

Definisi 1.11.

Jika suatu obyek diambil dari sekumpulan obyek yang berbeda sedemikian sehingga setiap obyek memiliki peluang yang sama untuk terpilih, maka kita katakan bahwa obyek tersebut dipilih secara acak.

Dengan cara yang sama apabila suatu gugus obyek diambil dari sekumpulan gugus obyek yang berbeda sedemikian sehingga setiap gugus obyek memiliki peluang yang sama untuk terpilih, maka kita katakan bahwa setiap gugus obyek tersebut dipilih secara acak.

Teorema 1.1.

Jika *A* adalah suatu peristiwa dan *A'* adalah komplemennya, maka *P(A)=1-P(A')*

Bukti

Karena *A'* merupakan komplemen dari *A* relatif terhadap *S*, maka $S = A \cup A'$. Karena juga $A \cap A' = \emptyset$, maka *A* dan *A'* adalah dua peristiwa yang saling lepas, maka berdasarkan properti peluang diperoleh $1=P(S)=P(A\cup A')=P(A)+P(A')$ □

Teorema 1.2.

Untuk sembarang peristiwa A dan B, maka $P(A\cup B)=P(A)+P(B)-P(A\cap B)$

Bukti

Dari properti himpunan, kita peroleh bahwa $A\cup B=(A\cap B')\cup B$ dan juga berlaku bahwa $A=(A\cap B)\cup(A\cap B')$. Peristiwa $(A\cap B')$ dan B merupakan peristiwa yang saling lepas, karena $(A\cap B')\cap B=\varnothing$. Sehingga $P(A\cup B)=P(A\cap B')+P(B)$.

Juga diketahui bahwa $A\cap B$ dan $A\cap B'$ juga dua peristiwa yang saling lepas, sehingga $P(A)=P(A\cap B)+P(A\cap B')$. Dan akhirnya

$$\begin{aligned} P(A\cup B) &= P(A\cap B')+P(B) \\ &= [P(A)-P(A\cap B)]+P(B) \\ &= P(A)+P(B)-P(A\cap B) \end{aligned}$$

Teorema 1.3.

Untuk sembarang tiga peristiwa A, B dan C, berlaku

$$P(A\cup B\cup C)=P(A)+P(B)+P(C)-P(A\cap B)-P(A\cap C)-P(B\cap C)+P(A\cap B\cap C)$$

Bukti

Sebagai latihan.

Teorema 1. 4. Prinsip Inklusi-Eksklusi.

Jika $A_1, A_2, \ldots A_n$ merupakan deret peristiwa, maka

$$P\left(\bigcup_{i=1}^{n} A_i\right)=\sum_{i=1}^{n} P(A_i)-\sum_{k_1<k_2} P(A_{k_1}\cap A_{k_2})+\sum_{k_1<k_2<k_3} P(A_{k_1}\cap A_{k_2}\cap A_{k_3})$$

$$-\sum_{k_1<k_2<k_3<k_4} P(A_{k_1}\cap A_{k_2}\cap A_{k_3}\cap A_{k_4})+\ldots+(-1)^{n+1}P(\bigcap_{k=1}^{n} A_k)$$

Bukti

Sebagai latihan.

Teorema 1.5.

Jika $A \subset B$, maka $P(A) \leq P(B)$

Bukti

Sebagai latihan. Petunjuk: $B = (A \cap B) \cup (B \cap \overline{A})$

Teorema 1.6.

Pertidaksamaan Boole. Jika $A_1, A_2, \ldots$ merupakan deret peristiwa, maka

$$P\left(\bigcup_{i=1}^{\infty} A_i\right) \leq \sum_{i=1}^{\infty} P(A_i)$$

Bukti

Misalkan $B_1 = A_1$, $B_2 = A_2 \cap A_1^{'}$ dan secara umum $B_i = A_i \cap \left(\bigcup_{j=i}^{i-1} A_j^{'}\right)$. Hal ini berarti bahwa $\bigcup_{i=1}^{\infty} A_i = \bigcup_{i=1}^{\infty} B_i$ dan B_1, $B_2, \ldots$ saling lepas. Karena $B_i \subset A_i$ sebagai akibat dari Teorema 1.4. maka $P(\bigcup_{i=1}^{\infty} A_i) = P(\bigcup_{i=1}^{\infty} B_i) = \sum_{i=1}^{\infty} P(B_i) \leq \sum_{i=1}^{\infty} P(A_i)$.

Hasil yang sama juga berlaku untuk gabungan peristiwa terhingga. Secara khusus diperoleh $P(A_1 \cup A_2 \cup \ldots \cup A_k) \leq P(A_1)+P(A_2)+\ldots+P(A_k)$.

Teorema 1.7.

Pertidaksamaan Bonferoni. Jika $A_1, A_2, \ldots, A_k$ merupakan peristiwa-peristiwa, maka

$$P\left(\bigcap_{i=1}^{k} A_i\right) \geq 1 - \sum_{i=1}^{k} P(A_i^{'})$$

Bukti

Gunakan Teorema 1.1. dengan menggunakan $\bigcap_{i=1}^{k} A_i = \left(\bigcup_{i=1}^{k} A_i^{'}\right)^{'}$ bersamaan dengan penggunaan Teorema 1.5.

Teorema 1. 8.

Misalkan $\{A_n\}$ merupakan sekuen peristiwa tidak menurun dalam $\mathscr{S}$ atau dengan kata lain $A_n \in \mathscr{S}$, n = 1, 2, … dan $A_n \supseteq A_{n-1}$, n = 2, 3, …, maka $\lim_{n\to\infty} P(A_n) = P(\lim_{n\to\infty} A_n) = P\left(\bigcup_{n=1}^{\infty} A_n\right)$

Bukti

Misalkan $A = \bigcup_{j=1}^{\infty} A_j$, maka $A = A_n \bigcup_{j=n}^{\infty}\left(A_{j+1} \cap \overline{A}_j\right)$. Dengan menggunakan sifat aditif terhingga, kita peroleh bahwa

$P(A) = P(A_n) + \sum_{j=n}^{\infty} P(A_{j+1} - A_j)$. Sebagaimana n→∞ maka

$\sum_{j=n}^{\infty} P(A_{j+1} - A_j) \to 0$, karena $\sum_{j=1}^{\infty} P(A_{j+1} - A_j) \leq 1$ sehingga

$$\lim_{n\to\infty} P(A_n) = P(\lim_{n\to\infty} A_n) = P\left(\bigcup_{n=1}^{\infty} A_n\right)$$

Teorema 1. 9.

Misalkan $\{A_n\}$ merupakan sekuen peristiwa tidak menaik dalam $\mathscr{S}$ atau dengan kata lain $A_n \in \mathscr{S}$, n = 1, 2, … dan $A_n \subseteq A_{n-1}$, n = 2, 3, …, maka $\lim_{n\to\infty} P(A_n) = P(\lim_{n\to\infty} A_n) = P\left(\bigcap_{n=1}^{\infty} A_n\right)$

Bukti

Misalkan $\{A_n^c\}$ merupakan sekuen peristiwa tak menurun dalam $\mathscr{S}$,

maka $\lim\limits_{n\to\infty} A_n^c = \bigcup\limits_{j=1}^{\infty} A_j^c = A^c$. Dari **Teorema 1. 8.**

diperoleh hasil bahwa

$\lim\limits_{n\to\infty} P(A_n^c) = P(\lim\limits_{n\to\infty} A_n^c) = P\left(\bigcup\limits_{j=1}^{\infty} A_j^c\right) = P(A^c)$. Dengan perkataan lain $\lim\limits_{n\to\infty}(1 - P(A_n)) = 1 - P(A)$. Dengan argumentasi demikian, pembuktian dapat dilakukan.

Peluang Bersyarat

Tujuan utama dalam pemodelan peluang adalah untuk menentukan seberapa sering peristiwa *A* akan terjadi bilamana sebuah percobaan tertentu dilakukan. Namun demikian, terdapat banyak kasus dimana peluang terjadinya suatu peristiwa *A* akan dipengaruhi oleh informasi terjadi atau tidak terjadinya peristiwa lain, *B* misalnya. Terminologi yang sering digunakan adalah 'peluang bersyarat peristiwa *A* terjadi setelah peristiwa *B*' dan dinotasikan dengan *P(A|B)* yang akan membedakan dengan peluang biasa *P(A).*

Teladan 1.10.

Dalam percobaan dengan setumpuk kartu bridge standard, sebuah kartu diambil secara acak. Jika *A* adalah peristiwa terambilnya kartu ♥, maka *A* = *{* A♥ ,K♥ ,Q♥ ,J♥ ,10♥ ,9♥ ,8♥ ,7♥ ,6♥ ,5♥ ,4♥ ,3♥ ,2♥ *}*. Jika *B* adalah peristiwa terambilnya kartu K, maka *B* = *{* K♠ ,K♥ , K♣, K♦ *}*. Dengan demikian *P(A)* = *13/52* = ¼ dan *P(B)* = *4/52* = *1/13.* Namun apabila kita ingin menghitung *P(A|B)* yaitu peluang munculnya *A* setelah *B* kita seolah-olah hanya memandang anggota *B* saja yang dapat terjadi, dan dari apa yang ada disini berapa frekuensi relatif terambilnya kartu ♥. Jadi, $P(A|B) = ¼$.

Misalkan kita melakukan suatu percobaan dengan ruang contoh *S*, dan misalkan pula peristiwa *B* telah terjadi. Kita ingin mengetahui peluang peristiwa *A* terjadi dengan syarat peristiwa *B* telah terjadi, yang dinotasikan dengan $P(A|B)$. Ini sama artinya dengan kita menghitung peluang bahwa peristiwa *A* terjadi apabila kita hanya memperhatikan ruang contoh tereduksi *B.* Kita tahu bahwa $B=(A\cap B)\cup(A'\cap B)$. Peristiwa $(A\cap B)$ adalah sub peristiwa *B* dimana peristiwa *A* benar, sehingga peluang *A* setelah *B* harus proporsional terhadap $P(A\cap B)$, sebut saja $P(A|B)=k\,P(A\cap B)$. Dengan cara yang sama $P(A'|B)=k\,P(A'\cap B)$. Apabila kedua peluang ini, $P(A|B)$ dan $P(A'|B)$ dijumlahkan adlah total peluang relatif terhadap peristiwa *B*. Maka

$$\begin{aligned} P(A|B)+P(A'|B) &= k.[P(A\cap B)+P(A'\cap B)] \\ &= k.P[(A\cap B)\cup(A'\cap B)] \\ &= k.P(B) \\ &= 1 \end{aligned}$$

Dengan demikian nilai $k=1/P(B)$, yang juga merupakan konstanta proporsional yang membuat peluang dalam ruang contoh tereduksinya berjumlah satu.

Definisi 1.12.

Misalkan $(\Omega,\mathscr{S},P)$ merupakan ruang peluang, dan misalkan $B\in\mathscr{S}$ dengan $P(B)>0$. Maka, Peluang Bersyarat suatu peristiwa $A\in\mathscr{S}$, terjadi setelah peristiwa *B*, didefinisikan sebagai $P(A\mid B)=\dfrac{P(A\cap B)}{P(B)}$ apabila $P(B)\neq 0$.

Relatif terhadap ruang contoh tereduksi *B*, peluang bersyarat ini juga mengikuti aturan-aturan peluang lainnya, seperti

$$P(A|B)=1-P(A'|B)$$
$$0 \le P(A|B) \le 1$$
$$P(A_1 \cup A_2|B)=P(A_1|B)+P(A_2|B)-P(A_1 \cap A_2|B)$$

Teorema 1.10.

Untuk sembarang peristiwa *A* dan *B*,

$$P(A \cap B)=P(B)P(A|B)=P(A)P(B|A)$$

Teorema ini sering disebut dengan **Teorema Perkalian Peluang**. Dalam pengambilan contoh tanpa pengembalian, teorema ini sangat bermanfaat. Sebagai ilustrasi, apabila sebuah kartu diambil satu per satu tanpa pengembalian dari setumpuk kartu bridge standard yang sudah diacak, berapakah peluang terambilnya 3 kartu As berturut-turut ? Jika A_j adalah peristiwa terambilnya kartu As pada penarikan ke-j maka

$$P(A_1 \cap A_2 \cap A_3)=P(A_1)P(A_2|A_1)P(A_3|A_1 \cap A_2)=(4/52)(3/51)(2/50)$$

Peristiwa $A_1 \cap A_2 \cap A_3$ adalah peristiwa terambilnya kartu As pada pengambilan pertama, kedua, dan ketiga yang ingin dicari peluangnya. Bila setumpuk kartu masih lengkap akan terdapat 52 buah dan didalamnya terdapat 4 kartu As. Dengan syarat bahwa pada pengambilan pertama ada satu kartu As yang terambil, maka dalam tumpukan sekarang hanya ada 3 kartu As dari 51 kartu tersisa. Peristiwa ini dinotasikan dengan $A_2|A_1$. Dan akhirnya akibat terambilnya kartu As pada penarikan kedua, maka jumlah kartu As tersisa ada 2 lagi dari 50 buah kartu tersisa, dan peristiwanya dapat dituliskan sebagai $A_3|A_1 \cap A_2$.

Teorema 1. 11. Aturan Perkalian

Misalkan $(\Omega, \mathscr{S}, P)$ merupakan ruang peluang dan $A_1, A_2, \ldots, A_n \in \mathscr{S}$, dengan $P\left(\bigcap_{j=1}^{n-1} A_j\right) > 0$, maka

$$P\left\{\bigcap_{j=1}^{n} A_j\right\} = P(A_1)P(A_2 \mid A_1)P(A_3 \mid A_1 \cap A_2)...P\{A_n \mid \bigcap_{j=1}^{n-1} A_j\}$$

Bukti: sebagai latihan.

Total Peluang dan Aturan Bayes

Kadang sangat bermanfaat membagi suatu peristiwa menjadi dua atau lebih peristiwa yang saling lepas. Secara umum jika $B_1, B_2, \ldots, B_k$ adalah peristiwa-peristiwa yang saling lepas dan jenuh (*mutually exclusive and exhaustive*) yang mana $B_1 \cup B_2 \cup \ldots \cup B_k = S$, maka

$$A = (A \cap B_1) \cup (A \cap B_2) \cup \ldots \cup (A \cap B_k)$$

Teorema 1.12.

Jika $B_1, B_2, \ldots, B_k$ adalah sekumpulan peristiwa yang saling lepas dan jenuh, maka untuk sembarang peristiwa A,

$$P(A) = \sum_{i=1}^{k} P(B_i)P(A \mid B_i)$$

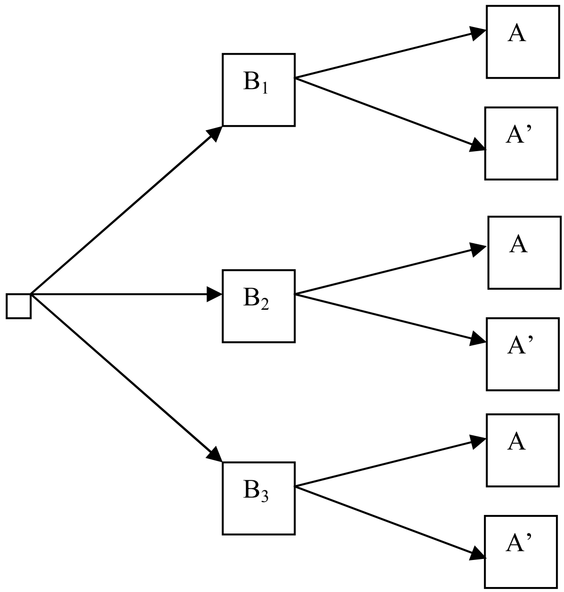
Bukti

Karena peristiwa-peristiwa $A \cap B_1, A \cap B_2, \ldots A \cap B_k$ saling lepas, maka kita peroleh

$$P(A) = \sum_{i=1}^{k} P(A \cap B_i)$$

Dan dengan menggunakan $P(A \cap B) = P(B)P(A|B)$ yang diaplikasikan untuk setiap term dalam penjumlahan tersebut. Teorema ini sering disebut dengan **Teorema Total Peluang**.

Diagram pohon diatas dapat dipergunakan untuk memudahkan konsep total peluang seperti tertera pada Teorema 1.8. Peluang yang berkenaan dengan peristiwa setiap B_i adalah $P(B_i)$, dan peluang yang berkenaan dengan anak cabang yang ditandai dengan A adalah peluang bersyarat $P(A|B_i)$.



Teladan 1.11.

Jika B_1 adalah peristiwa microprocessor bikinan pabrik X, B_2 adalah peristiwa microprocessor bikinan pabrik Y, dan B_3 adalah peristiwa microprocessor bikinan pabrik Z. Misalkan dari pabrik X, Y, dan Z masing-masing diambil sampel berturut-turut sebanyak 35, 25, dan 40 dan kemudian diperiksa apakah terdapat microchip yang cacat (peristiwa A) atau tidak (peristiwa A’), hasilnya diperoleh seperti pada tabel berikut:

	X	**Y**	**Z**	Total
Cacat, A	10	5	5	20
Tidak Cacat, A’	25	20	35	80
Total	35	25	40	100

Dari tabel diatas tampak bahwa *$P(B_1)$=35/100, $P(B_2)$=25/100* dan *$P(B_3)$=40/100* serta *$P(A|B_1)$=10/35, $P(A|B_2)$=5/25* dan *$P(A|B_3)$= 5/40.* Sehingga dengan menggunakan total peluang, kita peroleh

$$
\begin{aligned}
P(A) &= P(B_1)P(A|B_1) + P(B_2)P(A|B_2) + P(B_3)P(A|B_3) \\
&= (35/100)(10/35)+(25/100)(5/25)+(40/100)(5/40) \\
&= 0{,}10 + 0{,}05 + 0{,}05 = 0{,}20
\end{aligned}
$$

Dari teladan ini dapat dilihat bahwa cara mendapatkan nilai peluang diperoleh micropocessor yang cacat dapat digunakan tabel ataupun dengan menggunakan total peluang.

Sekarang perhatikan modifikasi dari teladan ini. Misalkan microprocessor tersebut disortir dalam tiga kotak dimana kotak I terdiri dari microprocessor dari pabrik X, kotak II dari pabrik Y, dan kotak III dari pabrik Z. Bila percobaan baru adalah memilih satu kotak secara acak, kemudian memilih sebuah microprocessor dari kotak terpilih, maka berapakah peluang mendapatkan sebuah microprocessor dalam keadaan cacat ? Dalam hal ini tidaklah bisa digunakan dari tabel secara langsung dalam perhitungan mencari peluang tersebut. Kita harus meredifinisi peluang terambilnya masing-masing kotak atau *$P(B_1)$=1/3, $P(B_2)$=1/3* dan *$P(B_3)$=1/3* serta *$P(A|B_1)$=10/35, $P(A|B_2)$=5/25* dan *$P(A|B_3)$= 5/40.* Dengan demikian

$$
\begin{aligned}
P(A) &= P(B_1)P(A|B_1) + P(B_2)P(A|B_2) + P(B_3)P(A|B_3) \\
&= (1/3)(10/35)+(1/3)(5/25)+(1/3)(5/40) \\
&= 10/105+5/75+5/120 \\
&= 57/280
\end{aligned}
$$

Teorema 1.13. Aturan Bayes

Jika kita asumsikan kondisi seperti pada Teorema 1.8, maka untuk setiap *j = 1,2,…,k*

$$P(B_j \mid A) = \frac{P(B_j)P(A \mid B_j)}{\sum_{i=1}^{k} P(B_i)P(A \mid B_i)}$$

Bukti

Dengan menggunakan definisi peluang bersyarat $P(B \mid A) = \dfrac{P(A \cap B)}{P(A)}$ dan $P(A \cap B) = P(B)P(A|B)$ juga dari total peluang $P(A) = \sum_{i=1}^{k} P(B_i)P(A \mid B_i)$ kita peroleh

$$P(B_j \mid A) = \frac{P(B_j \cap A)}{P(A)} = \frac{P(A \cap B_j)}{P(A)} = \frac{P(B_j)P(A \mid B_j)}{\sum_{i=1}^{k} P(B_i)P(A \mid B_i)}$$

Teladan 1.12.

Dengan menggunakan teladan sebelumnya, apabila sebuah microprocessor didapatkan cacat, berapakah peluang bahwa microprocessor yang cacat tersebut datangnya dari pabrik X ? atau kita ingin mencari $P(B_2|A)$

$$P(B_2 \mid A) = \frac{(1/3)(10/35)}{(1/3)(10/35) + (1/3)(5/25) + (1/3)(5/40)} = \frac{80}{171} = 0{,}468$$

Dengan cara yang sama bisa diperoleh bahwa $P(B_1|A) = 0{,}327$ dan $P(B_3|A) = 0{,}205$.

Peristiwa Bebas

Pengetahuan tentang peristiwa *A* telah terjadi, dalam beberapa situasi, tidak akan mempengaruhi peluang peristiwa *B* akan terjadi. Secara notasi statistik ini dituliskan dengan $P(B|A)=P(B)$. Sebagai akibatnya Teorema Perkalian akan menghasilkan $P(A \cap B)=P(A)P(B|A)=P(A)P(B)$. Secara umum, apabila hal ini terjadi, dua peristiwa disebut saling bebas atau saling bebas stokastik.

Definisi 1.13.

Dua peristiwa A dan B dikatakan peristiwa-peristiwa bebas atau saling bebas (*independent*) jika $P(A \cap B)=P(A)P(B)$. Selain itu, peristiwa A dan B dinyatakan saling terkait (*dependent*).

Teorema 1.14.

Jika A dan B adalah peristiwa-peristiwa sedemikian rupa sehingga $P(A)>0$ dan $P(B)>0$, maka peristiwa A dan B saling bebas jika dan hanya jika salah satu hal berikut dipenuhi:

$$P(A|B)=P(A) \text{ atau } P(B|A)=P(B)$$

Teorema 1.15.

Dua peristiwa A dan B saling bebas jika dan hanya jika pasangan peristiwa berikut juga saling bebas

1. A dan B'
2. A' dan B
3. A' dan B'

Bukti

Sebagai latihan.

Definisi 1.14.

Sebanyak k peristiwa $A_1, A_2, \ldots, A_k$ dikatakan bebas atau saling bebas jika untuk setiap $j=2,3,\ldots,k$ dan setiap anak gugus dari indeks $i_1,i_2,\ldots,i_j$

$$P(A_{i_1} \cap A_{i_2} \cap \ldots \cap A_{i_j}) = P(A_{i_1})P(A_{i_2})\ldots P(A_{i_j})$$

Teknik Penghitungan

Dalam berbagai hal apabila kita menggunakan ruang contoh terbatas, seperti permainan, mungkin lebih beralasan untuk mengasumsikan bahwa semua kemungkinan muncul dengan kesempatan yang sama. Dengan demikian, model peluang realistis menggunakan metode klasik dan peluang suatu peristiwa *A* dinotasikan dengan $P(A) = n(A)/N$, dimana *N* adalah total seluruh kemungkinan keluaran dan $n(A)$ adalah banyaknya kemungkinan peristiwa *A* terjadi. Menghitung banyaknya cara suatu peristiwa dapat terjadi dapat saja merupakan masalah yang rumit dalam percobaan yang kompleks. Beberapa teknik menghitung yang mungkin berguna akan disajikan disini.

Prinsip Multiplikasi

Apabila suatu operasi dapat dilakukan dalam n_1 cara dan operasi berikutnya dapat dilakukan dalam n_2 cara, maka secara keseluruhan terdapat sebanyak n_1n_2 cara dimana kedua operasi tersebut dilakukan.

Teladan 1.13.

Misalkan Antyo memiliki 5 baju lengan panjang warna terang yang diperbolehkan dipakai dikantor serta 4 celana panjang warna gelap untuk kegunaan yang sama. Dengan demikian, Antyo dapat memakai (5)(4) = 20 kombinasi baju dan celana panjang yang dapat dipakai bekerja.

Prinsip multiplikaksi ini dapat diperluas untuk lebih dari dua operasi. Lebih khusus lagi, jika sebanyak *r* operasi ke-j dapat dilaksanakan dalam n_j cara, maka keseluruhan *r* operasi tersebut akan menghasilkan sebanyak

$$\prod_{j=1}^{r} n_j = (n_1)(n_2)...(n_r)$$

Teorema 1.16.

Jika terdapat *N* kemungkinan keluaran dari tiap *r* tindakan dalam suatu percobaan, maka akan didapatkan sebanyak N^r kemungkinan keluaran dalam ruang contohnya.

Teladan 1.14.

Misalkan ada 15 soal pilihan berganda dalam suatu ujian, dimana setiap soal memiliki 5 jawaban. Dengan demikian, total seluruh kemungkinan jawaban yang terjadi adalah 5^{15}.

Permutasi dan Kombinasi

Dalam satu hal terambilnya 5 kartu {A♠, K♠, Q♠, J♠, 10♠} dan {10♠, A♠, K♠, J♠, Q♠} dapat merupakan peristiwa yang sama, tetapi juga dapat merupakan peristiwa yang tidak sama. Apabila kita inginkan keluaran tersebut berdasarkan urutan keluarnya, maka sudah jelas kedua peristiwa tersebut tidak sama. Namun apabila urutan keluarnya tidak dipentingkan, melainkan apa-apa saja yang menjadi anggota dalam peristiwa tersebut, maka kedua peristiwa tersebut dikatakan sama.

Hasil kali dari bilangan-bilangan bulat positif dari 1 sampai dengan n, yaitu (1)(2)(3) … (n-2)(n-1)(n) = n! (dibaca n faktorial). Misalnya 3! = (3)(2)(1) = 6 dan 5! = (5)(4)(3)(2)(1) = 120 dan sebagainya.

Teorema 1.17.

Banyaknya permutasi dari sebanyak *n* obyek yang dapat dibedakan adalah *n!*

Bukti

Dengan menggunakan teknik multiplikasi. Untuk mengisi *n* posisi dengan *n* obyek yang berbeda, maka posisi pertama dapat diisi dengan *n* cara, posisi kedua dengan *n-1* cara, dan seterusnya sampai obyek terakhir ditempatkan pada posisi terakhir. Dengan demikian, operasi ini dapat dilakukan dalam *n(n-1)(n-2)…1= n!* cara.

Teorema 1.18.

Banyaknya permutasi dari *n* obyek yang berbeda diambil sebanyak *r* sekaligus adalah

$$ {}_nP_r = P_r^n = \frac{n!}{(n-r)!} $$

Teorema ini dipakai apabila seseorang tertarik pada banyaknya cara memilih *r* obyek dari sebanyak *n* obyek yang berbeda dan kemudian mengurutkan *r* obyek tersebut.

Teladan 1.15.

Dari keempat pemain Pencak Silat terbaik yang dimilikinya (A, B, C, dan D), IPSI harus memilih dua teratas diantaranya berdasarkan ranking. Oleh karenanya seluruh kemungkinan susunan dua pemain terbaik tersebut adalah:

AB	AC	AD	BC	BD	CD
BA	CA	DA	CB	DB	DC

Soal. Suatu kotak terdiri dari *n* tiket yang diberi nomor 1, 2, 3, …, n. Jika tiga tiket dipilih secara acak tanpa pengembalian, berapakah peluang mendapatkan tiket dengan nomor yang berurut ?

Teorema 1.19.

Banyaknya kombinasi *n* obyek yang berbeda dan diambil sebanyak *r* sekaligus adalah

$$\binom{n}{r} = C_r^n = \frac{n!}{r!(n-r)!}$$

Bukti

Sebagaimana telah disarankan dalam teladan sebelumnya, bahwa P_r^n mungkin diinterpretasikan sebagai banyaknya cara memilih *r* obyek dari *n* obyek dan kemudian melakukan permutasi *r* obyek dengan *r!* cara dan menghasilkan

$${}_nP_r = \binom{n}{r} r! = \frac{n!}{(n-r)!}$$

Dengan membagi *r!* kita peroleh hasil yang dimaksud.

Teorema 1.20.

Banyaknya permutasi yang dapat dibedakan dari sebanyak *n* obyek dimana sebanyak *r* darinya adalah sejenis dan *n-r* adalah jenis lain adalah

$$\binom{n}{r} = C_r^n = \frac{n!}{r!(n-r)!}$$

Beberapa catatan penting yang berkaitan dengan penghitungan kombinasi

1. $C_0^n = 1$
2. $C_1^n = n$
3. $C_{n-1}^n = n$

4. $C_n^n = 1$

Teorema 1.21.

Banyaknya permutasi dari *n* obyek yang terdiri dari *k* jenis dimana masing-masing jenis berturut-turut banyaknya $r_1, r_2, \ldots, r_k$ adalah

$$\frac{n!}{r_1!r_2!\ldots r_k!}$$

Teladan 1.16.

Banyaknya susunan huruf (belum tentu merupakan kalimat) dari huruf-huruf penyusun BARBARA adalah

$$\frac{7!}{2!3!2!} = \frac{(7)(6)(5)(4)(3)(2)(1)}{(2)(1)(3)(2)(1)(2)(1)} = 210$$

Latihan

1. Lakukan verifikasi untuk hal-hal berikut ini, untuk $n > r$:
   a. $C_{n-r}^n = C_r^n$
   b. $C_r^{n+1} = C_{r-1}^n + C_r^n$
   c. $(a+b)^n = \sum_{k=0}^{n} C_k^n a^k b^{n-k}$
2. Tentukan nilainya
   a. $C_0^4 + C_2^4 + C_4^4$
   b. $C_0^6 + C_2^6 + C_4^6 + C_6^6$
   c. $\sum_{i=0}^{n} C_{2i}^{2n}$
3. Nyatakan, dengan metode yang mana peluang dari tiap pernyataan di bawah ini ditentukan:

   a. Berdasarkan data pemasaran masa lalu, maka peluang menjual lebih dari 5000 tiket bus dalam satu bulan adalah 0,15
   b. Seorang teknisi menduga bahwa terdapat 50 persen kemungkinan penggunaan jasa telepon menurun sebesar 10 persen dalam 3 tahun terakhir.
   c. Walaupun terdapat 20 dari 135 barang rusak, peluang barang yang rusak diasumsikan sebesar 0,20.
   d. Tidaklah benar untuk menyatakan bahwa 1 dari 4 kartu yang diambil dari setumpuk kartu bridge standar adalah ♥.
4. Ada berapa banyak anak gugus yang dapat dibuat dari suatu gugus yang memiliki anggota sebanyak *m* ?
5. Ada berapa kemungkinan apabila 5 buah kartu diambil dengan pengembalian dari setumpuk kartu bridge standar ?
6. Tentukan banyaknya kemungkinan susunan huruf yang dapat dibuat dari huruf-huruf penyusun kata seperti di bawah ini:
   a. MAHAKAM
   b. MISSISSIPPI
   c. LONDON
   d. ABRACADABRA
   e. MARTAPURA
   f. MATAHARI
   g. KOMODO
7. Dari sebuah pelatihan yang diikuti oleh 24 peserta, berapa peluang terdapat tiga orang atau lebih yang memiliki bulan kelahiran yang sama ?
8. Apabila terdapat 5 orang (3 laki-laki dan 2 perempuan) yang dinyatakan lolos tahap pertama seleksi ujian pegawai di suatu institusi, dan apabila tahap kedua atau terakhir hanya akan diambil 2 orang saja, berapakah peluang yang akan lolos 1 laki-laki dan 1 perempuan ?

9. Sebuah kotak berisikan tiga kartu Joker dan dua kartu As. Si A mengambil sebuah kartu dari kotak tersebut secara acak, kemudian diikuti si B. hitunglah
   a. P(A dapat kartu Joker)
   b. P(B kartu Joker|A kartu Joker)
   c. P(B kartu Joker|A kartu As)
   d. P(B kartu Joker $\cap$ A kartu Joker)
   e. P(A kartu As|B kartu Joker)
10. Misalkan $P(A_j)=1/(3+j)$ untuk $j=1,2,3,4$. Carilah batas atas untuk $P(A_1 \cup A_2 \cup A_3 \cup A_4)$.
11. Kantong A terdiri dari 5 kelereng merah dan 3 kelereng biru, dan kantong B terdiri dari 4 kelereng merah dan 4 kelereng biru. Seluruh kelereng tersebut berukuran sama sehingga susah sekali dibedakan jika diraba. Sebuah tindakan mengambil dua kelereng sekaligus dari kantong A, dan tanpa melihat dimasukkan ke kantong B. Kemudian dua buah kelereng diambil secara acak dari kantong B. Hitungkah:
    a. P(terambilnya 1 merah dan 1 biru dari kantong B)
    b. P(terambilnya 2 biru dari kantong B)
12. Plat nomor kendaraan untuk suatu daerah menggunakan dua huruf dan diikuti dengan 4 digit-angka.
    a. Berapa banyak kemungkinan susunan plat nomor yang dapat dibuat apabila huruf dan angka diperbolehkan diulang ?
    b. Berapa banyak kemungkinan susunan plat nomor yang dapat dibuat apabila huruf tak boleh diulang tetapi angka diperbolehkan diulang ?
    c. Berapa banyak kemungkinan susunan plat nomor yang dapat dibuat apabila huruf boleh diulang tetapi angka tidak diperbolehkan diulang ?
13. Sepuluh orang yang semuanya bisa mengendarai mobil ingin bepergian bersama dengan menggunakan tiga jenis mobil Alfa Romeo, Bugatti, dan Lamborghini. Apabila setiap

mobil hanya boleh diisi paling banyak 4 orang sesuai peraturan yang berlaku di suatu daerah

a. Berapakah banyaknya kemungkinan susunan mereka duduk dalam mobil ?
b. Berapakah banyaknya kemungkinan susunan mereka duduk dalam mobil jika si A dan si B harus berada dalam satu mobil ?
c. Berapakah banyaknya kemungkinan susunan mereka duduk dalam mobil jika si C dan si D harus berada dalam satu mobil ?

14. Untuk $(x + y + z)^7$ maka tentukan besarnya koeffisien dari masing-masing suku-suku berikut:
    a. x^4yz^2
    b. x^5yx
    c. $x^3y^2z^2$
    d. y^2z^5
15. Suatu kelas memiliki sebanyak *r* mahasiswa. Tanggal lahir dari *r* mahasiswa ini membentuk suatu contoh berukuran *r* dari 365 dalam setahun. Tentukan peluang bahwa semua mahasiswa memiliki tanggal lahir yang tidak sama
16. Tentukan peluang seorang sekretaris melakukan kesalahan bahwa *n* surat yang diketiknya semuanya salah dimasukkan kedalam *n* amplop yang seharusnya ?
17. Misalkan $(\Omega, \mathscr{S}, P)$ merupakan ruang peluang, dan misalkan $H \in \mathscr{S}$ dengan $P(H) > 0$. Tunjukkan bahwa $(\Omega, \mathscr{S}, P_H)$ merupakan ruang peluang, jika $P_H(A) = P\{A \mid H\}$ untuk semua $A \in \mathscr{S}$.
18. Sebanyak *r* kelereng didistribusikan secara acak kedalam *n* kotak. Misalkan A_k merupakan kejadian bahwa kotak tertentu berisikan sebanyak *k* kelereng. Tentukan peluang A_k tersebut.

Peubah Acak

Pendahuluan

Tujuan kita disini adalah mengembangkan model matematis untuk menggambarkan peluang terjadinya suatu keluaran atau suatu peristiwa dalam suatu ruang contoh. Karena persamaan matematis diekspresikan dalam nilai-nilai numerik daripada munculnya Gambar, Warna, atau hal lainnya; karenanya kita definisikan suatu fungsi, yang dikenal dengan nama **Peubah Acak** yang menghubungkan setiap keluaran dari suatu percobaan dengan suatu bilangan riel. Dengan demikian kita dapat mengekspresikan peluang model suatu percobaan dengan menggunakan peubah acak ini. Sudah tentu, dalam beberapa percobaan hasilnya memang sudah berupa angka, dan dalam kasus itu fungsi natural yang digunakan sebagai peubah acak adalah fungsi identitas.

Definisi 2.1.

Suatu peubah acak, misalkan saja *X*, adalah suatu fungsi yang terdefinisi pada suatu ruang contoh, *S*, yang dihubungkan dengan suatu bilangan nyata, *X(e)*=*x* dengan tiap kemungkinan keluaran *e* dalam *S.*

Atau dapat didefinisikan seperti berikut

Definisi 2. 2.

Misalkan $(\Omega, \mathscr{S})$ merupakan ruang contoh. Suatu fungsi bernilai tunggal terhingga yang memetakan Ω kedalam $\mathscr{R}$ disebut dengan peubah acak jika $X^{-1}(B) = \{\omega : X(\omega) \in B\} \in \mathscr{S}$ untuk semua B $\in$ B (gugus Borel dalam bilangan riil, $\mathscr{R}$).

Teorema 2. 1.

X merupakan peubah acak jika dan hanya jika untuk setiap $x \in \mathscr{R}$, $\{\omega: X(\omega) \leq x\} = \{X \leq x\} \in \mathscr{S}$.

Teorema 2. 2.

Misalkan X merupakan peubah acak yang terdefinisi pada $(\Omega, \mathscr{S})$, dan *a*, *b* konstanta, maka *aX*+*b* juga merupakan peubah acak yang tersefinisi pada $(\Omega, \mathscr{S})$.

Bukti: sebagai latihan.

Teladan 2. 1.

Misalkan Ω = {H, T}, dan $\mathscr{S}$ merupakan kelas semua anak gugus Ω. Dengan mendefinisikan *X(H) =1* dan *X(T)=0*, maka

$$X^{-1}(-\infty, x] = \begin{cases} \varnothing & , \quad x < 0 \\ \{T\} & , \quad 0 \leq x < 1 \\ \{H, T\} & , \quad 1 \leq x \end{cases}$$

Dan sesuai definisi maka *X* adalah peubah acak, karena $\varnothing \in \mathscr{S}$, $\{T\} \in \mathscr{S}$, dan $\{H, T\} \in \mathscr{S}$.

Teladan 2.2.

Sebuah dadu empat muka (tetrahedral) memiliki nomor yang berbeda yaitu 1, 2, 3, atau 4. Misalkan bahwa dadu tetrahedral tersebut seimbang, artinya setiap muka dadu memiliki kesempatan yang sama untuk muncul. Untuk kasus ini yang dimaksud dengan muncul adalah yang mengahadap kebawah.

Suatu permainan dengan dadu ini dilakukan dengan cara melempar dadu dua kali berurutan, dan skor yang diambil adalah maksimum dari angka yang muncul dari tiap dadu. Meskipun skor tak dapat diprediksi, kita dapat menentukan gugus nilai yang mungkin dan mendefinisikan peubah acak. Misalkan $e = (i, j)$ dimana $i,j \in \{1,2,3,4\}$, maka $X(e)=max(i,j)$

Ruang contoh dari percobaan diatas adalah {(1,1), (1,2), (1,3), (1,4), (2,1), (2,2), (2,3), (2,4), (3,1), (3,2), (3,3), (3,4), (4,1), (4,2), (4,3), (4,4)}. Dan X ={1, 2, 3, 4, 2, 2, 3, 4, 3, 3, 3, 4, 4, 4, 4, 4}.

Apabila peubah acak lain didefinisikan, misalnya $Y(e)=i+j$ maka Y={2,3,4,5,3,4,5,6,4,5,6,7,5,6,7,8}

Konsep peubah acak juga dapat dipakai apabila ruang contohnya adalah bilangan nyata, dan dengan demikian bahwa peristiwanya adalah anak gugus dari bilangan nyata. Seandainya peristiwa bernilaikan bilangan nyata tersebut adalah *A*, maka gugus $B = \{e \mid e \in S \quad and \quad X(e) \in A\}$ adalah peristiwa yang berada dalam ruang contoh *S.* Meskipun peristiwa *A* dan *B* adalah anak gugus dari ruang yang berbeda, namun biasanya dapat dikatakan merupakan peristiwa-peristiwa yang ekuivalen, dan $P[X \in A]=P(B)$.

Notasi $P_X(A)$ biasanya digunakan daripada $P[X \in A]$. Ini mendefinisikan fungsi gugus pada kumpulan peristiwa-peristiwa bernilaikan bilangan bulat, dan dapat ditunjukkan memenuhi tiga kondisi dasar suatu fungsi gugus peluang.

Pendekatan yang lebih umum dalam pemberian nilai peluang untuk peristiwa-peristiwa dalam ruang contoh bilangan nyata dapat berdasarkan pemberian peluang pada interval dalam bentuk $(-\infty, x)$ untuk semua bilangan nyata x. Untuk semua bilangan nyata x, gugus dalam bentuk B = $[\,X \le x\,] = \{\,e \mid e \in S$ dan $X(e) \in (-\infty, x)\}$ adalah peristiwa dalam ruang contoh *S*.

Peubah Acak Diskrit

Definisi 2.3.

Jika gugus semua nilai yang mungkin dari peubah acak *X* merupakan gugus terhitung $x_1, x_2, \ldots, x_n$ atau $x_1, x_2, \ldots$ maka *X* disebut dengan peubah acak diskrit. Fungsi $f(x)$ = P[X=x] untuk x = $x_1, x_2, \ldots$ mengalokasikan peluang untuk setiap kemungkinan nilai x yang disebut dengan Fungsi Kepekatan Peluang Diskrit atau Fungsi Densitas Peluang Diskrit.

Pendefinisian lain dari yang tersebut diatas

Definisi 2. 4.

Suatu peubah acak *X* yang terdefinisi pada $(\Omega, \mathscr{S}, P)$ dikatakan diskrit, jika terdapat suatu gugus terhingga $E \subseteq R$ sedemikian rupa sehingga $P(X \in E\} = 1$.

Titik-titik atau elemen *E* yang memiliki peluang positif disebut dengan titik lompatan atau titik kenaikan fungsi sebaran *X*, dan peluangnya merupakan besarnya lompatan atau kenaikan fungsi sebaran tersebut.

Definisi 2. 5.

Sekumpulan bilangan $\{p_i\}$ yang memenuhi $P\{X = x_i\} = p_i \geq 0$, untuk semua i dan $\sum_{i=1}^{\infty} p_i = 1$ disebut dengan fungsi kepekatan peluang. Notasi lain yang digunakan adalah $f(x_i)$.

Teorema 2.3.

Suatu fungsi *f(x)* merupakan fungsi kepekatan peluang diskrit (*fkp* diskrit) jika dan hanya jika memenuhi kedua sifat berikut : untuk setiap anggota gugus bilangan terbilang tak terhingga $x_1, x_2, \ldots$ $f(x_i) \geq 0$ untuk semua x_i, dan $\sum_{\forall x_i} f(x_i) = 1$

Teladan 2.3.

Fungsi kepekatan peluang diskret maksimum nilai yang muncul dari pelemparan dua kali dadu tetrahedral dapat disajikan seperti berikut:

x	1	2	3	4
f(x)	1/16	3/16	5/16	7/16

Sedangkan bila peubah acak yang kita inginkan adalah jumlah nilai yang muncul dari dua kali pelemparan dadu tersebut, maka fungsi kepekatan peluangnya adalah seperti berikut

y	2	3	4	5	6	7	8
f(y)	1/16	2/16	3/16	4/16	3/16	2/16	1/16

Definisi 2. 6.

Fungsi Sebaran Kumulatif dari suatu peubah acak *X* didefinisikan untuk sembarang bilangan nyata *x* dengan $F(x) = P[X \leq x]$

Atau dapat didefinisikan seperti berikut

Definisi 2. 7.

Misalkan peubah acak *X* terdefinisi pada $(\Omega, \mathscr{S}, P)$. Suatu fungsi *F(.)* pada $\mathscr{R}$ didefinisikan dengan $F(x)=P\{\omega: X(\omega) \leq x\}$ untuk semua $x \in \mathscr{R}$.. Fungsi *F* disebut dengan fungsi sebaran peubah acak *X*.

Fungsi *F(x)* sering cukup disebut dengan Fungsi Sebaran peubah acak *X*, juga notasinya sering menggunakan $F_X(x)$.

Teorema 2.4.

Misalkan *X* adalah peubah acak diskret dengan fungsi kepekatan peluang *f(x)* dan fungsi sebaran kumulatif *F(x)*. Jika nilai yang mungkin dari peubah acak *X* diindeks dalam urutan menaik, $x_1<x_2<x_3<\ldots$ maka $f(x_1) = F(x_1)$, dan untuk sembarang $i>1$,

$$f(x_i)=F(x_i)-F(x_{i-1})$$

Lebih lanjut, jika $x<x_1$ maka $F(x)=0$ dan untuk sembarang nilai bilangan nyata *x* lainnya

$$F(x)=\sum_{x_i\le x} f(x_i)$$

dimana penjumlahan dilakukan untuk seluruh indeks *i* sedemikian rupa sehingga $x_i \le x$.

Fungsi Sebaran Kumulatif untuk sembarang peubah acak harus memenuhi sifat-sifat seperti yang ada pada teorema berikut.

Teorema 2. 5.

Untuk suatu peluang *Q* pada ($\mathscr{R}$, $\mathscr{B}$), terdapat suatu fungsi sebaran *F* yang memenuhi $Q(-\infty,x] = F(x)$ untuk semua $x \in \mathscr{R}$, dan begitupula sebaliknya bila diberikan suatu fungsi sebaran *F*, akan terdapat peluang yang khas *Q* terdefinisi pada ($\mathscr{R}$, $\mathscr{B}$) yang memenuhi $Q(-\infty,x] = F(x)$ untuk semua $x \in \mathscr{R}$.

Teorema 2.6.

Suatu fungsi F(x) adalah fungsi sebaran kumulatif dari peubah acak X, jika dan hanya jika persyaratan berikut dipenuhi:

1. $\lim_{x\to-\infty} F(x)=0$

2. $\lim_{x \to +\infty} F(x) = 1$
3. $\lim_{h \to 0^+} F(x+h) = F(x)$
4. Jika *a*<*b*, maka *F(a)* $\leq$ *F(b)*

Syarat pertama sebagai akibat bahwa fungsi peluang suatu titik atau peristiwa elementer tidak negatif, dan untuk nilai x yang sangat negatif nilai kumulatifnya sama dengan nol, sedangkan sebagai akibat bahwa total peluang adalah 1, maka konsekuensi dari kumulatif peluang tertuang dalam syarat nomor dua. Syarat nomor 3 sering disebut dengan kontinu dari sebelah kanan, sedangkan syarat nomor 4 disebut dengan fungsi yang tidak menurun.

Definisi 2.8.

Jika X adalah peubah acak diskret dengan fungsi kepekatan peluang *f(x)*, maka **nilai harapan** dari X didefinisikan sebagai

$$E(X) = \sum_x x f(x) .$$

Penjumlahan bisa saja dilakukan untuk range nilai x yang terhingga jumlahnya atau bisa jadi penjumlahan tersebut untuk seluruh kemungkinan x yang jumlahnya terbilang tak terhingga sehingga perlu pengetahuan tentang deret suatu bilangan.

Teladan 2.4.

Sebagai mana teladan sebelumnya tentang percobaan dengan dadu tetrahedral diperoleh

x	1	2	3	4
f(x)	1/16	3/16	5/16	7/16

Dengan demikian nilai harapan dari maksimum munculnya mata dadu dari pelemparan dua kali dadu tetrahedral adalah : (1)(1/16)+(2)(3/16)+(3)(5/16)+(4)(7/16) = 50/16 = 3,125

Hal yang lebih umum lagi bisa dikembangkan untuk mencari nilai harapan dari suatu fungsi peubah acak, misalnya *u(X)*. Jika X adalah peubah acak diskret dengan fungsi kepekatan peluang *f(x)*, serta *u(X)* adalah suatu fungsi dari X, maka nilai harapan dari *u(X)* didefinisikan sebagai

$$E[u(X)] = \sum_x u(x) f(x).$$

Ragam atau Varian peubah acak X merupakan salah satu nilai harapan penting yang dapat diperoleh dengan cara ini, dengan $u(X)=(X-\mu)^2$ Notasi yang umum dipakai untuk menyatakan ragam adalah σ^2, σ^2_X atau V(X) yang nilainya adalah $E[u(X)]=E[(X-\mu)^2]$

$$\begin{aligned}
\text{Var(X)} &= \sum_x (x-\mu)^2 f(x) \\
&= \sum_x (x^2 - 2\mu x + \mu^2) f(x) \\
&= \sum_x x^2 f(x) - 2\mu \sum_x x f(x) + \mu^2 \sum_x f(x) \\
&= E(X^2) - 2\mu E(X) + \mu^2 \\
&= E(X^2) - 2\mu\mu + \mu^2 \\
&= E(X^2) - \mu^2
\end{aligned}$$

Catatan : Nilai harapan dari X adalah juga merupakan nilai rata-rata dari X yang dilambangkan dengan μ.

Sebaran Bernoulli

Performans dari suatu percobaan dengan hanya memiliki dua macam keluaran disebut dengan tindakan Bernoulli. Bila kemungkinan keluaran tersebut kita sebut dengan 'Berhasil' dan 'Gagal', maka peubah acak Bernoulli adalah

$$X(e) = \begin{cases} 1 \; jika \; e \in E \\ 0 \; jika \; e \in E^c \end{cases}$$

Sedangkan fungsi kepekatan peluangnya adalah *f(0)=q* dan *f(1)=p*. Sebaran dimaksud sering disebut dengan Sebaran Bernoulli, dan fungsi kepekatan peluangnya dapat diekspresikan sebagai

$$f(x) = p^x q^{1-x} \quad \text{untuk x = 0 atau 1}$$

Besarnya q = 1-p dan 0 < p < 1.

Nilai harapan dari X atau *E(X) = (0)(q)+(1)(p) = p* sedangkan $E(X^2) = (0^2)(q)+(1^2)(p) = p$. Oleh karenanya $Var(X) = E(X^2)-(E(X))^2 = p-p^2 = p(1-p) = pq$.

Sebaran Binomial

Bila percobaan terdiri dari sederetan tindakan Bernoulli yang saling bebas, dimana kuantitas yang diamati adalah banyaknya 'Berhasil' dari sebanyak n tindakan tersebut. Jika peluang 'Berhasil' pada setiap tindakan Bernoulli tersebut adalah *p*, dan *X* melambangkan banyaknya 'Berhasil' tersebut, maka fungsi kepekatan peluang dari *X* ini adalah

$$b(x;n,p) = C_x^n p^x q^{n-x} \quad x = 0,1,2,...,n$$

Peristiwa [X=x] terjadi apabila terdapat sebanyak *x* 'Berhasil' dan *n-x* 'Gagal' dalam keseluruhan *n* tindakan Bernoulli yang saling bebas tersebut. Seluruhnya akan ada sebanyak C_x^n cara. Sehingga diperolehlah fungsi kepekatan peluang Bernoulli seperti *b(x;n,p)* . Notasi ini digunakan sebagai pengganti *f(x)*, yang sekaligus mengindikasikan bahwa *b* singkatan dari Bernoulli, dengan argumen *x* serta fungsi tersebut sangat tergantung dari besaran parameter *n* dan *p*.

Hal-hal yang harus diperhatikan adalah

$$\sum_{x=0}^{n} b(x;n,p) = \sum_{x=0}^{n} C_x^n p^x q^{n-x} = (p+q)^n = 1^n = 1$$

Sedangkan fungsi sebaran kumulatifnya

$$B(x;n,p) = \sum_{k=0}^{x} b(k;n,p) \quad \text{untuk } x=0,1,2,\ldots,n$$

Catatan untuk Sebaran binomial

$$B(x;n,p) = 1 - B(n-x-1;n,1-p)$$

$$b(x;n,p) = B(x;n,p) - B(x-1;n,p)$$

Binomdist(number_s;trials;probability_s;cumulatif) adalah suatu fungsi dalam Microsoft Excel yang dapat digunakan untuk menghitung nilai fungsi kepekatan peluang dan fungsi sebaran peluang dari peubah acak Binomial. Parameter number_s adalah banyaknya 'Berhasil' (x), trials adalah total tindakan (n); probability_s adalah peluang 'Berhasil' dalam setiap tindakan Bernoulli (p), dan cumulatif diisikan nilai Logika TRUE apabila kita inginkan fungsi sebaran kumulatif dan FALSE apabila kita inginkan fungsi kepekatan peluangnya.

Dengan menggunakan Microsoft Excel kita dapat mencari *b(3;10, 0,35)* artinya secara rumus kita ingin mencari $C_3^{10}(0{,}35)^3(0{,}65)^7$ yang cukup dengan menuliskan =binomdist(3;10;0,35;**false**) pada lembar kerja Microsoft Excel yang akan memberikan hasil 0,2522. Namun apabila kita tuliskan =binomdist(3;10;0,35;**true**) yang hasilnya adalah 0,5138 yang tidak lain adalah sebagai akumulasi dari *b(**0**;10, 0,35)*+ *b(**1**;10, 0,35)*+ *b(**2**;10, 0,35)*+ *b(**3**;10, 0,35)*

Teladan 2.5.

Dari pelemparan sebuah mata uang logam yang seimbang 20 kali, berapakah peluang mendapatkan sisi Angka sebanyak 16 kali ? Jawab: *b(16;20,½)* = $C_{16}^{20}\, 0{,}5^{16} 0{,}5^{4} = 0{,}0046$.

Derivasi Nilai harapan X apabila X memiliki sebaran Binomial dengan parameter *n* dan *p* atau disingkat Bin(*n,p*)

$$
\begin{aligned}
E(X) &= \sum_{x=0}^{n} x C_x^n p^x q^{n-x} \\
&= \sum_{x=0}^{n} x \frac{n!}{x!(n-x)!} p^x q^{n-x} \\
&= \sum_{x=1}^{n} \frac{n(n-1)!}{(x-1)!(n-x)!} p p^{x-1} q^{(n-1)-(x-1)} \\
&= np \sum_{k=0}^{n-1} C_k^{n-1} p^k q^{(n-1)-k} \\
&= np
\end{aligned}
$$

$$
\begin{aligned}
E(X(X-1)) &= \sum_{x=0}^{n} x(x-1) C_x^n p^x q^{n-x} \\
&= \sum_{x=0}^{n} x(x-1) \frac{n!}{x!(n-x)!} p^x q^{n-x} \\
&= \sum_{x=2}^{n} \frac{n(n-1)(n-2)!}{(x-2)!(n-x)!} p^2 p^{x-2} q^{(n-2)-(x-2)} \\
&= n(n-1) p^2 \sum_{k=0}^{n-2} C_k^{n-1} p^k q^{(n-2)-k} \\
&= n(n-1)p^2
\end{aligned}
$$

Karena E[X(X-1)] = E(X^2) – E(X), maka E(X^2) = E[X(X-1)]+E[X] = n(n-1)p^2 + np = n^2p^2 – np^2 + np = n^2p^2 + np – np^2 = np(np) + np(1-p) = np(np+q).

Dan dengan demikian

$$Var(X) = E(X^2) - [E(X)]^2 = np(np+q)-(np)^2 = npq.$$

Sebaran Hypergeometrik

Suatu populasi atau kumpulan obyek yang terdiri dari *N* item, dan dari sejumlah itu ada sebanyak *M* item dari kategori pertama, sedangkan sisanya sebanyak *N-M* dari kategori kedua. Misalkan sejumlah *n* item diambil secara acak tanpa dikembalikan, dan *X* adalah peubah acak banyaknya item dari kategori pertama terambil.

Fungsi Kepekatan Peluang Diskret dari peubah acak X ini adalah

$$h(x;n,M,N) = \frac{C_x^M C_{n-x}^{N-M}}{C_n^N} \text{ untuk } max(0,n-N+M) \le x \le min(n,M)$$

Ruang contoh dari tindakan diatas beranggotakan sebanyak C_n^N kemungkinan. Apabila banyaknya yang terambil dari sebanyak *n* contoh tersebut ada sebanyak *x* dari kategori pertama, maka sisanya ada sebanyak *n-x* dari kategori kedua. Banyaknya kemungkinan terpilihnya pada kondisi tersebut adalah menggunakan perhitungan kombinasi terpilihnya sebanyak *x* dari kategori pertama, atau sebesar C_x^M kemungkinan dan dari kategori kedua ada sejumlah C_{n-x}^{N-M}. Banyak cara peristiwa terjadi diperoleh dengan menggunakan prinsip multiplikasi. Akhirnya fungsi kepekatan peluang hypergeometrik diturunkan dengan konsep peluang dengan menggunakan metode klasik atau apriori.

Teladan 2.6.

Sebuah kantong berisikan 100 busi, dimana 20 diantaranya rusak atau cacat. Banyaknya busi yang rusak atau cacat tersebut tidak diketahui oleh pembeli yang ingin memutuskan untuk mengambil 10 busi secara acak tanpa pengembalian. Apabila busi yang diambilnya secara acak tersebut tidak mengandung busi yang rusak lebih dari tiga, maka ia akan menerimanya. Berapakah peluang pembeli tersebut menerima busi yang diambil secara acak tersebut ?

Jawab:

$$P[X \le 3] = \sum_{x=0}^{3} \frac{C_x^{20} C_{10-x}^{80}}{C_{10}^{100}} = \frac{C_0^{20} C_{10}^{80}}{C_{10}^{100}} + \frac{C_1^{20} C_9^{80}}{C_{10}^{100}} + \frac{C_2^{20} C_8^{80}}{C_{10}^{100}} + \frac{C_3^{20} C_7^{80}}{C_{10}^{100}} = 0{,}890$$

Hypgeomdist(sample_s;number_sample;population_s;number_pop) adalah fungsi kepekatan peluang dari sebaran hypergeometrik yang disediakan oleh Microsoft Excel dimana **sample_s** adalah besarnya contoh yang 'Berhasil" atau *x*, **number_sample** adalah besarnya contoh yang diambil, **population_s** menunjuk ke besarnya populasi 'Berhasil' atau *M* dan **number_pop** sebagai ukuran populasi atau *N*.

Sebagai misal =hypgeomdist(**0**;10;20;100) = 0,0951
= hypgeomdist(**1**;10;20;100) = 0,2679
= hypgeomdist(**2**;10;20;100) = 0,3182
= hypgeomdist(**3**;10;20;100) = 0,2092

Fungsi Sebaran Kumulatif Hypergeometri dapat dituliskan sebagai

$$H(x;n,M,N) = \sum_{i=0}^{x} h(i;n,M,N)$$

Teorema 2.7.

Jika peubah acak X memiliki sebaran HYP(*n,M,N*) maka untuk setiap nilai *x*=*0,1,…,n* dan untuk $N \to \infty$ dan $M \to \infty$ dengan $M/N \to p$, suatu konstanta positif, maka

$$\lim_{N\to\infty} \frac{C_x^M C_{n-x}^{N-M}}{C_n^N} = C_x^n p^x (1-p)^{n-x}$$

Bukti : sebagai latihan.

Sebaran Geometrik

Disini kita sekali lagi memakai sekuen atau deret dari tindakan Bernoulli yang saling bebas dengan peluang 'Berhasil' untuk tiap tindakan Bernoulli sebesar *p*. Dalam kasus Binomial, jumlah tindakan yang akan dilakukan adalah tetap, dan peubah yang diamatinya adalah banyaknya peristiwa 'Berhasil'.

Kali ini akan dipelajari peubah acak yang menunjuk pada banyaknya tindakan yang harus dilakukan untuk mencapai 'Berhasil' yang pertama kali dari sederetan peristiwa yang harus dilakukan dari tindakan Bernoulli. Dengan demikian, bisa saja sekali tindakan langsung memperoleh 'Berhasil', dua kali tindakan baru memperoleh 'Berhasil' pertama, tiga kali, empat kali, dan seterusnya hingga bila digambarkan G sebagai 'Gagal' dan B sebagai 'Berhasil' adalah seperti berikut

X=1	B
X=2	GB
X=3	GGB
…	
X=x	$\underbrace{GG...G}_{x-1}B$

Fungsi Kepekatan Peluang Geometri ini dapat dituliskan seperti berikut

$$g(x;p) = q^{x-1}p \text{ untuk } x = 1, 2, 3, \ldots$$

Dapat dilihat bahwa fungsi diatas adalah fungsi kepekatan peluang yang apabila dijumlahkan untuk seluruh kemungkinan nilai x hasilnya akan sama dengan satu.

$$\sum_{x=1}^{\infty} g(x;p) = p\sum_{x=1}^{\infty} q^{x-1} = p(1+q+q^2+q^3+\ldots) = p\left(\frac{1}{1-q}\right) = \frac{p}{p} = 1$$

Nama lain dari Sebaran Geometrik ini adalah **Sebaran Pascal**. Fungsi Sebaran Kumulatif dari peubah acak ini adalah

$$G(x;p)=\sum_{i=1}^{x} pq^{i-1}=1-q^{x} \quad \text{untuk } x = 1, 2, 3, \ldots$$

Teladan 2.7.

Seorang pemanah dapat mengenai sasaran lingkar paling tengah dengan peluang 0,3 setiap melepaskan anak panah dari busurnya. Peluang ia pertama kali mengenai lingkar sasaran paling tengah pada tembakan ke lima adalah $g(5;0,3)=0,7^4(0,3)$. Setelah itu peluang ia baru mengenai sasaran dimaksud pada tembakan kesepuluh juga masih sama yaitu $0,7^4(0,3)$. Namun peluang bahwa pertama kali ia menembak mengenai sasaran pertama kali sebelum atau pada tembakan ke lima adalah $G(5;0,3) = 1-(0,7)^5 = 0,83193$.

Teorema 2.8.

Jika X memiliki sebaran Geometrik dengan parameter p, maka

$$P[X>j+k \mid X>j]=P[X>k]$$

Bukti

$$\begin{aligned} P[X>j+k \mid X>j] &= \frac{P[X>j+k]}{P[X>j]} \\ &= \frac{(1-p)^{j+k}}{(1-p)^{j}} \\ &= (1-p)^{k} \\ &= P[X>k] \end{aligned}$$

Nilai harapan peubah acak X yang menyebar menurut sebaran Geometrik ini dapat dicari seperti berikut:

$$
\begin{aligned}
E(X) &= \sum_{x=1}^{\infty} xpq^{x-1} \\
&= \sum_{x=1}^{\infty} p\left(\frac{d}{dq} q^{x}\right) \\
&= p\frac{d}{dq}\sum_{k=0}^{\infty} q^{k} \\
&= p\frac{d}{dq}(1-q)^{-1} \\
&= p(1-q)^{-2} \\
&= \frac{1}{p}
\end{aligned}
$$

Perlu diperhatikan walaupun sedikit agak berbeda, misalkan peubah acak Y adalah menunjuk pada banyaknya peristiwa 'Gagal' sebelum 'Berhasil' pertama kali. Dengan demikian sebetulnya Y = X – 1, dan

$$P[Y = y] = (1-p)^{y}\, p \qquad \text{untuk } y = 0, 1, 2, \ldots$$

Peubah acak Y tersebut juga memiliki sebaran yang disebut dengan Geometrik.

Sebaran Geometrik ini merupakan bentuk khusus dari Sebaran Negatif Binomial yang akan kita bahas dalam sub bab berikut ini.

Sebaran Negatif Binomial

Dalam tindakan Bernoulli serupa seperti dalam sebaran Geometrik, misalkan X melambangkan peubah acak banyaknya tindakan yang diperlukan hingga tercapai *r* 'Berhasil'. Peubah acak ini dikatakan memiliki sebaran Negatif Binomial yang fungsi kepekatan peluangnya dapat dituliskan sebagai berikut.

$$nb(x;r,p) = C_{r-1}^{x-1}\, p^{r} q^{x-r} \qquad \text{untuk } x=r, r+1, r+2, \ldots$$

Agar peristiwa [X=x] terjadi, kita pasti memiliki sederetan tindakan Bernoulli dimana 'Berhasil' yang ke-r ('Berhasil' pada tindakan ke-r) terjadi pada tindakan ke-x. Dengan demikian sampai dengan tindakan ke x-1 sudah terdapat sebanyak r-1 'Berhasil' dalam urutan apapun, dan untuk tindakan yang ke x harus 'Berhasil'. Misalkan 'Berhasil' ke-3 terjadi pada peristiwa ke 5, maka kemungkinan susunan urutan terjadinya peristiwa dimaksud adalah

BBGGB BGBGB BGGBB

GBBGB GBGBB GGBBB

Jadi ada sebanyak 6 atau kombinasi 2 dari 4 susunan peristiwa dimaksud.

Microsoft Excel juga memberikan fungsi untuk menghitung fungsi kepekatan peluang dari sebaran Negatif Binomial.

=negbinomdist(number_f;number_s;probability_s)

Number_f adalah banyaknya 'Gagal'; number_s adalah banyaknya 'Berhasil' dan probability_s adalah peluang 'Berhasil' pada setiap tindakan Bernoulli. Jadi misalkan untuk menghitung

nb(6;4, 0,6) =negbinomdist(2;4;0,6)

akan menghasilkan 0,20736

nb(5;3, 0,4) =negbinomdist(2;3;0,4)

akan menghasilkan 0,13824

g(5;0,3) =negbinomdist(4;1;0,3)

akan menghasilkan 0,07203

g(4;0,7) =negbinomdist(3;1;0,7)

akan menghasilkan 0,01890

Karena *0* < *p* < *1* maka

$$\sum_{x=r}^{\infty} C_{r-1}^{x-1} p^r q^{x-r} = p^r \sum_{i=0}^{\infty} C_{r-1}^{i+r-1} q^i = p^r (1-q)^{-r} = 1$$

sebagai akibat dari $\sum_{i=0}^{\infty} C_{r-1}^{i+r-1} q^i = (1-q)^{-r}$. Nama Negatif Binomial diambil dari hubungan ekspansi deret binomial dengan pangkat negatif –*r*.

Ada beberapa yang membuat peubah acak Y yang didefinisikan sebagai banyaknya 'Gagal' yang terjadi sebelum 'Berhasil' ke r kalinya. Dengan demikian X = Y + r dan fungsi kepekatannya adalah sebagai berikut

$$nb(y;r,p) = P[Y = y] = C_{r-1}^{y+r-1} p^r (1-p)^y \qquad \text{untuk } y=0,1,2,\ldots$$

Terdapat hubungan antara Negatif Binomial dan Binomial. Kadang kala permasalahan negatif binomial juga disebut dengan penarikan contoh invers binomial. Misalkan X memiliki sebaran Negatif Binomial (*r,p*) dan W memiliki sebaran Binomial (*n,p*). Maka

$$P[X \le n] = P[W \ge r]$$

Peristiwa [W ≥ r] adalah peristiwa terjadinya sebanyak r atau lebih 'Berhasil' dalam n tindakan, dan ini berarti bahwa n atau kurang tindakan akan diperlukan untuk memperoleh r 'Berhasil' pertama. Jelaslah bahwa sebaran negatif binomial dapat diekspresikan dalam sebaran binomial dalam hubungan seperti berikut:

$$F(x;r,p) = P[X \le x] = 1 - B(r-1;x,p) = B(x-r;x;q)$$

Teladan 2.8.

Sebuah pertandingan final Bola Basket yang dimainkan dengan sistem 'The Best of Seven' dilakukan oleh Tim A dan Tim B. Apabila peluang tim A memenangkan pertandingan di setiap pertandingan adalah 0,6 berapakah peluang Tim A akan menjadi juara paling lama harus bertanding 6 kali ?

Jawab: Tim A harus memenangkan pertandingan ke-empat kalinya paling banyak pada pertandingan ke-6. Jika X adalah peubah acak Negatif Binomial dengan x = 6 dan r = 4 serta p = 0,6, maka P(Tim A menang selambat-lambatnya pada pertandingan ke-6 tersebut)

= $P[X \le 6]$

= *F(6;4, 0,6)*

$$= \sum_{x=4}^{6} C_3^{x-1} (0{,}6)^4 (0{,}4)^{x-4}$$

$= B(2;6,\ 0{,}4)$

$$= \sum_{w=0}^{2} C_w^6 (0{,}4)^w (0{,}6)^{6-w}$$

= 0,5443

Sebaran Poisson

Suatu peubah acak diskret X dikatakan memiliki sebaran Poisson dengan parameter $\mu > 0$ jika memiliki fungsi kepekatan peluang seperti berikut

$$f(x;\mu) = \frac{e^{-\mu}\mu^x}{x!} \qquad x = 0,1,2,\ldots$$

Kita dapat periksa bahwa

$$\sum_{x=0}^{\infty} f(x;\mu) = e^{-\mu} \sum_{x=0}^{\infty} \frac{\mu^x}{x!} = e^{-\mu} e^{\mu} = 1$$

Fungsi Sebaran Kumulatif dari peubah acak Poisson didefinisikan sebagai

$$F(x;\mu) = \sum_{k=0}^{x} f(k;\mu)$$

Bentuk sederhana dari fungsi sebaran kumulatif ini tidak dapat dibuat, tetapi biasanya ditabulasikan.

Nilai harapan dari peubah acak X atau E(X) dapat dicari sebagai berikut:

$$E(X) = \sum_{x=0}^{\infty} x f(x;\mu)$$

$$= \sum_{x=0}^{\infty} x \frac{e^{-\mu}\mu^x}{x!}$$

$$= \quad e^{-\mu}\sum_{x=1}^{\infty}\frac{\mu\mu^{x-1}}{(x-1)!}$$

$$= \quad \mu e^{-\mu}\sum_{k=0}^{\infty}\frac{\mu^{k}}{k!}$$

$$= \quad \mu e^{-\mu}e^{\mu}$$

$$= \quad \mu$$

Teorema 2.9.

Jika X menyebar menurut sebaran Binomial(*n,p*) maka untuk setiap nilai *x=0,1,2,…* dan bilamana $p\to 0$ dengan $np=\mu$ konstan maka

$$\lim_{n\to\infty} C_x^n p^x(1-p)^{n-x} = \frac{e^{-\mu}\mu^x}{x!}$$

Bukti

$$C_x^n p^x(1-p)^{n-x} \quad = \quad \frac{n!}{x!(n-x)!}\left(\frac{\mu}{n}\right)^x\left(1-\frac{\mu}{n}\right)^{n-x}$$

$$= \quad \frac{\mu^x}{x!}\left(\frac{n}{n}\right)\left(\frac{n-1}{n}\right)\ldots\left(\frac{n-x+1}{n}\right)\left(1-\frac{\mu}{n}\right)^n\left(1-\frac{\mu}{n}\right)^{-x}$$

Hasil pembuktian diperoleh karena

$$\lim_{n\to\infty}\left(1-\frac{\mu}{n}\right)^n = e^{-\mu} \quad \text{dan} \quad \lim_{n\to\infty}\left(1-\frac{\mu}{n}\right)^{-x} = 1$$

Teladan 2.9.

Apabila besarnya resistor yang cacat yang diproduksi pabrik A sebesar 1%. Suatu jenis peralatan elektronik baru membutuhkan 100 resistor didalamnya, dan 100 resistor ini dipilih secara acak dari sekumpulan resistor yang diproduksi pabrik A tersebut. Berapakah peluang mendapatkan 3 buah resistor yang diambil tersebut cacat?

Jawab:
Dengan menggunakan sebaran Binomial *b(3;100, 0,01) = 0,0610*
Dan bila digunakan sebaran Poisson *f(3;1) = 0,0613*

Sebaran Seragam Diskret

Terdapat beberapa permasalahan yang umumnya menyangkut pemberian nilai peluang secara klasik, dapat dibuat modelnya dengan sebaran seragam diskret ini. Biasanya dimungkinkan menghubungkan permasalahan tersebut dengan menggunakan gugus bilangan bulat *1, 2, 3, …, N.*

Peubah acak X memiliki sebaran seragam diskret pada bilangan bulat *1,2,3,…,N* jika memiliki fungsi kepekatan peluang dalam bentuk

$$f(x) = \frac{1}{N} \qquad \text{untuk } x=1,2,3,\ldots,N$$

Nilai harapan peubah acak X dapat dicari seperti berikut:

$$\begin{aligned} E(X) &= \sum_{x=1}^{N} \frac{x}{N} = \frac{1}{N}(1+2+\ldots+N) \\ &= \frac{1}{N}\,\frac{N(N+1)}{2} \\ &= \frac{N+1}{2} \end{aligned}$$

Sedangkan

$$\begin{aligned} E(X^2) &= \sum_{x=1}^{N} \frac{x^2}{N} = \frac{1}{N}(1^2+2^2+\ldots+N^2) \\ &= \frac{1}{N}\,\frac{N(N+1)(2N+1)}{6} \end{aligned}$$

$$= \frac{(N+1)(2N+1)}{6}$$

Sehingga $Var(X) = \dfrac{N^2 - 1}{12}$

Peubah Acak Kontinu

Telah kita bahas peubah acak diskret yang sangat banyak dijumpai dalam dunia kehidupan, termasuk diantaranya, atau kalau dapat dikatakan kebanyakan peubah acak yang nilainya didapat dengan cara menghitung tergolong dalam peubah acak diskret. Namun demikian, apabila situasi tidak memungkinkan untuk menggunakan model dari peubah acak diskrit ini, tentu kita akan gunakan peubah acak kontinu. Pada umumnya, peubah acak yang cara mendapatkan nilainya dengan cara mengukur atau menggunakan alat ukur, seperti alat pengukur panjang, berat, kecepatan, dan masih banyak lagi, tergolong dalam peubah acak kontinu. Fungsi Sebaran Kumulatif yang kita definisikan pada peubah acak diskret juga berlaku untuk peubah acak kontinu.

Teladan 2.10.

Setiap hari kerja seorang pekerja naik bus untuk mencapai tempat kerjanya. Meskipun bus akan datang tepat setiap 5 menit, pekerja tersebut umumnya datang di tempat pemberhentian bus sembarang waktu secara acak diantara waktu kedatangan bus. Dengan demikian, waktu tunggu pekerja tersebut di setiap pagi merupakan suatu peubah acak (kontinu).

Cara lain untuk mempelajari sebaran tersebut adalah mengamati frekuensi relatif kedatangan bus pada interval waktu yang pendek berukuran sama, tetapi menyebar dalam interval waktu tunggu [0,5]. Bisa saja frekuensi kedatangan bus selama interval waktu dalam bentuk $(x, x+\Delta x)$ untuk nilai Δx yang kecil dan proporsional terhadap interval waktu dan tidak tergantung dari *x*. Dengan demikian kita dapatkan

$$P[x < X \leq x + \Delta x] = F(x + \Delta x) - F(x) = c\Delta x$$

Untuk semua $0 \leq x < x+\Delta x \leq 5$ dan beberapa $c > 0$. Sudah tentu bahwa *F(x) differensiabel* (memiliki turunan atau dapat diturunkan) pada *x*, dan turunannya konstan, $F'(x)=c>0$. Catatan bahwa untuk $x < 0$ atau $x > 5$ turunannya juga ada namun nilainya $F'(x) = 0$, karena $P[x < X \leq x+\Delta x] = 0$ apabila *x* dan $x+\Delta x$ bukan nilai kemungkinan peubah acak *X*, dan tidak ada turunan untuk $x=0$ dan $x=5$.

Secara umum, jika *F(x)* adalah fungsi sebaran kumulatif peubah acak kontinu *X,* maka turunan atau differensial dari fungsi tersebut adalah *f(x)*, dan dengan kondisi tertentu kita sebut sebagai fungsi kepekatan peluang dari peubah acak *X*. Dalam teladan kita tadi, *F(x)* dapat direpresentasikan untuk nilai *x* dalam interval [0, 5] sebagai integral dari turunannya:

$$F(x) = \int_{-\infty}^{x} f(t)dt = \int_{0}^{x} \frac{1}{5} dt = \frac{x}{5}$$

Definisi 2.9.

Peubah acak X dikatakan sebagai peubah acak kontinu jika memiliki fungsi *f(x)*, yang disebut sebagai fungsi kepekatan peluang dari X, sehingga fungsi sebaran kumulatifnya dapat dituliskan sebagai

$$F(x) = \int_{-\infty}^{x} f(t)dt$$

Definisi 2. 10.

Misalkan *X* terdefinisi pada (Ω,S ,P) dengan fungsi sebaran *F*. Maka *X* dikatakan kontinu jika *F* kontinu mutlak, yaitu jika terdapat fungsi tak negatif *f(x)* sedemikian rupa sehingga untuk setiap bilangan riil *x*, $F(x) = \int_{-\infty}^{x} f(t)dt$. Fungsi *f* disebut dengan fungsi kepekatan peluang atau fungsi densitas peluang peubah acak *X*.

Mengikuti teorema dasar Kalkulus, fungsi kepekatan peluang peubah acak kontinu ini dapat diperoleh dari fungsi sebaran kumulatifnya dengan cara melakukan differensiasi atau turunannya.

$$f(x) = \frac{d}{dx}F(x) = F'(x)$$

asalkan turunannya ada.

Dalam peubah acak kontinu, peristiwa [*X* = *c*] dimana *c* adalah konstanta, memiliki peluang nol. Untuk peubah acak kontinu berlaku : Jika $a < b$

$$\begin{aligned} P[\, a < X \leq b \,] &= P[\, a \leq X < b \,] \\ &= P[\, a < X < b \,] \\ &= P[\, a \leq X \leq b \,] \\ &= F(b) - F(a) \end{aligned}$$

Teorema 2.10.

Suatu fungsi *f(x)* adalah fungsi kepekatan peluang untuk beberapa peubah acak kontinu *X* jika dan hanya jika memenuhi syarat bahwa untuk semua bilangan nyata *x* berlaku $f(x) \geq 0$ dan $\int_{-\infty}^{+\infty} f(x)dx = 1$

Definisi 2.11.

Jika *X* adalah peubah acak kontinu dengan fungsi kepekatan peluang *f(x)*, maka **nilai harapan** dari *X* didefinisikan sebagai $E(X) = \int_{-\infty}^{+\infty} xf(x)dx$ jika inetgralnya ada atau konvergen. Jika tidak, maka dikatakan *E(X)* tidak ada.

Definisi 2.12.

Jika *0 < p < 1*, maka persentil ke-p dari sebaran peubah acak X adalah nilai terkecil x_p, sedemikian rupa sehingga $F(x_p) \geq p$.

Jika *X* kontinu, maka x_p adalah jawaban dari persamaan $F(x_p) = p$

Teladan 2.11.

Misalkan sebaran umur suatu lampu pijar memiliki fungsi sebaran seperti berikut: $F(x) = 1 - e^{-\left(x/5\right)^2}$, untuk *x > 0*, dan nol untuk *x* selainnya. Maka median umur lampu dapat dicari sebagai berikut:

$$m = 5[-\ln(1-0{,}5)]^{1/2} = 5\sqrt{\ln 2} = 4{,}163 \text{ bulan}$$

Jika diinginkan untuk mencari waktu *t* sedemikian rupa sehingga 10% komponen gagal berfungsi sebelum waktu *t*, maka nilai persentil ke-10 tersebut adalah

$$x_{0,10} = 5[-\ln(1-0{,}10)]^{1/2} = 5\sqrt{-\ln(0{,}9)} = 1{,}623 \text{ bulan}$$

Definisi 2.13.

Jika fungsi kepekatan peluang memiliki nilai maksimum yang khas pada $x = m_0$, sebut saja $\max_x f(x) = f(m_0)$ maka m_0 kita sebut dengan **modus** dari peubah acak *X*.

Definisi 2.14.

Suatu sebaran dengan fungsi kepekatakan peluang *f(x)* dikatakan simetris terhadap *c* jika *f(c-x)=f(c+x)* untuk semua *x*.

Sebaran Seragam Kontinu

Misalkan peubah acak X dapat memiliki nilai dalam interval (*a,b*), dan juga fungsi kepekatan peluangnya konstan, katakanlah *f(x)=c* pada interval tersebut. Untuk memenuhi syarat sebagai fungsi kepekatan peluang, hendaklah nilai *c=1/(b-a)*.

Fungsi kepekatan peluang peubah acak kontinu X yang menyebar menurut sebaran seragam dapat dituliskan sebagai berikut:

$$f(x) = \begin{cases} \dfrac{1}{b-a} & jika \quad a < x < b \\ 0 & jika \quad x \quad lainnya \end{cases}$$

Sedangkan fungsi sebaran kumulatifnya adalah

$$F(x;a,b) = \begin{cases} 0 & jika \quad x \le a \\ \dfrac{x-a}{b-a} & jika \quad a < x < b \\ 1 & jika \quad b \le x \end{cases}$$

Nilai harapan dari beberapa kuantiti peubah acak sebaran seragam dapat dicari seperti berikut:

$$E(X) = \int_a^b x\left(\frac{1}{b-a}\right)dx$$

$$= \frac{b^2 - a^2}{2(b-a)}$$

$$= \frac{(b+a)(b-a)}{2(b-a)}$$

$$= \frac{a+b}{2}$$

$$E(X^2) = \int_a^b x^2\left(\frac{1}{b-a}\right)dx$$

$$= \frac{b^3 - a^3}{3(b-a)}$$

$$= \frac{(b^2 + ab + a^2)(b-a)}{3(b-a)}$$

$$= \frac{b^2 + ab + a^2}{3}$$

Akhirnya diperoleh

$$\mathrm{Var}(X) = E(X^2) - [E(X)]^2$$

$$= \frac{b^2 + ab + a^2}{3} - \frac{(a+b)^2}{4}$$

$$= \frac{(b-a)^2}{12}$$

Sebaran Gamma

Banyak aplikasi yang memiliki peubah acak dengan sebaran Gamma. Nama Gamma ini ada hubungannya dengan fungsi gamma.

Definisi 2.15.

Fungsi Gamma, dinotasikan dengan $\Gamma(\kappa)$ untuk semua $\kappa > 0$, didefinisikan sebagai

$$\Gamma(\kappa) = \int_0^\infty t^{\kappa-1} e^{-t}\, dt$$

Teorema 2.11.

Fungsi Gamma memiliki sifat-sifat seperti berikut:

$$\Gamma(\kappa) = (\kappa - 1)\Gamma(\kappa - 1) \qquad \textit{untuk} \quad \kappa > 1$$
$$\Gamma(n) = (n-1)! \qquad \textit{untuk} \quad n = 1,2,3,...$$
$$\Gamma\left(\frac{1}{2}\right) = \sqrt{\pi}$$

Peubah acak kontinu *X* dikatakan memiliki sebaran Gamma dengan parameter $\kappa > 0$ dan $\theta > 0$ jika fungsi kepekatan peluangnya memiliki bentuk

$$f(x;\theta,\kappa) = \frac{1}{\theta^{\kappa}\Gamma(\kappa)} x^{\kappa-1} e^{-\frac{x}{\theta}} \qquad x > 0$$

dan nol untuk *x* lainnya.

Parameter κ sering disebut dengan parameter bentuk (***shape parameter***) karena besarnya menentukan bentuk dasar dari grafik fungsi kepekatan peluangnya. Lebih khusus lagi, terdapat tiga bentuk kurva untuk $\kappa < 1$, $\kappa = 1$, atau $\kappa > 1$.

Fungsi sebaran kumulatif dari peubah acak yang menyebar dengan sebaran gamma adalah seperti berikut

$$F(x;\theta,\kappa)=\int_0^x \frac{1}{\theta^\kappa\Gamma(\kappa)}t^{\kappa-1}e^{-\frac{t}{\theta}}dt \qquad x>0$$

Substitusi $u=t/\theta$ dalam integral ini akan menghasilkan

$$F(x;\theta,\kappa)=F(\frac{x}{\theta};1;\kappa)$$

yang hanya tergantung θ melalui peubah x/θ. Parameter θ ini sering disebut dengan parameter skala (***scale parameter***).

Fungsi sebaran kumulatif diatas secara umum tak dapat diselesaikan secara eksplisit, namun jika κ merupakan bilangan bulat positif *n*, maka integralnya dapat diselesaikan sebagai jumlah.

Teorema 2.12.

Jika *X* memiliki sebaran Gamma(θ,n) dimana n adalah bilangan bulat positif, maka fungsi sebaran kumulatif

$$F(x;\theta,n)=1-\sum_{i=0}^{n-1}\frac{\left(x/\theta\right)^i}{i!}e^{-x/\theta}$$

Teladan 2.12.

Besarnya presipitasi harian terukur di suatu lembah sungai menyebar menurut Sebaran Gamma dengan parameter θ = 0,2 dan κ = 6. Berapakah peluang presipitasi melebihi 2 inchi ?

$$P[X>2] = \int_2^\infty \frac{1}{(0{,}2)^6\Gamma(6)}x^{6-1}e^{-(x/0{,}2)}dx$$
$$= 1-F(2;0{,}2\ ,6)$$

$$= \sum_{i=0}^{5} \frac{10^i}{i!} e^{-10} = 0{,}067$$

Nilai harapan dari X diperoleh seperti berikut

$$E(X) = \int_0^\infty x \frac{1}{\theta^\kappa \Gamma(\kappa)} x^{\kappa-1} e^{-x/\theta} dx$$

$$= \frac{1}{\theta^\kappa \Gamma(\kappa)} \int_0^\infty x^{(1+\kappa)-1} e^{-x/\theta} dx$$

$$= \frac{\theta^{1+\kappa}\Gamma(1+\kappa)}{\theta^\kappa \Gamma(\kappa)} \int_0^\infty \frac{1}{\theta^{1+\kappa}\Gamma(1+\kappa)} x^{(1+\kappa)-1} e^{-x/\theta} dx$$

$$= \frac{\theta^{1+\kappa}\Gamma(1+\kappa)}{\theta^\kappa \Gamma(\kappa)}$$

$$= \theta \frac{\kappa\Gamma(\kappa)}{\Gamma(\kappa)}$$

$$= \kappa\theta$$

Dengan cara yang sama akan diperoleh $E(X^2) = \theta^2 \kappa(1+\kappa)$ dan dengan demikian kita peroleh bahwa $Var(X) = \kappa\theta^2$.

Sebaran Eksponensial

Suatu peubah acak kontinu *X* memiliki sebaran eksponensial dengan parameter $\theta > 0$ jika fungsi kepekatan peluangnya memiliki bentuk

$$f(x;\theta) = \frac{1}{\theta} e^{-x/\theta} \qquad untuk \quad x > 0$$

dan nol untuk *x* lainnya.

Fungsi sebaran kumulatif dari X

$$F(x;\theta) = 1 - e^{-x/\theta} \quad untuk \quad x > 0$$

Dengan demikian kita bisa lihat bahwa θ adalah parameter skala.

Bila kita perhatikan, sebaran eksponensial ini merupakan bentuk khusus dari sebaran Gamma dengan $\kappa = 1$.

Teorema 2.13.

Untuk peubah acak kontinu *X,* yang menyebar menurut sebaran eksponensial dengan parameter θ jika dan hanya jika $P[X > a+t \mid X > a] = P[X > t]$ untuk semua $a > 0$ dan $t > 0$.

Bukti : sebagai latihan.

Teladan 2.13.

Misalkan umur bola lampu LCD Projector merk Gendhon memiliki sebaran eksponensial dengan parameter $\theta = 100$ jam. Berapakah peluang bahwa sebuah bola lampu tersebut apabila dipilih secara acak memiliki umut sedikitnya 50 jam ?

$$P[X \geq 50] = 1 - F(50;100) = e^{-0,5} = 0,6065$$

Karena sebaran eksponensial adalah bentuk khusus dari sebaran gamma dengan parameter $\kappa=1$, maka $E(X) = \theta$ dan $E(X^2) = \theta^2$.

Sebaran Weibull

Seorang ahli fisika W. Weibull menyarankan menggunakan suatu sebaran dari peubah acak yang seringkali muncul dalam permasalahan fisika, khususnya ilmu bahan disamping sering juga dipakai untuk mempelajari sebaran umur suatu barang.

Suatu peubah acak kontinu *X* dikatakan memiliki distribusi Weibull dengan parameter $\beta > 0$ dan $\theta > 0$ jika memiliki fungsi kepekatan peluang dalam bentuk

$$f(x;\theta,\beta) = \frac{\beta}{\theta^\beta} x^{\beta-1} e^{-\left(x/\theta\right)^\beta} \qquad untuk \quad x > 0$$

dan nol untuk *x* lainnya.

Dalam sebaran ini β merupakan parameter bentuk, tergantung apakah $\beta<1$, $\beta=1$, atau $\beta>1$.

Salah satu keuntungan sebaran Weibull ini adalah bahwa Fungsi Sebaran kumulatifnya dapat diperoleh secara eksplisit dengan mengintegralkan fungsi kepekatan peluangnya menjadi

$$F(x;\theta,\beta) = 1 - e^{-\left(x/\theta\right)^\beta} \quad untuk \quad x > 0$$

Bentuk yang lebih khusus dari sebaran Weibull ini adalah apabila $\beta=2$ yang disebut dengan **Sebaran Rayleigh**.

Teladan 2.14.

Sebaran jarak antara tempat jatuhnya penerjun dengan titik pusat penerjunan diketahui menyebar menurut sebaran Weibull dengan parameter $\theta=10$ dan $\beta=2$. Berapakah peluang seorang penerjun jatuhnya berjarak kurang dari 5 meter dari titik pusat penerjunan ?

$$P[X \le 5] = F(5;10,2) = 1 - e^{-\left(5/10\right)^2} = 0{,}221$$

Nilai harapan peubah acak *X* dapat dicari seperti berikut

$$E(X) \quad = \quad \int_0^\infty x \frac{\beta}{\theta^\beta} x^{\beta-1} e^{-\left(x/\theta\right)^\beta} dx$$

$$= \frac{\beta}{\theta^{\beta}} \int_{0}^{\infty} x^{(1+\beta)-1} e^{-\left(x/\theta\right)^{\beta}} dx$$

Dengan menggunakan substitusi $t = \left(x/\theta\right)^{\beta}$ dan beberapa simplifikasi, maka integral diatas menjadi

$$\theta \int_{0}^{\infty} t^{(1+1/\beta)-1} e^{-t} dt = \theta \Gamma\left(1 + \frac{1}{\beta}\right)$$

Cara yang sama dapat dipergunakan untuk mendapatkan $E(X^2) = \theta^2 \Gamma\left(1 + \frac{2}{\beta}\right)$

sehingga $Var(X) = \theta^2 \left[\Gamma\left(1 + \frac{2}{\beta}\right) - \Gamma^2\left(1 + \frac{1}{\beta}\right)\right]$.

Sebaran Normal

Pada tahun 1733 Abraham de Moivre mempublikasikan sebaran Normal ini sebagai pendekatan dari sebaran jumlah dari peubah acak Binomial. Sebaran Normal adalah sebaran paling penting dalam teori peluang dan statistika.

Suatu peubah acak *X* dikatakan mengikuti sebaran Normal dengan rata-rata μ dan simpangan baku σ jika memiliki fungsi kepekatan peluang

$$f(x;\mu,\sigma) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} \qquad -\infty < x < +\infty; -\infty < \mu < +\infty; 0 < \sigma$$

Dengan menggunakan substitusi $Z = \frac{X-\mu}{\sigma}$, fungsi kepekatan peluangnya menjadi

$$\phi(z) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}z^2} \quad untuk \ -\infty < z < +\infty$$

Fungsi Sebaran Kumulatif dari sebaran *Z* ini adalah

$$\Phi(z) = \int_{-\infty}^{z} \phi(t)dt$$

Beberapa sifat geometri dasar dari fungsi kepekatan peubah acak *Z* atau sebaran normal baku ini adalah

- $\phi(z) = \phi(-z)$
- $\phi'(z) = -z\phi(z)$
- $\phi''(z) = (z^2 - 1)\phi(z)$
- ϕ*(z)* memiliki maksimum mutlak pada *z=0*
- ϕ*(z)* memiliki titik belok pada *z=-1* dan *z=+1*
- $\phi(z) \to 0$ apabila $|z| \to \infty$
- $\phi'(z) \to 0$ apabila $|z| \to \infty$

Dengan menggunakan sifat-sifat tersebut diatas, dapat ditunjukkan bahwa

$$E(Z) = \int_{-\infty}^{\infty} z\phi(z)dz = -\int_{-\infty}^{\infty} \phi'(z)dz = -\phi(z)\Big|_{-\infty}^{\infty} = 0$$

Var(X)=E(Z²) =

$$\int_{-\infty}^{\infty} z^2\phi(z)dz = \int_{-\infty}^{\infty} [\phi''(z) + \phi(z)]dz = \phi'(z)\Big|_{-\infty}^{\infty} + \int_{-\infty}^{\infty} \phi(z)dz = 0 + 1 = 1$$

Teorema 2.14.

Jika peubah acak kontinu *X* memiliki sebaran Normal dengan parameter rata μ dan simpangan baku σ, maka fungsi sebaran kumulatif dari *X* dapat diekspresikan sebagai $F_X(x) = \Phi\left(\frac{x-\mu}{\sigma}\right)$

Kumulatif peluang standar normal yang dinotasikan dengan $\Phi(z)$ pada umumnya banyak ditabelkan dalam buku-buku statistika. Microsoft Excel memberikan fungsi statistika untuk ini yaitu

=**normsdist(z)**

misalnya

=normsdist(-2) = 0,0228
=normsdist(+2) = 0,9772
=normsdist(0) = 0,5000

Untuk sebaran normal dengan rata-rata μ dan simpangan baku σ juga difasilitasi oleh Microsoft Excel. Berikut adalah hal-hal yang dapat diperoleh:

Probability Mass Function atau besarnya nilai *f(x)*

=normdist(x;mean;standard_dev;FALSE)

misalnya

=normdist(48;60;6;false) = 0,0090
=normdist(25;10;12;false) = 0,0152

Cumulative Distribution Function atau kumulatif peluang sampai dengan *x*

=normdist(x;mean;standard_dev;TRUE)

misalnya

=normdist(48;60;6;true) = 0,0228
=normdist(25;10;12;true) = 0,8944

Parameter Lokasi dan Skala

Dalam setiap definisi berikut, $F_0(z)$ menunjuk kepada fungsi sebaran kumulatif dan $f_0(z)$ fungsi kepekatan peluang.

Definisi 2.16.

Kuantiti η merupakan parameter lokasi untuk suatu sebaran peubah acak *X* jika fungsi sebaran kumulatifnya memiliki bentuk $F(x;\eta) = F_0(x-\eta)$. Dengan perkataan lain, fungsi kepekatan peluangnya memiliki bentuk $f(x;\eta) = f_0(x-\eta)$.

Teladan 2. 15.

Perhatikan sebuah sebaran yang memiliki fungsi kepekatan peluang $f(x;\eta) = e^{-(x-\eta)}$ *untuk* $x > \eta$ dan nol untuk selainnya. Parameter lokasi disini kadang juga sering disebut dengan **threshold parameter**.

Definisi 2.17.

Suatu kuantiti positif θ disebut sebagai parameter skala untuk sebaran peubah acak *X* jika fungsi sebaran kumulatifnya memiliki bentuk $F(x;\theta) = F_0\left(\frac{x}{\theta}\right)$ dan dengan demikian fungsi kepekatan peluangnya dalam bentuk $f(x;\theta) = \frac{1}{\theta} f_0\left(\frac{x}{\theta}\right)$.

Definisi 2.18.

Kuantiti η dan kuantiti positif θ disebut sebagai parameter lokasi-skala untuk sebaran peubah acak *X* jika fungsi sebaran kumulatifnya memiliki bentuk $F(x;\theta,\eta) = F_0\left(\frac{x-\eta}{\theta}\right)$ dan dengan demikian fungsi kepekatan peluangnya dalam bentuk $f(x;\theta,\eta) = \frac{1}{\theta} f_0\left(\frac{x-\eta}{\theta}\right)$.

Sebaran Campuran

Sangat dimungkinkan suatu peubah acak memiliki sebaran yang tak sepenuhnya diskret ataupun kontinu, melainkan campuran. Fungsi sebaran kumulatifnya memiliki bentuk

$$F(x) = \alpha F_d(x) + (1-\alpha)F_c(x)$$

Dimana $F_d(x)$ dan $F_c(x)$ masing-masing berturut-turut adalah fungsi sebaran kumulatif dari tipe diskret dan kontinu.

Teladan 2.16.

Di sebuah perempatan jalan dengan rambu lalu lintas STOP, seorang pengemudi akan menghadapi dua hal, yaitu berhenti beberapa saat menunggu giliran jalan, atau berhenti langsung jalan lagi. Dengan demikian peubah acaknya mengambil nilai (karena waktu) nol atau positif, keduanya dengan peluang yang tidak nol. Misalkan funbgsi sebaran waktu tunggu tersebut adalah

$$F(x) = 0{,}4F_d(x) + 0{,}6F_c(x)$$

Misalkan $F_d(x)=1$ dan $F_c(x)=1-e^{-x}$ jika $x \geq 0$ dan nol selainnya. Dengan demikian peluang tanpa harus menunggu adalah P[X=0] = 0,4. Dan peluang waktu tunggunya kurang dari 0,5 menit adalah $P[X \leq 0{,}5]=0{,}4+0{,}6(1-e^{-0{,}5}) = 0{,}636$

Latihan

1. Bila sebuah dadu bermuka dua belas (*dodecahedral*) yang seimbang dilempar dua kali. Bila setiap muka diberi nomor 1 sampai dengan 12, maka setiap nomor ini memiliki kesempatan muncul yang sama. Buatlah fungsi kepekatan peluang dari peubah acak maksimum nilai dari munculnya kedua mata dadu.
2. Sebuah permainan dikatakan adil apabila nilai harapan permainan tersebut adalah nol. Verifikasi pernyataan berikut ini:

Dari sebuah permainan angka dari 00 sampai dengan 99, seseorang harus memasang sebesar Rp. X,-- untuk sebuah nomor yang dipilihnya. Apabila nomor yang dipilihnya keluar, maka ia akan mendapatkan hadiah sebesar 60 kali uang yang dipertaruhkannya atau Rp. 60X,-- namun apabila kalah maka uang untuk memasang nomor tersebut hilang. Apabila keluarnya nomor dari 00 sampai dengan 99 tersebut acak, maka si pemain akan memiliki nilai harapan yang negatif. Artinya, si pemain dalam jangka panjang akan mengalami kekalahan. Berapakah nilai harapan si pemain dalam soal ini ?

3. Suatu peubah acak diskret memiliki fungsi kepekatan peluang $f(x)$.
   a. Jika $f(x)=k(1/2)^x$ untuk $x = 1, 2,$ dan 3. Untuk x lain peluangnya nol. Tentukan nilai k agar $f(x)$ memenuhi definisi fungsi kepekatan peluang.
   b. Apakah fungsi dalam bentuk $f(x)=k[(1/2)^x-1/2]$ untuk $x = 0, 1,$ dan 2 merupakan suatu fungsi kepekatan peluang ?
4. Fungsi [x] adalah fungsi bilangan bulat yang tidak melebihi x. Fungsi Sebaran Kumulatif dari $f(x)=(2x-1)/144$ untuk $x=1, 2, \ldots, 12$ adalah $F(x)=([x]/12)^2$ untuk $0 < x < 13$, nol untuk $x < 1$ dan satu untuk $x > 13$.
5. Suatu peubah acak X memiliki fungsi kepekatan peluang $f(x) = c(8-x)$ untuk $x = 0, 1, 2, 3, 4, 5$ dan nol untuk x selainnya.
   a. Tentukan besarnya c.
   b. Cari Fungsi Sebaran Kumulatifnya.
   c. Berapakah $P[X>2]$
   d. Carilah E(X)
   e. Carilah Var(X)
6. Suatu peubah acak bilangan bulat tidak negatif X memiliki fungsi sebaran kumulatif $F(x)=1-(1/2)^{x+1}$ untuk $x=0, 1, 2, \ldots$ dan nol untuk $x < 0$
   a. Cari Fungsi Kepekatan Peluangnya.
   b. Tentukan $P[10<X\leq 20]$
   c. Carilah $P[X$ bilangan bulat$]$

7. Apabila X memiliki sebaran hypergeometrik dengan parameter *n, M,* dan *N*. Lakukan verifikasi bahwa
   a. $E(X) = \dfrac{nM}{N}$
   b. $Var(X) = n\dfrac{M}{N}\left(1 - \dfrac{M}{N}\right)\dfrac{N-n}{N-1}$
8. Apabila X memiliki sebaran Geometrik dengan parameter *p*, maka lakukan verifikasi bahwa
   a. $E(X^2) = \dfrac{1+q}{p^2}$
   b. $Var(X) = \dfrac{q}{p^2}$
9. Lakukan verifikasi ragam peubah acak peubah acak di bawah ini
   a. $Var(X) = \dfrac{rq}{p^2}$ jika X menyebar dengan sebaran Negatif Binomial(*r,p*)
   b. $Var(X) = \mu$ jika X menyebar dengan sebaran Poisson(μ)
10. Jika *X* adalah peubah acak kontinu yang menyebar menurut sebaran Weibull dengan parameter θ dan β, tunjukkan bahwa persentil ke-p nya adalah $x_p = \theta[-\ln(1-p)]^{1/\beta}$.
11. Suatu peubah acak kontinu *X* dikatakan memiliki sebaran Pareto dengan parameter $\theta > 0$ dan $\kappa > 0$ jika memiliki fungsi kepekatan peluang dalam bentuk : untuk *x>0*
    $f(x;\theta,\kappa) = \dfrac{\kappa}{\theta}\left(1 + \dfrac{x}{\theta}\right)^{-(\kappa+1)}$ dan nol untuk *x* yang lainnya.
    a. Tentukan $F(x;\theta,\kappa)$
    b. Carilah *E(X)*
    c. Carilah *Var(X)*
    d. Carilah persentil ke-p

12. Buktikan bahwa $\Gamma\left(\frac{1}{2}\right) = \sqrt{\pi}$. Petunjuk : gunakan langkah-langkah berikut

   a. Gunakan substitusi $x = \sqrt{t}$ dalam integral $\Gamma\left(\frac{1}{2}\right) = \int_0^\infty t^{-1/2} e^{-t}\, dt$

   b. Rubahlah kedalam koordinat polar (kutub) dalam integral ganda dua $\left[\Gamma\left(\frac{1}{2}\right)\right]^2 = \int_0^\infty \int_0^\infty 4e^{-\left(x^2+y^2\right)} dx dy$

Sebaran Bersama

Pendahuluan

Dalam penerapannya, akan banyak ditemui penggunaan lebih dari satu peubah, sebut saja $X_1, X_2, \ldots, X_k$. Secara matematis peubah-peubah acak ini sebagai komponen dari vektor berdimensi-*k*, dinotasikan dengan $\underline{\mathbf{X}} = (X_1, X_2, \ldots, X_k)$, yang dapat mengasumsikan nilai $\underline{\mathbf{x}} = (x_1, x_2, \ldots, x_k)$ dalam ruang Euclid berdimensi-*k*. Nilai pengamatan $\underline{x}$ bisa saja merupakan hasil pengukuran dari sebanyak *k* ciri (karakteristik) dari tiap individu, atau hasil pengukuran satu ciri sebanyak *k* kali. Kasus terakhir bisa juga keluaran tindakan yang diulang sebanyak *k* kali dari suatu percobaan.

Sebaran Diskret Bersama

Definisi 3.1. Fungsi kepekatan peluang bersama dari peubah acak diskret berdimensi-*k* $\underline{\mathbf{X}} = (X_1, X_2, \ldots, X_k)$ didefinisikan sebagai

$$f(x_1, x_2, \ldots, x_k) = P[X_1 = x_1, X_2 = x_2, \ldots, X_k = x_k]$$

untuk semua kemungkinan nilai $\underline{\mathbf{x}} = (x_1, x_2, \ldots, x_k)$ dari $\underline{\mathbf{X}}$.

Dalam konteks ini, notasi $[\, X_1=x_1, X_2=x_2, \ldots, X_k=x_k \,]$ melambangkan interseksi atau potongan dari sebanyak *k* peristiwa $[\, X_1=x_1 \,] \cap [\, X_2=x_2 \,] \cap \ldots \cap [\, X_k=x_k \,]$. Notasi lain untuk fungsi kepekatan peluang bersama mencakup subskrip yang menunjukkan peubah acak apa saja yang termasuk didalamnya, seperti berikut

$$f_{X_1, X_2, \ldots, X_k}(x_1, x_2, \ldots, x_k)$$

Teladan 3. 1.

Sebuah kantong yang berisi 100 kelereng yang terdiri dari 40 kelereng warna merah, 40 kelereng warna biru, dan 20 kelereng warna hijau. Jika sebanyak 10 kelereng diambil secara acak sekaligus atau satu-satu tanpa pengembalian, maka $\boldsymbol{X_1}$ yang menyatakan banyaknya kelereng merah terpilih, dan $\boldsymbol{X}_2$ yang menyatakan banyaknya kelereng biru terpilih adalah peubah acak diskret bersama yang memiliki fungsi kepekatan peluang bersama seperti berikut

$$f(x_1, x_2) = \frac{C_{x_1}^{40} C_{x_2}^{40} C_{10-x_1-x_2}^{20}}{C_{10}^{100}}$$

untuk semua $0 \le x_1$, $0 \le x_2$ dan $x_1 + x_2 \le 10$.

Sebaran X-Hypergeometrik

Sari sejumlah terbatas (N) item, terdiri dari (k+1) item yang berbeda, terdapat M_1 item 1, M_2 item 2, dan seterusnya.. Bila dipilih secara acak sebanyak *n* contoh dari *N* item yang ada tanpa pengembalian, dan misalkan X_i merupakan banyaknya item tipe *i* yang dipilih. Vektor $\underline{\boldsymbol{X}} = (X_1, X_2, \ldots, X_k)$ memiliki sebaran *X-Hypergeometrik* (*Extended Hypergeometric*) dan memiliki fungsi kepekatan peluang bersama dalam bentuk seperti berikut

$$f(x_1, x_2, \ldots, x_k) = \frac{C_{x_1}^{M_1} C_{x_2}^{M_2} \ldots C_{x_k}^{M_k} C_{x_{k+1}}^{M_{k+1}}}{C_n^N}$$

untuk semua $0 \le x_i \le M_i$, dimana $M_{k+1} = N - \sum_{i=1}^{k} M_i$ dan

$$x_{k+1} = n - \sum_{i=1}^{k} x_i \ .$$

Sebaran Multinomial

Bila terdapat *k+1* peristiwa yang saling lepas dan menenggang (*mutually exclusive and exhaustive*), sebut saja $E_1, E_2, \ldots, E_k, E_{k+1}$ yang dapat terjadi pada sembarang tindakan dari suatu percobaan, dan misalkan $p_i = P(E_i)$ untuk $i = 1, 2, \ldots, k+1$. Pada *n* tindakan yang saling bebas dari suatu percobaan, X_i merupakan banyaknya peristiwa E_i terjadi. Vektor $\underline{\boldsymbol{X}}$ **=** (X_1, X_2, …, X_k) dikatakan memiliki sebaran multinomial dengan fungsi kepekatan peluang

$$f(x_1, x_2, ..., x_k) = \frac{n!}{x_1!x_2!...x_{k+1}!} p_1^{x_1} p_2^{x_2} ... p_{k+1}^{x_{k+1}}$$

untuk semua $0 \leq x_i \leq n$, dimana $x_{k+1} = n - \sum_{i=1}^{k} x_i$ dan $p_{k+1} = 1 - \sum_{i=1}^{k} p_i$. Notasi khusus untuk sebaran multinomial ini adalah $\underline{\boldsymbol{X}}$ ~ Multinomial($n, p_1, p_2, \ldots, p_k$)

Rasionalisasi multinomial ini adalah sebagai pengembangan dari Binomial.

Teladan 3. 2.

Sebuah dadu empat muka yang seimbang dilempar sebanyak 20 kali. Dengan demikian, $p_1 = p_2 = p_3 = p_4$ = 0,25.
Peluang munculnya (menghadap ke bawah) 4 kali mata 1, 6 kali mata 2, 5 kali mata 3 (dan tentu sisanya 5 kali mata 4) adalah

$$\frac{20!}{4!6!5!5!} 0{,}25^{20} = 0{,}0089$$

Sedangkan kalau kita memiliki peubah munculnya mata 1, mata 3, dan mata genap, yang berimplikai bahwa $p_1 = p_3$ = 0,25 dan p_{genap} = $1-p_1-p_3$ = 0,5. Peluang munculnya 4 kali mata 1, 5 kali mata 3, dan sisanya mata genap adalah

$$\frac{20!}{4!5!11!}(0{,}25)^9 (0{,}50)^{11} = 0{,}0394$$

Teorema 3. 1.

Suatu fungsi $f(x_1, x_2,\ldots, x_k)$ merupakan fungsi kepekatan peluang bersama untuk beberapa vektor nilai peubah acak $\underline{\mathbf{X}} = (X_1, X_2, \ldots, X_k)$ jika dan hanya jika sifat-sifat berikut dipenuhi.

$f(x_1, x_2,\ldots, x_k) \geq 0$ untuk semua kemungkinan nilai $(x_1, x_2, \ldots, x_k)$ dan

$$\sum_{x_1} \ldots \sum_{x_k} f(x_1, x_2, \ldots, x_k) = 1$$

Teladan 3. 3.

Dalam beberapa permasalahan dua dimensi, akan lebih nyaman bila fungsi kepekatan peluang bersama disajikan dalam bentuk tabel, khususnya bentuk sederhana fungsi tersebut tidak diketahui. Untuk Multinomial(3; 0,4 ; 0,4) akan kita peroleh fungsi kepekatan peluang bersamanya seperti berikut

Tabel 3. 1. Multinomial(3; 0,4 ; 0,4)

		x_2				
		0	1	2	3	
x_1	0	0,008	0,048	0,096	0,064	**0,216**
	1	0,048	0,192	0,192	0,000	**0,432**
	2	0,096	0,192	0,000	0,000	**0,288**
	3	0,064	0,000	0,000	0,000	**0,064**
		0,216	**0,432**	**0,288**	**0,064**	**1,000**

Untuk Multinomial(3; 0,3 ; 0,4) hasilnya akan menjadi

Tabel 3. 2. Multinomial (3; 0,3 ; 0,4)

		x_2				
		0	1	2	3	
x_1	0	0,027	0,108	0,144	0,064	**0,343**
	1	0,081	0,216	0,144	0,000	**0,441**
	2	0,081	0,108	0,000	0,000	**0,189**
	3	0,027	0,000	0,000	0,000	**0,027**
		0,216	**0,432**	**0,288**	**0,064**	**1,000**

Sedangkan Multinomial(3; 0,4 ; 0,2) akan diperoleh

Tabel 3. 3. Multinomial (3; 0,4 ; 0,2)

		x_2				
		0	1	2	3	
x_1	0	0,064	0,096	0,048	0,008	**0,216**
	1	0,192	0,192	0,048	0,000	**0,432**
	2	0,192	0,096	0,000	0,000	**0,288**
	3	0,064	0,000	0,000	0,000	**0,064**
		0,512	**0,384**	**0,096**	**0,008**	**1,000**

Dari ketiga teladan fungsi kepekatan **Trinomial** itu, apa yang dapat anda simpulkan ?

Definisi 3. 2.

Jika pasangan (X_1, X_2) dari peubah acak diskret memiliki fungsi kepekatan peluang bersama $f(x_1,x_2)$, maka fungsi kepekatan peluang marjinal dari X_1 dan X_2 adalah

$$f_1(x_1) = \sum_{x_2} f(x_1, x_2)$$

dan

$$f_2(x_2) = \sum_{x_1} f(x_1, x_2)$$

Pada teladan terakhir, angka pada bagian bawah yang merupakan jumlah pada masing-masing kolom, menunjukkan fungsi kepekatan peluang marjinal untuk X_2. Sedangkan angka di bagian kolom paling kanan, merupakan jumlah pada masing-masing baris, yang merupakan fungsi kepekatan peluang marjinal X_1.

Seandainya, (X_1,X_2) ~ Multinomial(n,p_1,p_2), maka fungsi kepekatan peluang marjinal X_1 dapat diturunkan seperti beriku

$$f_1(x_1) = \sum_{x_2} f(x_1,x_2)$$

$$= \sum_{x_2=0}^{n-x_1} f(x_1,x_2)$$

$$= \sum_{x_2=0}^{n-x_1} \frac{n!\,p_1^{x_1} p_2^{x_2}[(1-p_1)-p_2]^{(n-x_1)-x_2}}{x_1!\,x_2![(n-x_1)-x_2]!}$$

$$= \frac{n!}{x_1!(n-x_1)!} p_1^{x_1} \sum_{x_2=0}^{n-x_1} \frac{(n-x_1)!\,p_2^{x_2}[(1-p_1)-p_2]^{(n-x_1)-x_2}}{x_2![(n-x_1)-x_2]!}$$

$$= \frac{n!}{x_1!(n-x_1)!} p_1^{x_1} \sum_{x_2=0}^{n-x_1} C_{x_2}^{n-x_1} p_2^{x_2}[(1-p_1)-p_2]^{(n-x_1)-x_2}$$

$$= C_{x_1}^{n} p_1^{x_1}[p_2+(1-p_1)-p_2]^{n-x_1}$$

$$= C_{x_1}^{n} p_1^{x_1}(1-p_1)^{n-x_1}$$

Dengan demikian, kita dapat katakan bahwa X_1 ~ Bin(n,p_1).

Definisi 3. 3.

Fungsi Sebaran Kumulatif Bersama dari k peubah acak $X_1, X_2, \ldots, X_k$ adalah fungsi yang didefinisikan sebagai berikut

$$F(x_1,x_2,\ldots,x_k) = P[X_1 \le x_1, X_2 \le x_2, \ldots, X_k \le x_k]$$

Fungsi ini menunjukkan peluang bahwa vektor peubah acak tersebut mengambil suatu nilai dalam ruang berdimensi-*k*, *A.* Sebagaimana dalam kasus ruang berdimensi satu, peristiwa lain dapat diekspresikan dengan bentuk peristiwa seperti *A*, sehingga fungsi sebaran kumulatif secara lengkap menggambarkan model peluang.

Seperti dalam kasus satu dimensi, terdapat beberapa persyaratan yang harus dipenuhi sebagai suatu fungsi sebaran peluang. Fungsi sebaran kumulatif bersama harus memenuhi sifat-sifat yang tentunya analog dengan apa yang berlaku untuk kasus satu dimensi. Untuk kasus bivariat atau peubah ganda dua, akan disajikan dalam teorema berikut. Untuk peubah ganda-*k* akan mirip hasilnya.

Teorema 3. 2.

Suatu fungsi $F(x_1,x_2)$ merupakan fungsi sebaran kumulatif bivariat, jika dan hanya jika

a. $\lim_{x_1 \to -\infty} F(x_1,x_2) = F(-\infty, x_2) = 0$ untuk semua x_2

b. $\lim_{x_2 \to -\infty} F(x_1,x_2) = F(x_1, -\infty) = 0$ untuk semua x_1

c. $\lim_{\substack{x_1 \to \infty \\ x_2 \to \infty}} F(x_1,x_2) = F(\infty,\infty) = 1$

d. $F(b,d) - F(b,c) - F(a,d) + F(a,c) \geq 0$ untuk semua $a < b$ dan $c < d$.

e. $\lim_{h \to 0^+} F(x_1 + h, x_2) = \lim_{h \to 0^+} F(x_1, x_2 + h) = F(x_1,x_2)$ untuk semua x_1 dan x_2.

Sifat ke-4 (*d*) merupakan kondisi monotonik versi dua dimensi. Hal ini diperlukan untuk mencegah adanya nilai negatif dari suatu peluang peristiwa dalam bentuk $A = (a,b] \times (c,d]$. Secara khusus,

$$P[a < X_1 \leq b,\ c < X_2 \leq d] = F(b,d) - F(b,c) - F(a,d) + F(a,c)$$

yang merupakan nilai disebelah kiri poin *d*.

Sebaran Kontinu Bersama

Definisi 3. 4.

Suatu vektor peubah acak berdimensi-*k* $\underline{\boldsymbol{X}} = (X_1, X_2, \ldots, X_k)$ dikatakan kontinu jika terdapat suatu fungsi $f(x_1, x_2, \ldots, x_k)$, yang disebut dengan fungsi kepekatan peluang bersama dari $\underline{\boldsymbol{X}}$, sedemikian rupa sehingga fungsi sebaran kumulatif bersamanya dapat dituliskan seperti berikut

$$F(x_1,\ldots,x_k) = \int_{-\infty}^{x_k} \cdots \int_{-\infty}^{x_1} f(t_1,\ldots,t_k)\,dt_1 \cdots dt_k$$

untuk semua $\underline{\boldsymbol{x}} = (x_1, x_2, \ldots, x_k)$.

Sebagaimana dalam kasus satu dimensi, fungsi kepekatan peluang bersama dapat diperoleh dari fungsi sebaran kumulatif bersama dengan menggunakan teknik diferensial (turunan), seperti berikut :

$$f(x_1,\ldots,x_k) = \frac{\partial^k}{\partial x_1 \cdots \partial x_k} F(x_1,\ldots,x_k)$$

bilamana turunan parsialnya ada.

Teorema 3. 3.

Sembarang fungsi $f(x_1, x_2, \ldots, x_k)$ merupakan fungsi kepekatan peluang bersama dari peubah acak berdimensi-*k* jika dan hanya jika

$$f(x_1,\ldots,x_k) \geq 0 \quad \text{untuk semua } x_1, \ldots, x_k$$

dan

$$\int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} f(x_1,\ldots,x_k)\,dx_1 \cdots dx_k = 1$$

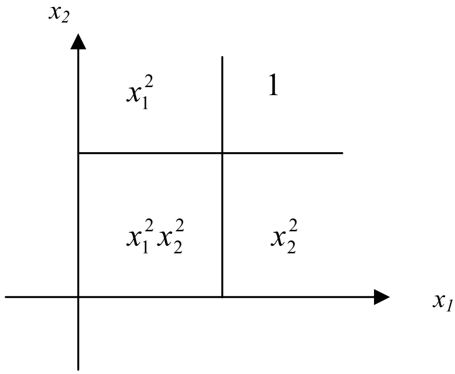
Banyak aplikasi yang dapat dimodelkan dengan peubah acak kontinu bersama.

Teladan 3. 4.

Misalkan X_1 merupakan konsentrasi suatu campuran dalam tindakan pertama, dan X_2 merupakan konsentrasi campuran dalam tindakan kedua. Misalkan fungsi kepekatan peluang bersama $f(x_1,x_2) = 4x_1x_2;\ 0<x_1<1,\ 0<x_2<1$.

Carilah fungsi sebaran kumulatif bersamanya !

Untuk menjawab itu semua, sebaiknya harus difahami terlebih dahulu daerah fungsi dari fungsi sebaran kumulatif bersamanya.



Sebagai teladan, untuk $0<x_1<1,\ 0<x_2<1$

$$F(x_1,x_2) = \int_{-\infty}^{x_2}\int_{-\infty}^{x_1} f(t_1,t_2)\,dt_1\,dt_2$$

$$= \int_{0}^{x_2}\int_{0}^{x_1} 4t_1t_2\,dt_1\,dt_2$$

$$= x_1^2x_2^2 \qquad 0<x_1<1\,;\,0<x_2<1$$

Untuk $1\le x_1<\infty,\ 0<x_2<1$

$F(x_1,x_2) = x_1^2$ karena $f(x_1,x_2) = 4x_1x_2;\ 0<x_1<1,\ 0<x_2<1$

Dengan cara yang sama, kita bisa peroleh nilai fungsi sebaran kumulatif bersama secara lengkap seperti berikut:

$$F(x_1, x_2) = \begin{cases} x_1^2 x_2^2 & untuk \quad 0 < x_1 < 1; 0 < x_2 < 1 \\ x_1^2 & untuk \quad 0 < x_1 < 1; 1 \le x_2 < \infty \\ x_2^2 & untuk \quad 1 \le x_1 < \infty; 0 < x_2 < 1 \\ 1 & untuk \quad 1 \le x_1 < \infty; 1 \le x_2 < \infty \\ 0 & untuk \quad selainnya \end{cases}$$

Secara umum, untuk peubah acak $\underline{\mathbf{X}}$ = ($X_1, X_2, \ldots, X_k$) dan peristiwa *A* berdimensi-*k* kita peroleh bahwa

$$P[\underline{X} \in A] = \int \underset{A}{\cdots} \int f(x_1, \ldots, x_k) dx_1 \cdots dx_k$$

Teladan 3. 5.

Dengan menggunakan permasalahan pada Teladan 3.4. peluang untuk mendapatkan rata-rata konsentrasi kurang dari 0,5 dapat dicari seperti berikut.

$$\begin{aligned} P[(X_1 + X_2)/2 < 0{,}5] &= P[(X_1, X_2) \in A] \\ &= \iint_A f(x_1, x_2) dx_1 dx_2 \\ &= \int_0^1 \int_0^{1-x_2} 4x_1 x_2 dx_1 dx_2 \\ &= \int_0^1 2x_2 (1 - x_2)^2 dx_2 \\ &= \frac{1}{6} \end{aligned}$$

Fungsi sebaran kumulatif marjinal X_1 dapat diperoleh dengan cara seperti berikut

$$\begin{aligned} F_1(x_1) &= P[X_1 \le x_1] \\ &= P[X_1 \le x_1, X_2 < \infty] \\ &= F(x_1, \infty) \end{aligned}$$

$$= \int_{-\infty}^{x_1}\left(\int_{-\infty}^{\infty} f_{X_1,X_2}(t_1,t_2)dt_2\right)dt_1$$

$$= \int_{-\infty}^{x_1} f_1(t_1)dt_1$$

Dengan demikian, untuk peubah acak kontinu, fungsi sebaran kumulatif marjinal bagi X_1 adalah $F(x_1,\infty)$ dan fungsi kepekatan peluang marjinalnya adalah

$$f_1(x_1) = \frac{d}{dx_1}F_1(x_1)$$

$$= \frac{d}{dx_1}\int_{-\infty}^{\infty}\int_{-\infty}^{x_1} f(t_1,x_2)dx_2$$

$$= \int_{-\infty}^{\infty} f(x_1,x_2)dx_2$$

Definisi 3. 5.

Jika pasangan peubah acak kontinu (X_1, X_2) memiliki fungsi kepekatan peluang bersama $f(x_1, x_2)$, maka fungsi kepekatan peluang marjinal X_1 dan X_2 masing-masing berturut-turut adalah

$$f_1(x_1) = \int_{-\infty}^{\infty} f(x_1,x_2)dx_2$$

$$f_2(x_2) = \int_{-\infty}^{\infty} f(x_1,x_2)dx_1$$

Teladan 3. 6.

Masih dengan kelanjutan permasalahan pada Teladan 3.4. maka fungsi kepekatan peluang marjinal X_1 adalah

$$f_1(x_1) = \int_0^1 4x_1x_2dx_2$$

$$= 4x_1\int_0^1 x_2dx_2$$

$$= 2x_1$$

untuk sembarang 0 < x_1 < 1, dan nol selainnya. Cara yang sama juga digunakan untuk mendapatkan fungsi kepekatan peluang marjinal X_2 adalah

$$f_2(x_2) = \int_0^1 4x_1x_2dx_1$$

$$= 4x_2\int_0^1 x_1dx_1$$

$$= 2x_2$$

untuk sembarang 0 < x_2 < 1, dan nol selainnya.

Definisi 3. 6.

Jika $\underline{\boldsymbol{X}}$ = ($X_1, X_2, \ldots, X_k$) merupakan suatu vektor peubah acak berdimensi-*k* dengan fungsi sebaran kumulatif $F(x_1,x_2,\ldots,x_k)$, maka fungsi sebaran kumulatif marjinal X_j adalah

$$F_j(x_j) = \lim_{\substack{x_i\to\infty \\ \forall i\neq j}} F(x_1,\ldots,x_j,\ldots,x_k)$$

Apabila $\underline{\boldsymbol{X}}$ diskret, fungsi kepekatan peluang marjinalnya adalah

$$f_j(x_j) = \sum \underset{\forall i\neq j}{\cdots} \sum f(x_1,\ldots,x_j,\ldots,x_k)$$

dan jika $\underline{\boldsymbol{X}}$ kontinu, fungsi kepekatan peluang marjinalnya adalah

$$f_j(x_j) = \int \underset{\forall i\neq j}{\cdots} \int f(x_1,\ldots,x_j,\ldots,x_k)dx_1\cdots dx_k$$

Teladan 3. 7.

Misalkan X_1, X_2, dan X_3 merupakan peubah acak kontinu dengan fungsi kepekatan peluang bersama $f(x_1,x_2,x_3) = c$; $0<x_1<x_2<x_3<1$; dan nol selainnya; karena total peluang harus sama dengan 1, maka nilai $c = 6$.

Bila kita tertarik untuk mendapatkan fungsi kepekatan peluang marjinal dari X_3, akan kita dapatkan dengan cara seperti berikut

$$f_3(x_3) = \int_0^{x_3}\int_0^{x_2} 6\,dx_1\,dx_2$$

$$= 6\int_0^{x_3} x_2\,dx_2$$

$$= 3x_3^2$$

untuk $0 < x_3 < 1$, dan nol selainnya.

Sedangkan untuk mendapatkan fungsi kepekatan peluang bersama (X_1,X_2) dilakukan dengan mengintegralkan fungsi kepekatan peluang $f(x_1,x_2,x_3)$ terhadap x_3 seperti berikut

$$f(x_1,x_2) = \int_{-\infty}^{\infty} f(x_1,x_2,x_3)\,dx_3$$

$$= \int_{x_2}^{1} 6\,dx_3$$

$$= 6(1-x_2)$$

untuk $0 < x_1 < x_2 < 1$, dan nol untuk selainnya.

Formula yang lebih umum untuk mendapatkan fungsi kepekatan peluang bersama dari sembarang gugus peubah acak akan melibatkan ekspresi yang lebih rumit, dan kita tak akan mencoba memberikan formula tersebut. Namun demikian, prosedur yang telah digambarkan diatas, dengan melibatkan pengintegralan menurut peubah yang “tak dikehendaki”, memberikan suatu pendekatan untuk penyelesaian masalah yang lebih umum.

Peubah Acak Independen

Misalkan kita memiliki peubah acak diskrit X_1 dan X_2 yang fungsi kepekatan peluang bersamanya dapat dilihat pada Tabel 3. 4. Jelas terlihat bahwa $f(1,1)=0,2=f_1(1)f_2(1)$. Dengan demikian $P[X_1 = 1$ dan $X_2 = 1] = P[X_1 = 1]P[X_2 = 1]$, dan akan kita katakan bahwa peristiwa $[X_1 = 1]$ dan $[X_2 = 1]$ adalah peristiwa yang saling bebas, sebagaimana yang telah disebutkan dalam bab awal dalam buku ini. Namun demikian, juga dapat diperiksa dengan cara yang sama $f(1,2)=0,1\neq f_1(1)f_2(2)$, maka kita katakan bahwa peristiwa $[X_1 = 1]$ dan $[X_2 = 2]$ adalah bukan peristiwa yang saling bebas. Sehingga, secara umum kita dapat katakan bahwa X_1 dan X_2 merupakan dua peubah yang saling **tidak bebas**. Jadi, jika $f(x_1,x_2)=f_1(x_1)f_2(x_2)$ untuk semua kemungkinan (x_1,x_2), maka hal ini baru kita dapat katakan bahwa kedua peubah X_1 dan X_2 saling **bebas**.

Tabel 3. 4.

Peluang bersama dua peubah acak X_1 dan X_2

		x_2			
		0	1	2	$f_1(x_1)$
x_1	0	0,1	0,2	0,1	**0,4**
	1	0,1	0,2	0,1	**0,4**
	2	0,1	0,1	0,0	**0,2**
$f_2(x_2)$		**0,3**	**0,5**	**0,2**	**1,0**

Hal yang sama, juga berlaku untuk peubah acak kontinu, misalkan $f(x_1,x_2) = f_1(x_1)f_2(x_2)$ untuk semua x_1 dan x_2. Sebagai akibatnya, jika $a < b$ dan $c < d$, maka

$$P[a \le X_1 \le b, c \le X_2 \le d] = \int_c^d \int_a^b f(x_1, x_2)dx_1 dx_2$$

$$= \int_c^d \int_a^b f_1(x_1) f_2(x_2) dx_1 dx_2$$

$$= \int_a^b f_1(x_1) dx_1 \int_c^d f_2(x_2) dx_2$$

$$= P[a \le X_1 \le b] P[c \le X_2 \le d]$$

Peristiwa $A=[a \le X_1 \le b]$ dan $B=[c \le X_2 \le d]$ merupakan dua peristiwa yang saling bebas, sebagaimana konsep dua peristiwa yang saling bebas yang telah dibahas dalam bab awal buku ini. Konsep saling bebas antar peristiwa ini dikembangkan untuk peubah acak-peubah acak. Dua peubah acak X_1 dan X_2 dikatakan saling bebas, apabila semua bentuk peristiwa *A* dan *B* saling bebas. Konsep ini berlaku untuk peubah acak baik diskrit maupun kontinu.

Definisi 3. 7.

Peubah acak $X_1, \ldots, X_k$ dikatakan saling bebas jika untuk setiap $a_i < b_i$

$$P[a_1 \le X_1 \le b_1, \ldots, a_k \le X_k \le b_k] = \prod_{i=1}^{k} P[a_i \le X_i \le b_i]$$

Ekspresi sisi sebelah kanan merupakan hasil kali peluang-peluang marjinalnya $P[a_1 \le X_1 \le b_1]$, $P[a_2 \le X_2 \le b_2], \ldots, P[a_k \le X_k \le b_k]$. Terminologi bebas stokastik juga sering digunakan dalam hal ini. Jadi, jika kondisi diatas tidak berlaku untuk semua $a_i < b_i$, maka peubah acak-peubah acak tersebut dikatakan **terkait** atau **saling tidak bebas**.

Teorema 3. 4.

Peubah acak $X_1, \ldots, X_k$ saling bebas jika dan hanya jika salah satu dari yang berikut dipenuhi

$$F(x_1, \ldots, x_k) = F_1(x_1) \cdots F_k(x_k)$$

$$f(x_1,...,x_k) = f_1(x_1)\cdots f_k(x_k)$$

dimana $F_i(x_i)$ dan $f_i(x_i)$ berturut-turut adalah fungsi sebaran kumulatif dan fungsi kepekatan peluang peubah acak X_i.

Teorema 3. 5.

Peubah acak X_1 dan X_2 dengan fungsi kepekatan peluang bersama $f(x_1,x_2)$ saling bebas jika dan hanya jika

1. "gugus pendukung" $\{(x_1,x_2)|\ f(x_1,x_2)>0\}$, merupakan hasil kali Cartesian, $A \times B$, dan
2. fungsi kepekatan peluang bersamanya dapat difaktorkan menjadi hasilkali fungsi x_1 dan x_2, $f(x_1,x_2)=g(x_1)h(x_2)$.

Teladan 3. 8.

Fungsi kepekatan peluang bersama X_1 dan X_2 adalah

$$f(x_1,x_2) = 8x_1x_2 \qquad 0 < x_1 < x_2 < 1$$

dan nol selainnya. Jelas bahwa fungsi ini dapat difaktorkan sesuai poin 2 Teorema 3. 5., namun gugus pendukungnya $\{(x_1,x_2)\ |\ 0 < x_1 < x_2 < 1\}$ merupakan segitiga yang tak dapat direpresentasikan sebagai hasil kali Cartesian. Dengan demikian X_1 dan X_2 dikatakan saling terkait.

Teladan 3. 9.

Sepasang peubah acak X_1 dan X_2 dengan fungsi kepekatan peluang bersama

$$f(x_1,x_2) = x_1 + x_2 \qquad 0 < x_1 < 1, 0 < x_2 < 1$$

dan nol selainnya. Meskipun gugus pendukung merupakan hasil kali Cartesian, namun fungsi kepekatan peluang bersamanya tidak dapat difaktorkan sebagaimana disyaratkan oleh poin 2 Teorema 3. 5. sehingga X_1 dan X_2 saling terkait.

Sebaran Bersyarat

Kebebasan atau independensi juga berhubungan dengan konsep peluang bersyarat, yang juga disarankan bahwa definisi peluang bersyarat suatu peristiwa dapat diperluas untuk konsep peubah acak bersyarat. Dalam teladan terdahulu, formula umum untuk menyatakan peluang bersyarat adalah seperti berikut:

$$P[T = t \mid X = x] = \frac{P[X = x, T = t]}{P[X = x]} = \frac{f_{X,T}(x,t)}{f_X(x)}$$

Definisi 3. 8.

Jika X_1 dan X_2 merupakan peubah acak diskrit atau kontinu dengan fungsi kepekatan peluang bersama $f(x_1,x_2)$, maka **fungsi kepekatan peluang bersyarat** dari X_2 bilamana $X_1 = x_1$ didefinisikan sebagai

$$f(x_2 \mid x_1) = \frac{f(x_1, x_2)}{f_1(x_1)}$$

untuk nilai-nilai x_1 sedemikian rupa sehingga $f_1(x_1) > 0$, dan nol selainnya.

Demikian pula fungsi kepekatan peluang bersyarat X_1 bilamana $X_2 = x_2$ adalah

$$f(x_1 \mid x_2) = \frac{f(x_1, x_2)}{f_2(x_2)}$$

untuk nilai-nilai x_2 sedemikian rupa sehingga $f_2(x_2) > 0$, dan nol selainnya.

Sebagaimana telah diberikan dalam teladan terdahulu, untuk peubah acak diskrit, suatu fungsi kepekatan peluang bersyarat sebenarnya adalah suatu peluang bersyarat. Misalnya, X_1 dan X_2 adalah diskrit, $f(x_2|x_1)$ merupakan peluang bersyarat $[X_2=x_2]$ bilamana $[X_1=x_1]$. Untuk peubah acak kontinu, interpretasi fungsi kepekatan peluang bersyarat tidak begitu saja, karena $P[X_1=x_1] = 0$.

Meskipun $f(x_2|x_1)$ tak dapat diinterpretasikan sebagai peluang bersyarat, namun dapat dipandang sebagai 'kepekatan peluang' bersyarat untuk sembarang interval yang cukup kecil $[x_2, x_2+\Delta x_2]$, yang kira-kira sama dengan fungsi kepekatan peluang marjinal $f_2(x_2)$ memberikan kepekatan peluang marjinal. Dengan demikian, untuk peubah acak kontinu, peluang bersyarat suatu peristiwa $[a \leq X_2 \leq b]$ bilamana $X_1=x_1$ adalah

$$P[a \leq X_2 \leq b \mid X_1 = x_1] = \int_a^b f(x_2 \mid x_1)dx_2$$

$$= \frac{\int_a^b f(x_1, x_2)dx_2}{\int_{-\infty}^{\infty} f(x_1, x_2)dx_2}$$

Nilai penyebut diatas merupakan total luas dimana fungsi kepekatan peluang bersama pada nilai $X_1 = x_1$, dan pembilangnya merupakan luasan dimana $a \leq X_2 \leq b$. Hal ini dapat dipandang sebagai peluang suatu peristiwa $[a \leq X_2 \leq b]$ pada suatu bidang atau "potongan tipis" pada $X_1 = x_1$ dari suatu ruang contoh bersama dari pasangan peubah acak (X_1, X_2).

Untuk melihat keabsahan persyaratan suatu fungsi kepekatan peluang, $f(x_2|x_1)$ harus memenuhi persyaratan dimaksud dalam nilai peubah x_2 untuk nilai x_1 yang tetap. Fakta bahwa $f(x_2|x_1) \geq 0$ diperoleh langsung seperti yang dinyatakan pada Definisi 3. 8., dan juga

$$\int_{-\infty}^{\infty} f(x_2 \mid x_1)dx_2 = \frac{1}{f_1(x_1)} \int_{-\infty}^{\infty} f(x_1, x_2)dx_2$$

$$= \frac{f_1(x_1)}{f_1(x_1)} = 1$$

Konsep sebaran bersyarat dapat diperluas untuk vektor peubah acak. Misalkan $\underline{\boldsymbol{X}} = (X_1,\ldots,X_r,\ldots,X_k)$ memiliki fungsi kepekatan peluang bersama $f(\underline{\boldsymbol{x}})$ dan $\underline{\boldsymbol{X}}_1 = (X_1,\ldots,X_r)$ memiliki fungsi kepekatan peluang bersama $f_1(\underline{\boldsymbol{x}}_1)$. Jika $\underline{\boldsymbol{X}}_2 = (X_{r+1},\ldots,X_k)$, maka fungsi kepekatan peluang bersyarat dari $\underline{\boldsymbol{X}}_2$ bilamana $\underline{\boldsymbol{X}}_1 = \underline{\boldsymbol{x}}_1$ adalah $f(\underline{\boldsymbol{x}}_2|\underline{\boldsymbol{x}}_1)=f(\underline{\boldsymbol{x}})/f_1(\underline{\boldsymbol{x}}_1)$ untuk semua $\underline{\boldsymbol{x}}_1$ yang memenuhi $f_1(\underline{\boldsymbol{x}}_1) > 0$.

Teladan 3. 10.

Perhatikan peubah acak sebagaimana dimaksud pada Teladan 3. 7. Fungsi kepekatan peluang bersyarat X_3 bilamana $(X_1,X_2) = (x_1,x_2)$ adalah

$$\begin{aligned} f(x_3 \mid x_1, x_2) &= \frac{f(x_1, x_2, x_3)}{f(x_1, x_2)} \\ &= \frac{6}{6(1-x_2)} \\ &= \frac{1}{1-x_2} \qquad 0 < x_1 < x_2 < x_3 < 1 \end{aligned}$$

dan nol selainnya.

Teorema 3. 6.

Jika X_1 dan X_2 merupakan peubah acak dengan fungsi kepekatan peluang bersama $f(x_1,x_2)$ dan fungsi kepekatan peluang marjinalnya $f_1(x_1)$ dan $f_2(x_2)$, maka

$$f(x_1, x_2) = f_1(x_1) f(x_2 \mid x_1) = f_2(x_2) f(x_1 \mid x_2)$$

dan jika X_1 dan X_2 saling bebas, maka

$$f(x_2 \mid x_1) = f_2(x_2) \text{ dan } f(x_1 \mid x_2) = f_1(x_1)$$

Notasi lain sering digunakan untuk menyatakan fungsi kepekatan peluang bersyarat. Misalnya, jika X dan Y merupakan peubah acak yang menyebar bersama, maka peluang bersyarat bagi Y bilamana $X = x$ sering dituliskan dengan notasi $f_{Y|X}(y|x)$ atau barangkali cukup dengan $f_{Y|x}(y)$. Dalam banyak aplikasi, tak akan terjadi kebingungan jika kita hilangkan penggunaan subskrip dan cukup dengan menggunakan *f(y|x)*. Kadang juga untuk menyatakan "peubah acak bersyarat" cukup digunakan *Y|X* atau *Y|X=x.*

Teladan 3. 11.

Misalkan $f(x,y) = \begin{cases} 1 & , \quad 0 < y < \dfrac{x}{2} \quad 0 < x < 1 \\ 0 & , \qquad selainnya \end{cases}$

Maka akan kita peroleh fungsi kepekatan marjinal untuk *X* dan *Y* seperti berikut

$$f_1(x) = \int_0^{x/2} dy = \frac{x}{2} \qquad 0 < x < 2$$ dan nol untuk *x* lainnya.

$$f_2(y) = \int_{2y}^{2} dx = 2(1-y) \qquad 0 < y < 1$$ dan nol untuk *y* selainnya.

Fungsi kepekatan peluang bersyarat *Y* bilamana *X=x* adalah

$$f(y \mid x) = \frac{1}{x/2} = \frac{2}{x} \qquad 0 < y < \frac{x}{2}$$ dan nol untuk *y* selainnya. Ini berarti bahwa dengan kondisi pada nilai *X=x, Y~Seragam(0,x/2)*.

Misalkan juga kita ingin menghitung *P[0,1≤Y≤0,7 | X=0,5]*. Maka dengan mudah kita dapat menghitungnya seperti berikut

$$P[0,1 \le Y \le 0,7 \mid X=0,5] = \int_{0,1}^{\min(0,7;\frac{x}{2})} f(y \mid 0,5)\,dy$$

$$= \int_{0,1}^{0,25} 4\,dy$$

$$= 0,6$$

Teladan 3. 12.

Apabila kita gunakan fungsi kepekatan peluang bersama seperti pada Teladan 3. 9.

$$f(x,y) = x + y \qquad 0 < x < 1,\, 0 < y < 1$$

Dalam hal ini, untuk sembarang nilai *x* diantara 0 dan 1,

$$f(y \mid x) = \frac{f(x,y)}{f_1(x)} = \frac{x+y}{x+0{,}5} \qquad 0 < y < 1$$

Sebagai teladan

$$P[0 < Y < 0{,}5 \mid X = 0{,}25] = \int_0^{0{,}5} \frac{0{,}25+y}{0{,}25+0{,}5}\,dy = \frac{1}{3}$$

Latihan

1. Lima kartu diambil secara acak tanpa pengembalian dari setumpuk kartu Bridge standar. Misalkan X merupakan peubah acak banyaknya kartu As (A) terambil, Y peubah acak banyaknya kartu King (K) terambil, dan Z peubah acak banyaknya kartu Queen (Q) terambil. Tentukan peluang dari hal-hal berikut:
   a. $A=[X=2]$
   b. $B=[Y=2]$
   c. $A \cap B$
   d. $A \cup B$
   e. $A|B$
   f. $[X=x]$
   g. $[X<1]$
   h. $[X\geq 1]$
   i. $[X=2, Y=2, Z=1]$
   j. Tuliskan ekspresi fkp bersama *X, Y,* dan *Z*.

2. Misalkan X_1 dan X_2 peubah acak diskrit dengan fungsi kepekatan peluang bersama seperti pada tabel berikut:

		x_2 = 1	2	3
x_1	1	1/12	1/6	0
	2	0	1/9	1/5
	3	1/18	1/4	2/15

a. Tentukan fungsi kepekatan peluang marjinal dari peubah acak X_1 dan X_2
b. Apakah peubah acak X_1 dan X_2 saling bebas ?
c. Dapatkan $P[X_1 \leq 2]$
d. Dapatkan $P[X_1 \leq X_2]$
e. $E(X_1 X_2)$
f. Tabelkan fungsi kepekatan peluang bersyarat $f(x_2|x_1)$ dan $f(x_1|x_2)$

3. Misalkan peubah acak X dan Y memiliki fungsi kepekatan peluang $f(x,y)=8xy$ untuk $0 \leq x \leq y \leq 1$ dan 0 untuk selainnya. Carilah:
a. Fungsi sebaran peluang bersama $F(x,y)$
b. $f(y|x)$
c. $P[X \leq 0{,}5 | Y = 0{,}75]$
d. $P[X \leq 0{,}5 | Y \leq 0{,}75]$

4. Asumsikan bahwa X dan Y peubah acak yang saling bebas, dimana X memiliki distribusi seragam kontinu (-1,1) dan Y memiliki distribusi seragam kontinu (0,1). Dapatkan peluang bahwa akar persamaan $h(t)=0$ adalah bilangan nyata, dimana

$$h(t) = t^2 + 2Xt + Y$$

5. Bila diketahui fungsi kepekatan peluang bersama $f(x,y)=x+y$ untuk $0<x<1$ dan $0<y<1$ dan nol selainnya. Tentukan
   a. $f(x)$
   b. F(x,y)

c. $f(y|x)$
d. $F(y|x)$

Sebaran Fungsi Peubah Acak

Pendahuluan

Dalam bab-bab sebelumnya telah kita pelajari apa yang dimaksud dengan peubah acak, termasuk fungsi kepekatan peluang dan fungsi sebaran kumulatifnya. Secara umum, peubah acak yang dinotasikan dengan huruf kapital, X misalnya akan memiliki fungsi kepekatan peluang $f_X(x)$ dan fungsi sebaran kumulatif $F_X(x)$.

Dalam aplikasinya, kadang kita sering menggunakan peubah acak yang lain dari pada apa yang kita amati. Peubah acak baru tersebut dapat merupakan suatu fungsi dari peubah acak yang kita miliki atau bahkan merupakan fungsi dari beberapa peubah acak serupa.

Jika X_1 adalah peubah acak yang menyatakan umur lampu pertama dalam hari dan X_2 adalah peubah acak yang menyatakan umur lampu kedua dalam hari, maka pertanyaan berapakah peluang lampu pertama akan hidup empat belas hingga duapuluh satu hari akan setara dengan peluang lampu pertama akan hidup dua hingga tiga minggu. Jika Y adalah peubah acak yang menyatakan umur lampu pertama dalam minggu, maka kita dapat tuliskan $P(14<X_1<21) = P(2<Y<3)$, karena $Y = 7X_1$. Di lain permasalahan, misalkan Z adalah total umur kedua lampu, atau $Z = X_1 + X_2$. Kita mungkin ingin mengetahui distribusi dari Z.

Terdapat beberapa teknik untuk mencari distribusi dari fungsi peubah acak yang kita akan pelajari.

Teknik Fungsi Sebaran Kumulatif

Misalkan suatu peubah acak X yang memiliki fungsi sebaran kumulatif $F_X(x)$ dan ada suatu fungsi $Y=u(X)$. Teknik Fungsi Sebaran Kumulatif ini bekerja dengan mengekspresikan fungsi sebaran kumulatif Y dalam bentuk fungsi sebaran kumulatif X. Untuk setiap bilangan nyata y, kita definisikan suatu gugus $A_y=\{x|u(x)\leq y\}$.

Sebagai akibatnya [$Y \le y$] dan [$X \in A_y$] merupakan dua kejadian yang ekuivalen, dan dengan demikian

$$F_Y(y) = P[u(X) \le y]$$

yang juga dapat diekspresikan sebagai P[$X \in A_y$], yang nilainya sama dengan

$$P[X \in A_y] = \int_{A_y} f_X(x)dx$$

jika *X* kontinu atau

$$P[X \in A_y] = \sum_{x \in A_y} P[X = x]$$

jika *X* diskrit.

Teladan 4. 1.

Misalkan $F_X(x) = 1 - e^{-2x}$, untuk $0<x<\infty$ dan misalkan juga $Y=e^X$.

Dengan demikian

$$\begin{aligned} F_X(y) &= P[Y \le y] \\ &= P[e^X \le y] \\ &= P[X \le \ln y] \\ &= F_X(\ln y) \\ &= 1 - y^{-2} \quad \text{untuk } 1<y<\infty \end{aligned}$$

dan fungsi kepekatan peluangnya

$$f_Y(y) = \frac{d}{dy} F_Y(y) = 2y^{-3}, \quad 1 < y < \infty$$

Teladan 4. 2.

Untuk peubah acak kontinu *X*, dan $Y=X^2$. Maka akan kita dapatkan

$$\begin{aligned} F_Y(y) &= P[Y \le y] \\ &= P[X^2 \le y] \\ &= P[-\sqrt{y} \le X \le \sqrt{y}] \end{aligned}$$

$$= F_X(\sqrt{y}) - F_X(-\sqrt{y})$$

dan

$$f_Y(y) = \frac{d}{dy}[F_X(\sqrt{y}) - F_X(-\sqrt{y})$$

$$= f_X(\sqrt{y})\frac{d}{dy}(\sqrt{y}) - f_X(-\sqrt{y})\frac{d}{dy}(-\sqrt{y})$$

$$= \frac{1}{2\sqrt{y}}[f_X(\sqrt{y}) + f_X(-\sqrt{y})] \text{ untuk y > 0.}$$

Teladan 4. 3.

Bila diketahui bahwa peubah acak *X* memiliki sebaran Seragam [0, 2π]. Jika *Y* = tan *X*. Maka fungsi sebaran kumulatif *Y* dapat dicari sbb:

$$F_Y(y) = P[Y \leq y]$$

$$= P[\tan X \leq y]$$

$$= P[0 < X < \tan^{-1}(y)] + P[\pi < X < \pi + \tan^{-1}(y)]$$

$$= \frac{\tan^{-1}(y)}{2\pi} + \frac{\pi + \tan^{-1}(y)}{2\pi} - \frac{\pi}{2\pi}$$

$$= \frac{\tan^{-1}(y)}{\pi}$$

dan

$$f_Y(y) = \frac{d}{dy}F_Y(y) = \frac{1}{\pi}\frac{1}{1+y^2}, \quad -\infty < y < \infty$$

yang juga dikenal dengan distribusi atau sebaran Cauchy(1,0).

Teknik Fungsi Sebaran Kumulatif dapat juga diperluas untuk fungsi dari beberapa peubah acak.

Teorema 4. 1.

Misalkan $\underline{\mathbf{X}}$ = (X_1, X_2, ..., X_k) adalah vektor peubah acak kontinu berdimensi-k, dengan fungsi kepekatan peluang bersama $f(x_1, x_2, ..., x_k)$. Jika Y = u(X_1, X_2, ..., X_k) = $u(\underline{\mathbf{X}})$ adalah suatu fungsi dari $\underline{\mathbf{X}}$, maka

$$F_Y(y) = P[u(\underline{X}) \le y] = \int ... \int f(x_1, x_2, ..., x_k)dx_1...dx_k$$

diintegralkan untuk $A_y = \{\underline{x} \mid u(\underline{x}) \le y\}$

Teladan 4. 4.

Kita ingin mencari sebaran dari jumlah dua peubah acak yang saling bebas, $Y = X_1 + X_2$ dimana X_i memiliki distribusi Eksponensial(1). Perlu dicari terlebih dahulu daerah pengintegralan, yaitu $A_y = \{(x_1, x_2) \mid 0 \le x_1 \le y - x_2; 0 \le x_2 \le y\}$. Oleh karenanya

$$\begin{aligned} F_Y(y) &= \int_0^y \int_0^{y-x_2} e^{-(x_1+x_2)} dx_1 dx_2 \\ &= \int_0^y (-e^{-(x_1+x_2)} \Big|_0^{y-x_2} dx_2 \\ &= \int_0^y (e^{-x_2} - e^{-y})dx_2 \\ &= (-e^{-x_2} - x_2 e^{-y} \Big|_0^y \\ &= 1 - e^{-y} - ye^{-y} \quad \text{untuk } y > 0 \end{aligned}$$

dan

$$\begin{aligned} f_Y(y) &= \frac{d}{dy} F_Y(y) \\ &= ye^{-y} \quad \text{untuk } y > 0 \end{aligned}$$

Teknik Transformasi

Pada sub bab ini akan dibahas sebaran fungsi peubah acak dengan menggunakan teknik transformasi. Pertama akan diawali dengan peubah acak diskrit.

Misalkan X adalah peubah acak diskrit dengan fungsi kepekatan peluang $f_X(x)$ dan $Y = u(X)$ adalah fungsi satu-satu. Maka, fungsi kepekatan peluang Y adalah $f_Y(y) = f_X(w(y))$ untuk $y \in \{y \mid f_Y(y) > 0\}$.

Teladan 4. 5.

Misalkan X memiliki sebaran Geometrik (p), sehingga $f_X(x) = pq^{x-1}$ untuk x = 1,2,3,... Jika $Y = X-1$, maka $u(x)=x-1$, $w(y)=y+1$ dan $f_Y(y) = f_X(y+1) = pq^y$ untuk y = 0,1,2,...

Teorema 4. 2.

Misalkan X adalah peubah acak kontinu dengan fungsi kepekatan peluang $f_X(x)$. Jika $A= \{x;\ f_X(x)>0\}$. Asumsikan bahwa : (i) $y = g(x)$ mendefinisikan suatu transformasi 1-1 dari A ke $B= \{y;\ f_Y(y)>0\}$. (ii) Turunan dari $x=g^{-1}(y)$ terhadap y kontinu dan tidak nol untuk $y \in Y$, dimana $g^{-1}(y)$ adalah fungsi kebalikan dari $g(x)$; yaitu, $g^{-1}(y)$ adalah nilai x yang membuat $g(x)=y$. Maka $Y = g(X)$ adalah peubah acak kontinu dengan fungsi kepekatan peluang

$$f_Y(y) = \left|\frac{d}{dy} g^{-1}(y)\right| f_X(g^{-1}(y)) I_Y(y)$$

Teladan 4. 6.

Misalkan X memiliki sebaran Beta. Jika $Y = -\ln X$, apakah sebaran peubah acak Y ? A= $\{x; f_X(x)>0\}=\{x; 0<x<1\}$. Fungsi $y = g(x) = -\ln X$ mendefinisikan transformasi fungsi 1-1 dari A ke B=$\{y; y>0\}$. Dengan demikian $x = g^{-1}(y) = e^{-y}$ dan $(d/dy)g^{-1}(y) = -e^{-y}$ yang kontinu dan tidak nol untuk . Dengan menggunakan teorema diatas,

$$f_Y(y) = \left|\frac{d}{dy}g^{-1}(y)\right| f_X(g^{-1}(y))I_\Psi(y)$$

$$= e^{-y}\frac{1}{B(a,b)}(e^{-y})^{a-1}(1-e^{-y})^{b-1}I_{(0,\infty)}(y)$$

$$= \frac{1}{B(a,b)}e^{-ay}(1-e^{-y})^{b-1}I_{(0,\infty)}(y)$$

Teladan 4. 7.

Misalkan X memiliki fungsi kepekatan peluang Pareto atau $f_X(x) = \theta x^{-\theta-1}I_{[1,\infty)}(x)$ dan ingin dicari sebaran dari $Y = \ln X$. Dengan menggunakan teorema diatas,

$$f_Y(y) = \left|\frac{d}{dy}g^{-1}(y)\right| f_X(g^{-1}(y))I_\Psi(y)$$

$$= e^{y}\theta(e^{y})^{-\theta-1}I_{[1,\infty)}(e^{y})$$

$$= \theta e^{-\theta y}I_{[0,\infty)}(y)$$

Teladan 4. 8.

Misalkan $F_X(x) = 1-e^{-2x}$, untuk $0<x<\infty$ dan misalkan juga $Y=e^X$. Dengan demikian $x = w(y) = \ln y$, dan $w'(y) = 1/y$, sehingga

$$f_Y(y) = \left|\frac{d}{dy}g^{-1}(y)\right| f_X(g^{-1}(y))I_\Psi(y)$$

$$= \left|\frac{1}{y}\right| f_X(\ln y)I_{(1,\infty)}(y)$$

$$= 2e^{-2\ln y}\frac{1}{y}I_{(1,\infty)}(y)$$

$$= 2y^{-3}I_{(1,\infty)}(y)$$

Jika transformasi bukan merupakan fungsi 1-1, teorema diatas tidak bisa langsung dipakai. Cara menyiasatinya adalah dengan jalan melakukan partisi daerah penyangga (daerah fungsi kepekatan peluang) menjadi beberapa gugus saling menenggang atau saling lepas sedemikian rupa sehingga pada masing-masing partisi transformasi sudah merupakan fungsi 1-1 sehingga teorema diatas dapat digunakan kembali. Jika $x = g_i^{-1}(y)$ melambangkan kebalikan dari $y = g(x)$ untuk $x \in \Xi_i$. Maka fungsi kepekatan peluang dari *Y=g(X)* adalah

$$f_Y(y) = \sum_i \left| \frac{d}{dy} g_i^{-1}(y) \right| f_X(g_i^{-1}(y)) I_\Psi(y)$$

dimana penjumlahan untuk seluruh nilai *i* dimana *g(x)* = *y* untuk beberapa nilai *x* di dalam Ξ_i.

Teladan 4. 9.

Misalkan *X* adalah peubah acak kontinu dengan fungsi kepekatan peluang $f_X(x)$ dan misalkan juga *Y* = *g(X)* = X^2. Sebagai catatan jika Ξ mencakup interval atau titik-titik negatif dan positif, maka fungsi *y* = *g(x)* = x^2 bukan merupakan fungsi 1-1. Namun demikian, jika Ξ dipisahkan menjadi $\Xi_1 = \{x; x \in \Xi, x < 0\}$ dan $\Xi_2 = \{x; x \in \Xi, x > 0\}$, maka *y* = *g(x)* mendefinisikan fungsi 1-1 pada tiap Ξ_i. Sebagai misal $g_1^{-1}(y) = -\sqrt{y}$ dan $g_2^{-1}(y) = \sqrt{y}$. Maka fungsi kepekatan peluang dari *Y=g(X)* adalah

$$f_Y(y) = \sum_i \left| \frac{d}{dy} g_i^{-1}(y) \right| f_X(g_i^{-1}(y)) I_\Psi(y)$$

$$= \left[\frac{1}{2} \frac{1}{\sqrt{y}} f_X(-\sqrt{y}) + \frac{1}{2} \frac{1}{\sqrt{y}} f_X(\sqrt{y}) \right] I_{(0,\infty)}(y)$$

Sebagai contoh khusus, apabila $f_X(x) = \left(\frac{1}{2}\right)e^{-|x|}$, maka

$$f_Y(y) = \frac{1}{2}\frac{1}{\sqrt{y}}e^{-\sqrt{y}}I_{(0,\infty)}(y)$$

Atau jika $f_X(x) = \frac{2}{9}(x+1)I_{(-1,2)}(x)$, maka

$$f_Y(y) = \left[\frac{1}{2}\frac{1}{\sqrt{y}}\frac{2}{9}(-\sqrt{y}+1) + \frac{1}{2}\frac{1}{\sqrt{y}}\frac{2}{9}(\sqrt{y}+1)\right]I_{(0,1)}(y)$$
$$+\left[\frac{1}{2}\frac{1}{\sqrt{y}}\frac{2}{9}(\sqrt{y}+1)\right]I_{[1,4)}(y)$$

Teorema 4. 3.

Probability Integral Transformation

Jika X adalah peubah acak dengan fungsi sebaran kumulatif $F_X(x)$, maka $U = F_X(X)$ memiliki distribusi Seragam (0,1). Sebagai kebalikannya, jika U memiliki distribusi Seragam (0,1), maka $X = F_X^{-1}(U)$ memiliki fungsi sebaran kumulatif $F_X(x)$.

$P[U \leq u] = P[F_X(X) \leq u] = P[X \leq F_X^{-1}(u)] = F_X(F_X^{-1}(u)) = u$ untuk 0 < u < 1.

Sebaliknya

$$P[X \leq x] = P[F_X^{-1}(U) \leq x] = P[U \leq F_X(x)] = F_X(x)$$

Teorema diatas sering disebut dengan Transformasi Peluang Integral atau Transformasi Peluang Menyeluruh (*Probability Integral Transform*).

Transformasi Bersama

Perhatikan peubah acak *X* dan *Y* yang saling bebas dan menyebar masing-masing menurut sebaran Geometrik (p). Fungsi kepekatan peluang dari *X* dan *T* = *X*+*Y* dapat dituliskan dalam bentuk

$$f_{X,T}(x,t) = f_{X,Y}(x,t-x)$$

dimana *x* dan *t-x* adalah jawaban dari transformasi bersama $x=u_1(x,y)$ dan $t=u_2(x,y)=x+y$.

Hal ini dapat secara umum dijelaskan seperti berikut. Misalkan $\underline{\mathbf{X}}=(X_1,X_2,...,X_k)$ merupakan vektor peubah acak berdimensi-k, dan misalkan $u_1(\underline{x}), u_2(\underline{x}), ..., u_k(\underline{x})$ adalah fungsi sebanyak k dari $\underline{x}$, sehingga $Y_i=u_i(\underline{\mathbf{X}})$ untuk *i* = 1,2,...,k mendefinisikan vektor peubah acak $\underline{\mathbf{Y}}=(Y_1,Y_2,...,Y_k)$. Secara ringkas dapat dituliskan $\underline{\mathbf{Y}}=\boldsymbol{u}(\underline{\mathbf{X}})$.

Teorema 4. 4.

Jika $\underline{\mathbf{X}}$ adalah vektor peubah acak diskret dengan fungsi kepekatan peluang bersama $f_{\mathbf{X}}(\underline{\mathbf{x}})$ dan $\underline{\mathbf{Y}}=\boldsymbol{u}(\underline{\mathbf{X}})$ mendefinisikan transformasi 1-1, maka fungsi kepekatan peluang bersama $\underline{\mathbf{Y}}$ adalah

$$f_{\underline{Y}}(y_1,y_2,...,y_k) = f_{\underline{X}}(x_1,x_2,...,x_k)$$

dimana $x_1, x_2, ..., x_k$ adalah jawaban dari $\underline{\mathbf{y}} = \boldsymbol{u}(\underline{\mathbf{x}})$ dan sebagai konsekuensinya tergantung pada $y_1, y_2, ..., y_k$.

Jika transformasi tidak 1-1, dan jika dapat dilakukan partisi, sehingga memungkinkan transformasi 1-1 pada tiap partisi, misalkan partisi tersebut adalah A_1, A_2, ... sehingga persamaan $\underline{\mathbf{y}} = \boldsymbol{u}(\underline{\mathbf{x}})$ memiliki jawaban yang khas $\underline{\mathbf{x}} = \underline{\mathbf{x}}_j = (x_{1j}, x_{2j}, ..., x_{kj})$ pada A_j, maka fungsi kepekatan peluang bersama $\underline{\mathbf{Y}}$ adalah

$$f_{\underline{Y}}(y_1,y_2,...,y_k) = \sum_j f_{\underline{X}}(x_{1j},x_{2j},...,x_{kj})$$

Untuk transformasi k peubah acak $\underline{\mathbf{y}} = \boldsymbol{u}(\underline{\mathbf{x}})$ dengan jawaban yang khas $\underline{\mathbf{x}} = \underline{\mathbf{x}}_j = (x_{1j}, x_{2j}, ..., x_{kj})$, **Jacobian** dari transformasi tersebut adalah matriks berukuran k x k yang bernilai turunan parsial seperti berikut

$$J = \begin{vmatrix} \dfrac{\partial x_1}{\partial y_1} & \dfrac{\partial x_1}{\partial y_2} & \cdots & \dfrac{\partial x_1}{\partial y_k} \\ \dfrac{\partial x_2}{\partial y_1} & \dfrac{\partial x_2}{\partial y_2} & \cdots & \dfrac{\partial x_2}{\partial y_k} \\ \cdots & \cdots & \cdots & \cdots \\ \dfrac{\partial x_k}{\partial y_1} & \dfrac{\partial x_k}{\partial y_2} & \cdots & \dfrac{\partial x_k}{\partial y_k} \end{vmatrix}$$

Teorema 4. 5.

Misalkan $\underline{\mathbf{X}}=(X_1,X_2,...,X_k)$ merupakan vektor peubah acak berdimensi-k dengan fungsi kepekatan peluang bersama $f_{\underline{\mathbf{X}}}(x_1, x_2, ..., x_k) > 0$ pada A, dan $\underline{\mathbf{Y}}=(Y_1,Y_2,...,Y_k)$ didefinisikan dengan transformasi 1-1 $Y_i=u_i(\underline{\mathbf{X}})$ untuk i = 1,2,...,k. Jika Jacobian kontinu dan tidak nol pada wilayah transformasi, maka fungsi kepekatan peluang bersama Y adalah

$$f_{\underline{Y}}(y_1, y_2, ..., y_k) = f_{\underline{X}}(x_1, x_2, ..., x_k)|J|$$

dimana $\underline{\mathbf{x}} = (x_1, x_2, ..., x_k)$ adalah jawaban dari $\underline{\mathbf{y}} = \mathbf{u}(\underline{\mathbf{x}})$.

Teladan 4. 10.

Misalkan X_1 dan X_2 menyebar saling bebas menurut sebaran Eksponensial (1), maka

$$f_{X_1,X_2}(x_1, x_2) = e^{-(x_1+x_2)} I_{(0,\infty)}(x_1) I_{(0,\infty)}(x_2)$$

Misalkan $Y_1 = X_1$ dan $Y_2 = X_1 + X_2$. Dengan demikian $y_1 = x_1$ dan $y_2 = x_1+x_2$ yang memiliki jawaban unik $x_1 = y_1$ dan $x_2 = y_2 - y_1$, sehingga menghasilkan Jacobian

$$J = \begin{vmatrix} \dfrac{\partial x_1}{\partial y_1} & \dfrac{\partial x_1}{\partial y_2} \\ \dfrac{\partial x_2}{\partial y_1} & \dfrac{\partial x_2}{\partial y_2} \end{vmatrix} = \begin{vmatrix} 1 & 0 \\ -1 & 1 \end{vmatrix} = 1$$

sehingga kita peroleh

$$f_{Y_1,Y_2}(y_1, y_2) = f_{X_1,X_2}(y_1, y_2 - y_1) I_{(0,y_2)}(y_1) I_{(y_1,\infty)}(y_2) = e^{-y_2} I_{(0,y_2)}(y_1) I_{(y_1,\infty)}(y_2)$$

Fungsi kepekatan peluang marjinal Y_1 dan Y_2 dapat dicari seperti berikut

$$f_{Y_1}(y_1) = \int_{y_1}^{\infty} e^{-y_2} dy_2 = e^{-y_1} I_{(0,\infty)}(y_1)$$

dan

$$f_{Y_2}(y_2) = \int_{0}^{y_2} e^{-y_2} dy_1 = y_2 e^{-y_2} I_{(0,\infty)}(y_2)$$

Teladan 4. 11.

Misalkan X_1 dan X_2 menyebar saling bebas menurut sebaran Eksponensial (1), maka

$$f_{X_1,X_2}(x_1, x_2) = e^{-(x_1+x_2)} I_{(0,\infty)}(x_1) I_{(0,\infty)}(x_2)$$

Misalkan $Y_1 = X_1 - X_2$ dan $Y_2 = X_1 + X_2$. Dengan demikian $y_1 = x_1 - x_2$ dan $y_2 = x_1 + x_2$ yang memiliki jawaban unik $x_1 = (y_2 + y_1)/2$ dan $x_2 = (y_2 - y_1)/2$, sehingga menghasilkan Jacobian

$$J = \begin{vmatrix} \dfrac{\partial x_1}{\partial y_1} & \dfrac{\partial x_1}{\partial y_2} \\ \dfrac{\partial x_2}{\partial y_1} & \dfrac{\partial x_2}{\partial y_2} \end{vmatrix} = \begin{vmatrix} \dfrac{1}{2} & \dfrac{1}{2} \\ -\dfrac{1}{2} & \dfrac{1}{2} \end{vmatrix} = \frac{1}{2}$$

sehingga kita peroleh

$$\begin{aligned} f_{Y_1,Y_2}(y_1, y_2) &= f_{X_1,X_2}((y_1 + y_2)/2, (y_2 - y_1)/2) \,|\, J \,|\, I_{(-y_2,y_2)}(y_1) I_{(0_1,\infty)}(y_2) \\ &= \frac{1}{2} e^{-y_2} I_{(-y_2,y_2)}(y_1) I_{(0_1,\infty)}(y_2) \end{aligned}$$

Fungsi kepekatan peluang marjinal Y_1 dan Y_2 dapat dicari seperti berikut

$$f_{Y_1}(y_1) = \int_{-y_1}^{\infty} \frac{1}{2} e^{-y_2} dy_2 = e^{y_1} I_{(-\infty,0)}(y_1)$$

$$f_{Y_1}(y_1) = \int_{y_1}^{\infty} \frac{1}{2} e^{-y_2} dy_2 = e^{-y_1} I_{(0,\infty)}(y_1)$$

atau bila digabung

$$f_{Y_1}(y_1) = e^{-|y_1|} I_{(-\infty,\infty)}(y_1)$$

dan

$$f_{Y_2}(y_2) = \int_{-y_2}^{y_2} \frac{1}{2} e^{-y_2} dy_1 = y_2 e^{-y_2} I_{(0,\infty)}(y_2)$$

Formula Konvolusi

Jika kita hanya tertarik fungsi kepekatan peluang $S = X_1 + X_2$, dimana X_1 dan X_2 merupakan peubah acak peubah acak kontinu dengan fungsi kepekatan peluang bersama $f(x_1,x_2)$, maka formula umum dapat diturunkan dengan menggunakan pendekatan seperti

$$f_S(s) = \int_{-\infty}^{\infty} f(t, s-t) dt$$

Jika X_1 dan X_2 peubah acak yang saling bebas, maka biasanya menggunakan formula konvolusi.

$$f_S(s) = \int_{-\infty}^{\infty} f_1(t) f_2(s-t) dt$$

Diperlukan latihan yang cermat, khususnya dalam penentuan batas-batas pengintegralan.

Teladan 4. 12.

Misalkan X_1 dan X_2 merupakan peubah acak yang saling bebas dari sebaran seragam, X_i ~ Seragam(0,1), dan misalkan $S = X_1 + X_2$. Carilah fungsi kepekatan peluang dari S !

Untuk itu perlu kita misalkan $t = x_1$ dan $s = x_1+x_2$. Bila A = $\{(x_1,x_2)|\ 0 < x_1 < 1,\ 0 < x_2 < 1\}$ merupakan daerah dimana fungsi kepekatan peluang bersama X_1 dan X_2 lebih besar dari nol. Melalui transformasi tersebut, maka daerah fungsi kepekatan peluang bersama antara S dan T adalah $\{(s,t)|\ 0 < t < s < t+1 < 2\}$

Sehingga,

$$
\begin{aligned}
f_S(s) &= \int_0^s dt = s && 0 < s < 1 \\
&= \int_{s-1}^1 dt = 2 - s && 1 \leq s < 2 \\
&= 0 && s \text{ selainnya}
\end{aligned}
$$

atau hasilnya dapat disingkat seperti berikut

$$
\begin{aligned}
f_S(s) &= 1 - |s-1| && 0 < s < 2 \\
&= 0 && s \text{ selainnya.}
\end{aligned}
$$

Sebaran Statistik Tataan

Konsep contoh acak berukuran n telah kita bahas sebelum bagian ini, dan fungsi kepekatan bersama yang berhubungan dengan n peubah acak yang saling bebas $X_1, X_2, ..., X_n$ adalah

$$f(x_1, x_2, ..., x_n) = f(x_1)f(x_2)...f(x_n)$$

Sebagai teladan misalnya, suatu contoh acak lima bola lampu yang diukur umurnya, diperoleh data seperti berikut [dalam bulan] $(x_1,\ x_2,\ x_3,\ x_4,\ x_5)$ = (4, 11, 17, 5, 19). Nilai pengamatan sebetulnya adalah $x_1 = 4$, $x_4 = 5$, $x_2 = 11$, $x_3 = 17$, dan $x_5 = 19$.

Kadang, cara penulisan data dari yang terkecil hingga terbesar sangat berguna dalam suatu hal. Kita dapat tuliskan $x_{(1)} = 5$, $x_{(2)} = 5$, $x_{(3)} = 11$, $x_{(4)} = 17$, dan $x_{(5)} = 19$. Seseorang bisa saja tak begitu peduli bola lampu mana yang akan diberi label nomor 1, label nomor 2, dan seterusnya. Karenanya, kita bisa menuliskannya sesuai urutan tanpa melihat nomor label. Dalam beberapa kasus, kita dapat menghentikan hingga pengamatan tertata ke *r* dari sebanyak *n* pengamatan. Penghentian percobaan hingga lampu mati ke *r* dari sebanyak yang diamati sering dikenal dengan Penarikan Contoh Tersensor Tipe II. (*Type II Censored Sampling*).

Sebaran bersama peubah tertata tidak sama dengan sebaran bersama dari peubah acak tak tertata. Dari teladan diatas, terdapat 120 = 5! permutasi contoh acak berukuran 5 yang hanya akan berkorespondensi hanya dengan satu nilai tataan.

Suatu transformasi yang akan mengurutkan nilai nilai $x_1, x_2, ..., x_n$. Sebagai teladannya:

$$y_1 = u_1(x_1, x_2, ..., x_n) = \min(x_1, x_2, ..., x_n)$$

$$y_n = u_n(x_1, x_2, ..., x_n) = \max(x_1, x_2, ..., x_n)$$

Dan secara umum, untuk $i = 1, 2, ..., n$. $y_i = u_i(x_1, x_2, ..., x_n)$ menggambarkan nilai terkecil-*i* dari $x_1, x_2, ..., x_n$.

Teorema 4. 6.

Jika $X_1, X_2, \ldots, X_n$ adalah contoh acak dari suatu populasi dengan fungsi kepekatan peluang kontinu $f(x)$, maka fungsi kepekatan peluang bersama dari statistik tataan $Y_1, Y_2, \ldots, Y_n$ adalah

$$g(y_1, y_2, ..., y_n) = n!\, f(y_1) f(y_2) \ldots f(y_n)$$

dimana $y_1 < y_2 < \ldots < y_n$, dan nol untuk selainnya.

Ini merupakan teladan dari suatu transformasi peubah acak kontinu yang tidak bersifat 1 – 1.

Teladan 4. 13.

Misalkan *X1, X2,* dan *X3* merupakan suatu contoh acak berukutan 3 dari suatu populasi dengan fungsi kepekatan peluang

$$f(x) = 2x \quad 0 < x < 1$$

Dengan demikian fungsi kepekatan peluang bersama statistik tataan Y_1, Y_2 dan Y_3 adalah

$$g(y_1, y_2, y_3) = 3!(2y_1)(2y_2)(2y_3) = 48y_1y_2y_3 \qquad 0 < y_1 < y_2 < y_3 < 1$$

dan nol selainnya.

Fungsi kepekatan peluang peubah acak marginal juga bisa diperoleh dengan mengintegralkan seluruh peubah lainnya. Misalkan untuk peubah acak statistik terkecil atau statistik minimum $X_{(1)} = Y_1$ memiliki fungsi kepekatan peluang yang dapat dicari seperti berikut:

$$\begin{aligned}
g_1(y_1) &= \int_{y_1}^{1}\int_{y_2}^{1} 48y_1y_2y_3dy_3dy_2 \\
&= \int_{y_1}^{1} (24y_1y_2y_3^2 \,|_{y_2}^{1}\; dy_2 \\
&= \int_{y_1}^{1} 24y_1y_2(1-y_2^2)dy_2 \\
&= \int_{y_1}^{1} 24y_1(y_2 - y_2^3)dy_2 \\
&= \int_{y_1}^{1} (24y_1y_2 - 24y_1y_2^3)dy_2 \\
&= (12y_1y_2^2 - 6y_1y_2^4 \,|_{y_1}^{1} \\
&= 6y_1 - 12y_1^3 + 6y_1^5 \\
&= 6y_1(1-y_1^2)^2 \qquad 0 < y_1 < 1
\end{aligned}$$

Untuk melihat peluang nilai terkecil berada pada kisaran nilai tertentu, dapat digunakan fungsi kepekatan peluang diatas. Sebagai misal, untuk mengetahui seberapa besar peluang statistik terkecil itu nilainya kurang dari 0,1 diperoleh dengan cara seperti berikut:

$$P[Y_1 < 0{,}1] = \int_0^{0{,}1} 6y_1(1-y_1^2)^2\,dy_1 = 0{,}271$$

Teorema 4. 7.

Misalkan *X1, X2, …, Xn* melambangkan suatu contoh acak berukuran *n* dari suatu fungsi kepekatan peluang kontinu *f(x)*, dimana *f(x)>0* untuk *a* < *x* < *b*. Oleh karenanya, fungsi kepekatan peluang statistik tataan ke *k* yang dinotasikan dengan $X_{(k)} = Y_k$ adalah

$$g_k(y_k) = \frac{n!}{(k-1)!(n-k)!}\left[F(y_k)\right]^{k-1}\left[1-F(y_k)\right]^{n-k} f(y_k)$$

Jika $a < y_k < b$, dan nol selainnya.

Untuk itu, dengan mudah kita bisa memperoleh kasus khusus apabila *k* = *1* atau *k* = *n*.

$$g_1(y_1) = n\left[1-F(y_1)\right]^{n-1} f(y_1) \qquad a < y_1 < b$$

$$g_n(y_n) = n\left[F(y_n)\right]^{n-1} f(y_n) \qquad a < y_n < b$$

Baik untuk peubah acak kontinu ataupun diskrit, fungsi sebaran kumulatif statistik minimum dan maksimum dapat dicari langsung dengan menggunakan teknik fungsi sebaran.

$$\begin{aligned} G_1(y_1) &= P\left[Y_1 \le y_1\right] \\ &= 1 - P\left[Y_1 > y_1\right] \\ &= 1 - P\left[semua\ X_i > y_1\right] \\ &= 1 - \left[1 - F(y_1)\right]^n \end{aligned}$$

Untuk statistik maksimumnya,

$$G_n(y_n) = P[Y_n \le y_n]$$
$$= P[semua\ X_i \le y_n]$$
$$= [F(y_n)]^n$$

Argumen yang sama dimungkinkan untuk mengekspresikan fungsi sebaran kumulatif dari statistik tataan ke-*k*. Dalam hal ini $Y_k \le y_k$ jika terdapat *k* atau lebih X_i bernilai tidak lebih dari y_k, dimana banyaknya X_i yang tidak melebihi y_k mengikuti sebaran binomial dengan parameter *n* dan $p = F(y_k)$.

Teorema 4. 8.

Untuk suatu peubah acak berukuran *n* dari fungsi sebaran kumulatif diskrit atau kontinu, *F(x)*, fungsi sebaran kumulatif marjinal dari statistik tataan ke-*k* adalah

$$G_k(y_k) = \sum_{j=k}^{n} C_j^n [F(y_k)]^j [1 - F(y_k)]^{n-j}$$

Argumen yang sama dengan mudah dapat digunakan untuk memperoleh fungsi kepekatan peluang bersama dari gugus statistik tataan. Statistik tataan $X_{(i)}=Y_i$ dan $X_{(j)}=Y_j$ memiliki fungsi kepekatan peluang bersama

$$g_{ij}(y_i, y_j) = \frac{n!}{(i-1)!(j-i-1)!(n-j)!}[F(y_i)]^{i-1} f(y_i)$$
$$\times [F(y_j) - F(y_i)]^{j-i-1} [1 - F(y_j)]^{n-j} f(y_j)$$

jika $a < y_i < y_j < b$, dan nol selainnya.

Teladan 4. 14.

Suatu contoh acak berukuran *n* dengan fungsi kepekatan dan fungsi sebaran peluang berturut-turut $f(x) = 2x$ dan $F(x) = x^2$ untuk *0*

< *x* < *1.* Fungsi kepekatan peluang range atau jangkauan contoh $R = Y_n - Y_1$ adalah

$$g(y_1, y_n) = \frac{n!}{(n-2)!}(2y_1)\left[y_n^2 - y_1^2\right]^{n-2}(2y_n) \qquad 0 < y_1 < y_n < 1$$

Dengan transformasi $R=Y_n-Y_1$ dan $S=Y_1$ akan menghasilkan invers transformasi $y_1=s$ dan $y_n = r+s$, dan $|J| = 1$. Dengan demikian, fungsi kepekatan peluang bersama antara *R* dan *S* adalah

$$h(r,s) = \frac{n!}{(n-2)!}(2s)\left[(r+s)^2 - s^2\right]^{n-2}(2(r+s))$$

$$= \frac{4n!}{(n-2)!}s(r+s)\left[r^2 + 2rs - s^2 + s^2\right]^{n-2}$$

$$= \frac{4n!}{(n-2)!}s(r+s)\left[r^2 + 2rs\right]^{n-2}$$

untuk $0 < s < 1-r, 0 < r < 1$

Fungsi kepekatan peluang dari Range (R) adalah

$$h_1(r) = \int_0^{1-r} h(r,s)ds$$

Apabila *n* = 2 akan kita dapatkan

$$h_1(r) = \int_0^{1-r} 8s(r+s)ds = (4/3)(r+2)(1-r)^2 \qquad 0 < r < 1$$

Ekspresi fungsi sebaran marginal dari R dapat diperoleh sebagai berikut

$$H_1(r) = \int_{-\infty}^{r}\int_{-\infty}^{+\infty} g(s, r+s)dsdr$$

$$= \int_{-\infty}^{+\infty}\int_{-\infty}^{r} \frac{n!}{(n-2)!} f(s)\left[F(r+s) - F(s)\right]^{n-2} f(r+s)drds$$

$$= \int_{-\infty}^{+\infty} nf(s)\left[F(r+s)-F(s)\right]^{n-1} ds$$

Teladan 4. 15.

Dalam teladan sebelumnya, dimana *F(s+r) = 1* jika *s > 1-r*, akan didapatkan

$$H_1(r) = \int_0^{1-r} n(2s)\left[(r+s)^2 - s^2\right]^{n-1} ds + \int_{1-r}^{1} n(2s)\left[1-s^2\right]^{n-1} ds$$

Untuk *n* = 2

$$H_1(r) = \int_0^{1-r} 4s(r^2+2rs)ds + \int_{1-r}^{1} 4s(1-s^2)ds$$

$$= \frac{8r}{3} - 2r^2 + \frac{r^4}{3}$$

yang ternyata konsisten dengan fungsi kepekatan peluangnya $h_1(r)$.

Penarikan Contoh Tersensor

Dalam percobaan pengujian umur suatu komponen elektronik misalnya, pengamatan tertata dapat terjadi secara alami. Bila hal ini terjadi, cara untuk menghemat waktu dan biaya percobaan adalah hanya dengan mengamati sebanyak *r* komponen pertama yang gagal berfungsi daripada harus mengamati keseluruhan *n* komponen. Hal seperti ini sering disebut dengan penarikan contoh tersensor tipe II.

Teorema 4. 9.

Fungsi kepekatan peluang marjinal bersama dari *r* statistik tataan pertama yang diperoleh dari contoh acak berukuran *n* dari suatu populasi dengan fungsi kepekatan peluang kontinu *f(x)* adalah

$$g(y_1,...,y_r) = \frac{n!}{(n-r)!}\left[1-F(y_r)\right]^{n-r}\prod_{i=1}^{r} f(y_i)$$

jika $-\infty < y_1 < ... < y_r < +\infty$ dan nol selainnya.

Dalam **penarikan contoh tersensor tipe II**, *r* merupakan nilai yang tetap atau sudah ditentukan, sedangkan peubah acaknya adalah lamanya percobaan Y_r.

Jika penghentian percobaan atau pensensoran dilakukan pada waktu t_0 setelah percobaan dimulai, maka prosedur ini dikenal dengan **penarikan contoh tersensor tipe I**. Dengan demikian banyaknya pengamatan (*R*) merupakan peubah acak. Jika peluang sebuah komponen akan gagal berfungsi sebelum waktu penyensoran untuk sembarang tindakan adalah $p = F(t_0)$, maka untuk contoh acak berukuran *n* peubah acak *R* berdistribusi binomial dengan parameter *n* dan $F(t_0)$.

Teorema 4. 10.

Jika $Y_1, \ldots, Y_r$ merupakan nilai amatan suatu contoh acak berukuran *n* dari *f(x)* yang tersensor dengan tipe I disebelah kanan pada t_0, maka fungsi kepekatan peluang bersama $Y_1, ..., Y_R$ adalah

$$f_{Y_1,...,Y_R}(y_1,...,y_R) = \frac{n!}{(n-R)!}\left[1-F(t_0)\right]^{n-R}\prod_{i=1}^{R} f(y_i)$$

jika $y_1 < ... < y_R < t_0$ dan $R = 1, 2, ..., n$ serta

$$P[R=0] = [1-F(t_0)]^n$$

Latihan

1. Misalkan X dan Y peubah acak kontinu yang saling bebas identik dengan distribusi Seragam (0,1). Tentukan distribusi dari
   a. X+Y
   b. X-Y
   c. XY
   d. X/Y
   e. X^2
2. Jari-jari sebuah lingkaran R memiliki fungsi kepekatan peluang $f(r) = 6r(1-r)$ untuk *0*<*r*<*1*.
   a. Carilah fungsi kepekatan peluang dari keliling lingkaran.
   b. Carilah fungsi kepekatan peluang dari luas lingkaran.
3. Misalkan X_1 dan X_2 adalah contoh acak berukuran 2 dari distribusi Gamma(*θ*=*2, κ*=*1/2*), tentukan
   a. Fungsi kepekatan peluang dari $Y = \sqrt{X_1 + X_2}$
   b. Fungsi kepekatan peluang dari $W = X_1 / X_2$

4. Misalkan X adalah peubah acak kontinu yang menyebar Seragam (0,1). Dengan menggunakan teknik fungsi sebaran kumulatif, tentukan fungsi kepekatan peluang peubah acak berikut ini berikut ini
   a. $Z = \ln X$
   b. $U = X(X-1)$
   c. $V = e^{-X}$

5. Suatu contoh acak berukuran n yang diambil dari distribusi Geometrik dengan parameter *p*. Tentukan fungsi sebaran kumulatif dari
   a. $Y_{(1)} = \min(X_1, X_2, ..., X_n)$
   b. $Y_{(k)} = terkecil\ ke-k$

c. $Y_{(n)} = \max(X_1, X_2, ..., X_n)$

6. Misalkan X_1 dan X_2 adalah peubah acak kontinu yang saling bebas dan identik menyebar Seragam (0,1). Bila $Y=\min(X_1,X_2)$ dan $Z=\max(X_1,X_2)$, Tentukan
   a. *F(y,z)*
   b. *f(y,z)*
   c. *f(y)*
   d. E(Z|Y)
   e. E(sin(YZ)|Y=y)

Sifat-sifat Peubah Acak

Pendahuluan

Cara mengekspresikan model matematis dari suatu fenomena fisik yang bersifat probabilistik adalah dengan menggunakan peubah acak dan sebaran distribusinya. Sebagaimana telah kita ketahui bahwa peubah acak mungkin dapat berasosiasi dengan beberapa karakteristik numerik dari populasi sesungguhnya ataupun konseptual, dan fungsi kepekatan peluangnya merepresentasikan sebaran populasi untuk semua kemungkinan nilai karakteristik tersebut.

Sangat sering kita jumpai bahwa fungsi kepakatan sesungguhnya dari suatu karakteristik populasi tidak diketahui. Satu cara pendekatan model dalam hal ini adalah dengan mempertimbangkan suatu famili fungsi kepekatan peluang yang diindeks dengan parameter yang tak diketahui.

Penekanan utama dalam statistika adalah mengembangkan atau memperoleh nilai estimasi dari parameter yang tak diketahui berdasarkan data contoh (sampel). Dalam beberapa hal, suatu parameter dapat merepresentasikan suatu kuantitas fisik yang bermakna, seperti rata-rata atau mean populasi.

Dengan demikian, akan sangat bermanfaat apabila perlu dipelajari berbagai sifat yang dimiliki atau properti dari berbagai peubah acak untuk merepresentasikan dan menginterpretasikan populasi awal, dan juga berguna dalam estimasi (pendugaan) atau pemilihan model yang sesuai.

Sifat-sifat Nilai Harapan

Pada sub bab ini akan disajikan beberapa sifat nilai harapan (nilai ekspektasi).

Teorema 5. 1.

Jika $\boldsymbol{X}$ = (X_1, ..., X_k) memiliki fungsi kepekatan peluang bersama $f(x_1, ..., x_k)$, dan jika $\boldsymbol{Y}$ = u(X_1, ..., X_k) merupakan fungsi dari $\boldsymbol{X}$, maka $E(\boldsymbol{Y})=E_X[u(X_1,...,X_k)]$, dimana

$$E_X\left[u(X_1,...,X_k)\right]=\sum_{x_1}...\sum_{x_k}u(x_1,...,x_k)f(x_1,...,x_k)$$

jika $\boldsymbol{X}$ diskrit, dan

$$E_X\left[u(X_1,...,X_k)\right]=\int_{-\infty}^{+\infty}...\int_{-\infty}^{+\infty}u(x_1,...,x_k)f(x_1,...,x_k)dx_1...dx_k$$

jika $\boldsymbol{X}$ kontinu.

Teladan 5. 1.

Misalkan peubah acak diskrit X dengan fungsi kepekatan peluang seperti berikut:

x	-2	0	2	3
$f_X(x)$	1/8	1/8	1/4	½

Dan misalkan $Y = X^2$.

Salah satu cara penyelesaiannnya adalah dengan mencari nilai-nilai Y dan fungsi kepekatan peluang dari Y, yaitu:

y	0	4	9
$f_Y(y)$	1/8	3/8	½

Nilai harapan dari Y dicari dengan

$$E(Y)=\sum_y yf_Y(y)=(0)(1/8)+(4)(3/8)+(9)(1/2)=6$$

Yang juga dapat dicari dengan

$$E(X^2)=(-2)^2(1/8)+(0)^2(1/8)+(2)^2(1/4)+(3)^2(1/2)=6$$

Teladan 5. 2.

Dalam suatu permainan lempar dadu, tiga buah dadu bermuka 6 seimbang dilemparkan sekaligus. Seorang pemain akan membayar

Rp. 1000 per permainan, dan akan memperoleh Rp. 1000 untuk setiap mata 6 yang keluar. Dengan demikian, seorang pemain akan memiliki fungsi kemungkinan perolehan Rp. 1000, Rp 2000, Rp 3000 atau –Rp. 1000. sesuai dengan keluar banyaknya mata 6 dari ketiga dadu yang dilempar sekaligus tersebut, yaitu 1, 2, 3, atau 0. dengan demikian *Y=u(X)*, dimana

x	0	1	2	3
u(x)	-1000	1000	2000	3000

Peubah acak *X* yang merupakan banyaknya mata dadu 6 keluar, memiliki distribusi Binomial dengan parameter n = 3 dan p = 1/6. Dengan demikian,

$$E(Y) = E_X[u(X)]$$

$$= \sum_{x=0}^{3} u(x) C_x^3 \left(\frac{1}{6}\right)^x \left(\frac{5}{6}\right)^{3-x}$$

$$= -1000\left(\frac{125}{216}\right) + 1000\left(\frac{75}{216}\right) + 2000\left(\frac{15}{216}\right) + 3000\left(\frac{1}{256}\right)$$

$$= -\frac{17000}{216} \doteq -80$$

Ini berarti bahwa nilai harapan kemenangannya bernilai negatif, yang memberikan interpretasi bahwa jika seorang pemain bermain pada kondisi yang sama dan berulang-ulang, maka dalam jangka panjang pemain tersebut secara rata-rata akan mengalami kekalahan sebesar Rp 80 untuk setiap Rp 1000 yang dibelanjakannya.

Ini merupakan salah satu cara untuk melihat apakah suatu permainan itu adil atau tidak, tetapi sayangnya, tidak setiap permainan dapat dianalisis adil atau tidaknya dengan cara ini. Permainan dapat dikatakan **adil** jika E(Y) = 0.

Terlihat jelas bahwa nilai harapan memiliki sifat linier sehubungan dengan penggunaan notasi integral dan penjumlahan.

Teorema 5. 2.

Jika *X* merupakan peubah acak dan *a* serta *b* merupakan konstanta, maka $E(aX+b)=aE(X)+b$

Bukti

Andaikan *X* kontinu. Dengan demikian

$$E(aX+b) = \int_{-\infty}^{+\infty}(ax+b)f_X(x)dx$$

$$= a\int_{-\infty}^{+\infty}xf_X(x)dx + b\int_{-\infty}^{+\infty}f_X(x)dx$$

$$= aE(X)+b$$

Untuk kasus diskrit, caranya sama dengan kasus kontinu.

Teorema 5. 3.

Jika X_1 dan X_2 merupakan peubah acak dengan fungsi kepekatan peluang bersama $f(x_1,x_2)$, maka

$$E_{X_1,X_2}(X_1+X_2)=E_{X_1}(X_1)+E_{X_2}(X_2)$$

Bukti

Untuk peubah acak X_1 dan X_2 kontinu,

$$E_{X_1,X_2}(X_1+X_2)$$

$$= \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}(x_1+x_2)f(x_1,x_2)dx_1dx_2$$

$$= \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}x_1f(x_1,x_2)dx_1dx_2 + \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}x_2f(x_1,x_2)dx_1dx_2$$

$$= \int_{-\infty}^{+\infty}x_1\int_{-\infty}^{+\infty}f(x_1,x_2)dx_1dx_2 + \int_{-\infty}^{+\infty}x_2\int_{-\infty}^{+\infty}f(x_1,x_2)dx_1dx_2$$

$$= \int_{-\infty}^{+\infty} x_1 f_{X_1}(x_1) dx_1 + \int_{-\infty}^{+\infty} x_2 f_{X_2}(x_2) dx_2$$

$$= E_{X_1}(X_1) + E_{X_2}(X_2)$$

Lebih jauh, jika *a1, ... ak* merupakan konstanta-konstanta, dan *X1, ...Xk* adalah peubah acak-peubah acak yang menyebar bersama, maka

$$E\left(\sum_{i=1}^{k} a_i X_i\right) = \sum_{i=1}^{k} a_i E(X_i)$$

Teorema 5. 4.

Jika *X* dan *Y* merupakan peubah acak peubah acak yang saling bebas, dan *g(x)* dan *h(y)* merupakan fungsi, maka

$$E[g(X)h(Y)] = E[g(X)]E[h(Y)]$$

Bukti

Untuk peubah acak kontinu,

$$E[g(X)h(Y)] = \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} g(x)h(y) f(x,y)\, dx\, dy$$

$$= \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} g(x)h(y) f_X(x) f_Y(y)\, dx\, dy$$

$$= \left[\int_{-\infty}^{+\infty} g(x) f_X(x) dx\right]\left[\int_{-\infty}^{+\infty} h(y) f_Y(y) dy\right]$$

$$= E[g(X)]E[h(Y)]$$

Generalisasi teorema diatas dimungkinan dapat dibuat untuk lebih dari dua peubah acak. Lebih khususnya, jika $X_1, ..., X_k$ merupakan peubah acak-peubah acak, dan $u_1(x_1), ...u_k(x_k)$ merupakan fungsi, maka

$$E\left(\prod_{i=1}^{k} u_i(X_i)\right) = \prod_{i=1}^{k} E\left[u_i(X_i)\right]$$

Definisi 5. 1.

Momen ke-*k* disekitar titik pusat dari suatu peubah acak *X* didefinisikan sebagai

$$\mu_k^{'} = E(X^k)$$

dan momen ke-*k* disekitar nilai tengah dari suatu peubah acak *X* didefinisikan sebagai

$$\mu_k = E[X - E(X)]^k = E(X - \mu)^k$$

Dengan demikian jelas terlihat bahwa momen pertama disekitar titik pusat dari peubah acak *X* adalah mean, yang notasi lebih mudahnya adalah μ. Momen pertama disekitar nilai tengahnya adalah nol, sedangkan momen kedua disekitar nilai tengah adalah varian dari peubah acak tersebut.

Teorema 5. 5.

Jika *X* adalah peubah acak, maka $Var(X) = E(X^2) - \mu^2$

Bukti

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

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

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

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

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

Dengan juga dapat diperoleh hubungan bahwa $E(X^2) = \sigma^2 + \mu^2$.

Teladan 5. 3.

Jika peubah acak *X* memiliki distribusi Seragam(a,b), atau dapat ditulis secara singkat bahwa *X ~ Seragam(a,b)*, maka momen ke-*k* disekitar titik pusat adalah

$$E(X^k) = \frac{b^{k+1} - a^{k+1}}{(k+1)(b-a)}$$

Teorema 5. 6.

Jika *X* peubah acak, dan *a* dan *b* adalah konstanta, maka

$$Var(aX+b) = a^2 Var(X)$$

Bukti

$$\begin{aligned} Var(aX+b) &= E\left[(aX+b)^2\right] - \left[E(aX+b)\right]^2 \\ &= E[(aX+b)^2 - (a\mu+b)^2] \\ &= E[a^2X^2 + 2abX + b^2 - a^2\mu^2 - 2ab\mu - b^2] \\ &= a^2E(X^2) + 2ab\mu - a^2\mu^2 - 2ab\mu \\ &= a^2[E(X^2) - \mu^2] \\ &= a^2 Var(X) \end{aligned}$$

Teorema ini menunjukkan kepada kita bahwa ragam atau varian dipengaruhi oleh perubahan skala, tetapi tidak dipengaruhi oleh translasi. Alat lain untuk mengukur keragaman adalah simpangan mutlak terhadap rataan atau sering disebut dengan **mean absolute deviation.**

Teorema 5. 7.

Jika sebaran peubah acak X simetris disekitar nilai tengah $\mu = E(X)$, maka momen ketiga disekitar nilai tengahnya adalah nol, $\mu_3 = 0$.

Definisi 5. 2.

Kovarian dari pasangan peubah acak X dan Y didefinisikan sebagai

$$Cov(X,Y) = E[(X-\mu_X)(Y-\mu_Y)]$$

Notasi kovarian tersebut juga sering dituliskan dengan σ_{XY}.

Teorema 5. 8.

Jika X dan Y adalah peubah acak peubah acak , maka

$$Cov(X,Y) = E(XY) - E(X)E(Y)$$

dan $Cov(X,Y)=0$, bilamana peubah acak X dan Y saling bebas.

Teorema 5. 9.

Jika X_1 dan X_2 adalah peubah acak peubah acak dengan fungsi kepekatan peluang bersama $f(x_1,x_2)$, maka

$$Var(X_1+X_2) = Var(X_1) + Var(X_2) + 2Cov(X_1,X_2)$$

dan

$$Var(X_1+X_2) = Var(X_1) + Var(X_2)$$

bilamana X_1 dan X_2 saling bebas.

Dapat juga dilakukan verifikasi bahwa jika $X_1, \ldots, X_k$ merupakan peubah acak peubah acak dan $a_1, \ldots, a_k$ merupakan konstanta-konstanta, maka

$$Var\left(\sum_{i=1}^{k} a_i X_i\right) = \sum_{i=1}^{k} a_i^2 Var(X_i) + 2\sum_{i}\sum_{<j} a_i a_j Cov(X_i, X_j)$$

dan

$$Var(\sum_{i=1}^{k} a_i X_i) = \sum_{i=1}^{k} a_i^2 Var(X_i)$$

jika $X_1, \ldots, X_k$ merupakan peubah acak peubah acak yang saling bebas.

Menduga Rata-rata dan Ragam

Bilamana memilih model peluang, perlu diperoleh mean (rata-rata) dan varian (ragam), dan mungkin momen-momen yang lebih tinggi dari model tersebut sehingga mendekati rata-rata dan ragam populasinya. Namun, biasanya nilai rata-rata dan ragam populasi ini tidak diketahui, tetapi dapat diestimasi atau diduga berdasarkan suatu contoh acak. Teknik pendugaan tentunya akan dibahas tersendiri pada beberapa bab berikutnya, namun teknik pendugaan secara umum disini digunakan untuk memberikan ilustrasi aplikasi penggunaan rata-rata, ragam dan beberapa sifat nilai harapan.

Rata-rata dan Varian Contoh

Misalkan $X_1, \ldots, X_n$ melambangkan contoh acak dari suatu populasi dengan fungsi kepekatan peluang *f(x)*. Suatu fungsi contoh acak yang tidak tergantung dengan sembarang parameter yang tak diketahui disebut dengan **statistik**. Salah satu statistik contoh yang penting adalah rata-rata contoh, yang tidak lain adalah rata-rata peubah dalam contoh acak, dan dinotasikan dengan

$$\overline{X} = \frac{\sum_{i=1}^{n} X_i}{n}$$

Statistik $\overline{X}$ adalah peubah acak. Jika suatu contoh benar-benar diamati, maka nilai amatan dari $\overline{X}$ biasanya dinotasikan dengan $\overline{x}$. Nilai ini berguna sebagai penduga bagi rata-rata populasi, $\mu = E(X)$.

Teorema 5. 10.

Jika $X_1, \ldots, X_n$ merupakan contoh acak dari suatu populasi dengan fungsi kepekatan peluang *f(x)* dengan $E(X)=\mu$ dan $Var(X)=\sigma^2$, maka $E(\overline{X}) = \mu$ dan $Var(\overline{X}) = \frac{\sigma^2}{n}$.

Sifat pertama dari teorema diatas memberikan arti bahwa, jika banyak peubah acak diambil, dan rata-rata contoh digunakan untuk menduga atau mengestimasi rata-rata populasi, maka dalam jangka panjang, nilai dugaan rata-rata akan memiliki rata-rata yang nilainya akan sama dengan nilai rata-rata populasi yang sebenarnya. Sudah tentu, setiap contoh $\overline{x}$ secara substansial bisa berbeda dengan μ.

Definisi 5. 3.

Suatu penduga $\hat{\theta}$ dikatakan sebagai penduga tak bias dari parameter θ jika $E(\hat{\theta}) = \theta$

Berdasarkan definisi ini, maka $\overline{x}$ merupakan penduga tak bias bagi μ untuk setiap fungsi kepekatan dimana rata-rata (mean) nya ada. Sifat berikutnya mengindikasikan bahwa ragam dari $\overline{X}$ menjadi kecil, bilamana n membesar, sehingga untuk ukuran contoh yang besar, nilai amatan $\overline{x}$ biasanya akan memberikan nilai dugaan yang sangat dekat dengan μ untuk n yang sangat besar.

Jika rata-rata populasi μ diketahui dan σ^2 tak diketahui, maka penduga alami bagi $\sigma^2 = E(X-\mu)^2$ adalah

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

Karena σ^2 merupakan mean (rata-rata) dari *(X-μ)2*. Sudah tentu, secara mudah dapat ditunjukkan bahwa $E(V) = \sigma^2$.

Dalam kebanyakan kasus, tidak dimungkinkan untuk mendapatkan nilai rata-rata populasi, μ., bilamana σ^2 tak diketahui, yang akan menghasilkan penduga

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

Namun, penduga diatas **bukan** merupakan penduga tak bias bagi ragam populasi, σ^2. Untuk itu, sebagai penduga tak bias baginya, digunakan

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

yang merupakan modifikasi dari penduga biasnya.

Untuk keperluan kemudahan dalam menghitung, formula lain dari penduga tak bias bagi ragam poluasi adalah

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

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

Teorema 5. 11.

Jika $X_1, \ldots, X_n$ merupakan contoh acak berukuran n dari suatu fungsi kepekatan peluang $f(x)$ dengan $E(X)=\mu$, $Var(X)=\sigma^2$, maka

$$E(S^2) = \sigma^2$$

$$Var(S^2) = \left(\mu_4^{'} - \frac{n-3}{n-1}\sigma^4\right) / n$$ untuk **n > 1.**

Bukti

sebagai latihan.

Batas-batas Peluang

Sangatlah dimungkinkan untuk memperoleh batas-batas peluang berdasarkan momen.

Teorema 5. 12.

Jika X adalah peubah acak dan *u(x)* merupakan fungsi bernilai-riil tak negatif, maka untuk sembarang konstanta *c > 0*,

$$P[u(X) \geq c] \leq \frac{E[u(X)]}{c}$$

Bukti

Jika $A = \{x \mid u(x) \geq c\}$, maka untuk peubah acak kontinu,

$$\begin{aligned}
E[u(X)] &= \int_{-\infty}^{+\infty} u(x) f(x) dx \\
&= \int_A u(x) f(x) dx + \int_{A^c} u(x) f(x) dx \\
&\geq \int_A u(x) f(x) dx \\
&\geq \int_A c f(x) dx \\
&= cP[X \in A] \\
&= cP[u(X) \geq c]
\end{aligned}$$

Untuk peubah acak kontinu, pembuktian dapat dilakukan dengan cara yang sama.

Jika $u(x) = |x|^r$ untuk nilai *r > 0*, kita dapatkan

$$P[|x| \geq c] \leq \frac{E(|x|^r)}{c^r}$$

Teorema 5. 13.

Jika X adalah peubah acak dengan nilai mean μ dan ragam σ^2, maka untuk sembarang nilai $k > 0$,

$$P[|X-\mu| \geq k\sigma] \leq \frac{1}{k^2}$$

Bukti

Jika $u(X)=(X-\mu)^2$, $c=k^2\sigma^2$, maka

$$P[(X-\mu)^2 \geq k^2\sigma^2] \leq \frac{E(X-\mu)^2}{k^2\sigma^2} \leq \frac{1}{k^2}$$

dan hasilnya menunjukkan pembuktian.

Teorema diatas dikenal dengan **Chebychev Inequality** atau pertidaksamaan Chebychev. Sebagai alternatif bentuk pertidaksamaan tersebut adalah

$$P[|X-\mu| < k\sigma] \geq 1-\frac{1}{k^2}$$

Dan jika kita misalkan $\varepsilon=k\sigma$, maka

$$P[|X-\mu| < \varepsilon] \geq 1-\frac{\sigma^2}{\varepsilon^2}$$

serta

$$P[|X-\mu| \geq \varepsilon] \leq \frac{\sigma^2}{\varepsilon^2}$$

Juga dimungkinkan untuk menunjukkan bahwa jika ragamnya nol, maka sebaran peubah acak terkonsentrasi pada sebuah nilai saja. Sebaran yang demikian disebut dengan sebaran **degenerate**.

Teorema 5. 14.

Misalkan $\mu = E(X)$ dan $\sigma^2 = Var(X)$. Jika $\sigma^2 = 0$, maka $P[X=\mu] = 1$.

Bukti

Jika $x \neq \mu$ untuk beberapa nilai amatan x, maka $| x-\mu | \geq 1/i$ untuk beberapa bilangan bulat $i \geq 1$, dan sebaliknya. Sehingga,

$$[X \neq \mu] = \bigcup_{i=1}^{\infty}\left[|X-\mu| \geq \frac{1}{i}\right]$$

dan dengan menggunakan pertidaksamaan Boole

$$P[X \neq \mu] \leq \sum_{i=1}^{\infty} P\left[|X-\mu| \geq \frac{1}{i}\right].$$

Namun,

$$\sum_{i=1}^{\infty} P\left[|X-\mu| \geq \frac{1}{i}\right] \leq \sum_{i=1}^{\infty} i^2\sigma^2 = 0$$

Yang berarti bahwa $P[X=\mu] = 1$.

Korelasi

Manfaat rata-rata dan ragam dalam mencirikan suatu sebaran peubah acak telah dibahas, dan disebutkan bahwa kovarian berguna untuk mengukur keterkaitan antara dua peubah acak.

Teladan 5. 4.

Misalkan X_1 dan X_2 merupakan peubah acak Eksponensial(1), dan $S = X_1 + X_2$ dan $T = X_1 - X_2$. Tampak jelas bahwa $E(T) = 0$, dan

$$\begin{aligned} Cov(S,T) &= E(ST)-E(S)E(T) \\ &= E[(X_1+X_2)(X_1-X_2)]-0 \\ &= E(X_1^2 - X_2^2) = E(X_1^2) - E(X_2^2) = 0 \end{aligned}$$

Meskipun nilai kovarian antara S dan T disini sama dengan nol, ini bukan berarti bahwa S dan T saling bebas. Kita telah tunjukkan pada bab sebelumnya bahwa S dan T saling terkait.

Peubah acak *X* dan *Y* dikatakan **tidak berkorelasi** jika nilai kovarian antara *X* dan *Y* sama dengan nol, sebaliknya dikatakan bahwa *X* dan *Y* **berkorelasi**.

Teladan 5. 5.

Jika *X* dan *Y* adalah peubah acak, serta *a* dan *b* konstanta, maka

$$Cov(aX, bY) = ab\,Cov(X, Y)$$
$$Cov(X + a, Y + b) = Cov(X, Y)$$
$$Cov(X, aX + b) = a\ Var(X)$$

Kovarian positif menunjukkan hubungan positif dua peubah, sedangkan kovarian negatif menunjukkan hubungan terbalik diantara keduanya, sedangkan kovarian yang mendekati nol mengindikasikan kurangnya keterkaitan, meskipun kovarian nol bukan berarti suatu indikasi ketidaktergantungan antara peubah-peubah tersebut. Kesulitan interpretasi kovarian terjadi apabila nilainya besar dan bahkan mungkin sangat besar. Hal ini dapat dihindari dengan menggunakan ukuran yang disebut dengan korelasi.

Definisi 5. 4.

Jika *X* dan *Y* adalah peubah acak dengan ragam berturut-turut σ_X^2 dan σ_Y^2 dan kovarian $\sigma_{XY} = Cov(X, Y)$, maka **koeffisien korelasi** peubah acak *X* dan *Y* adalah

$$\rho = \frac{\sigma_{XY}}{\sqrt{\sigma_X^2 \sigma_Y^2}} = \frac{\sigma_{XY}}{\sigma_X \sigma_Y}$$

Untuk penulisan notasi korelasi, sering juga digunakan ρ_{XY} .

Teorema 5. 15.

Jika ρ adalah koeffisien korelasi antara *X* dan *Y*, maka -1 $\leq \rho \leq$ 1 dan $\rho = \pm$ 1 jika dan hanya jika *Y* =*aX* + *b* dengan peluang 1 untuk beberapa nilai $a \neq 0$ dan *b*.

Bukti

Misalkan $W = \dfrac{Y}{\sigma_Y} - \rho \dfrac{X}{\sigma_X}$, maka

$$
\begin{aligned}
Var(W) &= \left(\frac{1}{\sigma_Y}\right)^2 Var(Y) + \left(\frac{\rho}{\sigma_X}\right)^2 Var(X) - 2\rho\, Cov(X,Y) \\
&= \left(\frac{1}{\sigma_Y}\right)^2 \sigma_Y^2 + \left(\frac{\rho}{\sigma_X}\right)^2 \sigma_X^2 - 2\rho \frac{\sigma_{XY}}{\sigma_X \sigma_Y} \\
&= 1 + \rho^2 - 2\rho^2 \\
&= 1 - \rho^2 \\
&\geq 0.
\end{aligned}
$$

Andaikan $\rho = \pm 1$ memberikan arti bahwa *Var(W)* = *0*, atau *P*[*W*=μ_W] = *1.* Sehingga dengan peluang 1, $Y/\sigma_Y - X\rho/\sigma_X = \mu_Y/\sigma_Y - \mu_X\rho/\sigma_X$, atau *Y* = *aX*+*b* dimana $a=\rho\sigma_Y/\sigma_X$ dan $b=\mu_Y-\mu_X\rho\sigma_Y/\sigma_X$.

Jika *Y*=*aX* + *b*, maka $Var(Y) = \sigma_Y^2 = a^2\, Var(X) = a^2\sigma_X^2$ serta $Cov(X,Y) = Cov(X, aX+b) = a\, Var(X) = a\sigma_X^2$, sehingga $\rho = \dfrac{a}{|a|}$, sehingga ρ = *1* jika *a* > 0 dan ρ = -1 jika *a* < 0.

Teladan 5. 6.

Jika *X* adalah suhu udara disuatu lokasi (dalam derajat Celcius) memiliki sebaran Seragam(20;30), dan *Z* adalah kesalahan pengukuran juga mengikuti sebaran Seragam(-1;+1) yang bebas dari suhu udara. Sehingga sebenarnya yang kita amati adalah peubah acak *Y* = *X* + *Z.* Dengan demikian fungsi kepekatan peluang bersamanya adalah

$$f(x,y)=\begin{cases}\dfrac{1}{20} & , \quad 20<x<30 \ dan \quad x-1<y<x+1\\ 0 & , \quad selainnya\end{cases}$$

Mekipun *Y* tidak merupakan fungsi linier dari *X*, karena kesalahan acak *Z*, sebaran bersama *X* dan *Y* terkonsentrasi pada gugus $\{(x,y)\,|\,20<x<30, \quad x-1<y<x+1\}$ tergerombol disekitar garis *y*=*x*, dan kita berharap bahwa nilai ρ mendekati 1.

Ragam *X* adalah $\sigma_X^2=(10)^2/12=25/3$, sedangkan ragam *Y* adalah $\sigma_Y^2=(10)^2/12+(2)^2/12=26/3$, dan kovariannya $Cov(X,X+Z)=Var(X)+Cov(X,Z)=Var(X)=25/3$. Dengan demikian korelasi antara *X* dan *Y* adalah $\rho=\dfrac{25/3}{\sqrt{25/3}\sqrt{26/3}}=\sqrt{25/26}\doteq 0,981$ yang nilainya mendekati 1 seperti yang kita harapkan.

Kovarian Contoh

Suatu contoh acak yang diambil dari populasi bivariat (ganda-2) dengan fungsi kepekatan peluang bersama *f(x,y)* merupakan sekumpulan pasangan peubah acak yang saling bebas (X_1,Y_1), …, (X_n,Y_n). Dengan perkataan lain bahwa (X_i,Y_i) dan (X_j,Y_j) saling bebas jika $i\neq j$, namun komponen-komponen dalam suatu pasangan yang mungkin saja saling terkait.

Kovarian contoh merupakan penduga kovarian populasi σ_{XY} yang dinotasikan dengan

$$S_{XY}=\frac{1}{n}\sum_{i=1}^{n}(X_i-\overline{X})(Y_i-\overline{Y})$$

Nilai kovarian contoh berdasarkan data amatan (x_1,y_1), …, (x_n,y_n) dapat dituliskan dengan

$$s_{XY}=\frac{1}{n}\sum_{i=1}^{n}(x_i-\overline{x})(y_i-\overline{y})$$

yang merupakan dugaan nilai σ_{XY}.

Nilai Harapan Bersyarat

Sebagaimana telah kita gunakan bahwa notasi *f(y|x)* merupakan fungsi kepekatan peluang bersyarat peubah acak *Y* seandainya peubah acak *X* bernilai *x*. Kita ingin mencari nilai harapan atau nilai ekspektasi dari peubah bersyarat *Y|X*.

Definisi 5. 5.

Jika *X* dan *Y* merupakan peubah acak yang menyebar bersama, maka **harapan bersyarat** peubah acak *Y* bilamana *X* = *x* adalah

$$E(Y \mid x) = \sum_{y} yf(y \mid x)$$

jika *X* dan *Y* diskrit, dan

$$E(Y \mid x) = \int_{y} yf(y \mid x)dy$$

jika *X* dan *Y* kontinu.

Notasi lain yang digunakan untuk menyatakan hal yang sama adalah $E_{Y|x}(Y)$ dan $E(Y \mid X = x)$.

Seandainya kita ingin mencoba untuk memprediksi nilai koordinat vertikal *(X,Y)*, maka *E(Y|x)* mestinya lebih bermanfaat daripada harapan marjinalnya, *E(Y),* karena memanfaatkan informasi tentang koordinat horisontal. Tentunya, ini menggunakan asumsi bahwa informasi tersebut tersedia.

Perlu dicatat bahwa harapan bersyarat dari *Y* bilamana *X*=*x* merupakan fungsi dari *x*, katakanlah *u(X)*=*E(Y|X)*, memiliki harapan marjinal *Y*, yang dinotasikan dengan *E(Y).*

Teorema 5. 16.

Jika *X* dan *Y* merupakan peubah acak yang menyebar bersama, maka $E[E(Y \mid X)] = E(Y)$

Bukti

Dalam kasus peubah kontinu,

$$\begin{aligned} E[E(Y \mid X)] &= \int_{-\infty}^{\infty} E(Y \mid x) f_X(x) dx \\ &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} yf(y \mid x)\, dy f_X(x)\, dx \\ &= \int_{-\infty}^{\infty} y \int_{-\infty}^{\infty} f(x, y)\, dx\, dy \\ &= \int_{-\infty}^{\infty} y\, f_Y(y)\, dy \\ &= E(Y) \end{aligned}$$

Teorema 5. 17.

Jika *X* dan *Y* merupakan peubah acak yang menyebar bersama, dan *h(x,y)* merupakan suatu fungsi, maka

$$E[h(X,Y)] = E_X\{E[h(X,Y) \mid X]\}$$

Teorema diatas menyatakan bahwa harapan bersama, dapat diselesaikan pertama dengan mencari harapan bersyarat $E[h(x,Y) \mid x]$, dan kemudian mencari harapan hasilnya berdasarkan sebaran *X.*

Teorema 5. 18.

Jika *X* dan *Y* merupakan peubah acak yang menyebar bersama, dan *g(x)* merupakan suatu fungsi, maka

$$E[g(X)Y \mid x] = g(x)E(Y \mid x)$$

Bukti

Sebagai latihan

Teladan 5. 7.

Jika *(X,Y) ~ Multinomial*(n,p_1,p_2) maka dapat ditunjukkan dengan mudah bahwa *X ~ Binomial*(n,p_1), *Y~Binomial*(n,p_2), dan $Y|x$ ~ *Binomial*$(n\text{-}x,p_2/(1\text{-}p_1))$. Rata-rata dan ragam dari *X* dan *Y* dapat diperoleh dengan cara yang telah dibahas pada awal bab ini. Juga dapat diperoleh $E(Y|x) = (n\text{-}x)p_2/(1\text{-}p_1)$.

Dua teorema yang baru saja kita bahas dapat dipergunakan untuk mendapatkan kovarian antara *X* dan *Y*.

$$\begin{aligned}
E(XY) &= E[E(XY|X)] \\
&= E[XE(Y|X)] \\
&= E\left[\frac{X(n-X)p_2}{1-p_1}\right] \\
&= \left[\frac{p_2}{1-p_1}\right]\left[nE(X) - E(X^2)\right] \\
&= \left[\frac{p_2}{1-p_1}\right]\left[np - np_1\left[1 + (n-1)p_1\right]\right] \\
&= n(n-1)p_1p_2
\end{aligned}$$

Sehingga,

$$\begin{aligned}
Cov(X,Y) &= n(n-1)p_1p_2 - (np_1)(np_2) \\
&= -np_1p_2
\end{aligned}$$

Teladan 5. 8.

Banyaknya pertandingan, *X*, yang diperlukan oleh pemain inti membantu memenangkan pertandingan dalam *n* pertandingan mengikuti sebaran *Binomial(n,p)*.

Dalam suatu pertandingan yang imbang dengan sistem '*the best of five*', banyaknya pertandingan yang harus dimainkan merupakan peubah acak, *N*, dengan fungsi kepekatan peluang

$$f_N(n) = C_2^{n-1}\left(\frac{1}{2}\right)^{n-1} \qquad n = 3, 4, 5$$

Harapan pertandingan yang harus dimainkan pemain inti dalam 'the best of five' tersebut adalah

$$\begin{aligned} E(X) &= E_N\left[E(X \mid N)\right] \\ &= E_N(Np) \\ &= p\sum_{n=3}^{5} nf_N(n) \\ &= 4{,}125\,p \end{aligned}$$

Teorema 5. 19.

Jika *X* dan *Y* merupakan peubah yang saling bebas, maka *E(Y|x)* = *E(Y)* dan *(E(X|y)* = *E(X)*

Bukti

Jika *X* dan *Y* merupakan peubah yang saling bebas, maka $f(x,y)=f_1(x)f_2(y)$, sehingga $f(y|x)=f_2(y)$ dan $f(x|y) = f_1(x)$. Bilamana peubah acaknya kontinu, maka

$$E(Y \mid x) = \int_{-\infty}^{\infty} yf(y \mid x)dx = \int_{-\infty}^{\infty} yf_2(y)dy = E(Y)$$

Definisi 5. 6.

Ragam bersyarat peubah *Y* bilamana *X*=*x* adalah

$$Var(Y \mid x) = E\{[Y - E(Y \mid x)]^2 \mid x\}$$

Bentuk yang setara dengan pernyataan diatas adalah

$$Var(Y \mid x) = E(Y^2 \mid x) - [E(Y \mid x)]^2$$

Teorema 5. 20.

Jika *X* dan *Y* peubah acak yang menyebar bersama, maka

$$Var(Y) = E_X[Var(Y \mid X)] + Var_X[E(Y \mid X)]$$

Bukti

$E_X[Var(Y|X)]$ = $E_X\{E(Y^2 \mid X) - [E(Y \mid X)]^2\}$

= $\{E(Y^2) - E_X[E(Y \mid X)]^2\}$

= $E(Y^2) - [E(Y)]^2 - \{E_X[E(Y \mid X)]^2 - [E(Y)]^2\}$

= $Var(Y) - Var_X[E(Y \mid X)]$

Teorema diatas menunjukkan bahwa secara rata-rata (untuk keseluruhan nilai *X*) ragam bersyarat akan lebih kecil dari keragaman tak bersyarat. Sudah barang tentu jika *X* dan *Y* saling bebas, maka nilainya akan sama, karena *E(Y|X)* tak akan merupakan fungsi dari *X* dan *Var[E(Y|X)]* akan bernilai nol.

Teorema 5. 21.

Jika *E(Y|x)* merupakan fungsi linier dari *x*, maka

$$E(Y \mid x) = E(Y) + \rho\sqrt{\frac{Var(Y)}{Var(X)}}(x - E(X))$$

dan

$$E_X[Var(Y \mid X)] = Var(Y)(1 - \rho^2)$$

Bukti

Perhatikan $E(Y \mid x) = E(Y) + \rho\sqrt{\frac{Var(Y)}{Var(X)}}(x - E(X))$. Jika *E(Y|x)* = *ax* + *b*, $E(Y) = E_X[E(Y \mid X)] = E_X[aX + b] = aE(X) + b$

dan

$$
\begin{aligned}
Cov(X,Y) = \sigma_{XY} &= E[(X - E(X))(Y - E(Y)] \\
&= E[(X - E(X))Y] - 0 \\
&= E_X\{E[(X - E(X))Y \mid X]\} \\
&= E_X[(X - E(X))E(Y \mid X)] \\
&= E_X[(X - E(X))(aX + b)] \\
&= aVar(X)
\end{aligned}
$$

Dengan demikian,

$$a = \frac{\sigma_{XY}}{\sigma_X^2} = \frac{Cov(X,Y)}{Var(X)} = \rho\sqrt{\frac{Var(Y)}{Var(X)}}$$

$$b = E(Y) - \rho\sqrt{\frac{Var(Y)}{Var(X)}}E(X)$$

Sedangkan bagian kedua dari Teorema ini mengikuti Teorema 5. 20.

$$E_X[Var(Y \mid X)]$$

$$
\begin{aligned}
&= Var(Y) - Var_X[E(Y) + \rho\sqrt{\frac{Var(Y)}{Var(X)}}(X - E(X))] \\
&= Var(Y) - \rho^2\,\frac{Var(Y)Var(X)}{Var(X)} \\
&= Var(Y)(1 - \rho^2)
\end{aligned}
$$

Sebagai catatan bahwa, jika ragam bersyarat tidak tergantung pada nilai *x*, maka

$$Var(Y \mid x) = E_X[Var(Y \mid X)] = Var(Y)(1 - \rho^2)$$

Hal ini menunjukkan bahwa penurunan ragam populasi bersyarat dibandingkan populasi tak bersyarat tergantung pada korelasi populasi kedua peubah, ρ.

Kasus penting dimana $E(Y|x)$ merupakan fungsi linier dari *x* dan $Var(Y|x)$ tak tergantung pada *x*, akan kita bahas.

Sebaran Bivariat Normal

Pasangan peubah acak kontinu *X* dan *Y* dikatakan memiliki sebaran Bivariat Normal atau sebaran Normal Ganda Dua, jika memiliki fungsi kepekatan peluang bersama

$$f(x,y) = \frac{1}{2\pi\sigma_X\sigma_Y\sqrt{1-\rho^2}}$$

$$\exp\left\{-\frac{1}{2(1-\rho^2)}\left[\left(\frac{x-\mu_X}{\sigma_X}\right)^2 - 2\rho\left(\frac{x-\mu_X}{\sigma_X}\right)\right.\right.$$

$$\left.\left.\left(\frac{y-\mu_Y}{\sigma_Y}\right)+\left(\frac{y-\mu_Y}{\sigma_Y}\right)^2\right]\right\}$$

untuk $-\infty < x < \infty, \quad -\infty < y < \infty$.

Cara lain untuk menuliskan peubah acak *X* dan *Y* memiliki sebaran Bivariat Normal adalah

$$(X,Y) \sim BVN(\mu_X, \mu_Y, \sigma_X^2, \sigma_Y^2, \rho)$$

yang tergantung dari parameter-parameter $-\infty < \mu_X < \infty$, $-\infty < \mu_Y < \infty$, $\sigma_X > 0$, $\sigma_Y > 0$, $-1 < \rho < +1$.

Teorema 5. 22.

Jika $(X,Y) \sim BVN(\mu_X, \mu_Y, \sigma_X^2, \sigma_Y^2, \rho)$, maka $X \sim N(\mu_X, \sigma_X^2)$ dan $Y \sim N(\mu_Y, \sigma_Y^2)$ dan ρ adalah koeffisien korelasi peubah *X* dan *Y*.

Bukti

Sebagai latihan.

Teorema 5. 23.

Jika $(X,Y) \sim BVN(\mu_X, \mu_Y, \sigma_X^2, \sigma_Y^2, \rho)$, maka

$$Y \mid x \sim N\left[\mu_Y + \rho\frac{\sigma_Y}{\sigma_X}(x-\mu_X), \sigma_Y^2(1-\rho^2)\right]$$

$$X \mid y \sim N\left[\mu_X + \rho\frac{\sigma_X}{\sigma_Y}(y-\mu_Y), \sigma_X^2(1-\rho^2)\right]$$

Bukti

Sebagai latihan.

Aproksimasi Rata-rata dan Ragam

Jika fungsi peubah acak *H(X)* dapat diekspansi dengan deret Taylor, maka ekspresi aproksimasi rata-rata dan ragam *H(X)* dapat diperoleh sebagai fungsi dari rata-rata dan ragam *X*.

Dengan persyaratan bahwa *H(x)* memiliki turunan *H'(x), H"(x), ...* dalam interval terbuka yang mengandung μ=*E(X)*, maka pendekatan *H(x)* dengan deret Taylor disekitar μ adalah

$$H(x) \doteq H(\mu) + H'(\mu)(x-\mu) + \frac{1}{2}H''(\mu)(x-\mu)^2$$

yang memberikan nilai pendekatan

$$E[H(X)] \doteq H(\mu) + \frac{1}{2}H''(\mu)\sigma^2$$

dan dengan menggunakan dua suku pertamanya, diperoleh

$$Var[H(X)] \doteq [H'(\mu)]^2\sigma^2$$

Catatan : $\sigma^2 = Var(X)$

Teladan 5. 9.

Misalkan *X* adalah peubah acak bernilai positif, dan misalkan *H(x)* = *ln x,* sehingga *H'(x)* = *1/x* dan *H"(x)* = *-1/x2.* Sehingga dapat diperoleh

$$E[\ln X] \doteq \ln \mu + \frac{1}{2}\left(-\frac{1}{\mu}\right)\sigma^2 = \ln \mu - \frac{\sigma^2}{2\mu^2}$$

dan

$$Var[\ln X] \doteq \left(\frac{1}{\mu}\right)^2 \sigma^2 = \frac{\sigma^2}{\mu^2}$$

Hasil yang sama dapat dikembangkan untuk fungsi lebih dari satu peubah. Dengan menggunakan pendekatan Taylor, kita peroleh pendekatan rata-rata dan ragam fungsi peubah acak *H(X,Y)*

$$E[H(X,Y)] \doteq H(\mu_X,\mu_Y) + \frac{\partial^2 H}{\partial x^2}\sigma_X^2 + \frac{\partial^2 H}{\partial y^2}\sigma_Y^2$$

$$Var[H(X,Y)] \doteq \left(\frac{\partial H}{\partial x}\right)^2 \sigma_X^2 + \left(\frac{\partial H}{\partial y}\right)^2 \sigma_Y^2 + \left(\frac{\partial H}{\partial x}\right)\left(\frac{\partial H}{\partial y}\right)\sigma_{XY}$$

Fungsi Pembangkit Momen

Bentuk nilai harapan khusus yang sering bermanfaat adalah fungsi pembangkit momen.

Definisi 5. 7.

Jika *X* adalah peubah acak, maka nilai harapan

$$M_X(t) = E(e^{tX})$$

disebut dengan fungsi pembangkit momen (*moment generating function*) dari *X*, jika nilai harapan tersebut ada untuk semua nilai *t* dalam beberapa interval yang mencakup *0* dalam bentuk $-h < t < h$ untuk beberapa nilai $h > 0$.

Teladan 5. 10.

Bila diasumsikan bahwa *X* adalah peubah acak diskrit terhingga dengan nilai-nilai yang mungkin $x_1, \ldots, x_m$. Fungsi pembangkit momennya dapat dituliskan sebagai

$$M_X(t) = \sum_{i=1}^{m} e^{tx_i} f_X(x_i)$$

merupakan fungsi dari *t* yang diferensiabel hingga *r* kali, dimana turunan ke-*r* nya dapat dituliskan sebagai

$$M_X^{(r)} = \sum_{i=1}^{m} x_i^r e^{tx_i} f_X(x_i)$$

Jika kita evaluasi pada *t=0* maka akan kita peroleh

$$M_X^{(r)}(0) = \sum_{i=1}^{m} x_i^r f_X(x_i) = E(X^r)$$

yang merupakan momen ke-*r* disekitar titik pusat.
Dengan menggunakan deret kuasa (*power series*) disekitar *t=0* kita dapat tuliskan

$$M_X(t) = \sum_{r=0}^{\infty} \frac{E(X^r)}{r!} t^r$$

Sifat-sifat diatas berlaku untuk semua peubah acak yang memiliki fungsi pembangkit momen, meskipun pembuktian secara umum lebih sulit.

Teorema 5. 24.

Jika fungsi pembangkit momen ada, maka

$$E(X^r) = M_X^{(r)}(0) \text{ untuk } r = 1,2,\ldots$$

dan

$$M_X(t) = 1 + \sum_{r=1}^{\infty} \frac{E(X^r)}{r!} t^r$$

Teladan 5. 11.

Jika *X ~ Bin(n,p),* maka

$$\begin{aligned} M_X(t) &= \sum_{x=0}^{n} e^{tx} C_x^n p^x q^{n-x} \\ &= \sum_{x=0}^{n} C_x^n (pe^t)^x q^{n-x} \\ &= (pe^t + q)^n \qquad -\infty < t < \infty \end{aligned}$$

Kita dapat cari $M'_X(t) = n(pe^t + q)^{n-1} pe^t$, dan dengan demikian $M'_X(0) = np.$

Teladan 5. 12.

Jika *X~Poi(μ)*, maka

$$\begin{aligned} M_X(t) &= \sum_{x=0}^{\infty} e^{tx} e^{-\mu} \frac{\mu^x}{x!} \\ &= e^{-\mu} \sum_{x=0}^{\infty} \frac{(\mu e^t)^x}{x!} \\ &= e^{-\mu} e^{\mu e^t} \\ &= e^{\mu(e^t - 1)} \end{aligned}$$

Dapat dengan mudah diperoleh $M'_X(t) = \mu e^{t+\mu(e^t-1)}$ dan $M'_X(0) = \mu$ serta $M''_X(t) = M'_X(t)[1 + \mu e^t]$ sehingga

$M''_X(0) = M'_X(0)[1+\mu e^0] = \mu(1+\mu)$. Dengan demikian $Var(X) = E(X^2) - (E(X))^2 = \mu$.

Teladan 5. 13.

Jika *X~Gam(θ,κ)*, maka

$$M_X(t) = \int_0^\infty e^{tx} \frac{x^{\kappa-1} e^{-x/\theta}}{\theta^\kappa \Gamma(\kappa)} dx$$

$$= \frac{1}{\theta^\kappa \Gamma(\kappa)} \int_0^\infty x^{\kappa-1} e^{(t-1/\theta)x} dx$$

Dengan menggunakan substitusi $u = -(t-1/\theta)x$

$$M_X(t) = \left(\frac{1}{\theta} - t\right)^{-\kappa} \frac{1}{\theta^\kappa \Gamma(\kappa)} \int_0^\infty u^{\kappa-1} e^{-u} du$$

$$= (1-\theta t)^{-\kappa} \qquad t < 1/\theta$$

Turunan ke-*r* dari fungsi pembangkit momen ini adalah

$$M_X^{(r)}(t) = (\kappa+r-1)...(\kappa+1)\kappa\theta^r(1-\theta t)^{-\kappa-r}$$

$$= \frac{\Gamma(\kappa+r)}{\Gamma(\kappa)} \theta^r (1-\theta t)^{-\kappa-r}$$

Sehingga kita dapatkan momen ke-*r* disekitar titik pusat sebagai berikut

$$E(X^r) = \frac{\Gamma(\kappa+r)}{\Gamma(\kappa)} \theta^r$$

berlaku untuk semua nilai *r* bilangan bulat positif, tetapi masih dimungkinkan absah untuk sembarang nilai $r > -\kappa$.

Secara umum fungsi pembangkit momen Gamma dapat dituliskan dalam bentuk *power series*

$$M_X(t) = 1 + \sum_{r=1}^\infty \frac{\Gamma(\kappa+r)}{\Gamma(\kappa)} \frac{\theta^r}{r!} t^r$$

Sifat-sifat Fungsi Pembangkit Momen

Terdapat beberapa sifat penting dari fungsi pembangkit momen peubah acak

Teorema 5. 25.

Jika $Y = aX+b$, maka $M_Y(t) = e^{bt} M_X(at)$

Bukti

$$
\begin{aligned}
M_Y(t) &= E(e^{tY}) \\
&= E(e^{t(aX+b)}) \\
&= E(e^{atX} e^{bt}) \\
&= e^{bt} E(e^{atX}) \\
&= e^{bt} M_X(at)
\end{aligned}
$$

Salah satu manfaat dari teorema ini adalah didalam menghitung momen ke-*r* disekitar nilai tengah, $E[(X-\mu)^r]$. Karena $M_{X-\mu}(t) = e^{-\mu t} M_X(t)$, maka

$$E[(X-\mu)^r] = \frac{d^r}{dt^r}\left[e^{-\mu t} M_X(t) \right]\Big|_{t=0}$$

Teorema 5. 26.

Jika $X_1, \ldots, X_n$ merupakan peubah-peubah acak yang saling bebas dengan fungsi pembangkit momen $M_{X_i}(t)$, maka fungsi pembangkit momen $Y = \sum_{i=1}^{n} X_i$ adalah

$$M_Y(t) = M_{X_1}(t) \ldots M_{X_n}(t) = \prod_{i=1}^{n} M_{X_i}(t)$$

Bukti

Karena $e^{tY} = e^{tX_1+...+tX_n}$, maka berdasarkan Generalisasi teorema yang telah kita bahas diawal bab ini

$$\begin{aligned} M_Y(t) &= E(e^{tY}) \\ &= E(e^{tX_1+...+tX_n}) \\ &= E(e^{tX_1})...E(e^{tX_n}) \\ &= M_{X_1}(t)...M_{X_n}(t) \\ &= \prod_{i=1}^{n} M_{X_i}(t) \end{aligned}$$

Apabila $X_1, \ldots, X_n$ merupakan contoh acak dari suatu populasi dengan fungsi kepekatan peluang *f(x)* dan fungsi pembangkit peluang *M(t)*, maka

$$M_Y(t) = [M(t)]^n$$

Teorema 5. 27.

Jika X_1 dan X_2 masing-masing memiliki fungsi sebaran kumulatif berturut-turut $F_1(x)$ dan $F_2(x)$, dan fungsi pembangkit momen $M_1(t)$ dan $M_2(t)$, maka $F_1(x)=F_2(x)$ untuk semua bilangan nyata *x* jika dan hanya jika $M_1(t)=M_2(t)$ untuk semua *t* dalam beberapa interval yang mencakup nilai *0*, $-h < t < h$ untuk beberapa nilai $h > 0$.

Teladan 5. 14.

Misalkan $X\sim Exp(\theta)$. Karena sebaran eksponensial merupakan bentuk khusus dari sebaran Gamma dengan $\kappa=1$ maka fungsi pembangkit momennya adalah $M_X(t) = (1-\theta t)^{-1}$ untuk $t < 1/\theta$. Dari teorema Teorema 5. 25. fungsi pembangkit momen dari $Y = aX$ adalah $M_Y(t) = (1-\theta at)^{-1}$. Ini merupakan fungsi pembangkit momen peubah acak eksponensial dengan parameter $a\theta$. Dengan demikian *Y~Eks(*$a\theta$*).*

Sudah jelas bahwa fungsi pembangkit momen dapat digunakan untuk menentukan sebaran fungsi peubah acak, dan tak diragukan lagi juga dapat digunakan untuk menghitung momen-momen peubah acak. Pendekatan fungsi pembangkit momen berguna khususnya dalam penentuan sebaran jumlah peubah acak yang saling bebas, dan ini jauh lebih mudah daripada dilakukan dengan cara menggunakan transformasi bersama.

Teladan 5. 15.

Misalkan $X_1, \ldots, X_k$ merupakan contoh acak berukuran k dari sebaran Bin(*n,p*), dan $Y = \sum_{i=1}^{k} X_i$. Dengan demikian

$M_Y(t) = [M_X(t)]^k = (pe^t + q)^{nk}$ yang merupakan fungsi pembangkit momen dari sebaran Bin(*nk,p*). *Y~Bin(nk,p)*.

Teladan 5. 16.

Misalkan $X_1, \ldots, X_n$ merupakan peubah-peubah acak dari sebaran $Poi(\mu_i)$ dan misalkan $Y=X_1+\ldots+X_n$. Fungsi pembangkit momen peubah acak X_i adalah $M_{X_i}(t) = e^{\mu_i(e^t-1)}$ dan dengan demikian

$$
\begin{aligned}
M_Y(t) &= e^{[\mu_1(e^t-1)]} \ldots e^{[\mu_n(e^t-1)]} \\
&= e^{(\mu_1+\ldots+\mu_n)(e^t-1)}
\end{aligned}
$$

yang berarti bahwa *Y~Poi(*$\mu_1 + \ldots + \mu_n$*)*.

Momen Faktorial dan Fungsi Pembangkit Momen Faktorial

Bentuk khusus lain dari nilai harapan peubah acak adalah momen faktorial.

Definisi 5. 8.

Momen faktorial ke-*r* peubah acak *X* didefinisikan sebagai

$$E[X(X-1)...(X-r+1)]$$

Dan **fungsi pembangkit momen faktorial** peubah acak *X* didefinisikan sebagai

$$G_X(t) = E(t^X)$$

jika nilai harapannya ada untuk semua *t* dalam beberapa interval terbuka yang mencakup nilai *1* dalam bentuk *1-h* < *t* < *1+h.*

Fungsi pembangkit momen faktorial lebih dapat diperoleh dalam bentuk yang sederhana daripada fungsi pembangkit momen untuk beberapa peubah acak *X*.

Teorema 5. 28.

Jika *X* memmiliki fungsi pembangkit momen faktorial, $G_X(t)$, maka

$$G_X^{'}(1) = E(X)$$
$$G_X^{''}(1) = E[X(X-1)]$$
$$G_X^{(r)}(1) = E[X(X-1)...(X-r+1)]$$

Bukti : sebagai latihan.

Teladan 5. 17.

Misalkan *X~Bin(n,p)*, maka

$$G_X(t) = \sum_{x=0}^{n} t^x C_x^n p^x q^{n-x}$$
$$= \sum_{x=0}^{n} C_x^n (pt)^x q^{n-x}$$
$$= (pt+q)^n$$
$$G_X^{'}(t) = n(pt+q)^{n-1} p$$
$$G_X^{''}(t) = (n-1)n(pt+q)^{n-2} p^2$$

Dalam beberapa kasus fungsi pembagkit momen faktorial dan fungsi pembangkit momen memiliki bentuk yang lebih mudah, namun demikian,kita bisa peroleh bahwa

$$G_X(t) = E(t^X) = E(e^{X \ln t}) = M_X(\ln t)$$

Fungsi Pembangkit Momen Bersama

Konsep fungsi pembangkit momen dapat digeneralisasikan untuk peubah acak berdimensi-*k*.

Definisi 5. 9.

Fungsi Pembangkit Momen bersama peubah acak $\boldsymbol{X}=(X_1, \ldots, X_k)$ didefinisikan sebagai

$M_{\underline{X}}(\underline{t}) = E[e^{\sum_{i=1}^{k} t_i X_i}]$ dimana $\underline{t} = (t_1, \ldots, t_k)$ dan $-h < t_i < h$ untuk beberapa nilai $h > 0$.

Sifat-sifat fungsi pembangkit momen bersama analog dengan sifat-sifat fungsi pembangkit momen univariat. Momen campuran seperti $E[X_i^r X_j^s]$ dapat diperoleh dengan mendeferensialkan fungsi pembangkit momen bersama *r* kali terhadap t_i dan *s* kali terhadap t_j

dan mengaturnya untuk semua nilai t_i dan t_j bernilai nol.

Juga dimungkinkan untuk mendapatkan fungsi pembangkit momen sebaran marjinal dari fungsi pembangkit momen bersama.

Teorema 5. 29.

Jika $M_{X,Y}(t_1,t_2)$ ada, maka peubah acak *X* dan *Y* dikatakan saling bebas, jika dan hanya jika $M_{X,Y}(t_1,t_2) = M_X(t_1)M_Y(t_2)$

Teladan 5. 18.

Misalkan bahwa ***X*** = *(X1, …, Xk)~Mult(n,p1,…,pk)*. Kita telah bahas pada bagian terdahulu bahwa sebaran marjinalnya adalah binomial, *Xi~Bin(n,pi)*. Dengan demikian, fungsi pembangkit momen bersamanya adalah

$$M_{\underline{X}}(\underline{t}) = E[e^{\sum_{i=1}^{k} t_i X_i}]$$

$$= \sum \ldots \sum \frac{n!}{x_1! \ldots x_{k+1}!}(p_1 e^{t_1})^{x_1} \ldots (p_k e^{t_k})^{x_k} p_{k+1}^{x_{k+1}}$$

$$= (p_1 e^{t_1} + \ldots + p_k e^{t_k} + p_{k+1})^n$$

dimana $p_{k+1} = 1 - p_1 - \ldots - p_k$.

Sebaran marjinal bersamanya, sudah jelas, juga multinomial. Misalnya, jika *(X1, X2, X3) ~ Mult(n,p1,p2,p3)*, maka

$$M_{X_1,X_2}(t_1,t_2) = M_{X_1,X_2,X_3}(t_1,t_2,0)$$

$$= [p_1 e^{t_1} + p_2 e^{t_2} + p_3 + (1 - p_1 - p_2 - p_3)]^n$$

$$= [p_1 e^{t_1} + p_2 e^{t_2} + (1 - p_1 - p_2)]^n$$

Dengan demikian, *(X1, X2) ~ Mult(n,p1,p2)*

Latihan

1. Dalam bidang Fisika, dikenal distribusi Furry yang memiliki fungsi kepekatan peluang $f(x) = \frac{1}{1+\lambda}\left(\frac{\lambda}{1+\lambda}\right)^x$ untuk *x=0, 1, 2, …* dan λ>0. Carilah
   a. Fungsi Pembangkit Momen Faktorial dari X, $G_X(t)$
   b. Var(X)
2. Misalkan peubah acak X memiliki distribusi Poisson dengan parameter μ. Tentukan:
   a. Fungsi Pembangkit Momen Faktorial X, atau $G_X(t)$
   b. E[X], dengan menggunakan $G_X(t)$
   c. E[X(X-1)], dengan menggunakan $G_X(t)$
3. Jika peubah acak X memiliki distribusi Normal dengan parameter μ dan σ^2, tentukanlah:
   a. $M_X(t)$
   b. $E(X-\mu)^{2r}$ untuk *r=1, 2, …*
4. Misalkan *f(x,y)=6x* untuk *0<x<y<1* dan nol selainnya. Carilah:
   a. $f_1(x)$
   b. $f_2(y)$
   c. Cov(X,Y)
   d. ρ
   e. *f(y|x)*
   f. E(Y|x)
5. Bila diketahui fungsi kepekatan peluang bersama *f(x,y)=x+y* untuk *0<x<1* dan *0<y<1* dan nol selainnya. Tentukan
   a. *f(x)*

b. F(x,y)
c. *f(y|x)*
d. F(y|x)

6. Misalkan $X \sim Poi(\mu)$. Tentukan
    a. E(X)
    b. E[X(X-1)]
    c. Var(X)
7. Misalkan *X* dan *Y* merupakan peubah acak diskrit dengan fungsi kepekatan peluang bersama $f(x,y)=4/(5xy)$ jika $x = 1, 2$ dan $y = 2, 3$, dan bernilai *0* untuk nilai-nilai *x* dan *y* lainnya. Dapatkan
    a. E(X)
    b. E(Y)
    c. E(XY)
    d. Cov(X,Y)
    e. Korelasi antara X dan Y
8. Misalkan *X* dan *Y* merupakan peubah acak kontinu dengan fungsi kepekatan peluang bersama $f(x,y) = x+y$ untuk $o<x<1$ dan $0<y<1$ dan bernilai *0* untuk nilai-nilai lainnya. Tentukan
    a. E(X)
    b. E(X+Y)
    c. E(XY)
    d. Cov(2X,3Y)
    e. E(Y|x)

Sifat Beberapa Sebaran Kontinu

Pendahuluan

Berbagai macam populasi, terutama dari hasil pengamatan biologis, dapat dimodelkan dengan baik oleh fungsi kepekatan peluang normal

$$f(x;\mu,\sigma) = \frac{1}{\sqrt{2\pi}\sigma}\exp\left[-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2\right] \qquad -\infty < x < \infty$$

Dengan nilai-nilai μ dan σ yang sesuai. Sebagai akibat dari permasalahan ini, usaha besar telah dilakukan untuk mengembangkan analisis data statistik dai populasi-poluasi yang memiliki distribusi/sebaran normal. Dalam bab ini akan disajikan pengembangan hasil distribusi yang diperlukan dalam analisis data dari satu atau lebih populasi normal.

Sifat-sifat Sebaran Normal

Beberapa sifat dasar peubah acak yang memiliki sebaran normal akan disajikan seperti di bawah ini.

Teorema 6. 1.

Jika $X \sim N(\mu,\sigma^2)$, maka

$$Z = \frac{X-\mu}{\sigma} \sim N(0,1)$$

$$M_X(t) = e^{\mu t + \sigma^2 t^2/2}$$

$$E(X-\mu)^{2r} = \frac{(2r)!\sigma^{2r}}{r!2^r} \qquad r = 1,2,..$$

$$E(X-\mu)^{2r-1} = 0 \qquad r = 1,2,...$$

Bukti

Untuk membuktikan persamaan diatas, jika $X \sim N(\mu, \sigma^2)$ dan $Z = (X-\mu)/\sigma$, maka $dx/dz = \sigma$ dan

$$f_Z(z) = f(z\sigma + \mu; \mu, \sigma)\left|\frac{dx}{dz}\right| = \frac{1}{\sqrt{2\pi}}\exp\left[-\frac{1}{2}z^2\right]$$

Persamaan berikutnya dapat ditunjukkan dengan definisi fungsi pembangkit momen dari sebaran normal baku.

$$M_Z(t) = \int_{-\infty}^{+\infty} \frac{1}{\sqrt{2\pi}} e^{tz} e^{-z^2/2} dz$$

$$= \int_{-\infty}^{+\infty} \frac{1}{\sqrt{2\pi}} e^{-(z-t)^2/2+t^2/2} dz = e^{t^2/2}$$

Karena $X = Z\sigma + \mu$, maka

$$M_X(t) = M_{\sigma Z+\mu}(t) = e^{\mu t} M_Z(\sigma t) = e^{\mu t + \sigma^2 t^2/2}$$

Dua persamaan terakhir diturunkan dengan menggunakan

$$M_{X-\mu}(t) = e^{\sigma^2 t^2/2}$$

$$= \sum_{r=0}^{\infty} \frac{\sigma^{2r} t^{2r}}{2^r r!}$$

$$= \sum_{r=0}^{\infty} \frac{\sigma^{2r}(2r)! t^{2r}}{2^r r!(2r)!}$$

Ekspansi ini hanya mencakup pangkat bilangan bulat positif genap, dan koeffisien dari $t^{2r}/(2r)!$ adalah momen ke-$2r$ dari $(X-\mu)$.

Beberapa aplikasi sebaran normal juga mencakup kombinasi linier dari peubah acak peubah acak normal yang saling bebas.

Teorema 6. 2.

Jika $X_i \sim N(\mu_i, \sigma_i^2); \quad i = 1,2,...,n$ melambangkan peubah-acak peubah-acak normal yang saling bebas, maka

$$Y = \sum_{i=1}^{n} a_i X_i \sim N(\sum_{i=1}^{n} a_i \mu_i, \sum_{i=1}^{n} a_i^2 \sigma_i^2)$$

Bukti

$$\begin{aligned} M_Y(t) &= M_{\sum_{i=1}^{n} a_i X_i}(t) \\ &= \prod_{i=1}^{n} M_{X_i}(a_i t) \\ &= \prod_{i=1}^{n} e^{a_i \mu_i t + a_i^2 t^2 \sigma_i^2 / 2} \\ &= \exp\left[t \sum_{i=1}^{n} a_i \mu_i + t^2 \sum_{i=1}^{n} a_i^2 \sigma_i^2 / 2 \right] \end{aligned}$$

yang merupakan fungsi pembangkit momen dari sebaran normal dengan rata-rata $\sum_{i=1}^{n} a_i \mu_i$ dan ragam atau varian sebesar $\sum_{i=1}^{n} a_i^2 \sigma_i^2$.

Teladan 6. 1.

Jika $X_1, ..., X_n$ merupakan contoh acak dari sebaran $N(\mu, \sigma^2)$, maka statistik rata-rata contoh memiliki sebaran normal dengan rata-rata μ dan ragam atau varian σ^2/n.

Sebaran Kai-kuadrat

Pertimbangkan sebuah sebaran Gamma dengan parameter θ = *2* dan κ = *ν/2.* Suatu peubah acak *Y* dikatakan memiliki sebaran kai-kuadrat dengan derajat bebas ν jika *Y ~ Gamma(2; ν/2)*. Notasi khusus yang biasanya digunakan adalah $Y \sim \chi^2_{(\nu)}$

Teorema 6. 3.

Jika $Y \sim \chi^2_{(\nu)}$, maka

$$M_Y(t) = (1-2t)^{-\nu/2}$$
$$E(Y^r) = 2^r \frac{\Gamma(\nu/2+r)}{\Gamma(\nu/2)}$$
$$E(Y) = \nu$$
$$Var(Y) = 2\nu$$

Tabel sebaran kumulatif kai-kuadrat telah banyak ditabulasikan dalam berbagai literatur. Dalam banyak literatur, persentil $\chi^2_{\gamma;\nu}$ tersedia untuk berbagai taraf γ dan nilai ν yang berbeda pula. Dengan demikian, jika $Y \sim \chi^2_{(\nu)}$, maka $\chi^2_{\gamma;\nu}$ merupakan suatu nilai sedemikian rupa sehingga

$$P[Y \le \chi_{\gamma;\nu}] = \gamma$$

Teorema 6. 4.

Jika *X* ~ *Gamma(θ,κ)*, maka *Y* = *2X/θ* ~ $\chi^2_{2\kappa}$.

Bukti

$$M_Y(t) = M_{2X/\theta}(t)$$
$$= M_X(2t/\theta)$$

$$= (1-2t)^{-2\kappa/2}$$

yang merupakan fungsi pembangkit momen sebaran kai-kuadrat dengan derajat bebas *2κ*. Dengan kekhasan sifat fungsi pembangkit momen, maka *Y* = *2X/θ* ~ $\chi^2_{2\kappa}$

Fungsi sebaran kumulatif Gamma juga dapat diekspresikan dalam notasi kai-kuadrat. Jika *X ~ Gamma(θ,κ)*, dan jika *H(y;ν)* menotasikan fungsi sebaran kumulatif kai-kuadrat dengan derajat bebas ν, maka

$$F_X(x) = H(2x/\theta; 2\kappa)$$

Dengan menggunakan relasi diatas, dapat disimpulkan bahwa secara umum, persentil ke-p dari sebaran Gamma dapat diekspresikan dengan

$$x_p = \frac{\theta\chi^2_{p;2\kappa}}{2}$$

Teorema 6. 5.

Jika $Y_i \sim \chi^2_{\nu_i}; \quad i = 1,2,...,n$ adalah peubah acak –peubah acak kai-kuadrat yang saling bebas, maka

$$V = \sum_{i=1}^{n} Y_i \sim \chi^2_{\sum_{i=1}^{n}\nu_i}$$

Bukti

Dengan menggunakan fungsi pembangkit momen

$$M_V(t) = M_{\sum_{i=1}^{n} Y_i}(t)$$

$$= \prod_{i=1}^{n} (1-2t)^{-\nu_i/2}$$

$$= (1-2t)^{-\sum_{i=1}^{n}\nu_i/2}$$

adalah fungsi pembangkit momen dari sebaran kai-kuadrat dengan derajat bebas $\sum_{i=1}^{n} v_i$. Dengan kekhasan sifat fungsi pembangkit momen, maka $V = \sum_{i=1}^{n} Y_i \sim \chi^2_{\sum_{i=1}^{n} v_i}$.

Bagaimana hubungan antara sebaran kai-kuadrat dengan sebaran normal baku beserta sifat-sifatnya juga akan dibahas dalam bab ini. Semua itu akan dimulai dengan teorema berikut ini.

Teorema 6. 6.

Jika $Z \sim N(0,1)$, maka

$$Z^2 \sim \chi^2_1$$

Bukti

$$\begin{aligned}
M_{Z^2}(t) &= E\left[e^{tZ^2}\right] \\
&= \int_{-\infty}^{+\infty} \frac{1}{\sqrt{2\pi}} e^{tz^2 - z^2/2} dz \\
&= \frac{1}{\sqrt{1-2t}} \int_{-\infty}^{+\infty} \frac{\sqrt{1-2t}}{\sqrt{2\pi}} e^{-z^2(1-2t)/2} dz \\
&= (1-2t)^{-1/2}
\end{aligned}$$

yang merupakan fungsi pembangkit momen dari sebaran kai-kuadrat dengan derajat bebas 1. Dengan kekhasan sifat fungsi pembangkit momen, maka $Z^2 \sim \chi^2_1$.

Teladan 6. 2.

Jika $X_1, \ldots X_n$ merupakan contoh acak dari sebaran normal, $N(\mu, \sigma^2)$ maka

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

$$\frac{n(\overline{X} - \mu)^2}{\sigma^2} \sim \chi_1^2$$

Penduga tak bias untuk varian, σ^2, adalah

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

telah kita bahas pada bab terdahulu, dan untuk contoh dari populasi normal, sebarannya dapat dihubungkan dengan sebaran kai-kuadrat. Sebaran S^2 ini tidak langsung mengikuti persamaan (6.16), karena bagian $X_i - \overline{X}$ tidak saling bebas. Bahkan secara fungsi mereka saling tergantung, karena $\sum_{i=1}^{n} (X_i - \overline{X}) = 0$

Teorema 6. 7.

Jika $X_1, ..., X_n$ merupakan contoh acak dari sebaran $N(\mu, \sigma^2)$, maka
1. $\overline{X}$ dan $X_i - \overline{X}$; $i = 1, 2, ..., n$ saling bebas.
2. $\overline{X}$ dan S^2 saling bebas.
3. $\frac{(n-1)S^2}{\sigma^2} \sim \chi_{n-1}^2$

Bukti
Sebagai latihan.

Sebaran t-student

Statistik S^2 digunakan untuk melakukan pendugaan parameter σ^2 pada sebaran normal. Demikian halnya, statistik $\overline{X}$ sangat berguna pada pendugaan parameter μ; namun demikian, sebaran dari $\overline{X}$ juga sangat tergantung pada σ^2. Hal ini yang membuat tidak mungkin menggunakan $\overline{X}$ untuk tipe prosedur statistika tertentu yang berkenaan dengan μ bilamana σ^2 tidak diketahui. Bilamana σ digantikan dengan *S* dalam statistik $\sqrt{n}(\overline{X}-\mu)/\sigma$, maka sebaran statistik $\sqrt{n}(\overline{X}-\mu)/S$ bukan lagi normal baku.

Teorema 6. 8.

Jika *Z ~ N(0,1)* dan *V ~* χ^2_ν dan jika *Z* dan *V* saling bebas, maka sebaran dari fungsi peubah acak

$$T = \frac{Z}{\sqrt{V/\nu}}$$

disebut dengan sebaran ***t-student*** dengan derajat bebas ν, dan dinotasikan dengan *T ~* t_ν. Fungsi kepekatan peluangnya adalah

$$f(t;\nu) = \frac{\Gamma\left(\frac{\nu+1}{2}\right)}{\Gamma\left(\frac{\nu}{2}\right)} \frac{1}{\sqrt{\nu\pi}} \left(1+\frac{t^2}{\nu}\right)^{-(\nu+1)/2}$$

Bukti
Sebagai latihan.

Sebaran *t* simetris terhadap titik nol, dan bentuk umumnya menyerupai sebaran normal. Apabila derajat bebasnya semakin membesar, maka sebaran *t* ini akan mendekati sebaran normal. Untuk derajat bebas yang kecil atau lebih kecil, sebaran *t* ini lebih datar dengan ekor yang lebih tebal, dan jika derajat bebasnya sama dengan 1, maka *T ~ Cauchy(1,0).*

Teorema 6. 9.

Jika *T ~* t_ν, maka untuk ν > *2r* akan diperoleh

$$E(T^{2r}) = \frac{\Gamma((2r+1)/2)\left[\Gamma((\nu-2r)/2)\right]^r}{\Gamma(1/2)\Gamma(\nu/2)}$$

$$E(T^{2r-1}) = 0 \qquad r = 1,2,..$$

$$\mathrm{var}(T) = \frac{\nu}{\nu-2} \qquad 2 < \nu$$

Bukti

Momen ke-$2r$ dari T adalah

$$E(T^{2r}) = E\left(\frac{Z}{\sqrt{V/\nu}}\right)^{2r} = E\left(Z^{2r}\right)E\left[\left(V/\nu\right)^{-r}\right]$$

dimana Z ~ $N(0,1)$ dan V ~ χ^2_ν dan dengan menggunakan persamaan (6.4) dan (6.8), kita peroleh (6.21). Persamaan (6.22) merupakan bentuk khusus dari (6.21), dan persamaan (6.23) lebih khusus lagi.

Teorema 6. 10.

Jika $X_1, ..., X_n$ menotasikan contoh acak berukuran n dari sebaran $N(\mu, \sigma^2)$, maka

$$\frac{\overline{X}-\mu}{S/\sqrt{n}} \sim t_{n-1}$$

Bukti

Menurut persyaratan pada (6.18) dan (6.19) , $Z = \sqrt{n}(\overline{X}-\mu)/\sigma \sim N(0,1)$ dan $\frac{(n-1)S^2}{\sigma^2} \sim \chi^2_{n-1}$, serta $\overline{X}$ dan S^2 saling bebas. Dengan demikian,

$$\left[\sqrt{n}(\overline{X}-\mu)/\sigma\right] / \left[\sqrt{\left(\frac{(n-1)S^2}{\sigma^2}\right)/(n-1)}\right] = \frac{\overline{X}-\mu}{S/\sqrt{n}} \sim t_{n-1}$$

Sebaran F-Snedecor

Salah satu sebaran yang dapat diturunkan dan penting dalam statistika adalah sebaran ***F-Snedecor*** atau yang lebih dikenal dengan sebaran ***F***. Peubah acak yang merupakan rasio dari dua peubah acak kai-kuadrat (dibagi dengan masing-masing derajat bebasnya terlebih dahulu) yang saling bebas, memiliki sebaran ***F*** ini.

Teorema 6. 11.

Jika $V_1 \sim \chi^2_{\nu_1}$ dan $V_2 \sim \chi^2_{\nu_2}$ dan saling bebas, maka peubah acak

$$X = \frac{V_1/\nu_1}{V_2/\nu_2}$$

Memiliki fungsi kepekatan peluang untuk $x > 0$

$$g(x;\nu_1,\nu_2) = \frac{\Gamma\left(\frac{\nu_1+\nu_2}{2}\right)}{\Gamma\left(\frac{\nu_1}{2}\right)\Gamma\left(\frac{\nu_2}{2}\right)}\left(\frac{\nu_1}{\nu_2}\right)^{\nu_1/2} x^{(\nu_2/2)-1}\left(1+\frac{\nu_1}{\nu_2}x\right)^{-(\nu_1+\nu_2)/2}$$

Fungsi kepekatan diatas dikenal sebagai fungsi kepekatan sebaran ***F*** dengan derajat bebas ν_1 dan ν_2. Sedangkan notasi yang digunakan untuk peubah acak yang memiliki sebaran ini adalah, $X \sim F(\nu_1, \nu_2)$. Fungsi kepekatan peluang diatas dapat diturunkan seperti menurunkan fungsi kepekatan sebaran-***t***.

Teorema 6. 12.

Jika $X \sim F(\nu_1, \nu_2)$, maka

$$E(X^r) = \frac{\left(\frac{\nu_2}{\nu_1}\right)^r \Gamma\left(\frac{\nu_1}{2}+r\right)\Gamma\left(\frac{\nu_2}{2}-r\right)}{\Gamma\left(\frac{\nu_1}{2}\right)\Gamma\left(\frac{\nu_2}{2}\right)} \qquad \nu_2 > 2r$$

$$E(X) = \frac{\nu_2}{\nu_2 - 2} \qquad \nu_2 > 2$$

$$Var(X) = \frac{2\nu_2^2(\nu_1+\nu_2-2)}{\nu_1(\nu_2-2)^2(\nu_2-4)} \qquad \nu_2 > 4$$

Bukti

Sebagai latihan.

Persentil $f_\gamma(\nu_1,\nu_2)$ dari *X~F(*ν_1,ν_2*)* sedemikian rupa sehingga $P\left[X \le f_\gamma(\nu_1,\nu_2)\right] = \gamma$ biasanya ditabelkan untuk beberapa nilai γ, ν_1 dan ν_2 terpilih.

Jika *X~F(*ν_1,ν_2*)*, maka *Y = 1/X ~F(*ν_2,ν_1*)*. Dengan demikian,

$$\begin{aligned} 1-\gamma &= P\left[X < f_{1-\gamma}(\nu_1,\nu_2)\right] \\ &= P\left[Y > \frac{1}{f_{1-\gamma}(\nu_1,\nu_2)}\right] \\ &= 1 - P\left[Y \le \frac{1}{f_{1-\gamma}(\nu_1,\nu_2)}\right] \end{aligned}$$

sehingga

$$\frac{1}{f_{1-\gamma}(\nu_1,\nu_2)} = f_\gamma(\nu_2,\nu_1)$$

$$f_{1-\gamma}(\nu_1,\nu_2) = \frac{1}{f_\gamma(\nu_2,\nu_1)}$$

Sebaran Beta

Transformasi dari peubah acak yang memiliki sebaran ***F*** dapat dilakukan. Jika $X \sim F(\nu_1, \nu_2)$, maka peubah acak

$$Y = \frac{(\nu_1 / \nu_2)X}{1 + (\nu_1 / \nu_2)X}$$

memiliki fungsi kepekatan peluang

$$f(y;a,b) = \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} y^{a-1}(1-y)^{b-1} \qquad 0 < y < 1$$

dimana $a = \nu_1/2$ dan $b = \nu_2/2$. Fungsi kepekatan peluang ini mendefinisikan sebatan ***Beta*** dengan parameter $a > 0$ dan $b > 0$, yang dinotasikan dengan $Y \sim Beta(a,b)$.

Sifat-sifat yang dimiliki (nilai harapan dan ragam) peubah yang menyebar *Beta(a,b)* dapat ditunjukkan seperti di bawah ini.

$$E(Y) = \frac{a}{a+b}$$

$$Var(Y) = \frac{ab}{(a+b+1)(a+b)^2}$$

Persentil ke-γ dari sebaran Beta ini dapat diekspresikan dalm persentil sebaran *F* sebagai akibat dari persamaan sebelum ini.

$$B_\gamma(a,b) = \frac{af_\gamma(2a,2b)}{b + af_\gamma(2a,2b)}$$

Jika *a* dan *b* merupakan bilangan bulat positif, maka pengintegralan parsial terus menerus akan menghasilkan hubungan antara sebaran kumulatif Beta dengan sebaran Binomial. Jika $X \sim Bin(n,p)$ dan $Y \sim Beta(n-i+1,i)$, maka $F_X(i-1) = F_Y(1-p)$.

Sebaran Beta juga sering digunakan dalam kaitannya dengan sebaran statistik tataan (*order statistic*). Untuk sebuah peubah acak yang kontinu $X \sim f(x)$, fungsi kepekatan peluang statistik tataan ke-*k* dari suatu contoh acak berukuran *n* adalah

$$g_k(x_{(k)}) = \frac{n!}{(k-1)!(n-k)!}\left[F(x_{(k)})\right]^{k-1}\left[1-F(x_{(k)})\right]^{n-k} f(x_{(k)})$$

Dengan merubah menjadi $U_{(k)} = F(X_{(k)})$, maka $U_{(k)} \sim Beta(k, n-k+1)$.

Dengan menggunakan *Probability Integral Transform* kita peroleh bahwa *U = F(X) ~ Seragam(0,1).* Juga $U_{(k)}$ melambangkan statistik tataan ke-*k* dari sebaran Seragam. Fungsi sebaran kumulatif dari $X_{(k)}$ dapat diekspresikan dalam bentuk fungsi sebaran kumulatif Beta, karena

$$\begin{aligned} G_k(x_{(k)}) &= P\left[X_{(k)} \le x_{(k)}\right] \\ &= P\left[F(X_{(k)}) \le F(x_{(k)})\right] \\ &= H(F(x_{(k)}); k, n-k+1) \end{aligned}$$

dimana *H(y;a,b)* merupakan fungsi sebaran kumulatif dari *Y ~ Beta(a,b)*.

Teladan 6. 3.

Misalkan *X ~ Eksponensial*(θ), dan kita ingin menghitung peluang yang berkenaan dengan statistik tataan ke-*k*. Kita dapat peroleh bahwa $F(x) = 1-e^{-x/\theta}$ dan $U_{(k)} = F(X_{(k)}) \sim Beta(k,n-k+1)$, sehingga

$$\begin{aligned} P[X_{(k)} \le c] &= P[F(X_{(k)}) \le F(c)] \\ &= P[U_{(k)} \le F(c)] \\ &= P\left[\frac{kU_{(k)}}{(n-k+1)(1-U_{(k)})} \le \frac{kF(c)}{(n-k+1)(1-F(c))}\right] \end{aligned}$$

dimana peluang terakhir mencakup peubah yang menyebar menurut sebaran ***F(2k,2(n-k+1))***. Dengan demikian, untuk nilai-nilai θ, *c, k,* dan *n*, nilai peluang ini dapat diperoleh dari tabel kumulatif Beta atau tabel kumulatif *F* jika taraf α untuk tabel ini tersedia.

Teladan 6. 4.

Sebagai ilustrasi dari teladan sebelum ini, kita ingin mencari c_γ sedemikian rupa sehingga $P[X_{(k)} < c_\gamma] = \gamma$, maka

$$f_\gamma(2k,2(n-k+1)) = \frac{kF(c_\gamma)}{(n-k+1)[1-F(c_\gamma)]}$$

$$= \frac{k(1-\exp(-c_\gamma/\theta))}{(n-k+1)\exp(-c_\gamma/\theta)}$$

dan dengan demikian kita peroleh

$$c_\gamma = \theta \ln\left[1+\frac{(n-k+1)f_\gamma(2k,2(n-k+1))}{k}\right]$$

Kita telah pelajari bahwa sebaran Beta memiliki hubungan dengan sebaran *F* dan Binomial. Sebaran Beta merepresentasikan sebaran Seragam secara lebih umum, dan memberikan model dua parameter yang lebih fleksibel untuk berbagai tipe peubah acak yang mengambil nilai diantara 0 dan 1.

Dalam bab ini, juga telah kita pelajari sebaran Normal dengan berbagai sifat dan beberapa sebaran yang berhubungan atau dapat diturunkan dari sebaran normal.

Latihan

1. Misalkan *X* melambangkan berat dalam kilogram dari satu

kantong biji-bijian, dimana $X \sim N(110,9)$. Berapakah peluang bahwa 10 kantong beratnya akan melebihi 1 ton ?

2. Misalkan bahwa $X \sim N(2,9)$
   a. Carilah $E(X-2)4$
   b. Dapatkan $E(X4)$
3. Asumsikan bahwa X_1 dan X_2 merupakan peubah acak-peubah acak yang saling bebas, dengan $X_i \sim N(\mu,\sigma^2)$, dan misalkan $Y_1 = X_1 + X_2$ dan $Y_2 = X_1 - X_2$. Tunjukkan bahwa Y_1 dan Y_2 menyebar normal dan saling bebas.
4. Sebuah komponen elektronik mulai dipakai dan tujuh komponen pengganti untuk itu telah tersedia. Umur komponen T_i, (dalam hari) menyebar saling bebas Eksponensial(150).
   a. Tentukan apa sebaran dari $\sum_{i=1}^{8} T_i$?
   b. Hitunglah peluang bahwa alat yang menggunakan komponen tersebut dapat beroperasi seditnya satu tahun ?
   c. Berapa banyak komponen pengganti diperlukan untuk menjamin bahwa kita yakin 95% alat tersebut akan berfungsi selama 1,5 tahun ?
   d. Ulangi soal a, b, dan c jika $T_i \sim Gamma(150, 1,5)$?
5. Misalkan bahwa $X \sim \chi^2_m$ dan $Y \sim \chi^2_n$ serta X dan Y saling bebas. Apakah $X - Y \sim \chi^2$ jika $m>n$?
6. Jarak dalam meter seorang atlit terjun payung tidak mengenai target penerjunan merupakan peubah acak, yaitu $D = \sqrt{X_1^2 + X_2^2}$ dimana $X_i \sim N(0,16)$. Carilah peluang bahwa penerjun tersebut akan meleset dari sasaran paling jauh 5 meter.
7. Misalkan peubah acak-peubah acak $Z_i \sim N(0,1)$, $i = 1,2,...,9$, dan misalkan juga $\bar{Z}$ adalah rata-rata contoh. Carilah
   a. $P[\bar{Z} < 1/3]$
   b. $P[Z_1 - Z_2 < 1]$
   c. $P[Z_1 + Z_2 < 1]$

d. $P\left[\sum_{i=1}^{9} Z_i^2 < 18\right]$

e. $P\left[\sum_{i=1}^{9} (Z_i - \overline{Z})^2 < 18\right]$

8. Jika $T \sim t_\nu$, tentukan sebaran dari T^2.
9. Jika $X \sim Beta(a,b)$, carilah $E(X^n)$
10. Ingat kembali bahwa $Y \sim LogNormal(\mu,\sigma^2)$ jika $\ln Y \sim Normal(\mu,\sigma^2)$. Asumsikan bahwa $Y_i \sim LogNormal(\mu_i, \sigma_i^2)$ menyebar saling bebas, $i = 1, 2, ..., n$. Carilah sebaran dari peubah acak-peubah acak berikut:

    a. $\prod_{i=1}^{n} Y_i$

    b. $\prod_{i=1}^{n} Y_i^{a_i}$

    c. Y_1/Y_2

    d. Carilah juga $E\left[\prod_{i=1}^{n} Y_i\right]$

11. Misalkan bahwa $X_i \sim Normal(\mu,\sigma^2)$, $i = 1,2, ...,n$ dan $Z_i \sim N(0,1)$, $i = 1,2, ..., k$. Dan semua peubah acak saling bebas. Nyatakan sebaran dari setiap fungsi peubah acak berikut dengan memberi "nama" sebaran jika mungkin, jika tidak nyatakan dengan "tak diketahui namanya".

    a. $X_2 - X_4$

    b. $X_1 + 2X_3$

    c. $\dfrac{X_2 - X_1}{\sigma S_Z \sqrt{2}}$

    d. $\sum_{i=1}^{n} a_i X_i$ dimana a_i, $i = 1, 2, ..., n$ adalah konstanta.

    e. Z_{234}^2

f. $\dfrac{\sqrt{n}(\overline{X}-\mu)}{\sigma S_Z}$

g. $Z_1^2 + Z_2^2 + Z_3^2$

h. $Z_1^2 - Z_2^2 - Z_3^2$

i. $\dfrac{Z_7}{\sqrt{Z_{10}^2}}$

j. $\dfrac{Z_3^2}{Z_4^2}$

k. $\dfrac{Z_2}{Z_1}$

l. $\dfrac{\overline{X}}{\overline{Z}}$

m. $\dfrac{\sqrt{nk}(\overline{X}-\mu)}{\sigma\sqrt{\sum_{i=1}^{k} Z_i^2}}$

n. $k\overline{Z}^2$

o. $\dfrac{\sum_{i=1}^{n}(X_i-\mu)^2}{\sigma^2} + \sum_{i=1}^{k}(Z_i-\overline{Z})^2$

p. $\dfrac{\overline{X}}{\sigma^2} + \dfrac{\sum_{i=1}^{k} Z_i}{k}$

q. $\dfrac{(k-1)\sum_{i=1}^{n}(X_i-\overline{X})^2}{(n-1)\sigma^2\sum_{i=1}^{k}(Z_i-\overline{Z})^2}$

12. Jika *F(x;λ)* adalah merupakan fungsi sebaran kumulatif dari *X* ~ *Poisson(λ)*, dan jika *H(y;ν)* adalah merupakan fungsi

sebaran kumulatif kai-kuadrat dengan derajat bebas ν, maka $F(x;\lambda) = 1 - H[2\lambda; 2(x+1)]$.

Limit Sebaran

Dalam beberapa bab terdahulu telah dibicarakan metode yang umum untuk menurunkan sebaran dari fungsi peubah acak-peubah acak, dalam bentuk $Y_n = u(X_1, \ldots, X_n)$. Dalam beberapa kasus, fungsi sebaran peluang dari Y_n dengan mudah diperoleh, namun dalam banyak kasus penting, penurunan sangat rumit dilakukan. Dalam hal ini, dimungkinkan untuk mendapatkan hasil pendekatan yang berguna apabila n sangat besar. Tentu saja, hasil ini berdasarkan teori konvergensi dalam sebaran dan limit sebaran.

Deretan Peubah Acak

Pertimbangkan suatu deretan atau barisan peubah acak $Y_1, Y_2, \ldots$ dengan fungsi sebaran kumulatifnya $G_1(y), G_2(y), \ldots$ sehingga untuk setiap $n = 1, 2, \ldots$ berlaku

$$G_n(y) = P[Y_n \le y]$$

Definisi 7. 1.

Jika $Y_n \sim G_n(y)$ [dibaca: Y_n memiliki fungsi sebaran kumulatif $G_n(y)$] untuk setiap $n = 1, 2, \ldots.$, dan jika untuk beberapa fungsi sebaran kumulatif $G(y)$, diperoleh bahwa

$$\lim_{n \to \infty} G_n(y) = G(y)$$

untuk semua nilai y dimana $G(y)$ kontinu, maka deretan $Y_1, Y_2, \ldots.$, dikatakan **konvergen dalam sebaran** ke $Y \sim G(y)$, yang dinotasikan dengan $Y_n \xrightarrow{d} Y$. $G(y)$ sering disebut dengan **limit sebaran** dari Y_n.

Teladan 7. 1.

Misalkan $X_1, \ldots, X_n$ merupakan contoh acak dari suatu sebaran seragam, X_i ~ Seragam(0,1) dan misalkan $Y_n = X_{(n)} = max(X_1, X_2, \ldots, X_n)$. Dengan demikian fungsi sebaran kumulatif dari Y_n adalah

$$G_n(y) = y^n \qquad 0 < y < 1$$

nol jika $y \leq 0$ dan satu jika $y \geq 1$. Sudah tentu, apabila $0 < y < 1$, y^n akan mendekati 0 apabila n mendekati ∞, dan jika $y \leq 0$, $G_n(y) \rightarrow 0$, serta jika $y \geq 1$, $G_n(y) \rightarrow 1$. Dengan demikian,

$$\lim_{n\to\infty} G_n(y) = G(y) = \begin{cases} 0 & y < 1 \\ 1 & y \geq 1 \end{cases}$$

Definisi 7. 2.

Fungsi $G(y)$ merupakan fungsi sebaran kumulatif dari suatu sebaran **degenerate** pada nilai $y = c$, jika

$$G(y) = \begin{cases} 0 & y < c \\ 1 & y \geq c \end{cases}$$

Dengan perkataan lain, $G(y)$ merupakan fungsi sebaran kumulatif dari suatu sebaran diskrit, dimana peluang $y = c$ bernilai satu, dan nol untuk y selainnya.

Teladan 7. 2.

Misalkan $X_1, \ldots, X_n$ merupakan contoh acak dari suatu sebaran Eksponensial, $X_i \sim Exp(\theta)$, dan misalkan juga $Y_n = X_{(1)} = min(X_1, X_2, \ldots, X_n)$. Dengan demikian fungsi sebaran kumulatif dari Y_n adalah

$$G_n(y) = 1 - e^{-ny/\theta} \qquad y > 0$$

dan 0 untuk y selainnya. Untuk $y > 0$, maka $\lim_{n\to\infty} G_n(y) = 1$.

Dengan demikian, nilai limitnya 0 jika $y < 0$ dan 1 jika $y > 0$, yang merupakan sebaran degenerate pada $y = 0$. Perhatikan bahwa nilai limit pada $y = 0$ adalah 0, yang berarti bahwa fungsi sebaran tidak hanya diskontinu pada $y = 0$, tetapi juga tidak kontinu dari kanan pada $y = 0$, yang merupakan salah satu persyaratan dari fungsi sebaran kumulatif. Ini bukan merupakan masalah, karena berdasarkan definisi di awal bab ini hanya limit fungsi yang harus sesuai atau setuju dengan fungsi sebaran kumulatif pada semua titik-titik kontinu.

Definisi 7. 3.

Suatu deretan peubah acak $Y_1, Y_2, \ldots.$, dikatakan konvergen stokastik atau konvergen dalam peluang ke suatu konstanta c jika deretan tersebut memiliki limit sebaran yang degenerate pada $y = c$.

Tidak semua limit sebaran bersifat degenerate. Limit-limit berikut sangat berguna dalam beberapa penyelesaian permasalahan.

$$\lim_{n\to\infty}\left(1+\frac{c}{n}\right)^{nb} = e^{cb}$$

$$\lim_{n\to\infty}\left(1+\frac{c}{n}+\frac{d(n)}{n}\right)^{nb} = e^{cb} \quad \text{jika} \quad \lim_{n\to\infty} d(n) = 0$$

Teladan 7. 3.

Misalkan $X_1, \ldots, X_n$ merupakan contoh acak dari suatu sebaran Pareto, $X_i \sim Par(1,1)$ dan misalkan $Y_n = nX_{(1)}$. Fungsi sebaran kumulatif X_i adalah $F(x) = 1-(1+x)^{-1} \quad x > 0$, maka fungsi sebaran kumulatif Y_n adalah

$$G_n(y) = 1-\left(1+\frac{y}{n}\right)^{-n} \qquad y > 0$$

Untuk $n \to \infty$, diperoleh

$$\lim_{n\to\infty} G_n(y) = \lim_{n\to\infty} 1-\left(1+\frac{y}{n}\right)^{-n} = 1-e^{-y} = G(y)$$ dan 0 untuk y selainnya, yang merupakan sebaran Eksponensial, Exp(1).

Deretan peubah acak tidak harus memiliki limit sebaran. Perhatikan teladan berikut.

Teladan 7. 4.

Misalkan $X_1, \ldots, X_n$ merupakan contoh acak dari suatu sebaran Pareto, $X_i \sim Par(1,1)$ dan misalkan $Y_n = X_{(n)}$. Fungsi sebaran kumulatif bagi Y_n adalah

$$G_n(y) = \left(\frac{y}{1+y}\right)^n \qquad y > 0$$

dan nol selainnya. Karena $y/(1+y) < 1$, maka $\lim_{n\to\infty} G_n(y) = \lim_{n\to\infty}\left(\frac{y}{1+y}\right)^{-n} = 0 = G(y)$ untuk semua y. Dengan demikian ini bukan merupakan fungsi sebaran kumulatif, karena nilainya tidak mendekati 1 bilamana nilai y menuju tak hingga.

Teladan 7. 5.

Misalkan $X_1, \ldots, X_n$ merupakan contoh acak dari suatu sebaran Pareto, $X_i \sim Par(1,1)$ dan misalkan $Y_n = (1/n)X_{(n)}$. Fungsi sebaran kumulatif bagi Y_n adalah

$$G_n(y) = \left(1 + \frac{1}{ny}\right)^{-n} \qquad y > 0$$

dan nol selainnya. Selanjutnya, untuk $\lim_{n\to\infty} G_n(y) = \lim_{n\to\infty}\left(1 + \frac{1}{ny}\right)^{-n} = e^{-1/y} = G(y)$, untuk $y > 0$.

Teladan 7. 6.

Misalkan $X_1, \ldots, X_n$ merupakan contoh acak dari suatu sebaran Pareto, $X_i \sim Par(1,1)$ dan misalkan $Y_n = (1/\theta)X_{(n)} - \ln n$. Fungsi sebaran kumulatif bagi Y_n adalah

$$G_n(y) = \left[1 - \frac{1}{n}e^{-y}\right]^n \qquad y > -\ln n$$

dan nol selainnya. $\lim_{n\to\infty} G_n(y) = \exp(-e^{-y}) = G(y)$ untuk $-\infty<y<\infty$.

Teladan 7. 7.

Suatu contoh acak berukuran n diambil dari sebaran normal, $X_i \sim N(\mu, \sigma^2)$. Rata-rata contoh, $Y_n = \overline{X}_n$ memiliki distribusi Normal(μ, σ^2/n), dan

$$G_n(y) = \Phi\left(\frac{\sqrt{n}(y-\mu)}{\sigma}\right)$$

Limit fungsi sebaran kumulatif ini, bisa ditunjukkan, degenerate pada $y = \mu$, karena $\lim_{n\to\infty} G_n(y) = \lim_{n\to\infty} \Phi\left(\frac{\sqrt{n}(y-\mu)}{\sigma}\right) = \begin{cases} 0 & y<\mu \\ 1/2 & y=\mu \\ 1 & y>\mu \end{cases}$,

sehingga rata-rata contoh konvergen stokastik ke μ.

Pendekatan Fungsi Pembangkit Momen

Barusaja dibicarakan teladan-teladan fungsi sebaran kumulatif yang pasti dan diketahui untuk nilai *n* yang terhingga, dan limit sebaran diperoleh secara langsung dari deretan sebaran. Salah satu keuntungan dari limit sebaran adalah dimungkinkannya penentuan limit sebaran tanpa harus mengetahui bentuk pasti dari fungsi sebaran kumulatif untuk *n* terhingga. Dengan demikian limit sebaran dapat memberikan pendekatan yang bermanfaat bilamana peluang yang pasti tak mungkin diperoleh. Salah satu cara untuk itu adalah dengan menggunakan fungsi pembangkit momen.

Teorema 7. 1.

Misalkan $Y_1, Y_2, \ldots$ adalah deret peubah acak dengan fungsi sebaran kumulatif berturut-turut $G_1(y), G_2(y), \ldots$ dan fungsi pembangkit momen berturut-turut $M_1(t), M_2(t), \ldots$. Jika $M(t)$ adalah fungsi pembangkit momen dari fungsi sebaran kumulatif $G(y)$, dan jika $\lim_{n\to\infty} M_n(t) = M(t)$ untuk semua t dalam interval terbuka yang mencakup nol, $-h < t < h$, maka $\lim_{n\to\infty} G_n(y) = G(y)$ untuk semua titik kontinu $G(y)$.

Teladan 7. 8.

Misalkan $X_1, X_2, \ldots, X_n$ merupakan contoh acak dari sebaran Bernoulli, $X_i \sim Bin(1,p)$ dan $Y_n = \sum_{i=1}^{n} X_i$. Jika misalkan $p \to 0$ bilamana $n \to \infty$ sedemikian rupa sehingga $np = \mu$, untuk nilai $\mu > 0$ yang tetap, maka

$$M_n(t) = (pe^t + q)^n = \left(\frac{\mu e^t}{n} + 1 - \frac{\mu}{n}\right)^n = \left[1 + \frac{\mu(e^t - 1)}{n}\right]^n$$

dan sebagaimana diketahui, jika $n \to \infty$, diperoleh

$$\lim M_n(t) = e^{\mu(e^t - 1)}$$

yang merupakan fungsi pembangkit momen sebaran Poisson dengan parameter μ. Dengan demikian, dapat kita simpulkan bahwa $Y_n \xrightarrow{d} Y \sim$Poisson($\mu$).

Teladan 7. 9. .Hukum Bilangan Besar Bernoulli.

Misalkan jika kita pertahankan nilai p yang tetap dan kita perhatikan deretan proporsi contoh, $W_n = \hat{p}_n = Y_n / n$. Dengan menggunakan ekspansi deret $e^u = 1 + u + u^2/2 + \ldots$ dengan $u = t/n$, kita peroleh

$$M_{W_n}(t) = (pe^{t/n} + q)^n = \left[p\left(1 + \frac{t}{n} + \ldots\right) + q\right]^n = \left[1 + \frac{pt}{n} + \frac{d(n)}{n}\right]^n$$

dimana *d(n)/n* merupakan sisa deret yang bisa diabaikan, *d(n)* → *0* bilamana *n* → ∞. Dan dengan demikian, $\lim M_{W_n}(t) = e^{pt}$ yang merupakan fungsi pembangkit momen dari sebaran degenerate pada *y* = *p*, dan dengan demikian $\hat{p}_n$ → p dalam peluang bilamana *n* → ∞.

Teladan 7. 10.

Perhatikan deretan peubah yang "dibakukan", yaitu $Z_n = \frac{Y_n - np}{\sqrt{npq}} = \frac{Y_n}{\sigma_n} - \frac{np}{\sigma_n}$. Dengan menggunakan ekspansi deret,

$$\begin{aligned} M_{Z_n}(t) &= e^{-npt/\sigma_n}(pe^{t/\sigma_n} + q)^n \\ &= \left[e^{-pt/\sigma_n}(pe^{t/\sigma_n} + q)\right]^n \\ &= \left[\left(1 - \frac{pt}{\sigma_n} + \frac{p^2t^2}{2\sigma_n^2} + \ldots\right)\left(1 + \frac{pt}{\sigma_n} + \frac{p^2t^2}{2\sigma_n^2} + \ldots\right)\right]^n \\ &= \left[1 + \frac{t^2}{2n} + \frac{d(n)}{n}\right]^n \end{aligned}$$

dimana *d(n)*→*0* bilamana *n*→∞. Dengan demikian,

$$\lim_{n\to\infty} M_{Z_n}(t) = e^{t^2/2}$$

yang merupakan fungsi pembangkit momen sebaran normal baku, dan dengan demikian $Z_n \xrightarrow{d} Z$ ~ N(0,1).

Teorema 7. 2. Limit Pusat

Jika $X_1, \ldots, X_n$ merupakan contoh acak dari suatu sebaran dengan rata-rata μ dan ragam $\sigma^2 < \infty$, maka limit sebaran dari

$$Z_n = \frac{\sum_{i=1}^{n} X_i - n\mu}{\sqrt{n}\sigma}$$

Adalah normal baku. Atau dengan kata lain $Z_n \xrightarrow{d} Z \sim$ N(0,1) bilamana $n \rightarrow \infty$.

Bukti

Hasil limit ini cocok untuk contoh acak contoh acak dari sembarang sebaran dengan rata-rata dan ragam yang terhingga, namun pembuktian akan dikhususkan dengan adanya asumsi yang lebih kuat yaitu memiliki fungsi pembangkit momen.

Misalkan *m(t)* adalah fungsi pembangkit momen dari $X - \mu$, $m(t) = M_{X-\mu}(t)$, dan sebagai catatan bahwa $m(0) = 1$, $m'(0) = E(X - \mu) = 0$, dan $m''(0) = E(X - \mu)^2 = \sigma^2$. Dengan menggunakan ekspansi Taylor, kita peroleh

$$\begin{aligned} m(t) &= m(0) + m'(0)t + \frac{m''(\xi)t^2}{2} \qquad 0 < \xi < t \\ &= 1 + \frac{m''(\xi)t^2}{2} \\ &= 1 + \frac{\sigma^2 t^2}{2} + \frac{(m''(\xi) - \sigma^2)t^2}{2} \end{aligned}$$

Sekarang,

$$Z_n = \frac{\sum_{i=1}^{n} X_i - n\mu}{\sqrt{n}\sigma} = \frac{\sum_{i=1}^{n} (X_i - \mu)}{\sqrt{n}\sigma}$$

dan

$$\begin{aligned} M_{Z_n}(t) &= M_{\sum (X_i-\mu)}\left(\frac{t}{\sqrt{n}\sigma}\right) \\ &= \prod_{i=1}^{n} M_{X_i-\mu}\left(\frac{t}{\sqrt{n}\sigma}\right) \\ &= \left[M_{X-\mu}\left(\frac{t}{\sqrt{n}\sigma}\right)\right]^n \end{aligned}$$

$$= \left[m\left(\frac{t}{\sqrt{n}\sigma} \right) \right]^n$$

$$= \left[1 + \frac{\sigma^2 t^2}{2n\sigma^2} + \frac{(m''(\xi) - \sigma^2)t^2}{2n\sigma^2} \right]^n \qquad 0 < \xi < \frac{t}{\sqrt{n}\sigma}$$

Apabila $n \rightarrow \infty$, $t / \sqrt{n}\sigma \rightarrow 0$, $\xi \rightarrow 0$, dan $m''(\xi)-\sigma^2 \rightarrow 0$, maka

$$M_{Z_n}(t) = \left[1 + \frac{t^2}{2n} + \frac{d(n)}{n} \right]^n$$

dimana $d(n) \rightarrow 0$ apabila $n \rightarrow \infty$. Dengan demikian

$$\lim_{n \to \infty} M_{Z_n}(t) = e^{t^2/2}$$

atau

$$\lim_{n \to \infty} F_{Z_n}(z) = \Phi(z)$$

yang berarti bahwa $Z_n \xrightarrow{d} Z$ ~ N(0,1).

Sebagai catatan bahwa peubah $Z_n = \dfrac{\sum_{i=1}^{n} X_i - n\mu}{\sqrt{n}\sigma}$ dapat juga dihubungkan dengan peubah $Z_n = \dfrac{\overline{X}_n - \mu}{\sigma / \sqrt{n}}$.

Teorema 7. 3.

Jika $Y_\nu \sim \chi^2(\nu)$, maka $Z_\upsilon = \dfrac{Y_\upsilon - \upsilon}{\sqrt{2\upsilon}} \xrightarrow{d} Z$ ~ N(0,1) bilamana $\nu \rightarrow \infty$.

Bukti Dengan menggunakan Teorema limit pusat, dimana Y_ν adalah sebaran jumlah peubah acak yang menyebar bebas stokastik dan identik masing-masing kai-kuadrat dengan derajat bebas 1.

Teladan 7. 11.

Misalkan $X_1, X_2, \ldots, X_n$ merupakan contoh acak dari sebaran seragam kontinu, X_i ~ Seragam(0,1), dan misalkan juga $Y_n = \sum_{i=1}^{n} X_i$. Dari properti distribusi, diperoleh bahwa $E(X_i) = ½$ dan $Var(X_i) = 1/12$. Dengan demikian kita peroleh $Y_n \sim N(n/2,n/12)$.

Pendekatan Sebaran Binomial

Permasalahan yang harus diselesaikan dengan sebaran Binomial untuk nilai *p* yang tetap dan *n* yang cukup besar, dapat dikerjakan dengan menggunakan pendekatan sebaran normal. Pada kondisi demikian Y_n ~ Normal(*np*, *npq*). Pendekatan ini akan bagus untuk nilai *p* yang mendekati 0,5 karena untuk nilai p = 0,5 distribusi akan simetris. Tingkat akurasi yang diperlukan dalam suatu pendekatan tergantung pada aplikasinya. Pendekatan Normal digunakan apabila $np \geq 5$ dan $nq \geq 5$, dan sekali lagi hal ini tergantung pada tingkat akurasi yang dikehendaki.

Teladan 7. 12.

Peluang seorang pemain basket memasukkan bola ke dalam ring atau berhasil menembakkan bola ke sasaran yaitu *p = 0,5*. Jika ia melakukan 20 tembakan beruntun, berapakan peluang ia memasukkan sedikitnya 9 bola ?

Peluang pastinya, bisa dihitung sebagai berikut:

$$
\begin{aligned}
P[Y_{20} \geq 9] &= 1-P[Y_{20} \leq 8] \\
&= 1 - \sum_{y=0}^{8} C_y^{20}\, 0{,}5^{y}\, 0{,}5^{20-y} \\
&= 0{,}7483
\end{aligned}
$$

Apabila dicari dengan menggunakan pendekatan normal, maka akan diperoleh

$$P[Y_{20} \geq 9] = 1\text{-}P[Y_{20} \leq 8]$$
$$\cong 1 - \Phi\left(\frac{8-10}{\sqrt{5}}\right)$$
$$= 1 - \Phi(-0{,}89)$$
$$= 0{,}8133$$

Karena sebaran binomial tergolong sebaran diskrit dan sebaran normal tergolong sebaran kontinu, nilai pendekatan dapat diperbaiki dengan menggunakan apa yang dinamakan **koreksi kekontinuan** (*continuity correction*).

Misalkan untuk menghitung peluang memasukkan tepat 7 bola, atau tembakan yang mengenai sasaran sebanyak 7 kali dari 20 lemparan seperti teladan diatas, maka bila dihitung dengan menggunakan sebaran binomial akan diperoleh

$$P[Y_{20} = 7] = C_7^{20}\, 0{,}5^7\, 0{,}5^{13} = 0{,}0739$$

Sedangkan apabila didekati dengan sebaran normal, diperolah

$$P[Y_{20} = 7] = \Phi\left(\frac{7{,}5-10}{\sqrt{5}}\right) - \Phi\left(\frac{6{,}5-10}{\sqrt{5}}\right)$$
$$= \Phi(-1{,}12) - \Phi(-1{,}57) = 0{,}0732$$

Dan sebagaimana hal ini, maka

$$P[Y_{20} \geq 9] = 1\text{-}P[Y_{20} \leq 8]$$
$$\cong 1 - \Phi\left(\frac{8{,}5-10}{\sqrt{5}}\right)$$
$$= 1 - \Phi(-0{,}67)$$
$$= 0{,}7486$$

Jadi, secara umum jika Y_n ~ *Binomial(n,p)* dan bilangan bulat $a \leq b$, maka

$$P[a \leq Y_n \leq b] \cong \Phi\left(\frac{b + 0{,}5 - np}{\sqrt{npq}}\right) - \Phi\left(\frac{a - 0{,}5 - np}{\sqrt{npq}}\right)$$

Koreksi kekontinuan juga sangat bermanfaat dalam penyelesaian sebaran diskret lainnya, bila didekati dengan sebaran normal.

Sebaran Asimtotik Normal

Dengan mengikuti Teorema limit pusat bahwa rata-rata contoh yang dibakukan, maka akan memiliki limit sebaran yaitu normal baku, atau normal dengan rata-rata 0 dan ragam 1.

Definisi 7. 4.

Jika $Y_1, Y_2, \ldots$ merupakan deretan peubah acak dan *m* serta *c* adalah konstanta sedemikian rupa sehingga

$$Z_n = \frac{Y_n - m}{c/\sqrt{n}} \xrightarrow{d} Z \sim N(0,1)$$

Jika $n \to \infty$, maka Y_n dikatakan memiliki sebaran asimtotik normal dengan asimtot rata-rata *m* dan asimtot ragam c^2/n.

Teladan 7. 13.

Contoh acak sebanyak 50 komponen elektronik, dimana umur komponen tersebut masing-masing memiliki distribusi Eksponensial(80). Dengan menggunakan Teorema limit pusat, $\overline{X}_n$ memiliki sebaran asimtotik normal dengan asimtot rata-rata m = 80 dan ragam $c^2/n = (80)^2/50 = 128$.

Sifat-Sifat Konvergen Stokastik

Dalam beberapa bagian terdahulu kita peroleh teladan dimana suatu deret peubah acak konvergen stokastik atau konvergen dalam peluang ke suatu konstanta. Seperti misalnya, proporsi contoh konvergen stokastik ke proporsi populasi. Jelas kiranya, konsep umum ini sangat berguna untuk evaluasi penduga dari parameter yang tak diketahui. Salah satu penduga yang baik adalah penduga yang memiliki sifat konvergen stokastik ke nilai parameter bilamana ukuran sampel atau contoh menjadi sangat besar.

Teorema 7. 4.

Deret $Y_1, Y_2, \ldots$ konvergen stokastik ke c jika dan hanya jika untuk setiap $\varepsilon > 0$,

$$\lim_{n\to\infty} P\left[\left|Y_n - c\right| < \varepsilon\right] = 1$$

Deretan peubah acak yang memenuhi hal diatas juga disebut konvergen dalam peluang ke suatu konstanta c, yang dilambangkan dengan $Y_n \xrightarrow{P} c$.

Teladan 7. 14.

Dalam teladan Hukum Bilangan Besar Bernoulli kita gunakan pendekatan fungsi pembangkit momen. Tetapi dapat juga digunakan dengan menggunakan Teorema terakhir dan pertidaksamaan Chebychev. Kita dapatkan bahwa $E(\hat{p}_n) = p$ dan $Var(\hat{p}_n) = pq/n$, sehingga

$$P\left[\left|\hat{p}_n - p\right| < \varepsilon\right] \geq 1 - \frac{pq}{\varepsilon^2 n}$$

Untuk sembarang $\varepsilon > 0$, $\lim_{n\to\infty} P\left[\left|\hat{p}_n - p\right| < \varepsilon\right] = 1$

Teorema 7. 5.

Jika $X_1, \ldots, X_n$ adalah contoh acak dari suatu sebaran dengan rata-rata $\mu < \infty$ dan ragam $\sigma^2 < \infty$, maka deret rata-rata contoh konvergen dalam peluang ke μ, $\overline{X}_n \xrightarrow{p} \mu$.

Bukti

Hal ini mengikuti fakta bahwa $E(\overline{X}_n) = \mu$, $Var(\overline{X}_n) = \sigma^2 / n$ dan dengan demikian

$$P\left[\left|\overline{X}_n - \mu\right| < \varepsilon\right] \geq 1 - \frac{\sigma^2}{\varepsilon^2 n}$$

Sehingga $\lim\limits_{n\to\infty} P\left[\left|\overline{X}_n - \mu\right| < \varepsilon\right] = 1$

Hasil diatas menggambarkan bahwa rata-rata contoh merupakan penduga yang baik bagi rata-rata populasi, dalam artian bahwa peluangnya mendekati 1 pada saat $\overline{X}_n$ mendekati nilai μ apabila n → ∞.

Sisi kanan pertidaksamaan diatas memiliki informasi tambahan. Untuk sembarang nilai $\varepsilon > 0$ dan $0 < \delta < 1$, jika $n > \sigma^2/(\varepsilon^2\delta)$, maka berlaku

$$P\left[\mu - \varepsilon < \overline{X}_n < \mu + \varepsilon\right] \geq 1 - \delta$$

Teladan 7. 15.

Untuk ragam contoh, $S^2 = \dfrac{\sum_{i=1}^{n} X_i^2 - \left(\sum_{i=1}^{n} X_i\right)^2 / n}{n-1}$ diperoleh informasi bahwa $E(S^2) = \sigma^2$ dan $Var(S^2) = \left(\mu_4' - \dfrac{n-3}{n-1}\sigma^4\right) / n$. Dengan menggunakan pertidaksamaan Chebychev,

$$P\left[\left|S^2 - \sigma^2\right| < \varepsilon\right] \geq 1 - \frac{Var(S^2)}{\varepsilon^2}$$

Dengan demikian,

$$\lim_{n\to\infty} P\left[\left|S^2 - \sigma^2\right| < \varepsilon\right] = 1 \quad \text{dan} \quad S^2 \xrightarrow{p} \sigma^2$$

Apakah $S \xrightarrow{p} \sigma$? Kita dapat tunjukkan bahwa $E(S) \neq \sigma$, namun dengan beberapa teknik aljabar, dapat ditunjukkan bahwa dengan menggunakan pertidaksamaan Chebychev, $S \xrightarrow{p} \sigma$.

Teorema 7. 6.

Jika $Z_n = \sqrt{n}(Y_n - m)/c \xrightarrow{d} Z$ ~ N(0,1), maka $Y_n \xrightarrow{p} m$.

Teladan 7. 16.

Median contoh, X_M memiliki sebaran asimtotik normal dengan asimtot rata-rata $x_{0,5}$, yaitu median sebaran. Berdasarkan Teorema ini, $X_M \xrightarrow{p} x_{0,5}$ bilamana n → ∞, dengan *k/n* → 0,5. Dengan cara yang sama, jika *k/p* → *p*, Statistika terkecil ke-*k* akan konvergen stokastik ke persentil ke-*p*, $X_{(k)} \xrightarrow{p} x_p$.

Beberapa Teorema Limit Lainnya

Definisi 7. 5.

Deretan peubah acak Y_n dikatakan konvergen dalam peluang ke peubah acak *Y*, dituliskan dengan $Y_n \xrightarrow{p} Y$, jika $\lim_{n\to\infty} P\left[|Y_n - Y| < \varepsilon\right] = 1$.

Apa yang telah kita pelajari dalam beberapa sub bab terdahulu, konvergen stokastik sama dengan konvergen dalam peluang ke suatu konstanta *c*. Konvergen dalam peluang lebih kuat dari pada konvergen dalam sebaran. Tentu ini tidak menghereankan, karena konvergen dalam sebaran tidak memberikan persyaratan sebaran bersama dari peubah Y_n dan *Y*, sedangkan konvergen dalam peluang mempersyaratkan.

Teorema 7. 7.

Untuk suatu deretan peubah acak, jika $Y_n \xrightarrow{p} Y$, maka $Y_n \xrightarrow{d} Y$.

Teorema 7. 8.

Jika $Y_n \xrightarrow{p} c$, maka untuk sembarang fungsi *g(y)* yang kontinu pada *c*, $g(Y_n) \xrightarrow{p} g(c)$

Bukti

Karena *g(y)* kontinu pada *c*, maka untuk setiap $\varepsilon > 0$, akan terdapat sebuah $\delta > 0$ sedemikian rupa sehingga $|y-c| < \delta$ yang membuat $g(y)-g(c)| < \varepsilon$. Dengan demikian

$$P\left[|g(Y_n) - g(c)| < \varepsilon\right] \geq P\left[|Y_n - c| < \delta\right]$$

[karena P(B) ≥ P(A) bilamana A ⊂ B]. namun karena $Y_n \xrightarrow{p} c$, maka untuk setiap $\delta > 0$, maka

$$\lim_{n\to\infty} P\left[|g(Y_n) - g(c)| < \varepsilon\right] \geq \lim_{n\to\infty} P\left[|Y_n - c| < \delta\right] = 1$$

Sisi kiri nilainya tak dapat melampaui 1, sehingga harus sama dengan 1, dan dengan demikian $g(Y_n) \xrightarrow{p} g(c)$.

Teorema 7. 9.

Jika X_n dan Y_n adalah dua deretan peubah acak sedemikian rupa sehingga $X_n \xrightarrow{p} c$ dan $Y_n \xrightarrow{p} d$, maka

1. $aX_n + bY_n \xrightarrow{p} ac + bd$
2. $X_n Y_n \xrightarrow{p} cd$
3. $X_n / c \xrightarrow{p} 1$ untuk $c \neq 0$.
4. $1/X_n \xrightarrow{p} 1/c$ jika $P[X_n \neq 0] = 1$ untuk semua *n*, $c \neq 0$.
5. $\sqrt{X_n} \xrightarrow{p} \sqrt{c}$ jika $P[X_n \geq 0] = 1$ untuk semua *n.*

Teladan 7. 17.

Misalkan *Y ~ Bin(n,p)*. Telah kita ketahui bahwa $\hat{p} = Y/n \xrightarrow{p} p$. Dengan menggunakan Teorema diatas, kita peroleh bahwa $\hat{p}(1-\hat{p}) \xrightarrow{p} p(1-p)$.

Teorema 7. 10. Slutsky

Jika X_n dan Y_n adalah dua deretan peubah acak sedemikian rupa sehingga $X_n \xrightarrow{p} c$ dan $Y_n \xrightarrow{d} Y$, maka

1. $X_n + Y_n \xrightarrow{d} c + Y$
2. $X_n Y_n \xrightarrow{d} cY$
3. $Y_n / X_n \xrightarrow{d} Y/c$, $c \neq 0$.

X_n mungkin dapat merupakan deretan numerik seperti $X_n = n/(n-1)$.

Teladan 7. 18.

Dari sebaran $T_\nu \sim t(\nu)$, dimana $T_\nu = \dfrac{Z}{\sqrt{\chi^2_\nu/\nu}}$. Dari sifat-sifat yang dimiliki sebaran kai-kuadrat, kita peroleh $E(\chi^2_\nu/\nu) = 1$, $Var(\chi^2_\nu/\nu) = 2/\nu$ dan dengan menggunakan pertidaksamaan Chebychev, $P\left[\left|\chi^2_\nu/\nu - 1\right| < \varepsilon\right] \geq 1 - \dfrac{2}{\nu\varepsilon^2}$, sehingga $\chi^2_\nu/\nu \xrightarrow{p} 1$, bilamana $\nu \to \infty$. Dengan demikian, sebaran t-student memiliki limit sebaran, yaitu sebaran normal baku.

$$T_\nu = \frac{Z}{\sqrt{\chi^2_\nu/\nu}} \xrightarrow{d} Z \sim N(0,1)$$

Teladan 7. 19.

Suatu contoh acak berukuran *n* dari sebaran Bernoulli, X_i ~ *Bin(1,p)*. Kita tahu bahwa

$$\frac{\hat{p}-p}{\sqrt{p(1-p)/n}} \xrightarrow{d} Z \sim N(0,1)$$

dari informasi terdahulu juga diperoleh $\hat{p}(1-\hat{p}) \xrightarrow{p} p(1-p)$. Sebagai akibatnya, kita peroleh

$$\frac{\hat{p}-p}{\sqrt{\hat{p}(1-\hat{p})/n}} \xrightarrow{d} Z \sim N(0,1)$$

Teorema 7. 11.

Jika $Y_n \xrightarrow{d} Y$, maka untuk sembarang fungsi kontinu *g(y),* $g(Y_n) \xrightarrow{d} g(Y)$

Teladan 7. 20.

Telah kita tunjukkan bahwa sebaran t-student konvergen ke sebaran standar normal baku bilamana $\nu \rightarrow \infty$, $T_\nu = \dfrac{Z}{\sqrt{\chi^2_\nu / \nu}} \xrightarrow{d} Z \sim$ *N(0,1).* Jika kita misalkan $F_{1;\nu} = T_\nu^2$, maka dengan menggunakan Teorema diatas, kita peroleh bahwa

$$F_{1;\nu} = T_\nu^2 \xrightarrow{d} Z^2 \sim \chi^2(1)$$

Berikut diberikan kondisi dimana suatu fungsi dari peubah yang memiliki sebaran asimtotik normal juga menyebar asimtotik normal.

Teorema 7. 12.

Jika $\sqrt{n}(Y_n - m)/c \xrightarrow{d} Z$ ~ *N(0,1)*, dan jika *g(y)* memiliki turunan yang tidak nol pada $y = m$, $g'(m) \neq 0$, maka

$$\frac{\sqrt{n}\left[g(Y_n) - g(m)\right]}{\left[cg'(m)\right]} \xrightarrow{d} Z \sim N(0,1)$$

Sesuai interpretasi sebaran asimtotik normal sebelumnya, kita simpulkan bahwa untuk n yang sangat besar, jika $Y_n \sim N(m,c^2/n)$, maka kira-kira

$$g(Y_n) \sim N\left(g(m); \frac{c^2\left[g'(m)\right]^2}{n} \right)$$

Teladan 7. 21.

Teorema Limit Pusat mengatakan bahwa rata-rata contoh memiliki sebaran asimtotik normal,

$$\frac{\sqrt{n}(\overline{X}_n - \mu)}{\sigma} \xrightarrow{d} Z \sim N(0,1)$$

atau pendekatan untuk n yang besar,

$$\overline{X}_n \sim N(\mu; \frac{\sigma^2}{n}).$$

Teorema diatas menyebutkan bahwa fungsi yang differensiabel atau fungsi yang mempunyai turunan juga akan memiliki sebaran asimtotik normal. Jika $g(\overline{X}_n) = \overline{X}_n^2$, maka $g'(\mu)=2\mu$, dan kira-kira,

$$\overline{X}_n^2 \sim N\left[\mu^2, \frac{(2\mu)^2\sigma^2}{n} \right]$$

Teladan 7. 22.

Misalkan S_n^2 merupakan notasi ragam contoh dari suatu contoh acak berukuran n dari sebaran normal, $N(\mu,\sigma^2)$. Kita tahu bahwa

$$V_n = \frac{(n-1)S_n^2}{\sigma^2} \sim \chi^2(\text{n-1})$$

dan berdasarkan Teorema yang telah kita pelajari dalam bab ini

$$\frac{V_n - (n-1)}{\sqrt{2(m-1)}} \xrightarrow{d} Z \sim N(0,1)$$

Sehingga

$$\frac{\sqrt{n-1}\left[S_n^2 - \sigma^2\right]}{\sigma^2\sqrt{2}} \xrightarrow{d} Z$$

Atau kira-kira

$$S_n^2 \sim N\left(\sigma^2, \frac{2\sigma^4}{n-1}\right)$$

dan

$$S_n \sim N\left(\sigma, \frac{\sigma^2}{2(n-1)}\right)$$

Latihan

1. Sebuah contoh acak berukuran *n* dari suatu sebaran yang memiliki fungsi sebaran kumulatif $F(x) = 1-1/x$ jika $1 \le x < \infty$, dan nol selainnya.
   a. Carilah fungsi sebaran kumulatif Statistik Minimum, $X_{(1)}$.
   b. Carilah limit sebaran $X_{(1)}$
   c. Carilah limit sebaran $X_{(1)}^n$
2. Sebuah contoh acak berukuran *n* dari suatu sebaran yang memiliki fungsi sebaran kumulatif $F(x) = 1-x^{-2}$ jika x > 1, dan nol selainnya. Tentukan apakah setiap sekuen peubah acak berikut memiliki limit sebaran ? Jika ya, tentukan limit sebarannya.
   a. $X_{(1)}$
   b. $X_{(n)}$
   c. $n^{-1/2}X_{(n)}$
3. Misalkan $Y_n \sim \chi^2(n)$. Carilah limit sebaran dari

$(Y_n - n)/\sqrt{2n}$ bilamana *n* → ∞, dengan menggunakan fungsi pembangkit momen.

4. Misalkan X_i ~ Seragam(0,1), dimana $X_1, X_2, \ldots, X_{20}$ saling bebas. Carilah pendekatan normal untuk tiap pertanyaan berikut
   a. $P\left[\sum_{i=1}^{20} X_i \le 12\right]$
   b. Persentil ke-90 dari $\sum_{i=1}^{20} X_i$
5. Misalkan W_j merupakan berat begasi penumpang kapal ke-j. Asumsikan bahwa berat begasi seluruh penumpang saling bebas, dan maisng-masing memiliki fungsi kepekatan peluang $f(w) = \theta B^{-\theta} w^{\theta-1}$ jika $0 < w < B$, dan nol selainnya.
   a. Untuk n = 100, θ = 3, dan B = 80, tentukan nilai perkiraan $P\left[\sum_{i=1}^{100} W_i > 6025\right]$
   b. Jika $W_{(1)}$ merupakan bagasi teringan, tunjukkan bahwa $W_{(1)} \to 0$ bilamana $n \to \infty$.
   c. Jika $W_{(n)}$ merupakan bagasi terberat, tunjukkan bahwa $W_{(n)} \to \infty$ bilamana $n \to \infty$.
   d. Carilah limit sebaran dari $(W_{(n)}/B)^n$
6. Suatu contoh acak diambil dari sebaran Poisson, X_i ~ *Poi*(μ).
   a. Tunjukkan bahwa $Y_n = e^{-\overline{X}_n}$ konvergen stokastik ke $P[X=0] = e^{-\mu}$
   b. Carilah sebaran asimtotik normal dari Y_n
   c. Tunjukkan bahwa $\overline{X}_n e^{-\overline{X}_n}$ konvergen stokastik ke $P[X=1] = \mu e^{-\mu}$

Sebaran Nilai Ekstrim

Dalam bab sebelum ini, statistik tataan terpusat $X_{(k)}$ menyebar asimtotik normal bilamana $n \rightarrow \infty$ dan $k/n \rightarrow p$. Jika statistik tataan ekstrim seperti $X_{(1)}$, $X_{(2)}$ dan $X_{(k)}$ dibakukan sehingga mereka memiliki limit sebaran yang *nondegenerate*, limit sebaran ini bukan sebaran normal. Dapat ditunjukkan bahwa limit sebaran *nondegenerate* dari suatu peubah ekstrim akan menjadi salah satu dari tiga kemungkinan tipe sebaran. Dengan demikian, ketiga tipe sebaran ini sangat berguna dalam mempelajari nilai ekstrim, analog dengan cara sebaran normal yang berguna dalam mempelajari rata-rata melalui Teorema Limit Pusat.

Berbagai teladan penggunaan dalam kehidupan sehari-hari tentang statistik ekstrim ini, diantaranya adalah:

a. dalam hal mempelajari permasalahan banjir, peubah titik maksimum (*maximum flood stage*) dalam setahun. Peubah ini dapat berperilaku seperti maksimum dari jumlah tingkatan banjir yang saling bebas selama setahun;
b. kekuatan suatu rantai yang sama dengan kelemahan dari suatu penyambungnya;
c. umur dari suatu rangkaian elektronik yang disusun secara seri merupakan minimum dari umur komponen; sedangkan umur dari suatu rangkaian elektronik yang disusun secara paralel merupakan maksimum dari umur komponen;

Sebaran Asimtotik Statistik Tataan Ekstrim

Teorema-teorema berikut, yang dituliskan tanpa pembuktian, berguna dalam mempelajari tingkah laku asimtotik dari statistik tataan ekstrim.

Teorema 8. 1.

Jika limit suatu barisan (deret) suatu fungsi sebaran kumulatif merupakan fungsi sebaran kumulatif yang kontinu, $F(y) = \lim_{n\to\infty} F_n(y)$, maka untuk sembarang $a_n > 0$ dan b_n,

$$\lim_{n\to\infty} F_n(a_n y + b_n) = F(ay + b)$$

jika dan hanya jika $\lim_{n\to\infty} a_n = a > 0$ dan $\lim_{n\to\infty} b_n = b$

Teorema 8. 2.

Jika limit suatu barisan suatu fungsi sebaran kumulatif merupakan fungsi sebaran kumulatif yang kontinu, dan jika $\lim_{n\to\infty} F_n(a_n y + b_n) = G(y)$ untuk semua $a_n > 0$ dan semua bilangan nyata y, maka $\lim_{n\to\infty} F_n(\alpha_n y + \beta_n) = G(y)$ untuk $\alpha_n > 0$, jika dan hanya jika $\alpha_n/a_n \to 1$ dan $(\beta_n - b_n)/a_n \to 0$ bilamana $n \to \infty$.

Limit Sebaran Maksimum

Misalkan $X_{(1)},\ldots,X_{(n)}$ merupakan contoh acak tertata berukuran n dari suatu sebaran kontinu dengan fungsi sebaran kumulatif $F(x)$. Dalam hal teori nilai-ekstrim, maksimum $X_{(n)}$ dikatakan memiliki limit sebaran (*nondegenerate*) $G(y)$ jika terdapat barisan konstanta $\{a_n\}$ dan $\{b_n\}$ dengan $a_n > 0$, sedemikian rupa sehingga peubah terstandarisasi, $Y_n = (X_{(n)} - b_n)/a_n$ konvergen dalam sebaran $G(y)$.

$$Y_n = \frac{X_{(n)} - b_n}{a_n} \xrightarrow{d} Y \sim G(y)$$

Dengan demikian kita dapat katakan bahwa, $X_{(n)}$ memiliki limit sebaran tipe G, dimana yang kita maksudkan disini adalah limit sebaran peubah terstandarisasi (peubah baku) Y_n adalah sebaran nondegenerate $G(y)$. Jika $G(y)$ kontinu, barisan konstanta pembaku tidak akan khas; namun demikian, tidak mungkin mendapatkan suatu limit sebaran dari tipe yang berbeda dengan merubah konstanta pembaku.

Bila kita ingit kembali sebaran pasti dari statistik $X_{(n)}$ adalah

$$F_{(n)}(x) = [F(x)]^n .$$

Jika kita tertarik pada $Y_n = (X_{(n)} - b_n)/a_n$, maka sebaran pasti dari Y_n adalah

$$\begin{aligned} G_n(y) &= P[Y_n \le y] \\ &= F_{(n)}(a_n y + b_n) \\ &= [F(a_n y + b_n)]^n \end{aligned}$$

Dengan demikian, limit sebaran dari $X_{(n)}$ (atau lebih tepatnya Y_n) adalah

$$G(y) = \lim_{n\to\infty} G_n(y) = \lim_{n\to\infty}[F(a_n y + b_n)]^n$$

Formula diatas merupakan suatu pendekatan langsung dalam penentuan limit sebaran nilai-ekstrim, jika barisan $\{a_n\}$ dan $\{b_n\}$ dapat diperoleh sehingga menghasilkan limit nondegenetare.

Teladan 8. 1.

Misalkan X ~ Eksponensial(1). Dengan menggunakan $a_n = 1$ dan $b_n = \ln n$. Maka,

$$G_n(y) = [F(y + \ln n)]^n = \left[1 - \left(\frac{1}{n}\right)e^{-y}\right]^n$$

sehingga

$$G(y) = \lim_{n\to\infty}\left[1 - \left(\frac{1}{n}\right)e^{-y}\right]^n = e^{-e^{-y}} = \exp(-e^{-y})$$

Teorema 8. 3.

Jika $Y_n = (X_{(n)} - b_n)/a_n$ memiliki limit sebaran *G(y)*, maka *G(y)* harus merupakan salah satu dari tiga tipe sebaran nilai-ekstrim berikut:

1. Type I (untuk Maksimum). (**Tipe Eksponensial**)
$$G^{(1)}(y) = \exp(-e^{-y}) \qquad -\infty < y < \infty$$
2. Type II (untuk Maksimum). (**Tipe Cauchy**)
$$G^{(2)}(y) = \exp(-y^{-\gamma}) \qquad y > 0, \gamma > 0$$
3. Type III (untuk Maksimum). (**Tipe Limited**)
$$G^{(3)}(y) = \begin{cases} \exp[-(-y)^{\gamma}] & jika \quad y < 0, \gamma > 0 \\ 1 & jika \quad y \geq 0 \end{cases}$$

Limit sebaran maksimum dari kepekatan seperti sebaran-sebaran Normal, LogNormal, Logistik dan Gamma merupakan sebaran nilai-ekstrim Tipe I. Terdapat beberapa fungsi kepekatan dengan ekor tidak lebih tebal dari sebaran eksponensial. Kelas ini mencakup sejumlah besar sebaran yang umum diketahui, dan sebaran nilai-ekstrim Tipe I (untuk maksimum) harus memberikan model yang baik untuk berbagai tipe peubah yang ada hubungannya dengan maksimum. Sudah barang tentu, parameter lokasi dan parameter skala perlu dikenalkan dalam model bilamana diaplikasikan secara langsung dalam peubah yang tidak standar (tidak baku).

Limit sebaran Tipe II sebagai hasil dari maksimum dari kepekatan dengan ekor lebih tebal, seperti distribusi Cauchy. Sedangkan Tipe III dapat berasal dari kepekatan-kepekatan dimana batas atasnya terhingga pada range peubah.

Teorema berikut memberikan bentuk alternatif yang lebih memudahkan dalam penyelesaian nilai limit.

Teorema 8. 4.

Gnedenko. Dalam penentuan limit sebaran $Y_n = (X_{(n)} - b_n)/a_n$

$$\lim_{n\to\infty} G_n(y) = \lim_{n\to\infty} \left[F(a_n y + b_n)\right]^n = G(y)$$

jika dan hanya jika

$$\lim_{n\to\infty} n\left[1 - F(a_n y + b_n)\right] = -\ln G(y)$$

Dalam banyak kasus, kesulitan besar yang dialami dalam penentuan barisan pembakuan yang sesuai sedemikian rupa sehingga menghasilkan limit sebaran nondegenerate. Untuk suatu fungsi sebaran kumulatif *F(x)* dimungkinkan untuk menggunakan Teorema Gnedenko untuk menjawab a_n dan b_n dalam bentuk *F(x)* untuk setiap tiga kemungkinan limit sebaran. Dengan demikian, jika kita mengetahui tipe limit *F(x)*, maka a_n dan b_n dapat dihitung. Jika kita tidak mengetahui tipenya, kita dapat menghitung a_n dan b_n untuk tiap tipe dan kemudian mencoba menentukan tipe mana yang sesuai. Satu properti dari fungsi sebaran kumulatif yang bermanfaat dalam mengekspresikan konstanta pembaku adalah "nilai ciri terbesar" (*largest characteristic value*).

Definisi 8. 1.

Nilai ciri terbesar, u_n, dari atau fungsi sebaran kumulatif *F(x)* didefinisikan dengan persamaan $n[1-F(u_n)] = 1$.

Untuk contoh acak berukuran *n* dari *F(x)*, rata-rata banyaknya observasi yang akan melebih u_n adalah 1. Peluang bahwa satu observasi melebihi u_n adalah

$$p = P[X > u_n] = 1 - F(u_n)$$

dan rata-rata banyaknya *n* observasi saling bebas adalah

$$np = n[1 - F(u_n)]$$

Teorema 8. 5.

Misalkan *X* ~ *F(x)*, dan asumsikan bahwa $Y_n = (X_{(n)} - b_n)/a_n$ memiliki limit sebaran.

1. Jika *F(x)* kontinu dan menaik, limit sebaran Y_n akan memiliki tipe eksponensial jika dan hanya jika
$$\lim_{n\to\infty} n\left[1 - F(a_n y + b_n)\right] = e^{-y} \quad -\infty < y < \infty$$
dimana $b_n = u_n$ dan a_n merupakan jawaban dari
$$F(a_n + u_n) = 1 - (ne)^{-1}$$
2. *G(y)* memiliki Tipe Cauchy, jika dan hanya jika
$$\lim_{y\to\infty} \frac{1 - F(y)}{1 - F(ky)} = k^{\gamma} \qquad k > 0, \gamma > 0$$
dan dalam kasus ini, $a_n = u_n$ serta $b_n = 0$.
3. *G(y)* memiliki Tipe Limited, jika dan hanya jika
$$\lim_{y\to 0^-} \frac{1 - F(ky + x_0)}{1 - F(y + x_0)} = k^{\gamma} \qquad k > 0$$
dimana $x_0 = max\{x|F(x)<1\}$, batas atas dari *x*. Juga $b_n = 0$ dan $a_n = x_0 - u_n$.

Teladan 8. 2.

Misalkan *X* ~ Exponensial(θ), dan kita tertarik dengan Maksimum dari suatu contoh acak berukuran *n*. Nilai ciri terbesar u_n dapat diperoleh dari
$$n[1 - F(u_n)] = n[1 - (1 - e^{-u_n/\theta})] = 1$$
yang akan menghasilkan nilai $u_n = \theta \ln n$.
Kita berharap bahwa kepakatan Eksponensial akan berada dalam Tipe I, sehingga kita perlu mencobanya terlebih dahulu. Dengan $b_n = u_n = \theta \ln n$, nilai a_n dapat ditentukan dengan
$$F(a_n + u_n) = 1 - e^{-(a_n + \theta \ln n)/\theta} = 1 - (1/n)e^{-a_n/\theta} = 1 - 1/(ne)$$
yang akan menghasilkan nilai $a_n = \theta$.

Dengan demikian, jika kepekatan Eksponensial masuk ke dalam Tipe I, maka

$$Y_n = \frac{X_{(n)} - \theta \ln n}{\theta} \xrightarrow{d} Y \sim G^{(1)}(y)$$

Hal ini dapat dengan mudah diperiksa dengan menggunakan cara seperti berikut

$$\lim_{n\to\infty} n[1 - F(a_n y + b_n)] = \lim_{n\to\infty} n[e^{-(y+\ln n)}]$$
$$= \lim_{n\to\infty} e^{-y}$$
$$= e^{-y} \qquad -\infty < y < \infty$$

Teladan 8. 3.

Suatu peubah *X* dengan fungsi sebaran kumulatif $F(x) = 1 - x^{-\theta}, \quad x \geq 1$ memiliki ekor bagian atas yang tebal, sehingga ada kecenderungan bahwa limit sebarannya akan memiliki Tipe Cauchy. Dengan memeriksa kondisi

$$\lim_{y\to\infty} \frac{1 - F(y)}{1 - F(ky)} = \lim_{y\to\infty} \frac{y^{-\theta}}{(ky)^{-\theta}} = k^{\theta}$$

sehingga limit sebarannya Tipe Cauchy dengan $\gamma = \theta$. Juga kita dapat evaluasi bahwa $n[1 - F(u_n)] = nu_n^{-\theta} = 1$ yang akan menghasilkan $u_n = n^{1/\theta} = a_n$, dan kita misalkan b_n = 0 dalam hal ini.

Dengan demikian, kita tahu bahwa

$$Y_n = \frac{X_{(n)}}{n^{1/\theta}} \xrightarrow{d} Y \sim G^{(2)}(y)$$

Kita juga dapat melakukan verifikasi bahwa

$$\lim_{n\to\infty} n[1 - F(a_n y + b_n)] = \lim_{n\to\infty} y^{-\theta} = -\ln G(y)$$

sehingga $G(y) = \exp(-y^{-\theta})$ yang merupakan Tipe Cauchy dengan $\gamma = \theta$.

Teladan 8. 4.

Untuk *X* ~ Seragam(0,1), dimana $F(x) = x, \quad 0 < x < 1$. Kita berharap bahwa sebaran limit maksimumnya tergolong Tipe III. Kita punya

$$n[1 - F(u_n)] = n(1 - u_n) = 1$$

yang akan menghasilkan nilai $u_n = 1 - 1/n$. Dengan demikian $b_n = x_0 = 1$ dan $a_n = x_0 - u_n = 1$. Dengan memeriksa kondisi bahwa

$$\lim_{y \to 0^-} \frac{1 - F(ky + x_0)}{1 - F(y + x_0)} = \lim_{y \to 0^-} \frac{1 - (ky + x_0)}{1 - (y + x_0)} = \lim_{y \to 0^-} \frac{ky}{y} = k$$

sehingga limit sebaran $Y_n = n(X_{(n)} - 1)$ tergolong Tipe III dengan γ = *1*. Sebagai ilustrasi selanjutnya

$$\begin{aligned} \lim_{n \to \infty} n[1 - F(a_n y + b_n)] &= \lim_{n \to \infty} n\left[1 - \left(\frac{y}{n} + 1\right)\right] \\ &= -y \\ &= -\ln G(y) \end{aligned}$$

dan $Y_n = n(X_{(n)} - 1) \xrightarrow{d} Y \sim G(y) = \begin{cases} e^y & jika \quad y < 0 \\ 1 & jika \quad y \geq 0 \end{cases}$

Limit Sebaran Minimum

Seandainya limit sebaran nondegenerate ada untuk Minimum suatu contoh acak, maka akan limitnya akan masuk ke dalam salah satu dari tiga Tipe. Sudah barang tentu sebaran minimum dapat dihubungkan dengan sebaran maksimum, karena

$$\min(x_1, ..., x_n) = -\max(-x_1, ..., -x_n)$$

Dengan demikian, seluruh hasil apa yang telah kita bahas pada Maksimum dapat dimodifikasi untuk diaplikasikan jika detilnya dapat di atur kembali.

Misalkan *X* peubah kontinu, $X \sim F_X(x)$, dan misalkan $Z = -X \sim F_Z(z) = 1-F_X(-z)$. Juga perlu diingat bahwa $X_{(1)}=-Z_{(n)}$. Sekarang apabila $W_n = (X_{(1)} + b_n)/a_n$, kita punya

$$\begin{aligned} G_{W_n}(w) &= P\left[\frac{X_{(1)} + b_n}{a_n} \le w\right] \\ &= P\left[\frac{-Z_{(n)} + b_n}{a_n} \le w\right] \\ &= P\left[\frac{Z_{(n)} - b_n}{a_n} \ge -w\right] \\ &= P[Y_n \ge -w] \\ &= 1 - G_{Y_n}(-w) \end{aligned}$$

Limit sebaran W_n, sebut saja misalnya *H(w)* adalah

$$\begin{aligned} H(w) &= \lim_{n\to\infty} G_{W_n}(w) \\ &= \lim_{n\to\infty}[1 - G_{Y_n}(-w)] \\ &= 1 - G(-w) \end{aligned}$$

yang mana *G(y)* merupakan limit sebaran dari $Y_n = (Z_{(n)} - b_n)/a_n$. Untuk memperoleh *H(w),* yaitu limit sebaran minimum, langkah pertama adalah menentukan

$$F_Z(z) = 1 - F_X(-z)$$

baru kemudian menentukan a_n, b_n dan limit sebaran *G(y)* dengan menggunakan metode seperti apa yang telah digambarkan untuk Maksimum yang diaplikasikan pada $F_Z(z)$. Dengan demikian limit sebaran untuk W_n adalah

$$H(w) = 1 - G(-w)$$

Perlu dicatat bahwa bilamana $F_X(x)$ termasuk pada satu tipe limit, adalah mungkin bahwa $F_Z(z)$ akan masuk kedalam tipe yang berbeda. Sebagai teladan Maksimum dari Eksponensial(θ) memiliki sebaran limit Tipe I, namun demikian $F_Z(z)$ tergolong pada Tipe III.

Ringkasnya, prosedur langsung untuk menentukan a_n, b_n dan $H(w)$ adalah pertama kali kita harus mendapatkan $F_Z(z)$ dan aplikasikan metode untuk Maksimum dalam penentuan $G(y)$ untuk $Y_n = (Z_{(n)} - b_n)/a_n$. Kemudian dapatkan $H(w)$.

Definisi 8. 2.

Nilai ciri terkecil adalah s_n yang memenuhi $nF(s_n) = 1$.

Bila diperhatikan maka $s_n(x) = -u_n(z)$. Dengan cara yang sama, kondisi $F_Z(a_n + u_n(z)) = 1 - 1/(ne)$ menjadi $F_X(-a_n + s_n) = 1/(ne)$, dan sebagainya.

Teorema 8. 6.

Jika $W_n = (X_{(1)} + b_n)/a_n$ memiliki limit sebaran $H(w)$, maka $H(w)$ akan tergolong ke dalam satu dari tiga tipe sebaran nilai-ekstrim berikut:

1. Tipe I (untuk minimum). (**Tipe Eksponensial**)

Dalam hal ini b_n = -s_n, dan a_n diperoleh dari

$$F(s_n - a_n) = \frac{1}{ne} \qquad W_n = \frac{X_{(1)} - s_n}{a_n}$$

dan

$$H_W^{(1)}(w) = 1 - G^{(1)}(-w) = 1 - \exp(-e^w) \qquad -\infty < w < \infty$$

jika dan hanya jika $\lim_{n\to\infty} nF(-a_n y + s_n) = e^{-y}$

2. Tipe II (untuk minimum). (**Tipe Cauchy**)

Pilih nilai a_n = -s_n, b_n = 0, W_n = -$X_{(1)}/s_n$

dan

$$H_W^{(2)}(w) = 1 - G^{(2)}(-w) = 1 - \exp[-(-w)^{-\gamma}] \qquad w < 0, \gamma > 0$$

jika dan hanya jika

$$\lim_{y\to\infty} \frac{F(-y)}{F(-ky)} = k^{\gamma} \qquad k > 0, \gamma > 0$$

atau

$$\lim_{n\to\infty} nF(s_n y) = y^{-\gamma} \qquad y > 0$$

3. Tipe III (untuk minimum). **Tipe Limited**

Jika $x_1 = \min\{x \mid F(x) > 0\}$ merupakan batas bawah dari *x* (yang mana $x_1 = -x_0$), maka

$$b_n = -x_1 \qquad a_n = -x_1 + s_n \qquad w_n = \frac{X_{(1)} - x_1}{s_n - x_1}$$

dan

$$H_W^{(3)}(w) = 1 - G^{(3)}(-w) = 1 - \exp(-w^{\gamma}) \qquad w > 0, \gamma > 0$$

jika dan hanya jika

$$\lim_{y\to 0} \frac{F(ky + x_1)}{F(y + x_1)} = k^{\gamma} \qquad k > 0$$

atau

$$\lim_{n\to\infty} nF[(x_1 - s_n)y + x_1] = (-y)^{\gamma}$$

Sebaran minimum Tipe I dikenal juga sebagai **sebaran nilai-ekstrim Tipe I**. Tipe II dari sebaran minimum ini dikenal dengan sebaran Weibull. Dalam menentukan tipe limit sebaran minimum, perlu diperhatikan ketebalan ekor kanan dari $F_Z(z)$ dimana $Z=-X$. Dengan demikian, limit sebaran minimum beberapa kepekatan, seperti Eksponensial dan Gamma, tergolong Tipe III. Inilah mungkin suatu alasan bahwa sebaran Weibull kadang merupakan jawaban dalam beberapa aplikasi.

Teladan 8. 5.

Kita perhatikan Minimum dari suatu contoh acak berukuran *n* dari suatu peubah acak yang menyebar Eksponensial(θ). Kita sudah mengetahui bahwa $X_{(1)}$ ~ Eksponensial(θ/n), dan demikain juga $nX_{(1)}/\theta$ ~ Eksponensial(1). Dengan demikian, limit sebaran $nX_{(1)}/\theta$ juga Eksponensial(1), yang tergolong Tipe III dimana $\gamma = 1$. Jika kita tahu jawabnya, kita dapat menerka ('*guess*') bahwa limit sebarannya

Tipe III, karena range dari peubah Z=-X hanya ada di kanan. Dengan menggunakan teorema terakhir, kita dapatkan x_1 = 0 dan

$$\lim_{y\to 0^+}\frac{F(ky+x_1)}{F(y+x_1)}=\lim_{y\to 0^+}\frac{1-\exp(-ky)}{1-\exp(-y)}=\lim_{y\to 0^+}\frac{k\exp(-ky)}{\exp(-y)}=k$$

Dengan demikian, kita tahu bahwa

$$H_W(w)=1-e^{-w}$$

dimana

$$W_N=\frac{X_{(1)}-x_1}{s_n-x_1}=\frac{X_{(1)}}{s_n}$$

Dalam kasus ini, nilai s_n diperoleh dari

$$F(s_n)=1-e^{-s_n/\theta}=\frac{1}{n}\quad\text{atau}\quad s_n=-\theta\ln\left(1-\frac{1}{n}\right)$$

Hal tersebut tidak menghasilkan konstanta pembaku yang sama dan identik seperti apa yang disarankan sebelumnya; namun demikian, hasilnya konsisten karena

$$\frac{-\ln(1-1/n)}{1/n}\to 1$$

Secara ringkas, tujuan dari bab ini adalah untuk menunjukkan bahwa statistik tataan yang ekstrim, seperti Minimum dan Maksimum, apabila ditransformasi dengan pas (sesuai), memiliki satu dari tiga tipe sebaran ekstrim. Termasuk Weibull dan Tipe I Sebaran Nilai-Ekstrim. Teori nilai-ekstrim ini memberikan motivasi penggunaan sebaran dalam pemecahan masalah seperti fenomena banjir dan kekuatan bahan.

Latihan

Tentukan konstanta pembaku dan sebaran nilai ekstrim dari tiap soal berikut:

a. $X_{1:n}$ dan $X_{n:n}$ apabila $F(x)=(1+e^{-x})^{-1}$.
b. $X_{1:n}$ dan $X_{n:n}$ apabila $X_i \sim Wei(\theta,\beta)$
c. $X_{1:n}$ dan $X_{n:n}$ apabila $X_i \sim EV(\theta,\eta)$
d. $X_{1:n}$ dan $X_{n:n}$ apabila $X_i \sim Par(\theta,\kappa)$

Teori Pendugaan Titik

Pendahuluan

Beberapa bab terdahulu berkenaan dengan pengembangan konsep peluang dan peubah acak untuk membangun model matematik dari fenomena fisik yang nondeterministik. Beberapa ciri numerik fenomena fisik yang menjadi perhatian kita, tetapi nilai cirinya tak dapat dihitung secara langsung. Namun, dimungkinkan mengamati satu atau lebih peubah acak, sebaran yang tergantung ciri yang dipelajari. Tujuan utama dalam beberapa bab kedepan ini adalah membangun metode-metode untuk menganalisis nilai amatan peubah acak untuk memperoleh informasi mengenai ciri yang tak diketahui tersebut.

Proses untuk memperoleh nilai amatan suatu fenomena fisik disebut dengan suatu **percobaan** (*experiment*). Misalkan hasil suatu pecobaan adalah peubah acak X, dan $f(x;\theta)$ melambangkan fungsi kepekatan peluangnya. Umumnya X sering digunakan sebagai nilai ukuran yang diperoleh dari individu yang dipilih secara acak dari suatu populasi. Dalam konteks ini, $f(x;\theta)$ merupakan fungsi kepekatan peluang populasi, dan merupakan sebaran ukuran individu dalam populasi.

Tujuan pendugaan titik adalah memberikan nilai yang cocok untuk parameter θ berdasarkan data amatan dari populasi. Hasil amatan tindakan berulang dari suatu percobaan dapat dimodelkan secara matematik sebagai peubah acak dari fungsi kepekatan peluang populasi. Dengan kata lain, diasumsikan bahwa suatu gugus n peubah acak yang saling bebas $X_1, X_2, \ldots, X_n$ masing-masing dengan fungsi kepekatan peluang $f(x;\theta)$ akan diamati, menghasilkan suatu gugus data $x_1, x_2, \ldots, x_n$. Sudah tentu, dimungkinkan bagi kita untuk menuliskan fungsi kepekatan peluang bersamanya sebagai hasil kali fungsi kepekatan peluang masing-masing

$$f(x_1, x_2, \ldots, x_n; \theta) = f(x_1;\theta) f(x_2;\theta) \cdots f(x_n;\theta)$$

Fungsi kepekatan peluang bersama ini menyediakan hubungan antara data amatan dan model matematik populasinya. Kita akan perhatikan seberapa baik penggunaan data tersebut dalam pendugaan nilai yang tak diketahui dari parameter θ.

Dalam bab ini akan kita asumsikan bahwa sebaran populasi yang dipelajari dapat direpresentasikan dengan suatu famili fungsi kepekatan peluang $f(x;\theta)$ dengan index parameter θ. Dalam beberapa kasus, parameter dapat berbentuk vektor, dan akan dilambangkan dengan $\boldsymbol{\theta}$.

Kita akan misalkan Ω, disebut dengan **ruang parameter**, merupakan gugus dari semua nilai yang mungkin untuk parameter θ. Jika $\boldsymbol{\theta}$ adalah suatu vektor, maka Ω akan merupakan anak gugus dari ruang Euclid berdimensi sama dengan banyaknya parameter yang tak diketahui.

Suatu contoh acak berukuran n dari suatu populasi dengan fungsi kepekatan peluang $f(x;\theta)$. Terminologi contoh acak menunjuk pada suatu gugus peubah acak $X_1, X_2, \ldots, X_n$ atau data amatan $x_1, x_2, \ldots, x_n$.

Definisi 9. 1.

Suatu fungsi peubah acak, $T = t(X_1, X_2, \ldots, X_n)$, yang tidak tergantung pada sembarang parameter disebut dengan **statistik**.

Sebuah statistik juga merupakan peubah acak, sebarannya tergantung pada sebaran contoh acak dan bentuk fungsi $t(x_1, x_2, \ldots, x_n)$. Sebaran statistik kadang sering disebut dengan sebaran turunan atau **sebaran sampling**, sebagai lawan kontras dari **sebaran populasi**.

Teladan 9. 1.

Misalkan $X_1, X_2, \ldots, X_n$ merupakan contoh acak dari sebaran Normal, $X_i \sim N(\mu,\sigma^2)$. Rata-rata contoh merupakan salah satu statistik, karena $\overline{X} = T = t(X_1, X_2, ..., X_n)$ dimana $t(x_1, x_2, ..., x_n) = (x_1 + x_2 + \cdots + x_n)/n$. Ragam contoh atau varian contoh, $S^2 = \sum_{i=1}^{n} \frac{(X_i - \overline{X})^2}{n-1}$ juga merupakan contoh suatu statistik.

Esensinya, statistik-statistik digunakan untuk mereduksi suatu gugus *n* nilai amatan menjadi gugus nilai yang lebih kecil, sehingga mudah diinterpretasikan. Sebagai catatan, sebaran sampling rata-rata contoh dan ragam contoh dapat diperoleh dengan menggunakan Teorema-teorema yang telah disampaikan pada bab-bab terdahulu. Kita dapat tunjukkan bahwa $\overline{X} \sim N(\mu, \sigma^2/n)$ dan $(n-1)S^2/\sigma^2 \sim \chi^2_{(n-1)}$. Penggunaan rata-rata dan ragam contoh dalam pendugaan parameter populasi μ dan σ^2 kadang digunakan justifikasi secara intuitif.

Kita asumsikan bahwa $X_1, X_2, \ldots, X_n$ merupakan contoh acak dari $f(x;\theta)$ dan $\tau(\theta)$ merupakan fungsi dari θ.

Definisi 9. 2.

Sebuah statistik $T = t(X_1, X_2, \ldots, X_n)$ yang digunakan untuk menduga nilai $\tau(\theta)$ disebut sebagai **penduga** bagi $\tau(\theta)$, dan nilai pengamatan dari statistik $t(x_1, x_2, \ldots, x_n)$ disebut dengan nilai dugaan bagi $\tau(\theta)$.

Istilah selain penduga yang sering digunakan adalah **penaksir** atau terjemahan langsung tanpa merubah hurufnya adalah **estimator**. Hal serupa juga dengan penggunaan istilah pendugaan, yaitu penaksiran atau estimasi.

Digunakan tiga (atau setidaknya dua) macam huruf dalam notasi ini. Huruf besar, misalnya *T,* merepresentasikan statistik yang digunakan sebagai penduga, huruf kecil t yang merupakan nilai amatan atau dugaan, dan skrip *t* merepresentasikan fungsi contoh acak. Untuk menyatakan penduga, tidak jarang pula digunakan *caret* atau topi, $\hat{\theta}$, untuk membedakan antara parameter yang tak diketahui dengan penduganya. Notasi lain yang sering juga digunakan adalah *tilde*, $\tilde{\theta}$. Bila digunakan notasi ini, tidak perlu lagi menggunakan huruf kapital dan sudah biasa untuk menyatakan penduganya.

Beberapa Metode Pendugaan

Dalam beberapa kasus, penduga yang beralasan dapat dicari berdasarkan intuisi, tetapi berbagai metode umum telah dikembangkan untuk menurunkan penduga.

Metode Momen

Rata-rata contoh, $\overline{X}$, sebagaimana kita ketahui merupakan penduga bagi rata-rata populasi μ. Salah satu pendekatan umum dan menghasilkan penduga, dikenal dengan **penduga metode momen**. Selanjutnya dapat disingkat dengan PM.

Perhatikan suatu fungsi kepekatan peluang populasi $f(x;\theta_1,\theta_2,...,\theta_k)$ yang tergantung pada satu atau lebih parameter $\theta_1, \theta_2, \ldots, \theta_k$. Momen-momen di sekitar nilai tengah, $\mu_j^{'}$, telah didefinisikan pada bab-bab terdahulu. Momen-momen ini umumnya tergantung pada parameter-parameternya, seperti terlihat

$$\mu_j^{'}(\theta_1,...,\theta_k) = E(X^j) \qquad\qquad j = 1,2,...,k$$

Definisi 9. 3.

Jika $X_1, \ldots, X_n$ adalah contoh acak dari $f(x;\theta_1,\theta_2,...,\theta_k)$, **momen contoh** *k* pertama adalah

$$M_j^{'} = \frac{\sum_{i=1}^{n} X_i^j}{n} \qquad\qquad j = 1,2,...,k$$

Momen pertama adalah rata-rata populasi, $\mu_1' = \mu$. Dengan cara yang sama, momen contoh pertama adalah rata-rata contoh.

Perhatikan contoh sederhana dari satu parameter yang tak diketahui, misalnya saja $\theta = \theta_1$. Fakta bahwa $\bar{X} = M_1'$ umumnya merupakan penduga beralasan bagi $\mu = \mu_1'(\theta)$ menyarankan penggunaan jawaban $\hat{\theta}$ dari persamaan $M_1' = \mu_1'(\hat{\theta})$ sebagai penduga bagi θ. Dengan perkataan lain, karena M_1' cenderung dekat ke $\mu_1'(\theta)$, kita dapat berharap, dengan kondisi tertentu bahwa $\hat{\theta}$ akan cenderung mendekati θ.

Secara umum, $\hat{\theta}_1, \hat{\theta}_2, ..., \hat{\theta}_k$ merupakan jawaban dari persamaan

$$M_j' = \mu_j'(\hat{\theta}_1, \hat{\theta}_2, ..., \hat{\theta}_k) \qquad j = 1,2,...,k$$

Teladan 9. 2.

Suatu contoh acak dari suatu sebaran dengan dua parameter tak diketahui, rata-rata populasi μ dan ragam populasi σ^2. Dari apa yang telah kita bahas sebelumnya, kita tahu bahwa $\mu = \mu_1'$ dan $\sigma^2 = E(X^2) - \mu^2 = \mu_2' - (\mu_1')^2$, sehingga Penduga Metode Momennya adalah jawaban dari persamaan-persamaan $M_1' = \hat{\mu}$ dan $M_2' = \hat{\sigma}^2 + \hat{\mu}^2$, yaitu $\hat{\mu} = \bar{X}$ dan $\hat{\sigma}^2 = \sum_{i=1}^{n} \frac{X_i^2}{n} - \bar{X}^2 = \sum_{i=1}^{n} \frac{(X_i - \bar{X})^2}{n}$. Bila ragam contoh sebagaimana kita tahu $S^2 = \sum_{i=1}^{n} \frac{(X_i - \bar{X})^2}{n-1}$. Sehingga penduga metode momennya sangat dekat dengan ragam contohnya, $\hat{\sigma}^2 = [(n-1)/n]S^2$.

Teladan 9. 3.

Bila contoh acak diambil dari populasi yang menyebar menurut sebaran eksponensial dengan fungsi kepekatan peluang $f(x;\eta) = e^{-(x-\eta)} I_{(\eta,\infty)}(x)$. Dapat dicari dengan mudah bahwa rata-rata populasinya $\mu = \mu(\eta) = 1 + \eta$, dan seandainya kita atur $\overline{X} = 1 + \hat{\eta}$, maka $\hat{\eta} = \overline{X} - 1$ adalah penduga metode momen dari η.

Teladan 9. 4.

Misalkan contoh acak yang diambil dari sebaran eksponensial dengan fungsi kepekatan peluang $f(x;\theta) = \frac{1}{\theta} e^{-x/\theta} I_{(0,\infty)}(x)$ dan seandainya kita ingin menduga peluang $p(\theta) = P(X \geq 1) = e^{-1/\theta}$. Sebagai catatan bahwa, $\mu_1^{'}(\theta) = \mu = \theta$, maka penduga momen bagi θ adalah $\hat{\theta} = \overline{X}$. Jika model direparameterisasi dengan $p = p(\theta) = e^{-1/\theta} = e^{-1/\mu}$, maka $\mu = \mu(p) = -1/\ln p$, dan jika kita samakan, akan kita peroleh $\overline{X} = \mu(\hat{p}) = -1/\ln \hat{p}$, maka penduga momen dari *p* adalah $\hat{p} = e^{-1/\overline{X}}$. Dengan demikian, dalam kasus ini, $\hat{p} = p(\hat{\theta})$. Jika suatu kelas penduga memiliki sifat seperti ini, maka kelas penduga tersebut dikatakan memiliki sifat invarian.

Dengan demikian, untuk menduga $\tau(\theta)$, kita selesaikan dahulu $\overline{X} = \mu(\hat{\theta})$ untuk mendapatkan penduga momen bagi θ dan kita gunakan $\tau(\hat{\theta})$, atau mungkin mengekspresikan μ secara langsung dalam bentuk τ dan mencari jawaban $\overline{X} = \mu(\hat{\tau})$ untuk penduga momen bagi τ. Tidak jelas apakah kedua pendekatan akan selalu memberikan hasil yang sama, tetapi jika $\hat{\theta}$ adalah penduga momen bagi θ, kita akan merujuk $\tau(\hat{\theta})$ sebagai penduga momen bagi $\tau(\theta)$. Secara umum, jika penduga momen penduga momen parameter asli $\theta_1, \theta_2, \ldots, \theta_k$ telah diperoleh, maka $\hat{\tau}_j(\theta_1, \theta_2, \ldots, \theta_k) = \tau_j(\hat{\theta}_1, \hat{\theta}_2, \ldots, \hat{\theta}_k)$ akan digunakan untuk pendugaan fungsi parameter alami lainnya, daripada memerlukan persamaan momen yang diekspresikan langsung dalam bentuk τ_j.

Teladan 9. 5.

Perhatikan bila suatu contoh acak berasal dari sebaran Gamma, X_i ~ Gamma(θ, κ). Bisa diperlihatkan bahwa

$\mu_1^{'} = \mu = \kappa\theta$ dan $\mu_2^{'} = \sigma^2 + \mu^2 = \kappa\theta^2 + \kappa^2\theta^2 = \kappa(1+\kappa)\theta^2$

dengan demikian

$$\kappa\theta = \sum_{i=1}^{n} \frac{X_i}{n} = \overline{X} \quad \text{dan} \quad \kappa(1+\kappa)\theta^2 = \sum_{i=1}^{n} \frac{X_i^2}{n}$$

Penduga momen bagi kedua parameter tersebut adalah

$$\hat{\theta} = \sum_{i=1}^{n} \frac{(X_i - \overline{X})^2}{n\overline{X}} = \frac{[(n-1)/n]S^2}{\overline{X}} \quad \text{dan} \quad \hat{\kappa} = \frac{\overline{X}}{\hat{\theta}}$$

Metode Kemungkinan Maksimum

Metode ini seringkali menghasilkan penduga yang memiliki properti yang diinginkan, khususnya properti contoh berukuran besar. Idenya adalah, sebagai dugaan parameter yang tak diketahui, menggunakan suatu nilai dalam ruang parameter yang berkenaan dengan "kemungkinan" terbesar data amatan.

Teladan 9. 6.

Dari percobaan pelemparan sebuah koin yang tidak seimbang, diketahui bahwa rata-rata proporsi munculnya Gambar adalah salah satu dari tiga nilai berikut, p = *0,20 ; 0,30 ;* atau *0,80*. Sebuah percobaan melempar koin tersebut dua kali dan diamati jumlah munculnya Gambar. Secara matematis, ini dapat dimodelkan sebagai suatu contoh acak X_1, X_2 berukuran $n = 2$ dari sebaran Bernoulli, $X_i \sim Bin(1,p)$ dimana ruang parameternya adalah $\Omega = \{0{,}20; 0{,}30; 0{,}80\}$. Telah kita ketahui bahwa atau dapat ditunjukkan dengan mudah bahwa penduga momen bagi p adalah $\overline{X}$, tidak akan menghasilkan nilai yang beralasan atau diinginkan karena $\bar{x}$ = 0 ; 0,5 atau 1. Dan nilai-nilai ini tidak berada di dalam ruang parameter.

Fungsi kepekatan peluang bersama dari contoh acak tersebut adalah

$$f(x_1, x_2; p) = p^{x_1+x_2}(1-p)^{2-x_1-x_2}$$

untuk x_i = 0 atau 1. Nilai-nilai fungsi kepekatan peluang tersebut dapat dilihat pada tabel berikut.

	(x_1,x_2)			
p	(0,0)	(0,1)	(1,0)	(1,1)
0,20	0,64	0,16	0,16	0,04
0,30	0,49	0,21	0,21	0,09
0,80	0,04	0,16	0,16	0,64

Misalkan percobaan menghasilkan pasangan (x_1,x_2) = (0,0). Dari tabel diatas p yang akan dipilih adalah 0,20. Dengan cara yang sama, jika (x_1,x_2) = (0,1) atau (1,0) maka nilai p yang akan dipilih adalah 0,30, serta (x_1,x_2) = (1,1) mengakibatkan terpilihnya nilai p = 0,80. Dengan demikian nilai yang memaksimumkan “kemungkinan” untuk tiap pasangan nilai (x_1,x_2) adalah

$$\hat{p} = \begin{cases} 0{,}20 & jika \quad (x_1, x_2) = (0{,}0) \\ 0{,}30 & jika \quad (x_1, x_2) = (0{,}1), (1{,}0) \\ 0{,}80 & jika \quad (x_1, x_2) = (1{,}1) \end{cases}$$

Secara umum, untuk suatu gugus peubah acak diskrit, fungsi kepekatan peluang bersama dari contoh acak yang dievaluasi pada gugus data contoh tertentu, misalnya $f(x_1,\ldots,x_n;\theta)$, menunjukkan peluang bahwa gugus data amatan $x_1, \ldots, x_n$ akan muncul. Untuk peubah acak kontinu, $f(x_1,\ldots,x_n;\theta)$ bukan merupakan peluang namun tetap menggambarkan "kemungkinan" relatif bahwa gugus data tertentu akan terjadi, dan kemungkinan ini tergantung pada nilai parameter sesungguhnya.

Definisi 9. 4.

Fungsi kepekatan peluang dari *n* peubah acak $X_1, \ldots, X_n$ yang dievaluasi pada $x_1, \ldots, x_n$, sebut saja $f(x_1,\ldots,x_n;\theta)$ dikenal dengan **fungsi kemungkinan**. Untuk nilai-nilai $x_1, \ldots, x_n$ fungsi kemungkinan adalah suatu fungsi dari θ dan sering dinotasikan dengan $L(\theta)$. Jika $X_1, \ldots, X_n$ merupakan contoh acak dari $f(x;\theta)$, maka

$$L(\theta) = f(x_1;\theta)\ldots f(x_n;\theta) = \prod_{i=1}^{n} f(x_i;\theta)$$

Untuk suatu gugus data yang diberikan, $L(\theta)$ memberikan kemungkinan gugus tersebut sebagai fungsi dari θ. Prinsip pendugaan dengan kemungkinan maksimum ini adalah memilih dugaan θ, untuk gugus data yang diberikan, sehingga nilai untuk gugus data yang teramati akan hampir pasti terjadi. Degan demikian, jika kemungkinan mengamati suatu pengamatan akan jauh lebih tinggi bilamana $\theta = \theta_1$ dari pada bila $\theta = \theta_1$, maka sangat beralasan kita pilih θ_1 sebagai dugaan θ dari pada θ_2.

Definisi 9. 5.

Misalkan $L(\theta) = f(x_1,\ldots,x_n;\theta), \theta \in \Omega$ merupakan fungsi kepekatan peluang bersama dari $X_1, \ldots, X_n$. Untuk suatu gugus amatan yang diberikan, $(x_1, \ldots, x_n)$, nilai $\hat{\theta}$ dalam Ω bilamana $L(\theta)$ maksimum, disebut dengan dugaan kemungkinan maksimum dari θ. Dengan demikian, $\hat{\theta}$ adalah nilai θ yang memenuhi

$$f(x_1,...,x_n;\hat{\theta}) = \max_{\theta\in\Omega} f(x_1,...,x_n;\theta)$$

Perlu diketahui bahwa jika tiap gugus pengamatan ($x_1, \ldots, x_n$) berkenaan dengan suatu nilai $\hat{\theta}$ yang khas, maka prosedur ini mendefinisikan suatu fungsi, yaitu $\hat{\theta} = t(x_1,...,x_n)$. Fungsi yang sama, bilamana diaplikasikan untuk contoh acak, $\hat{\theta} = t(X_1,...,X_n)$ disebut dengan **penduga kemungkinan maksimum**, selanjutya disingkat dengan PKM. Biasanya, notasi yang sama, $\hat{\theta}$, digunakan untuk dugaan kemungkinan maksimum dan penduga kemungkinan maksimumnya.

Dalam banyak aplikasi $L(\theta)$ merepresentasikan fungsi kepekatan peluang bersama suatu contoh acak, meskipun prinsip-prinsip kemungkinan maksimum juga berlaku untuk kasus yang lain seperti gugus statistik tataan.

Jika Ω merupakan interval terbuka, dan jika $L(\theta)$ differensiabel dan dan memiliki maksimum pada Ω, maka PKM merupakan jawaban dari persamaan $\frac{d}{d\theta}L(\theta) = 0$. Jika terdapat satu atau lebih jawaban yang memenuhi, perlu diverifikasi, jika ada, yang memaksimumkan $L(\theta)$. Setiap nilai θ yang memaksimumkan $L(\theta)$ juga akan memaksimumkan fungsi log-kemungkinan, $\ln L(\theta)$. Sehingga untuk kenyamanan perhitungan, sering digunakan $\frac{d}{d\theta}\ln L(\theta) = 0$.

Teladan 9. 7.

Suatu contoh acak dari sebaran Poisson, $X_i \sim Poi(\theta)$. Maka fungsi kemungkinannya adalah

$$L(\theta) = \prod_{i=1}^{n} f(x_i;\theta) = \prod_{i=1}^{n} \frac{e^{-\theta}\theta^{x_i}}{x_i!} = \frac{e^{-n\theta}\theta^{\sum_{i=1}^{n} x_i}}{\prod_{i=1}^{n} x_i!}$$

dan fungsi log-kemungkinan nya adalah

$$\ln L(\theta) = -n\theta + \sum_{i=1}^{n} x_i \ln\theta - \ln\left(\prod_{i=1}^{n} x_i!\right)$$

sehingga diperoleh persamaan kemungkinan maksimumnya adalah

$$\frac{d}{d\theta}\ln L(\theta) = -n + \frac{\sum_{i=1}^{n} x_i}{\theta} = 0$$

yang tentunya akan menghasilkan $\hat{\theta} = \frac{\sum_{i=1}^{n} x_i}{n} = \bar{x}$. Dimungkinkan untuk memverifikasi bahwa nilai ini akan memaksimumkan fungsi log-kemungkinan dengan menggunakan uji turunan kedua. Karena $\frac{d^2}{d\theta^2}\ln L(\theta) = -\sum_{i=1}^{n} \frac{x_i}{\theta^2}$ < 0. Dengan demikian, memang benar bahwa $\hat{\theta} = \frac{\sum_{i=1}^{n} x_i}{n} = \bar{x}$ adalah PKM bagi θ.

Teladan 9. 8.

Masih berkenaan dengan teladan sebelum ini, seandainya kita ingin menduga τ = $\tau(\theta)$ = *P[X=0]* = $e^{-\theta}$. Dengan cara reparametrisasi dalam bentuk τ dengan memisalkan θ = *-ln* τ untuk memperoleh

$$L^*(\tau) = \frac{\tau^n(-\ln\tau)^{\sum_{i=1}^{n} x_i}}{\prod_{i=1}^{n} x_i!}$$

yang juga merepresentasikan fungsi kemungkinan relatif terhadap τ. Sehingga

$$\ln L^*(\tau) = n \ln \tau + \sum_{i=1}^{n} x_i \ln(-\ln \tau) - \ln \prod_{i=1}^{n} x_i$$

$$\frac{d}{d\tau} \ln L^*(\tau) = \frac{n}{\tau} + \frac{\sum_{i=1}^{n} x_i}{-\ln \tau}\left(\frac{-1}{\tau}\right)$$

Dengan demikian

$$\frac{n}{\hat{\tau}} + \frac{\sum_{i=1}^{n} x_i}{-\ln \hat{\tau}}\left(\frac{-1}{\hat{\tau}}\right) = 0$$

akan menghasilkan $\hat{\tau} = e^{-\bar{x}}$. Teladan ini juga menunjukkan bahwa $\hat{\tau} = \hat{\tau}(\theta) = \tau(\hat{\theta})$.

Teorema 9. 1.

Jika $\hat{\theta}$ adalah PKM bagi θ dan jika $u(\theta)$ merupakan fungsi dari θ, maka $u(\hat{\theta})$ adalah PKM bagi $u(\theta)$.

Teorema tersebut menjelaskan bahwa jika kita melakukan reparametrisasi $\tau = \tau(\theta)$, maka PKM bagi τ adalah $\hat{\tau} = \tau(\hat{\theta})$.

Teladan 9. 9.

Misalkan suatu contoh acak yang berasal dari sebaran eksponensial, X_i ~ Eksponensial(θ). Fungsi kemungkinan dari contoh acak berukuran n ini adalah

$$L(\theta) = \frac{1}{\theta^n} e^{-\sum_{i=1}^{n} x_i / \theta} \qquad 0 < x_i < \infty$$

Dengan demikian,

$$\ln L(\theta) = -n \ln\theta - \frac{\sum_{i=1}^{n} x_i}{\theta}$$ dan $$\frac{d}{d\theta}\ln L(\theta) = \frac{-n}{\theta} + \frac{\sum_{i=1}^{n} x_i}{\theta^2}.$$

Selanjutnya, $\frac{-n}{\hat{\theta}} + \frac{\sum_{i=1}^{n} x_i}{\hat{\theta}^2} = 0$ akan diperoleh bahwa $\hat{\theta} = \bar{x}$.

Apabila diinginkan untuk menduga $p(\theta) = P(X \geq 1) = e^{-1/\theta}$, dengan menggunakan teorema terakhir, bahwa $p(\hat{\theta}) = e^{-1/\bar{x}}$.

Terdapat beberapa kasus dimana PKM ada tetapi tak dapat diperoleh sebagai jawaban dari persamaan kemungkinan maksimum.

Teladan 9. 10.

Sebuah contoh acak berasal dari sebaran eksponensial dua parameter, X_i ~ Exp(1,η). Fungsi kemungkinan atau fungsi likelihood nya adalah

$$L(\eta) = \prod_{i=1}^{n} e^{-(x_i-\eta)} I_{[\eta,\infty)}(x_i) = e^{-\sum_{i=1}^{n}(x_i-\eta)} \prod_{i=1}^{n} I_{[\eta,\infty)}(x_i)$$

Jika kita notasikan minimum dari $x_1, \ldots, x_n$ dengan $x_{(1)}$ maka fungsi likelihood tadi dapat dituliskan dengan

$$L(\eta) = e^{n(\eta-\bar{x})} I_{[\eta,\infty)}(x_{(1)})$$

Kita tahu bahwa semua nilai amatan tidak lebih kecil dari η, sehingga $e^{n(\eta-\bar{x})}$ akan maksimum apabila $\eta - \bar{x}$ sekecil mungkin. Dengan demikian fungsi kemungkinan akan maksimum apabila $\hat{\eta} = x_{(1)}$, dan penduga kemungkinan maksimum (PKM) nya adalah statistik minimum.

Ini merupakan salah satu teladan yang menghasilkan penduga momen dan penduga kemungkinan maksimum berbeda.

Teladan 9. 11.

Umur suatu komponen elektronik mengikuti sebaran Eksponensial dengan parameter θ. Misalkan sejumlah n komponen diambil secara acak dan dilakukan pengujian, dan pengamatan dilakukan untuk sejumlah r komponen yang gagal berfungsi pertama yang dinotasikan dengan $x_{(1)}, x_{(2)}, \ldots, x_{(r)}$. Fungsi kepekatan peluang bersama dari $X_{(1)}, X_{(2)}, \ldots, X_{(r)}$ adalah

$$L(\theta) = f(x_{(1)},\ldots,x_{(r)};\theta)$$

$$= \frac{n!}{(n-r)!}\exp\left(\frac{-\sum_{i=1}^{r} x_{(i)}}{\theta}\right)\exp\left[\frac{-(n-r)x_{(r)}}{\theta}\right]\left(\frac{1}{\theta}\right)^r$$

$$= \frac{n!}{(n-r)!\theta^r}\exp\frac{-\left[\sum_{i=1}^{r} x_{(i)} + (n-r)x_{(r)}\right]}{\theta}$$

Sebagai catatan bahwa $T = \sum_{i=1}^{r} X_{(i)} + (n-r)X_{(r)}$ merupakan total *survival time* dari sebanyak n komponen yang diuji hingga percobaan dihentikan. Untuk mendapatkan PKM bagi θ berdasarkan data tersebut, maka

$$\ln L(\theta) = \ln\left(\frac{n!}{(n-r)!}\right) - r\ln\theta - \frac{T}{\theta}$$

$$\frac{d}{d\theta}\ln L(\theta) = \frac{-r}{\theta} + \frac{T}{\theta^2}$$

Dengan mengatur bahwa turunan atau diferensial sama dengan nol $\frac{-r}{\hat{\theta}} + \frac{T}{\hat{\theta}^2} = 0$ akan diperoleh $\hat{\theta} = \frac{T}{r}$.

Teladan-teladan terdahulu menggunakan sebaran dengan satu parameter yang tak diketahui. Definisi fungsi kemungkinan dan penduga kemungkinan maksimum dapat diterapkan dalam kasus jumlah parameter lebih dari satu jika θ merepresentasikan suatu vektor parameter, sebut saja $\boldsymbol{\theta} = (\theta_1, \ldots, \theta_k)$. Meskipun Ω secara umum dapat memiliki k-dimensi, dalam kebanyakan teladan merupakan hasil kali Cartesian *k* interval. Apabila Ω dalam bentuk ini dan juga turunan parsial $L(\theta_1, \ldots\theta_k)$ ada, dan PKM tidak terjadi pada batas Ω, maka PKM merupakan jawaban dari

$$\frac{\partial}{\partial\theta_j}\ln L(\theta_1,\ldots,\theta_k) = 0$$

untuk $j = 1, \ldots, k$. Ini yang disebut dengan **persamaan kemungkinan maksimum**, dan jawabannya dinotasikan dengan $\hat{\theta}_1,\ldots\hat{\theta}_k$. Sebagaimana dalam kasus satu parameter, umumnya perlu diverifikasi bahwa jawab dari persamaan kemungkinan maksimum memaksimumkan $L(\theta_1, \ldots\theta_k)$.

Teorema 9. 2.

Jika $\hat{\theta} = (\hat{\theta}_1,\ldots,\hat{\theta}_k)$ merupakan PKM bagi $\theta = (\theta_1,\ldots,\theta_k)$, maka PKM bagi $\tau = (\tau_1(\theta),\ldots,\tau_r(\theta))$ adalah $\hat{\tau} = (\hat{\tau}_1,\ldots,\hat{\tau}_r) = (\tau_1(\hat{\theta}),\ldots,\tau_r(\hat{\theta}))$ untuk $1 \le r \le k$.

Teladan 9. 12.

Untuk suatu gugus peubah acak $X_i \sim N(\mu,\sigma^2)$, kita ingin mencari penduga kemungkinan maksimum (PKM) bagi μ dan $\theta = \sigma^2$ berdasarkan contoh acak berukuran *n*. Kita punya

$$f(x;\mu,\theta) = \frac{1}{\sqrt{2\pi\theta}}e^{-(x-\mu)^2/2\theta}$$

$$L(\mu,\theta) = (2\pi\theta)^{-n/2} \exp\left[\frac{-\sum_{i=1}^{n}(x_i-\mu)^2}{2\theta}\right]$$

$$\ln L(\mu,\theta) = -\frac{n}{2}\ln(2\pi) - \frac{n}{2}\ln\theta - \frac{\sum_{i=1}^{n}(x_i-\mu)^2}{2\theta}$$

$$\frac{\partial \ln L(\mu,\theta)}{\partial\mu} = \frac{2\sum_{i=1}^{n}(x_i-\mu)}{2\theta} \quad \text{dan}$$

$$\frac{\partial \ln L(\mu,\theta)}{\partial\theta} = \frac{-n}{2\theta} + \frac{\sum_{i=1}^{n}(x_i-\mu)^2}{2\theta^2}$$

Dengan demikian, PKM bagi μ dan θ dapat dicari dengan cara menyelesaikan kedua persamaan berikut

$$\frac{2\sum_{i=1}^{n}(x_i-\hat{\mu})}{2\hat{\theta}} = 0 \quad \text{dan} \quad \frac{-n}{2\hat{\theta}} + \frac{\sum_{i=1}^{n}(x_i-\hat{\mu})^2}{2\hat{\theta}^2} = 0$$

Dengan demikian kita dapat peroleh

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

Teladan 9. 13.

Sebuah contoh acak dari sebaran Eksponensial dua parameter, $X_i \sim$ *Eksp*(θ,η) . Fungsi kepekatan peluang populasi

$$f(x;\theta,\eta) = \frac{1}{\theta}\exp[-(x-\eta)/\theta] I_{\eta,\infty)}(x)$$

Fungsi likelihood nya adalah

$$L(\theta,\eta)=\prod_{i=1}^{n}\frac{1}{\theta}\exp[-(x_i-\eta)/\theta]I_{\eta,\infty)}(x_i)=\left(\frac{1}{\theta}\right)^n\exp\left[\frac{\sum_{i=1}^{n}(x_i-\eta)}{\theta}\right]I_{[\eta,\infty)}(x_{(1)})$$

dan fungsi log-likelihood nya adalah

$$\ln L(\theta,\eta)=-n\ln\theta-\frac{\left[\sum_{i=1}^{n}x_i-n\eta\right]}{\theta}I_{[\eta,\infty)}(x_{(1)})$$

Fungsi likelihood akan maksimum untuk η dengan mengambil nilai $\hat{\eta}=x_{(1)}$. Untuk maksimissi relatif terhadap θ, kita perlu mencari turunan dari $\ln L(\theta,\hat{\eta})$ terhadap θ, dan jawab persamaan hasil turunannya setelah diatur nilainya sama dengan nol.

$$\frac{d\ln L(\theta,\hat{\eta})}{d\theta}=\frac{-n}{\theta}+\frac{\sum_{i=1}^{n}(x_i-\hat{\eta})}{\theta^2}$$

Kemudian

$$\frac{-n}{\hat{\theta}}+\frac{\sum_{i=1}^{n}(x_i-\hat{\eta})}{\hat{\theta}^2}=0$$

sehingga diperoleh

$$\hat{\theta}=\frac{\sum_{i=1}^{n}(x_i-\hat{\eta})}{n}=\overline{x}-\hat{\eta}=\overline{x}-x_{(1)}$$

Persentil ke-α, yang dinotasikan dengan x_α, adalah nilai yang sedemikian rupa sehingga $F(x_\alpha)=\alpha$. Dengan demikian, dalam kasus ini, $x_\alpha=-\theta\ln(1-\alpha)+\eta$. PKM bagi x_α adalah $\hat{x}_\alpha=-\hat{\theta}\ln(1-\alpha)+\hat{\eta}$, berdasarkan Teorema Invarian.

Kriteria untuk Mengevaluasi Penduga

Terdapat beberapa kriteria penduga yang dapat dipakai untuk mengevaluasi apakah penduga tersebut baik atau tidak, seperti misalnya ketidakbiasan.

Definisi 9. 6.

Sebuah penduga T dikatakan sebagai **penduga tak bias** bagi $\tau(\theta)$ jika $E(T) = \tau(\theta)$ untuk semua $\theta \in \Omega$. Jika tidak, kita katakan bahwa T merupakan **penduga bias** bagi $\tau(\theta)$.

Jika suatu penduga tak bias digunakan untuk memberikan nilai bagi $\tau(\theta)$, nilai $\tau(\theta)$ yang sebenarnya mungkin tak akan dapat dipenuhi dengan sembarang nilai dugaannya, t, tetapi nilai "rata-rata" T akan sama dengan $\tau(\theta)$.

Dimungkinkan untuk mendapatkan sebuah penduga beralasan yang bias, dan seringkali penduga tersebut dapat disesuaikan menjadi penduga yang tak bias.

Teladan 9. 14.

Sebuah contoh acak berukuran n diambil dari sebaran eksponensial, X_i ~ Eksponensial(θ). Karena θ adalah juga merupakan rata-rata (mean) sebaran Eksponensial, kita tahu bahwa PKMnya, $\overline{X}$, tidak bias bagi θ. Jika kita inginkan penduga bagi kebalikan rata-rata, $\tau(\theta) = 1/\theta$, maka dengan sifat invarian, PKM bagi 1/θ adalah $T_1 = 1/\overline{X}$. Namun demikian T_1 merupakan penduga yang bias bagi $1/\theta$. Dari pembahasan pada bab-bab terdahulu

$$Y = \frac{2n\overline{X}}{\theta} = \sum_{i=1}^{n} \frac{2X_i}{\theta} \sim \chi^2_{2n}$$

Selanjutnya $E(Y^{-1}) = 1/[2(n-1)]$, dan dengan demikian $E(T_1) = [n/(n-1)](1/\theta)$. Meskipun terlihat bahwa T_1 merupakan penduga yang bias terhadap $1/\theta$, penduga tak bias bagi $1/\theta$ dapat dicari dengan cara menyesuaikan koeffisien sedemikian rupa sehingga $E(T_1) = 1/\theta$. Dengan demikian, penduga tak bias bagi $1/\theta$ adalah $[(n-1)/n]T_1$.

Tidak selalu dimungkinkan untuk menyesuaikan penduga yang bias menjadi tak bias dengan cara ini.

Teladan 9. 15.

Misalkan diinginkan untuk mengestimasi $1/\theta$ dengan hanya menggunakan statistik terkecil atau $X_{(1)}$. Dapat juga diperlihatkan bahwa $X_{(1)}$ ~ Eksponensial(θ/n) dan sebagai konsekuensinya $nX_{(1)}$ adalah penduga yang tak bias bagi θ. Hal serupa menyarankan bahwa $T_2 = 1/(nX_{(1)})$ dapat juga digunakan untuk menduga $1/\theta$. Namun, statistik T_2 tak dapat disesuaikan dengan cara seperti agar tak bias untuk $1/\theta$, karena $E(T_2)$ tidak ada.

Statistik T_1 dan T_2 menggambarkan suatu kesalahan atau ketidaksempurnaan dalam konsep ketidakbiasan sebagai prinsip yang umum. Secara khusus, jika $\hat{\theta}$ sebagai penduga tak bias bagi θ, maka $\tau(\hat{\theta})$ belum tentu merupakan penduga tak bias bagi $\tau(\theta)$. Namun demikian, $\tau(\hat{\theta})$ mungkin merupakan penduga beralasan bagi $\tau(\theta)$.

Kadang-kadang dimungkinkan untuk menurunkan beberapa penduga potensial yang berbeda untuk suatu parameter. Sebagai teladan, dalam berbagai kasus, penduga momen dan penduga kemungkinan maksimum pada dasarnya memiliki bentuk dasar yang sama. Hal ini menimbulkan pertanyaan yang gampang seperti bagaimana memilih penduga mana yang "terbaik" dalam beberapa hal.

Ide yang sangat umum adalah memilih penduga yang cenderung terdekat atau "paling terkonsentrasi" disekitar nilai yang sebenarnya dari parameter yang diduga. Mungkin lebih beralasan untuk mengatakan bahwa T_1 **lebih terkonsentrasi** *(more concentrated)* daripada T_2 disekitar $\tau(\theta)$ jika

$$P[\tau(\theta)-\varepsilon < T_1 < \tau(\theta)+\varepsilon] \geq P[\tau(\theta)-\varepsilon < T_2 < \tau(\theta)+\varepsilon]$$

untuk semua $\varepsilon > 0$, dan juga sebuah penduga dikatakan **paling terkonsentrasi** *(most concentrated)* jika penduga tersebut lebih terkonsentrasi dari setiap penduga lainnya.

Tidaklah jelas bagaimana mendapatkan sebuah penduga yang paling terkonsentrasi, namun beberapa konsep akan dibahas yang mungkin secara parsial menjawab pertanyaan ini. Sebagai misal, jika T merupakan penduga yang tak bias bagi $\tau(\theta)$, maka menurut peridaksamaan Chebychev kita dapatkan

$$P[\tau(\theta)-\varepsilon < T < \tau(\theta)+\varepsilon] \geq 1 - Var(T)/\varepsilon^2$$

untuk semua $\varepsilon > 0$. Dapat disimpulkan bahwa penduga-penduga yang tak bias, dengan ragam yang lebih kecil cenderung akan lebih terkonsentrasi dan dengan demikian inilah yang lebih disukai (dipakai).

Teladan 9. 16.

Dua teladan terakhir menggambarkan bahwa $\hat{\theta}_1 = \overline{X}$ dan $\hat{\theta}_2 = nX_{(1)}$ merupakan dua penduga yang tak bias bagi θ, tetapi $Var(\hat{\theta}_1) = \theta^2/n$ dan $Var(\hat{\theta}_2) = \theta^2$. Dengan demikian, untuk $n > 1$, $Var(\hat{\theta}_1) = \theta^2/n <$ $Var(\hat{\theta}_2) = \theta^2$, dan $\hat{\theta}_1 = \overline{X}$ merupakan penduga yang lebih baik dengan kriteria ini.

Dalam beberapa kasus sebuah penduga dapat memiliki ragam yang lebih kecil untuk beberapa nilai θ dan lebih besar untuk nilai θ lainnya. Dengan demikian tak ada penduga yang dikatakan lebih baik dari penduga lainnya secara umum. Dalam kasus-kasus tertentu, dimungkinkan untuk memperlihatkan bahwa sebuah penduga tak bias memiliki ragam terkecil diantara semua penduga tak bias untuk semua kemungkinan nilai θ. Dalam kasus ini, kita perlu membatasi perhatian kita pada penduga tersebut.

Penduga Tak-Bias Ragam Minimum Seragam

Teorema 9. 3.

Misalkan $X_1, X_2, \ldots, X_n$ merupakan contoh acak berukuran n dari $f(x;\theta)$. Sebuah penduga T^* bagi $\tau(\theta)$ dikatakan sebagai **penduga tak bias ragam minimum seragam** atau **uniformly minimum variance unbiased estimator** (UMVUE) bagi $\tau(\theta)$ jika

1. T^* merupakan penduga tak bias bagi $\tau(\theta)$, dan
2. untuk sembarang penduga tak bias T bagi $\tau(\theta)$, berlaku

Dalam beberapa kasus, batas bawah dapat diturunkan dari ragam penduga tak biasnya. Jika suatu penduga tak bias yang diperoleh mencapai nilai batas bawah ini, maka penduga tersebut dapat dikatakan sebagai penduga tak-bias ragam minimum seragam (UMVUE). Batas bawah untuk ragam penduga tak bias dapat dibangun, apabila turunan fungsi parameter dan turunan fugsi kepekatan peluang terhadap parameter ada serta diperbolehkannya pertukaran notasi turunan dan integral (penjumlahan). Tentunya, juga diperlukan bahwa daerah fungsi (domain) integran harus tidak boleh tergantung pada parameter θ.

Jika T merupakan penduga tak bias bagi $\tau(\theta)$, maka batas bawah Cramer-Rao atau **Cramer-Rao Lower Bound** (CRLB) bagi varian penduga berdasarkan contoh acak adalah

$$Var(T) \geq \frac{[\tau'(\theta)]^2}{nE\left[\frac{\partial}{\partial\theta}\ln f(X;\theta)\right]^2}$$

Dengan asumsi kondisi differensiabilitas sebagaimana telah disebutkan terdahulu, CRLB dapat dibuat seperti berikut. Kita akan gunakan asumsi bahwa yang berikut ini apabila peubahnya bersifat kontinu. Untuk kondisi peubah diskrit, analog dengan apa yang ada di bawah ini, hanya dengan menggantikan notasi integral dengan sigma untuk penjumlahan.

Misalkan kita memiliki fungsi yang didefinisikan sebagai berikut

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

yang dapat juga dituliskan menjadi

$$u(x_1,...,x_n;\theta)=\frac{1}{f(x_1,...,x_n;\theta)}\frac{\partial}{\partial\theta}f(x_1,...,x_n;\theta)$$

Jika kita definisikan suatu peubah acak $U = u(X_1,...,X_n;\theta)$, maka

$$\begin{aligned} E(U) &= \int\cdots\int u(x_1,...,x_n;\theta)f(x_1,...,x_n;\theta)dx_1\cdots dx_n \\ &= \int\cdots\int \frac{\partial}{\partial\theta}f(x_1,...,x_n;\theta)dx_1\cdots dx_n \\ &= \frac{\partial}{\partial\theta}\int\cdots\int f(x_1,...,x_n;\theta)dx_1\cdots dx_n \\ &= \frac{\partial}{\partial\theta}1 \\ &= 0 \end{aligned}$$

Seandainya juga $T = t(X_1,...,X_n)$ merupakan penduga tak biasa bagi $\tau(\theta)$, maka kita akan peroleh

$$\tau(\theta) = E(T) = \int\cdots\int t(x_1,...,x_n)f(x_1,...,x_n;\theta)dx_1\cdots dx_n$$

Jika kita turunkan terhadap θ, maka

$$\begin{aligned}
\tau'(\theta) &= \frac{\partial}{\partial\theta}\int\cdots\int t(x_1,...,x_n)f(x_1,...,x_n;\theta)dx_1\cdots dx_n \\
&= \int\cdots\int t(x_1,...,x_n)\frac{\partial}{\partial\theta}f(x_1,...,x_n;\theta)dx_1\cdots dx_n \\
&= \int\cdots\int t(x_1,...,x_n)u(x_1,...,x_n;\theta)f(x_1,...,x_n;\theta)dx_1\cdots dx_n \\
&= E(TU)
\end{aligned}$$

Adanya fakta bahwa $E(U)=0$, maka $Var(U)=E(U^2)$ dan $Cov(T,U)=E(TU)$. Karena koeffisien korelasi selalu berada pada kisaran –1 sampai dengan +1, maka kita dapatkan hubungan bahwa $[Cov(T,U)]^2 \le Var(T)Var(U)$ dan sebagai konsekuensinya $Var(T)E(U^2) \ge [\tau'(\theta)]^2$ sehingga

$$Var(T) \ge \frac{[\tau'(\theta)]^2}{E\left[\frac{\partial}{\partial\theta}\ln f(X_1,...,X_n;\theta)\right]^2}$$

Bilamana *X1, …, Xn* merepresentasikan contoh acak,

$$f(x_1,...,x_n;\theta) = f(x_1;\theta)\cdots f(x_n;\theta)$$

sehingga

$$u(x_1,...,x_n;\theta) = \sum_{i=1}^{n}\frac{\partial}{\partial\theta}\ln f(x_i;\theta)$$

dimana kita dapatkan

$$E(U^2) = Var(U) = nVar\left[\frac{\partial}{\partial\theta}\ln f(X;\theta)\right] = nE\left[\frac{\partial}{\partial\theta}\ln f(X;\theta)\right]^2$$

Akhirnya, kita dapatkan batas bawah Cramer-Rao untuk ragam penduga tak bias bagi $\tau(\theta)$.

Teladan 9. 17.

Sebuah contoh acak dari sebaran Eksponensial, $f(x;\theta) = \frac{1}{\theta} e^{-\frac{x}{\theta}}$.

Dengan demikian

$$\ln f(x;\theta) = -x/\theta - \ln\theta$$

$$\frac{\partial}{\partial\theta} \ln f(x;\theta) = x/\theta^2 - 1/\theta = (x-\theta)/\theta^2$$

$$E\left[\frac{\partial}{\partial\theta} \ln f(x;\theta)\right]^2 = E\left[(X-\theta)^2/\theta^4\right] = \theta^2/\theta^4 = 1/\theta^2$$

dan CRLB untuk $\tau(\theta) = \theta$ adalah $1/[n(1/\theta^2)] = \theta^2/n$. Karena $Var(\overline{X}) = \theta^2/n$, maka $\overline{X}$ adalah UMVUE bagi θ.

Dimungkinkan untuk memperoleh informasi tentang tipe penduga, yang memiliki varian sama dengan CRLB. Batas bawah ini dapat dicapai hanya apabila koeffisien korelasi antara *T* dan *U* adalah ± 1, atau dengan kata lain bahwa antara *T* dan *U* memiliki hubungan linier, sebut saja $T = aU + b$ dengan peluang 1 untuk konstanta $a \neq 0$ dan sembarang *b*. Jadi, agar *T* mencapai CRLB bagi $\tau(\theta)$ haruslah merupakan fungsi linier dari $\sum_{i=1}^{n} \frac{\partial}{\partial\theta} \ln f(x;\theta)$.

Teladan 9. 18.

Suatu contoh acak dari sebaran Geometrik, $f(x;\theta) = \theta(1-\theta)^{x-1}$.

Kita ingin mencari UMVUE bagi $\tau(\theta) = 1/\theta$. Dengan demikian

$$\ln f(x;\theta) = \ln\theta + (x-1)\ln(1-\theta)$$

$$\frac{\partial}{\partial\theta} \ln f(x;\theta) = \frac{1}{\theta} - \frac{x-1}{1-\theta} = \frac{x-1/\theta}{\theta-1}$$

$$E\left[\frac{\partial}{\partial\theta} \ln f(x;\theta)\right]^2 = E\left[\frac{X-1/\theta}{\theta-1}\right]^2 = \frac{1-\theta}{\theta^2(1-\theta)^2} = \frac{1}{\theta^2(1-\theta)}$$

$$CRLB = \frac{\left(-\frac{1}{\theta^2}\right)^2}{n\frac{1}{\theta^2(1-\theta)}} = \frac{(1-\theta)}{n\theta^2}$$

Untuk membuat ragam atau varian penduga tak bias *T* mencapai CRLB, haruslah memiliki bentuk $T = a\sum_{i=1}^{n}(X_i - 1/\theta)/(\theta - 1) + b$ yang juga dapat diekspresikan sebagai fungsi linier rataan contoh, $T = c\overline{X} + d$ untuk konstanta *c* dan *d*. Karena $\overline{X}$ merupakan penduga tak bias bagi $1/\theta$, maka diperlukan *c = 1* dan *d = 0*, sehingga $T = \overline{X}$ hanya satu-satunya penduga tak bias bagi $1/\theta$. Ragam bagi $\overline{X}$ adalah $Var(\overline{X}) = (1-\theta)/(n\theta^2)$ yang juga sama dengan CRLB. Jadi $\overline{X}$ adalah UMVUE bagi $1/\theta$.

Teorema 9. 4.

Jika suatu penduga tak bias bagi $\tau(\theta)$ ada, dimana ragamnya sama dengan CRLB, maka hanya fungsi linier dari $\tau(\theta)$ yang akan memiliki penduga tak bias, yang ragamnya juga sama dengan CRLB yang sesuai.

Dengan demikian, teladan diatas tak akan ada penduga tak bias yang ragamnya sama dengan CRLB untuk penduga tak bias bagi θ, karena θ bukan merupakan funsi linier dari $1/\theta$. Tak dapat disimpulkan dari sini bahwa UMVUE untuk θ tak ada, tetapi tidak dapat diperoleh dengan menggunakan pendekatan CRLB. Dalam bab berikutnya kita akan bahas suatu metode yang memungkinkannya bilamana pendekatan saat ini tak berhasil.

Pembandingan antar ragam penduga sering digunakan untuk memutuskan metode mana dalam penggunaan data lebih effisien.

Definisi 9. 7.

Relatif efisiensi dari sebuah penduga tak bias *T* bagi $\tau(\theta)$ terhadap penduga tak bias lainnya, *T** bagi $\tau(\theta)$, adalah $re(T,T^*) = \dfrac{\operatorname{var}(T^*)}{\operatorname{var}(T)}$.

Suatu penduga tak bias *T** bagi $\tau(\theta)$ dikatakan efisien jika *re(T,T*)* $\leq$ *1* untuk semua penduga tak bias *T* bagi $\tau(\theta)$, dan $\theta \in \Omega$. Efisiensi suatu penduga tak bias *T* bagi $\tau(\theta)$ diberikan dengan rumus *e(T) = re(T,T*)* jika *T** adalah penduga efisien bagi $\tau(\theta)$.

Teladan 9. 19.

Dalam beberapa teladan terdahulu (contoh acak dari sebaran eksponensial dengan parameter θ), kita memiliki penduga tak bias $\hat{\theta}_1 = \overline{X}$ dan $\hat{\theta}_2 = nX_{(1)}$ bagi θ. Telah kita ketahui juga bahwa $\hat{\theta}_1 = \overline{X}$ adalah UMVUE. Dengan demikian $\hat{\theta}_1 = \overline{X}$ adalah penduga efisien bagi θ, dan efisiensi $\hat{\theta}_2 = nX_{(1)}$ adalah

$$e(\hat{\theta}_2) = re(\hat{\theta}_2, \hat{\theta}_1) = \frac{\theta^2 / n}{\theta^2} = \frac{1}{n}$$

dan dengan demikian $\hat{\theta}_2 = nX_{(1)}$ merupakan penduga yang tak baik bagi θ karena nilai efisiensinya akan kecil apabila ukuran sampelnya besar.

Penduga yang agak bias atau sedikit bias namun sangat terkonsentrasi disekitar parameter yang dipelajari mungkin lebih disukai daripada penduga tak bias namun kurang terkonsentrasi. Dengan demikian, diperlukan kriteria yang lebih umum yang memungkinkan baik penduga bias dan tak bias diperbandingkan.

Teorema 9. 5.

Jika *T* adalah penduga bagi $\tau(\theta)$, maka besarnya bias adalah

$$b_T = E(T) - \tau(\theta)$$

dan kuadrat tengah galat bagi *T* adalah

$$KTG(T) = E[T - \tau(\theta)]^2$$

Teorema 9. 6.

Jika *T* adalah penduga bagi $\tau(\theta)$, maka

$$KTG(T) = Var(T) + [b_T]^2$$

Bukti

$$\begin{aligned} KTG(T) &= E[T - \tau(\theta)]^2 \\ &= E[T - E(T) + E(T) - \tau(\theta)]^2 \\ &= E[T - E(T)]^2 + 2[E(T) - \tau(\theta)][E(T) - E(T)] + [E(T) - \tau(\theta)]^2 \\ &= Var(T) + [b_T]^2 \end{aligned}$$

Teladan 9. 20.

Suatu contoh acak dari sebaran dengan $f(x;\eta) = e^{-(x-\eta)} I_{[\eta,\infty)}(x)$. Kita ingin membandingkan Penduga Momen dan Penduga Kemungkinan Maksimum bagi η. Berdasarkan teladan yang sudah ada, penduga momennya adalah $\hat{\eta}_1 = \overline{X} - 1$ dan penduga kemungkinan maksimumnya adalah $\hat{\eta}_2 = X_{(1)}$. Dapat diperlihatkan dengan metode yang sudah ada bahwa $\overline{X} - \eta$ ~ Gamma(*1/n;n*) dan $X_{(1)} - \eta$ ~Eksp(*1/n*).

Dengan demikian kita dapatkan

$$E(\hat{\eta}_1) = E(\overline{X} - 1) = E(\overline{X}) - 1 = 1 + \eta - 1 = \eta$$

$$E(\hat{\eta}_2) = E(X_{(1)}) = E(X_{(1)} - \eta + \eta) = E(X_{(1)} - \eta) + \eta = \frac{1}{n} + \eta$$

Dengan demikian, $\hat{\eta}_1 = \overline{X} - 1$ adalah penduga tak bias dan $\hat{\eta}_2 = X_{(1)}$ merupakan penduga yang bias dengan besarnya bias $b_{\hat{\eta}_2} = 1/n$. Kuadrat tengah galat masing-masing dapat dicari sebagai berikut:

$$KTG(\hat{\eta}_1) = Var(\overline{X} - 1) = Var(\overline{X}) = 1/n$$

dan

$$KTG(\hat{\eta}_2) = Var(\hat{\eta}_2) + (1/n)^2 = Var(X_{(1)}) + (1/n)^2 = Var(X_{(1)} - \eta) + (1/n)^2$$

$$= (1/n)^2 + (1/n)^2 = 2/n^2$$

Dengan demikian, untuk *n* > 2 penduga yang bias memiliki lebih kecil kuadrat tengah galat daripada penduga tak biasnya.

Dimungkinkan untuk membuat penyesuaian dari peubah bias $\hat{\eta}_2 = X_{(1)}$ menjadi tak bias, misalnya dengan membuat $\hat{\eta}_3 = X_{(1)} - 1/n$, sehingga

$$E(\hat{\eta}_3) = E(X_{(1)} - 1/n) = E(X_{(1)}) - 1/n = \eta + 1/n - 1/n = \eta$$

$$KTG(\hat{\eta}_3) = Var(\hat{\eta}_3) = Var(X_{(1)}) = Var(X_{(1)} - \eta) = 1/n^2$$

Untuk *n* > *1*, $\hat{\eta}_3 = X_{(1)} - 1/n$ memiliki kuadrat tengah galat terkecil diantara ketiganya.

Apabila nilai dugaan berbeda dengan nilai sebenarnya dari parameter yang diestimasi, perlu dipertimbangkan kerugian yang juga merupakan fungsi dari perbedaan ini.

Definisi 9. 8.

Jika *T* adalah penduga bagi $\tau(\theta)$, maka fungsi kerugian adalah sembarang fungsi bilangan nyata $L(t;\theta)$ sedemikian rupa sehingga

$L(t;\theta) \geq 0$ untuk setia nilai *t,* dan

$L(t;\theta) = 0$ bilamana $t = \tau(\theta)$

Definisi 9. 9.

Fungsi resiko didefinisikan sebagai rata-rata kerugian,

$$R_T(\theta) = E[L(T;\theta)]$$

Definisi 9. 10.

Suatu penduga T_1 adalah penduga yang lebih baik dari penduga T_2 jika dan hanya jika $R_{T_1}(\theta) \le R_{T_2}(\theta)$ untuk semua $\theta \in \Omega$, dan $R_{T_1}(\theta) < R_{T_2}(\theta)$ sedikitnya untuk satu θ.

Suatu penduga *T* dikatakan **admisibel** jika dan hanya jika tak ada penduga yang lebih baik. Dengan demikian, jika satu penduga memiliki resiko yang lebih kecil dan seragam, maka penduga ini perlu dipertimbangkan, dan hilangkan yang lain karena tidak admisibel.

Teladan 9. 21.

Pada teladan sebelum ini $\hat{\eta}_3 = X_{(1)} - 1/n$ adalah merupakan penduga tak bias, yang cukup beralasan digunakan untuk menduga parameter lokasi η. Bila kita punya kelas penduga dalam bentuk $\hat{\eta}_4 = c\hat{\eta}_3$ untuk beberapa konstanta *c* > 0. Penduga-penduga semacam ini akan bias untuk menduga η, kecuali apabila nilai *c = 1*, dan kuadrat tengah galatnya adalah $KTG(\hat{\eta}_4) = Var(c\hat{\eta}_3) + [b_{c\hat{\eta}_3}]^2 = c^2/n^2 + (c-1)^2\eta^2$.

Diskusi lebih jauh dapat diperlihatkan bahwa kuadrat tengah galat mana yang lebih kecil dibanding yang lainnya. Hal ini sangat tergantung dengan besarnya sampel yang diambil dan nilai parameter yang diduga.

Definisi 9. 11.

Suatu penduga T_1 dikatakan sebagai penduga minimax jika $\max_{\theta} R_{T_1}(\theta) \leq \max_{\theta} R_T(\theta)$ untuk setiap penduga *T*.

Dengan perkataan lain, T_1 adalah penduga yang meminimumkan maksimum resiko, atau

$$\max_{\theta} R_{T_1}(\theta) = \min_{T} \max_{\theta} R_T(\theta)$$

Teladan 9. 22.

Melanjutkan permasalahan pada teladan sebelumnya,

$$\max_{\eta} KTG(\hat{\eta}_3) = 1/n^2$$

$$\max_{\eta} KTG(\hat{\eta}_4) = \max_{\eta}[c^2/n^2 + (c-1)^2\eta^2] = \infty$$

penduga tak bias $\hat{\eta}_3 = X_{(1)} - 1/n$ adalah penduga minimax dalam kelas penduga $\hat{\eta}_4 = c(X_{(1)} - 1/n)$.

Definisi 9. 12.

Untuk suatu contoh acak dari $f(x;\theta)$, resiko Bayes dari suatu penduga *T* relatif terhadap suatu fungsi resiko $R_T(\theta)$ dan fungsi kepekatan peluang $p(\theta)$ adalah resiko rata-rata berdasarkan $p(\theta)$,

$$A_T = E_{\theta}\left[R_T(\theta)\right] = \int_{\Omega} R_T(\theta)p(\theta)d\theta$$

Definisi 9. 13.

Untuk suatu contoh acak dari $f(x;\theta)$, penduga Bayes T^* relatif terhadap suatu fungsi resiko $R_T(\theta)$ dan fungsi kepekatan peluang $p(\theta)$ adalah penduga dengan rata-rata resiko paling kecil, $E_{\theta}\left[R_{T^*}(\theta)\right] \leq E_{\theta}\left[R_T(\theta)\right]$ untuk setiap penduga *T.*

Dalam berbagai permasalahan, sangatlah beralasan untuk mengasumsikan bahwa parameter bervariasi untuk berbagai kasus, dan dengan demikian kita perlakukan θ sebagai peubah acak. Dalam kasus lainnya, *p(θ)* dapat merupakan informasi awal atau yang dipercaya sebagai nilai parameter yang sebenarnya. Dalam kasus-kasus itu, pengenalan mengenai fungsi kepekatan peluang *p(θ)*, yang biasanya sering disebut dengan kepekatan awal (*prior density*) untuk parameter θ, berhubungan dengan asumsi tambahan yang mungkin berguna atau tidak tergantung kebenarannya. Dalam sembarang peristiwa, mengambil nilai rata-rata sehubungan dengan fungsi kepekatan peluang *p(θ)* merupakan prosedur yang memberikan jalan untuk membedakan dua penduga bilamana memiliki fungsi resiko yang lebih kecil dan seragam daripada penduga lainnya untul semua nilai θ. Jika suatu cara dapat dikembangkan untuk menentukan penduga Bayes untuk *p(θ)* tertentu, maka suatu kelas penduga dapat diperoleh dengan mempertimbangkan kemungkinan *p(θ)* yang berbeda.

Teladan 9. 23.

Masih dengan menggunakan $\hat{\eta}_3 = X_{(1)} - 1/n$ dan $\hat{\eta}_4 = 0{,}9(X_{(1)} - 1/n)$. Dengan fungsi kerugian kuadrat (*squared error loss*) kita peroleh bahwa $\hat{\eta}_3 = X_{(1)} - 1/n$ lebih baik menurut prinsip minimax, namun $\hat{\eta}_4 = 0{,}9(X_{(1)} - 1/n)$ lebih baik bilamana $\eta^2 < 2 + 1/n^2$, karena memiliki kuadrat tengah galat yang lebih kecil untuk η dalam anak gugus Ω. Kita asumsikan bahwa η ~ *N(0,1)*. $E_\eta[R_{\hat{\eta}_3}(\eta)] = E_\eta(1/n^2) = 1/n^2$ dan $E_\eta[R_{\hat{\eta}_4}(\eta)] = E_\eta[0{,}81/n^2 - 0{,}01\eta^2] = 0{,}81/n^2 + 0{,}01$.

Berdasarkan kriteria ini, $\hat{\eta}_3 = X_{(1)} - 1/n$ lebih baik jika $n \geq 5$, dan $\hat{\eta}_4 = 0{,}9(X_{(1)} - 1/n)$ jika $n \leq 4$.

Contoh Berukuran Besar

Kita telah bahas properti atau sifat-sifat yang dimiliki penduga seperti ketidakbiasan (*unbiasedness*) dan ragam minimum seragam (*uniformly minimum variance*). Keduanya didefinisikan untuk contoh atau sampel berukuran tetap *n*, dan kecil. Perlu dipertimbangkan asimtot atau sifat-sifat bilamana ukuran sampel besar. Suatu penduga bisa saja menjadi tidak diinginkan atau tidak baik, bilamana ukuran sampelnya kecil, tetapi mungkin masih bisa dipertimbangkan untuk dipakai pada beberapa aplikasi jika memiliki sifat asimtot sejalan dengan meningkatnya ukuran sampel.

Definisi 9. 14.

Misalkan $\{T_n\}$ merupakan sekuen atau deretan penduga bagi $\tau(\theta)$. Penduga-penduga ini dikatakan sebagai penduga-penduga yang konsisten bagi $\tau(\theta)$ jika untuk setiap $\varepsilon>0$,

$$\lim P[|T_n - \tau(\theta)| < \varepsilon] = 1$$

untuk setiap $\theta \in \Omega$.

Definisi diatas juga berarti bahwa T_n konvergen dalam peluang ke $\tau(\theta)$, $T_n \xrightarrow{p} \tau(\theta)$ bilamana $n \to \infty$. Hal ini sering disebut dengan **konsisten sederhana**. Yang dimaksud dengan **konsisten** adalah untuk ukuran contoh yang semakin besar, penduga semakin lebih terkonsentrasi disekitar $\tau(\theta)$, dan dengan membuat *n* cukup besar, T_n dapat dibuat atau disesuaikan hingga seberapa tingkatan konsentrasi yang diinginkan.

Definisi 9. 15.

Jika $\{T_n\}$ merupakan sekuen atau deretan penduga bagi $\tau(\theta)$, maka penduga-penduga tersebut dikatakan kuadrat tengah galat konsisten (*mean squared error consistent*) jika

$$\lim_{n\to\infty} E[T_n - \tau(\theta)]^2 = 0$$

untuk setiap $\theta \in \Omega$..

Definisi 9. 16.

Deretan $\{T_n\}$ disebut sebagai penduga tak bias secara asimtotik (*asymptotically unbiased*) bagi $\tau(\theta)$ jika

$$\lim_{n\to\infty} E(T_n) = \tau(\theta)$$

untuk semua $\theta \in \Omega$..

Dapat ditunjukkan bahwa deretan $\{T_n\}$ yang konsisten berdasarkan kuadrat tengah galat juga tak bias secara asimtotik dan konsisten sederhana.

Teorema 9. 7.

Deretan penduga $\{T_n\}$ bagi $\tau(\theta)$ konsisten berdasarkan kuadrat tengah galat jika dan hanya jika deretan tersebut tak bias secara asimtotik dan $\lim_{n\to\infty} Var(T_n) = 0$.

Bukti sebagai latihan.

Teladan 9. 24.

Dalam teladan di awal pembahasan bab ini, kita pertimbangkan $T_n = 1/\overline{X}$ sebagai penduga bagi $\tau(\theta) = 1/\theta$. Kita juga dapat perlihatkan $Y = 2n\overline{X}/\theta \sim \chi^2_{2n}$. Dengan demikian,

$$E(T_n) = [n/(n-1)](1/\theta)$$

dan

$$Var(T_n) = [n/(n-1)]^2 / [(n-2)\theta^2],$$

sehingga apabila $n \to \infty$, maka $E(T_n) \to 1/\theta$ dan $Var(T_n) \to 0$. Dengan demikian, meskipun T_n bias, namun T_n tak bias asimtotik dan konsisten kuadrat tengah galat untuk $\tau(\theta)$.

Sebagaimana telah kita bahas bahwa konsisten berdasarkan kuadrat tengah galat lebih kuat daripada konsisten sederhana.

Teorema 9. 8.

Jika $\{T_n\}$ merupakan sekuen atau deretan penduga bagi $\tau(\theta)$ yang konsisten berdasarkan kuadrat tengah galat, maka deretan tersebut juga konsisten sederhana.
Bukti :sebagai latihan.

Teladan 9. 25.

Misalkan $X_1, \ldots, X_n$ merupakan contoh acak dari sebaran dengan rata-rata μ dan ragam σ^2. Dalam beberapa bab terdahulu, $\overline{X} \xrightarrow{p} \mu$, dan jika momen ke-empat, $\mu_4^{'}$, terhingga, maka $S_n^2 \xrightarrow{p} \sigma^2$. Karena $\overline{X}_n$ dan S_n^2 tak bias dan ragam masing-masingnya mendekati nol, bilamana ukuran contoh semakin besar, sehingga keduanya adalah penduga yang konsisten sederhana dan konsisten kuadrat tengah galat.

Jika sebarannya Eksponensial dengan parameter θ, maka $\overline{X}_n$ konsisten kuadrat tengah galat, namun penduga $\hat{\theta}_n = nX_{(1)}$ bahkan tidak konsisten sederhana, karena $nX_{(1)}$ ~ Eksp(θ).

Teorema 9. 9.

Jika *{Tn}* konsisten sederhana bagi $\tau(\theta)$ and jika *g(t)* kontinu pada setiap nilai $\tau(\theta)$, maka *g(Tn)* konsisten sederhana pada $g(\tau(\theta))$.

Bukti : sebagai latihan.

Definisi 9. 17.

Misalkan $\{T_n\}$ dan $\{T_n^*\}$ merupakan dua deretan penduga yang tak bias asimtotik bagi $\tau(\theta)$. Efisiensi relatif asimtotik T_n terhadap T_n^* diberikan oleh

$$are(T_n, T_n^*) = \lim_{n\to\infty} \frac{Var(T_n^*)}{Var(T_n)}$$

Deretan $\{T_n^*\}$ dikatakan efisien asimtotik (*asymtotically efficient*) jika $are(T_n, T_n^*) \leq 1$ untuk semua deretan tak bias asimtotik $\{T_n\}$, dan semua $\theta \in \Omega$. Efisiensi asimtotik dari suatu deretan tak bias asimtotik $\{T_n\}$ diberikan oleh formula

$$ae(T_n) = are(T_n, T_n^*)$$

jika $\{T_n^*\}$ efisien asimtotik.

Batas bawah Cramer-Rao (CRLB) tidak selalu tercapai untuk *n* tertentu, tetapi kadang dapat tercapai secara asimtotik, dalam kasus ini sangatlah berguna dalam penentuan efisiensi asimtotik.

Teladan 9. 26.

Melanjutkan teladan yang menggunakan contoh acak dari sebaran Eksponensial dengan parameter θ, sekuen $T_n = 1/\overline{X}$ telah ditunjukkan merupakan penduga tak bias asimtotik bagi $1/\theta$.

Ragamnya adalah $Var(T_n) = [n/(n-1)]^2/[(n-2)\theta^2]$ dan CRLB nya adalah $[-1/\theta^2]^2/[n(1/\theta^2)] = 1/[n\theta^2]$. Karena

$$\lim_{n\to\infty}\frac{CRLB}{Var(T_n)} = \lim_{n\to\infty}\frac{1/[n\theta^2]}{[n/(n-1)]^2/[(n-2)\theta^2]} = 1$$

maka T_n efisien asimtotik untuk menduga $1/\theta$.

Teladan 9. 27.

Kembali lagi ke teladan yang menggunakan contoh acak dari populasi dengan $f(x;\eta) = e^{-(x-\eta)}I_{[\eta,\infty)}(x)$. Karena wilayah X_i tergantung pada η, CRLB tak dapat dipakai disini. Penduga $\hat{\eta}_2 = X_{(1)}$ dan $\hat{\eta}_3 = X_{(1)} - 1/n$ keduanya tak bias asimtotik, dan memiliki ragam yang sama, $Var(\hat{\eta}_2) = Var(\hat{\eta}_3) = 1/n^2$. Dengan demikian

$$are(\hat{\eta}_3, \hat{\eta}_2) = \lim_{n\to\infty}\frac{1/n^2}{1/n^2} = 1$$

Kita akan tunjukkan nanti bahwa $\hat{\eta}_3 = X_{(1)} - 1/n$ adalah UMVUE bagi η, dan dengan demikian $\hat{\eta}_2 = X_{(1)}$ juga efisien asimtotik untuk menduga η.

Penduga tak bias yang lain adalah $\hat{\eta}_1 = \overline{X} - 1$ yang memiliki $Var(\hat{\eta}_1) = 1/n$. Dengan demikian,

$$are(\hat{\eta}_1, \hat{\eta}_3) = \lim_{n\to\infty}\frac{1/n^2}{1/n} = 0$$

dengan demikian $\hat{\eta}_1 = \overline{X} - 1$ kurang bagus atau kurang diinginkan daripada $\hat{\eta}_3 = X_{(1)} - 1/n$.

Perlu dicatat bahwa ini merupakan teladan yang tidak seperti biasanya; dalam kebanyakan kasus, ragam suatu penduga memiliki bentuk *c/n*, namun ragam dari $\hat{\eta}_3 = X_{(1)} - 1/n$ ini adalah *1/n²*, dimana lebih kuasa daripada *1/n*. Suatu penduga dengan ragam ordo *1/n²* biasanya disebut sebagai **penduga super efisien**.

Sifat-sifat Asimtotik PKM

Dapat diperlihatkan, dalam beberapa kondisi, bahwa PKM memiliki sifat yang dikehendaki. Khususnya, jika kondisi umum dipenuhi, maka $\hat{\theta}_n$, jawaban persamaan kemungkinan maksimum memiliki sifat-sifat seperti berikut:

1. $\hat{\theta}_n$ ada dan unik (khas),
2. $\hat{\theta}_n$ adalah penduga konsisten bagi θ,
3. $\hat{\theta}_n$ memiliki sebaran normal asimtotik dengan rata-rata θ dan ragam $1/nE\left[\frac{\partial}{\partial\theta}\ln f(X;\theta)\right]^2$, dan
4. $\hat{\theta}_n$ efisien asimtotik.

Sudah tentu bahwa, agar PKM ada sebagai jawaban dari persamaan kemungkinan maksimum, diperlukan persyaratan bahwa turunan parsial dari $\ln f(x;\theta)$ terhadap θ ada, dan juga gugus $A = \{x : f(x;\theta) > 0\}$ tidak tergantung pada θ. Beberapa kondisi tambahan yang menyangkut $\ln f(x;\theta)$ dan $f(x;\theta)$ juga diperlukan, namun kita tak akan membahasnya disini.

Perlu dicatat bahwa efisiensi asimtotik dari $\hat{\theta}_n$ karena adanya fakta bahwa ragam asimtotiknya sama dengan CRLB penduga tak bias bagi θ. Dengan demikian untuk *n* yang cukup besar, kira-kira

$$\hat{\theta}_n \sim N(\theta, CRLB)$$

[$\hat{\theta}_n$ kira-kira memiliki sebaran Normal dengan rata-rata θ dan ragam CRLB]

Juga berdasarkan teorema limit sebaran, jika $\tau(\theta)$ merupakan fungsi dengan turunan yang tidak nol, maka $\hat{\tau}_n = \tau(\hat{\theta}_n)$ juga memiliki sebaran normal asimtotik dengan rata-rata asimtotik $\tau(\theta)$ dan ragam $[\tau(\theta)]^2$CRLB. Juga dapat diperlihatkan bahwa ragam asimtotik dari $\hat{\tau}_n = \tau(\hat{\theta}_n)$ sama dengan CRLB untuk ragam penduga tak bias bagi $\tau = \tau(\theta)$, sehingga $\hat{\tau}_n = \tau(\hat{\theta}_n)$ juga efisien asimtotik.

Teladan 9. 28.

Penduga Kemungkinan Maksimum bagi rata-rata θ sebaran Eksponensial adalah rata-rata contohnya, $\hat{\theta}_n = \bar{X}_n$. Dimungkinkan untuk melakukan inferensia sifat asimtotik yang sama baik dari yang telah kita bahas ataupun Teorema Limit Pusat (*Central Limit Theorem*). Secara khusus, $\hat{\theta}_n$ memiliki sebaran Normal asimtotik dengan rata-rata θ dan ragam θ^2/n. Juga telah diperoleh bahwa CRLB = θ^2/n. Kita juga tahu sebaran pasti dari $\hat{\theta}_n$ ini, karena

$$\frac{2n\hat{\theta}}{\theta} \sim \chi^2_{2n}$$

yang konsisten dengan sebaran normal asimtotik, sehingga diperoleh

$$\sqrt{n}\,\frac{\hat{\theta}_n - \theta}{\theta} \sim N(0,1).$$

Seandainya kita ingin menduga $R = R(t;\theta) = P(X > t) = \exp(-t/\theta)$. Suatu pendekatan bagi ragam $\hat{R} = \exp(-t/\hat{\theta})$ diberikan oleh ragam asimtotiknya

$$Var(\hat{R}) \cong \left[\frac{\partial}{\partial\theta} R(t;\theta)\right]^2 \left(\frac{\theta^2}{n}\right)$$

$$= [\exp(-t/\theta)(t/\theta^2)]^2(\theta^2/n)$$

$$= [\exp(-t/\theta)(t/\theta)]^2 / n$$
$$= [R(\ln R)]^2 / n$$

dan dengan demikian, untuk *n* yang cukup besar, kira-kira

$$\hat{R} \sim N(R,(R \ln(R))^2/n)$$

Teladan 9. 29.

Suatu contoh acak dari sebaran Pareto dengan $f(x;\kappa) = \kappa(1+x)^{-\kappa-1}$ untuk $x > 0$. Sehingga kita dapatkan

$$L(\kappa) = \prod_{i=1}^{n} \kappa(1+x_i)^{-(\kappa+1)} = \kappa^n \prod_{i=1}^{n}(1+x_i)^{-(\kappa+1)}$$

$$\ln L(\kappa) = n\ln\kappa - (\kappa+1)\sum_{i=1}^{n}\ln(1+x_i)$$

dan persamaan kemungkinan maksimumnya adalah

$$\frac{\partial}{\partial\kappa}\ln L(\kappa) = n/\kappa - \sum_{i=1}^{n}\ln(1+x_i) = 0$$

yang menghasilkan penduga kemungkinan maksimum

$$\hat{\kappa} = \frac{n}{\sum_{i=1}^{n}\ln(1+x_i)}$$

Untuk mendapatkan CRLB, maka

$$\ln f(x;\kappa) = \ln\kappa - (\kappa+1)\ln(1+x)$$

$$\frac{\partial}{\partial\kappa}\ln f(x;\kappa) = 1/\kappa - \ln(1+x)$$

dengan demikian,

$$CRLB = \frac{1}{nE\left[\frac{1}{\kappa} - \ln(1+X)\right]^2}$$

Untuk memperoleh hasil dari ekspresi terakhir, akan lebih mudah apabila kita melakukan transformasi $Y = \ln(1+X)$. Dengan metode transformasi, maka $Y = \ln(1+X)$ memiliki distribusi Eksponensial dengan parameter *1/κ.* Sehingga, dapat diperoleh dengan mudah

$$E[(1+X)] = 1/\kappa$$

$$E[1/\kappa - \ln(1+X)]^2 = Var[\ln(1+X)] = 1/\kappa^2$$

$$Var(\hat{\kappa}) \cong CRLB = \kappa^2 / n$$

dan kira-kira

$$\hat{\kappa} \sim N(\kappa, \kappa^2/n)$$

Penduga Bayes dan Minimax

Penduga Bayes adalah penduga yang meminimumkan rata-rata resiko, dimana fungsi resiko, $R_T(\theta)$, dirata-ratakan berdasarkan atau menggunakan fungsi kepekatan peluang prior $p(\theta)$. Prinsip pembandingan minimax adalah memilih penduga yang meminimumkan resiko maksimum.

Keseluruhan kelas penduga dapat dihasilkan dengan menggunakan $p(\theta)$ yang berbeda. Sangat baik untuk memiliki kelas penduga dalam suatu permasalahan, meskipun jika ada beberapa alasan fisik tertentu untuk memilih $p(\theta)$ yang sesuai, maka penduga yang berkenaan dengan $p(\theta)$ akan diasumsikan terlebih dahulu sebagai yang terbaik untuk digunakan dalam permasalahan tersebut. Terdapat filosofi yang berbeda sehubungan dengan pemilihan kepekatan prior $p(\theta)$, namun kita tak akan begitu memperhatikan bagaimana $p(\theta)$ dipilih. Dalam berbagai kasus θ mungkin bertindak seperti peubah acak, dan $p(\theta)$ akan mencerminkan fakta ini. Sebagai teladan, θ dapat merepresentasikan proporsi senjata dalam sebuah tumpukan yang berfungsi. Dilain pihak, $p(\theta)$ merepresentasikan derajat kepercayaan yang berhubungan dengan nilai θ yang datangnya dari informasi penarikan contoh (penyuplikan) sebelumnya, atau dengan cara lainnya. Dalam sembarang peristiwa, penduga potensial yang bermanfaat dapat dikembangkan melalui struktur ini.

Definisi 9. 18.

Kepekatan bersyarat θ bilamana observasi contoh *x* = *(x₁, x₂, …, xₙ)* disebut dengan **kepekatan posterior** atau fkp posterior, dan diberikan oleh

$$f_{\theta|x}(\theta) = \frac{f(x_1,...,x_n \mid \theta)\,p(\theta)}{\int f(x_1,...,x_n \mid \theta)\,p(\theta)\,d\theta}$$

Penduga Bayes adalah penduga yang meminimumkan rata-rata resiko untuk keseluruhan θ, $E_\theta[R_T(\theta)]$. Namun demikian,

$$E_\theta[R_T(\theta)] = E_\theta\{E_{X|\theta}[L(T;\theta)]\} = E_X\{E_{\theta|X}[L(T;\theta)]\}$$

dan penduga *T* yang meminimumkan $E_X\{E_{\theta|X}[L(T;\theta)]\}$ untuk setiap ***x*** juga meminimumkan rata-rata pada ***X***. Dengan demikian, penduga Bayes dapat diperoleh dengan meminimumkan rata-rata kerugian dengan memperhatikan sebaran posterior.

Teorema 9. 10.

Jika $X_1, \ldots, X_n$ melambangkan suatu contoh acak dari $f(x|\theta)$, maka penduga Bayes adalah penduga yang meminimumkan rata-rata kerugian dengan memperhatikan sebaran posterior $\theta|\boldsymbol{x}$,

$$E_{\theta|x}[L(T;\theta)]$$

Untuk beberapa tipe fungsi kerugian, ekspresi penduga Bayes dapat ditentukan lebih eksplisit dalam bentuk sebaran posterior.

Teorema 9. 11.

Penduga Bayes, *T*, bagi $\tau(\theta)$ dengan menggunakan fungsi kerugian galat kuadrat (*squared error loss function*), $L(T;\theta) = [T - \tau(\theta)]^2$ adalah rata-rata bersyarat bagi $\tau(\theta)$ berdasarkan sebaran posteriornya,

$$T = E_{\theta|X}[\tau(\theta)] = \int \tau(\theta) f_{\theta|X}(\theta) d\theta$$

Bukti : sebagai latihan.

Teorema 9. 12.

Penduga Bayes, *T*, bagi θ dengan menggunakan fungsi kerugian harga mutlak (*absolute error loss*), $L(T;\theta) = |T - \theta|$ adalah median dari sebaran posterior $f_{\theta|X}(\theta)$.

Bukti : sebagai latihan.

Teorema 9. 13.

Jika *T** adalah penduga Bayes dengan resiko konstan, $R_{T*}(\theta) = c$, maka *T** adalah penduga minimax.

Bukti:

Kita dapatkan $\max_{\theta} R_{T*}(\theta) = \max_{\theta} c = c = R_{T*}(\theta)$, namun karena $R_{T*}(\theta)$ konstan untuk keseluruhan θ, maka

$$R_{T*}(\theta) = E_{\theta}[R_{T*}(\theta)] \le E_{\theta}[R_T(\theta)]$$

untuk setiap *T* karena *T** adalah penduga Bayes. Rata-rata atau nilai harapan peubah tidak lebih besar dari nilai maksimum peubahnya, sehingga

$$E_{\theta}[R_T(\theta)] \le \max_{\theta} R_T(\theta)$$

dan

$$\max_{\theta} R_{T*}(\theta) \le \max_{\theta} R_T(\theta)$$

yang menunjukkan bahwa *T** adalah penduga minimax.

Teladan 9. 30.

Misalkan X_i ~ *Poi*(θ), dan kita ingin mendapatkan penduga Bayes bagi θ, dengan asumsi menggunakan fungsi kerugian galat kuadrat. Kita pilih fungsi kepekatan prior dari kelas Gamma, θ ~ *Gamma*(β, κ),

$$p(\theta) = \frac{1}{\beta^{\kappa}\Gamma(\kappa)}\theta^{\kappa-1}e^{-\theta/\beta}$$

dimana β dan κ adalah sembarang konstanta yang diketahui. Dengan demikian, sebaran posteriornya adalah

$$f_{\theta|X}(\theta) = \frac{\dfrac{e^{-n\theta}\theta^{\sum x_i}\theta^{\kappa-1}e^{-\theta/\beta}}{\prod(x_i!)\beta^{\kappa}\Gamma(\kappa)}}{\displaystyle\int\frac{e^{-n\theta}\theta^{\sum x_i}\theta^{\kappa-1}e^{-\theta/\beta}}{\prod(x_i!)\beta^{\kappa}\Gamma(\kappa)}d\theta}$$

yang tidak lain adalah

$$\theta \mid x \sim Gam\left[\frac{1}{n+\frac{1}{\beta}}, \sum x_i + \kappa\right]$$

Dengan demikian, penduga Bayes bagi θ adalah

$$T^* = E(\theta \mid X) = \frac{\sum X_i + \kappa}{n + \frac{1}{\beta}}$$

Fungsi kepekatan prior dengan nilai β yang besar dan κ yang kecil, membuat penduga Bayes ini mendekati penduga kemungkinan maksimum, $\hat{\theta} = \bar{x}$.

Besarnya resiko dalam hal ini

$$R_T(\theta) = E[T-\theta]^2 \quad = Var(T) + [E(T)-\theta]^2$$

$$= \frac{nVar(X)}{(n+1/\beta)^2} + \left[\frac{n\theta+\kappa}{n+1/\beta} - \theta\right]^2$$

$$= \frac{n\theta + [\kappa - \theta/\beta]^2}{(n+1/\beta)^2}$$

Tak ada nilai β atau κ yang membuat resiko konstan untuk keseluruhan nilai θ, sehingga penduga minimax, jikalau ada, tak dapat dihasilkan dari turunan ini. Namun demikian, pilihan fkp prior yang berbeda mungkin dapat menghasilkan penduga minimax.

Teladan 9. 31.

Misalkan suatu contoh acak berukuran *n* berasal dari sebaran Bernoulli,

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

dan misalkan θ ~ Seragam(0,1). Misalkan dalam permasalahan ini digunakan fungsi kerugian galat kuadrat 'tertimbang', $L(t;\theta) = \dfrac{(t-\theta)^2}{\theta(1-\theta)}$ yang memberikan bobot lebih nilai θ dekat ke nol atau satu. Tidaklah sulit untuk menunjukkan bahwa fungsi yang memaksimumkan $E_{\theta|x}[L(T;\theta]$ adalah

$$t^*(x) = \frac{E_{\theta|x}[(1-\theta)^{-1}]}{E_{\theta|x}[\theta^{-1}(1-\theta)^{-1}]}$$

Sebaran posterior dalam hal ini adalah Beta,

$$\theta\,|\,x \sim Beta\left(\sum x_i + 1, n - \sum x_i + 1\right)$$

yang berarti bahwa

$$E_{\theta|x}[(1-\theta)^{-1}] = \frac{n+1}{n-\sum x_i}$$

dan

$$E_{\theta|x}[\theta^{-1}(1-\theta)^{-1}] = \frac{n(n+1)}{(\sum x_i)(n-\sum x_i)}$$

sehingga $t^*(x) = \sum x_i / n = \bar{x}$ dan penduga Bayesnya adalah $T^* = \bar{X}$. Lebih jauh lagi

$$R_{T^*}(\theta) = \frac{E[(\bar{X}-\theta)^2}{\theta(1-\theta)} = \frac{\theta(1-\theta)/n}{\theta(1-\theta)} = \frac{1}{n}$$

yang konstan untuk keseluruhan θ. Dengan demikian $T^* = \bar{X}$ dalam teladan ini adalah penduga minimax.

Teladan 9. 32.

Suatu contoh acak dari sebaran eksponensial, Eksp(*1/θ*)

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

Misalkan juga fungsi kepekatan prior θ juga eksponensial, Eks(*1/β*), dimana nilai β diketahui. Maka

$$f_{\theta|x}(\theta) = \frac{\theta^n e^{-\theta \sum x_i} \beta e^{-\theta\beta}}{c(\beta,x)}$$

dimana $c(\beta,x)$ adalah fungsi yang tak tergantung pada θ, dan hanya merupakan konstanta yang membuat integral dari fkp tersebut sama dengan 1. Sudah jelas dalam kasus ini, bahwa

$$\theta\,|\,x \sim Gam[(\sum x_i + \beta)^{-1}, n+1]$$

Untuk fungsi kerugian galat kuadrat, penduga Bayes bagi θ adalah

$$T = E(\theta\,|\,X) = \frac{n+1}{\sum x_i + \beta}$$

Penduga Bayes bagi rata-rata, $\mu = 1/\theta$, dengan fungsi kerugian galat kuadrat adalah

$$\hat{\mu} = E_{\theta|x}(1/\theta) = \frac{\sum x_i + \beta}{n}$$

Penduga Kuadrat Minimum

Dalam beberapa tipe model statistika, prinsip-prinsip kuadrat minimum sangat penting. Kita asumsikan bahwa rata-rata suatu peubah acak *Y* merupakan fungsi linier dari *p* parameter yang tak diketahui $(\beta_1, \ldots, \beta_p)$ dan *p* faktor $\mathbf{x} = (x_1, \ldots, x_p)$ yang dapat dikatakan tetap atau tanpa kesalahan pengukuran,

$$E(Y_x) = \sum_{i=1}^{p} \beta_i x_i$$

Diasumsikan juga bahwa $Var(Y_x) = \sigma^2$, dimana σ^2 bukan merupakan fungsi dari x_i.

Notasi lain yang sering digunakan selain $E(Y_x)$ adalah $\mu_{Y|x}$ dan $E(Y|\mathbf{x})$, namun dalam beberapa kasus ini bukan merupakan harapan bersyarat seperti notasi biasanya, jika faktor-faktor tetap $x_1, \ldots, x_p$ bukan nilai dari gugus peubah acak.

Sebagai teladan, misalkan rata-rata hasil suatu reaksi kimia merupakan fungsi linier dari waktu reaksi x_2 dan temperatur x_3. Sebuah konstanta perlu dimasukkan dengan memisalkan $x_1 = 1$, untuk mendapatkan model linier

$$E(Y_x) = \beta_1 + \beta_2 x_2 + \beta_3 X_3$$

Perlu dicatat bahwa modelnya linier dalam parameter, tetapi tidak perlu linier dalam peubah *x* nya. Kita dapat memisalkannya $x_4 = x_3^2$ atau $x_4 = x_2 x_3$ dan sebagainya.

Parameter-parameter β_1, β_2, dan β_3 tidak diketahui, tetapi jika semua ini diduga, maka untuk suatu waktu reaksi tertentu x_2 dan temperatur tertentu x_3, maka hasil Y_{x_0} atau rata-rata $E(Y_{x_0})$ dapat diduga dengan

$$\widetilde{E}(Y_{x_0}) = \widetilde{\beta}_1 + \widetilde{\beta}_2 x_2 + \widetilde{\beta}_3 x_3 \qquad x_0 = (x_2, x_3)$$

Sudah barang tentu, nilai dugaan parameter harus didapatkan dari data contoh hasil y_i yang telah diamati dari berbagai nilai faktor *x*.

Model Linier Sederhana

Misalkan

$$E(Y_x) = \alpha + \beta x \qquad\qquad Var(Y_x) = \sigma^2$$

Subskrip pada peubah *Y* kadang perlu ditekankan untuk kenyamanan. Kita amati *n* nilai respon peubah *Y* untuk *n* nilai faktor *x*, sehingga diperoleh *n* pasang data $(y_1, x_1), \ldots, (y_n, x_n)$. Kita asumsikan bahwa pengamatan-pengamatan ini sedikitnya tidak saling berkorelasi, sehingga kita memiliki

$$E(Y_i) = \alpha + \beta x_i \qquad Var(Y_i) = \sigma^2 \qquad i = 1,2,\ldots,n$$

$$Cov(Y_i, Y_j) = 0 \qquad i \neq j$$

Asumsi tambahan yang umum adalah bahwa Y_i menyebar normal, $Y_i \sim N(\alpha+\beta x_i, \sigma^2)$, tetapi asumsi sebaran tak diperlukan untuk mendapatkan dugaan titik bagi parameter-parameternya.

Jika nilai amatan, y_i, selalu berada pada nilai tengah atau rata-rata, $E(Y_i)$, maka semua titik-titik (y_i, x_i) akan berada pada sebuah garis lurus, dan persamaan garusnya dapat ditentukan dengan cara aljabar. Karena nilai bervariasi disekitar nilai tengahnya, maka

$$y_i \neq \alpha + \beta x_i$$

namun

$$Y_i = E(Y_i) + \varepsilon_i = \alpha + \beta x_i + \varepsilon_i$$

dimana $E(\varepsilon_i) = 0$ dan

$$y_i = \alpha + \beta x_i + e_i$$

Situasi ideal tentunya apabila semua pasangan (y_i, x_i) berada dalam satu garis lurus, $y = \tilde{\alpha} + \tilde{\beta} x$, dengan semua $e_i = 0$. Kemudian kita dapat memprediksi *y* tanpa kesalahan. Tahapan terbaik berikutnya adalah membuat cocok sebuah garis lurus yang melalui semua titik (y_i, x_i) sedemikian rupa sehingga meminimumkan penyimpangan nilai pengamatan y_i dari garis yang telah dibuat. Yang dimaksudkan adalah, kita memilih sebuah garis $y = \tilde{\alpha} + \tilde{\beta} x$ yang meminimumkan beberapa fungsi $\tilde{e}_i = y_i - (\tilde{\alpha} + \tilde{\beta}x_i)$. Kriteria yang berbeda untuk kesesuaian model mengharuskan penggunaan fungsi untuk $\tilde{e}_i$ yang berbeda pula, tetapi kita akan menggunakan prinsip meminimumkan kuadrat penyimpangan dari garis yang dibuat. Dengan demikian, kita ingin mencari nilai-nilai α dan β, sebut saja $\tilde{\alpha}$ dan $\tilde{\beta}$ sebagai hasil atau nilai minimum dari

$$\sum \tilde{e}_i^2 = \sum \left[y_i - (\tilde{\alpha} + \tilde{\beta}x_i) \right]^2$$

Dengan demikian, kita ingin mencari nilai $\tilde{\alpha}$ dan $\tilde{\beta}$ yang meminimumkan

$$Q = \sum \left[y_i - (\alpha + \beta x_i) \right]^2$$

Dengan menurunkan Q terhadap α dan β dan menyamakan hasilnya dengan nol, akan menghasilkan nilai dugaan $\tilde{\alpha}$ dan $\tilde{\beta}$, yaitu

$$2\sum_{i=1}^{n} \left[y_i - (\tilde{\alpha} + \tilde{\beta}x_i) \right](-1) = 0$$

$$2\sum_{i=1}^{n} \left[y_i - (\tilde{\alpha} + \tilde{\beta}x_i) \right](-x_i) = 0$$

Kedua persamaan diatas secara bersama-sama akan menghasilkan

$$\tilde{\beta} = \frac{\sum x_i y_i - (\sum x_i)(\sum y_i)/n}{\sum x_i^2 - (\sum x_i)^2/n} = \frac{\sum (x_i - \bar{x})(y_i - \bar{y})}{\sum (x_i - \bar{x})^2}$$

$$\tilde{\alpha} = \bar{y} - \tilde{\beta}\bar{x}$$

Dengan demikian, jika kita ingin membuat garis lurus melalui sekumpulan titik-titik, persamaan garis $y = \tilde{\alpha} + \tilde{\beta}x$ merupakan suatu garis yang memenuhi kriteria meminimumkan kuadrat tengah galat antara nilai amatan dan titik pada garis. Jumlah kuadrat galat dapat diekspresikan, sekali lagi, dengan

$$JK(Galat) = \sum \tilde{e}_i^2 = \sum \left[y_i - (\tilde{\alpha} + \tilde{\beta}x_i) \right]^2 = \sum \left(y_i - \tilde{y}_i \right)^2$$

Prinsip kuadrat minimum ini tidak secara langsung dapat digunakan untuk memperoleh dugaan σ^2, tetapi besarnya ragam tercermin dalam kuantitas JK(Galat), karena ragam merupakan alasan bahwa semua nilai amatan *Y* tidak berada pada garis rata-rata yang sebenarnya. Akan ditunjukkan nanti bahwa dugaan bagi σ^2 akan diberikan oleh

$$\tilde{\sigma}^2 = \frac{JK(Galat)}{n-2}$$

Juga

$$\tilde{y}_0 = \tilde{\alpha} + \tilde{\beta} x_0$$

dapat digunakan untuk memprediksi nilai *Y* pada $x = x_0$, dan kuantitas yang sama dapat digunakan untuk menduga nilai rata-rata *Y* pada $x = x_0$. Yaitu, nilai dugaan dari $E(Y_{x_0}) = \alpha + \beta x_0$ yang diberikan oleh

$$\tilde{E}(Y_{x_0}) = \tilde{\alpha} + \tilde{\beta} x_0$$

Kombinasi linier lain dari α dan β dapat diduga dengan cara yang serupa, dan penduga-penduga ini tidak bias. Penduga-penduga kuadrat minimum merupakan fungsi linier dari Y_i, dan dapat ditunjukkan bahwa diantara semua penduga tak bias linier, penduga-penduga kuadrat minimum memiliki ragam minimum. Dengan demikian, penduga-penduga kuadrat minimum sering disebut dengan **penduga-penduga tak bias linier terbaik** atau *best linear unbiased estimators* (BLUE).

Berikut ini diberikan teorema yang berkaitan dengan penduga kuadrat terkecil. Pembuktian digunakan sebagai latihan.

Teorema 9. 14.

Jika $E(Y_i) = \alpha + \beta x_i$, $Var(Y_i) = \sigma^2$ dan $Cov(Y_i, Y_j) = 0$ untuk $i \neq j$ dan $i = 1, \ldots, n$, maka penduga-penduga kuadrat minimum memiliki sifat-sifat seperti berikut:

1. $E(\widetilde{\beta}) = \beta, \quad Var(\widetilde{\beta}) = \sigma^2 / \sum (x_i - \overline{x})^2$
2. $E(\widetilde{\alpha}) = \alpha, \quad Var(\widetilde{\alpha}) = \sigma^2 \sum x_i^2 /[n \sum (x_i - \overline{x})^2]$
3. $E(c_1\widetilde{\alpha} + c_2\widetilde{\beta}) = c_1\alpha + c_2\beta$
4. $c_1\widetilde{\alpha} + c_2\widetilde{\beta}$ adalah BLUE untuk $c_1\alpha + c_2\beta$

Hasil dari teorema diatas dapat dikembangkan atau diperluas untuk multifaktor, termasuk model polinomial.

Model Linier Umum

Tidaklah mungkin untuk mengembangkan model multifaktor dengan mudah tanpa harus mengenalkan notasi matriks. Berbagai hasil sederhana dapat dimulai dengan notasi matriks berikut.

Misalkan model linier dengan

$$E(Y_x) = \sum_{j=1}^{p} \beta_j x_j \qquad Var(Y_x) = \sigma^2$$

Suatu respon y_i diamati pada nilai $x_{i1}, \ldots, x_{ip}$, $i = 1,\ldots,n$. Dengan demikian, asumsikan bahwa

$$E(Y_i) = \sum_{j=1}^{p} \beta_j x_{ij} \qquad Var(Y_i) = \sigma^2 \qquad Cov(Y_i, Y_j) = 0$$

Jika kita misalkan matriks varian-kovarian adalah ***V*** maka

$$\boldsymbol{V} = \{\sigma_{ij}\} = \{Cov(Y_i, Y_j)\} = \sigma^2 I$$

juga

$$E(\boldsymbol{Y}) = \boldsymbol{X\beta}$$

dimana

$$\boldsymbol{Y} = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} \quad \boldsymbol{\beta} = \begin{pmatrix} \beta_1 \\ \vdots \\ \beta_n \end{pmatrix} \quad \boldsymbol{X} = \begin{pmatrix} x_{11} & \cdots & x_{1p} \\ \vdots & \ddots & \vdots \\ x_{n1} & \cdots & x_{np} \end{pmatrix}$$

Nilai dugaan $\boldsymbol{\beta}$ adalah nilai-nilai β_i yang meminimumkan

$$\sum_{i=1}^{n}(y_i - E(Y_i))^2 = (Y\text{-}X\beta)'(Y\text{-}X\beta)$$

Dengan mencari turunannya terhadap β_i dan membuat nilainya sama dengan nol, kita dapat menunjukkan dugaan kuadrat minimum diberikan oleh persamaan

$$\tilde{\beta} = (X'X)^{-1}X'Y$$

Nilai dugaan ini merupakan fungsi linier dari pengamatan dan tidak bias, karena

$$E(\tilde{\beta}) = (X'X)^{-1}X'Y = (X'X)^{-1}(X'X)\beta = \beta$$

Matriks ragam peragam bagi $\tilde{\beta}_i$ adalah

$$C = \{Cov(\tilde{\beta}_i, \tilde{\beta}_j)\} = \sigma^2(X'X)^{-1}$$

Dapat ditunjukkan bahwa BLUE dari kombinasi linier $\tilde{\beta}_i$ sebut saja $r'\beta = \sum r_i\beta_i$ adalah $r'\tilde{\beta}$. Teorema dibawah ini dikenal dengan nama Teorema Gauss-Markov.

Teorema 9. 15.

Jika $E(Y) = X\beta$ dan $V = \{\mathrm{Cov}(Y_i, Y_j)\} = \sigma^2 I$, maka $r'\tilde{\beta}$ adalah penduga tak bias linier terbaik (BLUE) bagi $r'\beta$, dimana

$$\tilde{\beta} = (X'X)^{-1}X'Y$$

Penduga Kuadrat Tengah Invarian Minimum

Kita telah lihat bahwa tak mungkin untuk mendapatkan penduga yang meminimumkan MSE (*Mean Square Error*) atau Kuadrat Tengah Galat dalam kelas semua penduga, meskipun MSE mungkin berguna dalam pembandingan dua penduga. Salah satu cara pendekatannya adalah membatasi pembahasan pada penduga-penduga tak bias, dan penduga dengan ragam minimum mungkin ada dalam kelas tereduksi ini. Prinsip kuadrat terkecil memberikan penduga-penduga dengan ragam minimum dalam kelas penduga tak bias linier dalam aturan tertentu.

Suatu penduga $T(x_1, \ldots, x_n)$ merupakan penduga skala invarian jika memiliki sifat bahwa $T(kx_1, \ldots, kx_n) = k\, T(x_1, \ldots, x_n)$ untuk semua nilai $k > 0$.

Jika suatu permasalahan memiliki struktur skala invarian, maka sangatlah beralasan untuk membatasi perhatian kita pada kelas penduga skala invarian.

Teorema 9. 16.

Misalkan $x_1, \ldots, x_n$ merupakan suatu contoh acak berukuran n dari $f(x;\theta)$, $0 < x < \infty$, dimana $\theta > 0$ adalah parameter skala. Dengan menggunakan fungsi kerugian $L(\hat{\theta},\theta) = (\hat{\theta}-\theta)^2/\theta^2$, penduga dengan resiko minimum seragam dalam kelas penduga skala invarian adalah

$$\hat{\theta} = \frac{\int_0^\infty (1/\theta^2)\prod f(x_i;\theta)d\theta}{\int_0^\infty (1/\theta^3)\prod f(x_i;\theta)d\theta}$$

Hasil serupa dapat digunakan untuk model parameter lokasi. Perhatikan suatu model dengan parameter lokasi μ dalam bentuk

$$f(x;\mu) = g(x-\mu)$$

dan fungsi kerugian lokasi invarian

$$L(\hat{\mu},\mu) = L(\hat{\mu}-c, \mu-c)$$

Selanjutnya, sebuah penduga dikatakan lokasi invarian jika model hasil transformasi $Y = X - c$ dan $\mu^* = \mu - c$,

$$L(\hat{\mu},\mu) = L(\hat{\mu}^*, \mu^*)$$

Sebuah penduga dikatakan lokasi invarian, jika *T(x_1-c, …, x_n-c)* = *T(x_1, …, x_n)* – *c* untuk semua nilai *c.*

Teorema 9. 17.

Misalkan $x_1, \ldots, x_n$ merupakan suatu contoh acak berukuran *n* dari $f(x;\mu)$, $-\infty < \mu < \infty$, dimana θ adalah parameter lokasi. Dengan menggunakan fungsi kerugian kuadrat galat $L(\hat{\mu},\mu) = (\hat{\mu}-\mu)^2$, penduga dengan resiko kuadrat tengah galat minimum seragam dalam kelas penduga lokasi invarian adalah

$$\hat{\mu} = \frac{\int_{-\infty}^{\infty} \mu \prod f(x_i;\mu)\,d\mu}{\int_{-\infty}^{\infty} \prod f(x_i;\mu)\,d\mu}$$

Latihan

1. Dengan menggunakan metode momen, carilah penduga bagi θ berdasarkan contoh acak $X_1, \ldots, X_n$ dari setiap fungsi kepekatan peluang berikut:
   a. $f(x;\theta) = \theta x^{\theta-1} \quad ;0 < x < 1,\ 0 < \theta$
   b. $f(x;\theta) = \theta^2 x e^{-\theta x} \quad ;0 < x,\ 0 < \theta$
2. Carilah penduga momen berdasarkan contoh acak berukuran *n* dari tiap sebaran berikut:
   a. $X_i \sim f(x;\theta,\eta) = \frac{1}{2\theta} e^{-|x-\eta|/\theta} \quad ;-\infty < x < +\infty,\ 0 < \theta$
   b. $X_i \sim f(x;\theta,\eta) = \frac{1}{\theta}\exp\left\{\frac{x-\eta}{\theta} - \exp\left[\frac{x-\eta}{\theta}\right]\right\} \quad ;-\infty < x < +\infty,\ 0 < \theta$
3. Carilah penduga kemungkinan maksimum berdasarkan suatu contoh acak $X_1, \ldots, X_n$ dari setiap sebaran berikut:
   a. $X_i \sim f(x;p) = pq^{x-1} \quad ;x = 1,2,\ldots,0 < p < 1,\ q = 1-p$
   b. $X_i \sim f(x;p) = C_3^{x-1} p^4 q^{x-4} \quad ;x = 4,5,\ldots\ 0 < p < 1,\ q = 1-p$

c. $X_i \sim f(x;\theta) = \frac{1}{\sqrt{2\pi\theta}} e^{-\frac{x^2}{2\theta}} \quad ;0 < \theta$

d. $X_i \sim f(x;\theta) = (1+x)^{-(\theta+1)} \quad ;0 < \theta$

e. $X_i \sim f(x;\theta_1,\theta_2) = 1/(\theta_2 - \theta_1) \quad ;\theta_1 \le x \le \theta_2$

f. $X_i \sim f(x;\theta,\eta) = \theta\eta^\theta x^{-(\theta+1)} \quad ;\eta \le x, 0 < \theta$

4. Misalkan $X_1, \ldots, X_n$ adalah contoh acak dari sebaran Geometrik dengan parameter p. Carilah penduga kemungkinan maksimum untuk kuantitas-kuantitas berikut
   a. $E(X) = 1/p$.
   b. $Var(X) = (1-p)/p^2$
   c. $P[X<k] = 1-(1-p)^k$ untuk sembarang $k = 1, 2, \ldots$
5. Berdasarkan contoh acak berukuran n dari sebaran Normal, $X \sim N(\mu,\sigma2)$, carilah penduga kemungkinan maksimum untuk
   a. $P[X < c]$ untuk sembarang nilai c.
   b. Persentil ke-95 dari X
6. Misalkan $x_{(1)}$ dan $x_{(n)}$ adalah nilai amatan terkecil dan terbesar dari suatu contoh acak berukuran n dari sebaran dengan fungsi kepekatan peluang $f(x;\theta); 0 < \theta$.
   a. Jika $f(x;\theta) = 1$ untuk $\theta - 0,5 \le x \le \theta + 0,5$ dan nol selainnya, tunjukkanlah bahwa sembarang nilai $\hat{\theta}$ sedemikian rupa sehingga $x_{(n)} - 0,5 \le \hat{\theta} \le x_{(1)} + 0,5$ adalah penduga kemungkinan maksimum bagi θ.
   b. Jika $f(x;\theta) = 1/\theta$ untuk $\theta \le x \le 2\theta$ dan nol selainnya, tunjukkanlah bahwa $\hat{\theta} = 0,5x_{(n)}$ adalah penduga kemungkinan maksimum bagi θ, jika $x_{(n)} < 2x_{(1)}$.
7. Misalkan $X \sim Bin(n,p)$ dan $\hat{p} = X/n$
   a. Carilah konstanta c sedemikian rupa sehingga $E[c\hat{p}(1-\hat{p})] = p(1-p)$
   b. Carilah penduga tak bias bagi $Var(X)$
   c. Apabila sebuah contoh acak acak berukuran N berasal dari populasi yang menyebar Bin(n,p). Carilah penduga tak bias bagi p dan $Var(X)$ berdasarkan contoh acak tersebut.

8. Sebuah contoh acak berukuran *n* berasal dari sebaran Seraagam, X_i ~ Ser($-\theta,\theta$) dimana $\theta > 0$. Carilah konstanta *c* sedemikian rupa sehingga $c(X_{(n)}-X_{(1)})$ merupakan penduga tak bias bagi θ.
9. Jika suatu contoh acak berukuran *n* berasal dari sebaran Bernoulli, $X_i \sim Bin(1,p)$.
   a. Carilah CRLB bagi ragam penduga tak bias untuk *p.*
   b. Carilah CRLB bagi ragam penduga tak bias untuk $p(1-p)$.
   c. Carilah UMVUE bagi *p.*
10. Andaikan sebuah contoh berukuran *n* berasal dari sebaran Normal, $X_i \sim N(\mu,9)$.
    a. Carilah CRLB ragam penduga tak bias untuk μ.
    b. Apakah penduga kemungkinan maksimum $\hat{\mu} = \overline{X}$, UMVUE bagi μ ?
    c. Apakah PKM dari persentil ke-99 juga UMVUE ?
11. Andaikan sebuah contoh acak berukuran *n* berasal dari sebaran dengan fkp $f(x;\theta) = 1/\theta$ jika $0<x\leq\theta$ dimana $0 < \theta$.
    a. Carilah penduga kemungkinan maksimumnya, $\hat{\theta}$.
    b. Carilah penduga momennya, $\tilde{\theta}$.
    c. Apakah $\hat{\theta}$ tak bias ?
    d. Apakah $\tilde{\theta}$ tak bias ?
    e. Bandingkan kuadrat tengah galat dari $\hat{\theta}$ dan $\tilde{\theta}$.
    f. Tunjukkan bahwa $\hat{\theta}$ adalah konsisten kuadrat tengah galat.
    g. Tunjukkan bahwa $\tilde{\theta}$ adalah konsisten kuadrat tengah galat.
12. Misalkan sebuah contoh acak berukuran $n = 2$ dari sebaran normal, $X_i \sim N(\theta,1)$ dimana $\Omega = \{\theta \mid 0 \leq \theta \leq 1\}$. Definisikan penduga-penduga seperti berikut : $\hat{\theta}_1 = (1/2)X_1 + (1/2)X_2$; $\hat{\theta}_2 = (1/4)X_1 + (3/4)X_2$; $\hat{\theta}_3 = (2/3)X_1$; dan $\hat{\theta}_4 = (2/3)\hat{\theta}_1$. Fungsi kerugian yang digunakan adalah $L(t;\theta) = (t-\theta)^2$.

a. Bandingkan seluruh fungsi resiko untuk semua penduga.
b. Bandingkan semua penduga dengan prinsip minimax.
c. Dapatkan resiko Bayes bagi penduga jika θ ~ *Ser(0,1)*
d. Dapatkan resiko Bayes bagi penduga jika θ ~ *Beta(2,1)*

13. Misalkan *X* ~ *Poi*(μ) dan misalkan fungsi kerugian yang dipakai $L(\hat{\mu};\mu) = (\hat{\mu} - \mu)^2 / \mu$. Diasumsikan juga misalnya μ ~ *Gam*(θ,κ), dimana θ dan κ diketahui.
    a. Carilah penduga Bayes bagi μ.
    b. Tunjukkan bahwa $\hat{\mu} = X$ adalah penduga minimax.
14. Suatu contoh acak berukuran *n* dari sebaran Poisson, X_i ~ *Poi*(μ).
    a. Dapatkan CRLB ragam penduga tak bias bagi μ.
    b. Dapatkan CRLB ragam penduga tak bias bagi $\theta = e^{-\mu}$.
    c. Carilah UMVUE bagi μ.
    d. Carilah PKM $\hat{\theta}$ bagi θ.
    e. Apakah $\hat{\theta}$ penduga tak bias bagi θ?
    f. Apakah $\hat{\theta}$ tak bias asimtotik ?
    g. Tunjukkan bahwa $\tilde{\theta} = \left[\dfrac{n-1}{n}\right]^{\sum X_i}$ merupakan penduga tak bias bagi θ.
    h. Carilah $Var(\tilde{\theta})$ dan bandingkan dengan CRLB pada bagian (b).
15. Suatu contoh acak berukuran *n* dari sebaran dengan fungsi kepekatan peluang $f(x;p) = p(1-p)^x \quad ; x = 0,1,...$
    a. Carilah PKM bagi *p*.
    b. Carilah PKM bagi $\theta = (1-p)/p$
    c. Dapatkan CRLB ragam penduga tak bias bagi θ.
    d. Apakah PKM bagi θ merupakan penduga tak bias dengan ragam minimum seragam (UMVUE) ?
    e. Apakah PKM bagi θ juga konsisten kuadrat tengah galat (MSE consistent) ?

f. Carilah sebaran asimtotik dari PKM bagi θ.

g. Misalkan $\tilde{\theta} = \frac{n}{n+1}\bar{X}$. Carilah fungsi resiko untuk $\tilde{\theta}$ dan $\bar{X}$ dengan menggunakan fungsi kerugian $L(t;\theta) = (t-\theta)^2/(\theta^2+\theta)$.

16. Misalkan $\hat{\theta}_i, i = 1,...,n$ merupakan penduga-penduga tak bias yang saling bebas bagi θ dengan $Var(\hat{\theta}_i) = \sigma_i^2$. Seandainya, ada penduga gabungan $\hat{\theta}_g = \sum_{i=1}^{n} a_i\hat{\theta}_i$ dimana $\sum_{i=1}^{n} a_i = 1$.

    a. Tunjukkan bahwa $\hat{\theta}_g$ adalah penduga tak bias bagi θ.

    b. Tunjukkan bahwa $Var(\hat{\theta}_g)$ akan minimum jika $a_i = (1/\sigma_i^2)/\sum_{i=1}^{n}(1/\sigma_i^2)$

17. Misalkan $Y_1, \ldots, Y_n$ saling bebas, dimana $Y_i \sim N(\beta x_i, \sigma^2)$.

    a. Jika $y_1, \ldots, y_n$ adalah nilai amatan, carilah dugaan kuadrat terkecil $\tilde{\beta}$ berdasarkan (x_i, y_i).

    b. Tunjukkan bahwa $\tilde{\sigma}^2 = \sum_{i=1}^{n}(Y_i - \tilde{\beta}x_i)^2/(n-1)$ merupakan penduga tak bias bagi σ^2.

18. Asumsikan bahwa seluruh peubah acak saling bebas $Y_i \sim N(\alpha+\beta x_i, \sigma^2)$, $i=1,\ldots,n$. Andaikan $\tilde{\alpha}$, $\tilde{\beta}$, dan $\tilde{\sigma}^2$ adalah penduga-penduga kuadrat terkecil (*LS estimators*)

    a. Carilah fungsi kepekatan peluang bagi $\tilde{\alpha}$.

    b. Carilah fungsi kepekatan peluang bagi $\tilde{\beta}$.

19. Suatu contoh acak dari sebaran Eksponensial, $Y_i \sim$ *Eksponensial*(βx_i); $i = 1, \ldots, n$

    a. Carilah penduga kuadrat terkecil bagi β.

    b. Carilah penduga kemungkinan maksimum bagi β.

20. Suatu contoh acak berukuran n dari sebaran Eksponensial, X_i ~ *Eksponensial*(θ).
    a. Carilah penduga Pitman bagi θ.
    b. Tunjukkanlah bahwa kuadrat tengah galat penduga Pitman ini lebih kecil dan seragam dibandingkan kuadrat tengah galat dari UMVUE, $\overline{X}$.

Statistik Cukup dan Lengkap

Pendahuluan

Dalam bab sebelum ini telah disajikan teori-teori untuk menurunkan penduga-penduga titik, dengan menggunakan contoh acak, untuk menduga parameter sebaran populasi yang tidak diketahui. Dalam beberapa kasus dimungkinkan untuk membuat statistik atau gugus statistik yang mencakup seluruh "informasi" dalam contoh yang berkenaan dengan parameter-parameter populasinya. Sehingga, sangatlah beralasan untuk membatasi pembahasan kita pada statistik-statistik dimaksud bilamana melakukan pendugaan atau melakukan inferensi tentang parameter-parameter tersebut.

Secara umum, ide **cukup** mencakup reduksi suatu gugus data menjadi suatu gugus statistik yang lebih ringkas tanpa hilangnya informasi tentang parameter yang tak diketahuinya. Kasarnya, suatu statistik *S* dikatakan sebagai suatu statistik "cukup" untuk suatu parameter θ jika sebaran bersyarat dari sembarang statistik *T* apabila nilai dari *S* diberikan tidak mencakup θ. Dengan perkataan lain bahwa jika nilai suatu statistik cukup diketahui, maka nilai amatan dari sembarang statistik yang lainnya tidak akan memiliki informasi yang lebih tentang parameter tersebut.

Teladan 10. 1.

Dalam percobaan pelemparan koin sebanyak *n* kali dan munculnya keluaran dicatat untuk setiap lemparan. Secara statistika hal ini dapat dimodelkan dengan suatu contoh acak $X_1, \ldots, X_n$ dari sebaran Bernoulli. Misalkan bahwa koin tersebut tidak seimbang, dan kita ingin menduga besarnya θ = *P(*Gambar*)*. Banyaknya gambar muncul, atau $S = \sum_i X_i$ harus memberikan informasi sebanyak mungkin tentang nilai θ dari seluruh keluaran.

$$f(x_1,\ldots,x_n;\theta) = \theta^{\sum x_i}(1-\theta)^{n-\sum x_i} \qquad x_i = 0, 1$$

Kita juga dapat tunjukkan bahwa $S = \sum_i X_i$ ~ *Bin(n, θ)*, sehingga

$$f_S(s;\theta) = C_s^n \theta^s (1-\theta)^{n-s} \qquad s = 0,1,\ldots,n$$

Jika $\sum_i x_i = s$, maka peristiwa-peristiwa $[X_1 = x_1,\ldots,X_n = x_n, S = s]$ dan $[X_1 = x_1,\ldots,X_n = x_n]$ ekuivalen, dan

$$\begin{aligned} f_{X_1,\ldots,X_n|s}(x_1,\ldots,x_n) &= \frac{P[X_1 = x_1,\ldots,X_n = x_n, S = s]}{P[S = s]} \\ &= \frac{f(x_1,\ldots,x_n;\theta)}{f_S(s;\theta)} \\ &= \frac{\theta^{\sum x_i}(1-\theta)^{n-\sum x_i}}{C_s^n \theta^s (1-\theta)^{n-s}} \\ &= \frac{1}{C_s^n} \end{aligned}$$

Jika $\sum_i x_i \neq s$, maka fungsi kepekatan peluang bersyarat diatas sama dengan nol. Keduanya, tidak mengandung θ.

Statistik Cukup

Seperti halnya pada bab-bab terdahulu, suatu gugus data $x_1, \ldots, x_n$ akan dimodelkan secara matematik sebagai nilai amatan gugus peubah acak $X_1, \ldots, X_n$. Untuk kenyamanan, akan digunakan notasi vektor $\boldsymbol{X} = (X_1, \ldots, X_n)$ dan $\boldsymbol{x} = (x_1, \ldots, x_n)$ sebagai notasi untuk peubah acak yang diamati dan nilai-nilai amatannya. Dimungkinkan juga penggunaan vektor parameter $\boldsymbol{\theta}$ dan vektor statistik seperti $\boldsymbol{S}$ dan $\boldsymbol{T}$.

Definisi 10. 1. Statistik Cukup Bersama.

Misalkan $\boldsymbol{X} = (X_1, \ldots, X_n)$ memiliki fungsi kepekatan peluang bersama $f(\mathbf{x};\boldsymbol{\theta})$, dan misalkan $\boldsymbol{S} = (S_1, \ldots, S_k)$ adalah statistik berdimensi-*k*. Selanjutnya $S_1, \ldots, S_k$ adalah **statistik cukup bersama** untuk $\boldsymbol{\theta}$ jika untuk sembarang vektor statistik $\boldsymbol{T}$, fungsi kepekatan peluang bersama $\boldsymbol{T}$ apabila $\boldsymbol{S} = \boldsymbol{s}$, yang dinotasikan dengan $f_{\boldsymbol{T}|\boldsymbol{s}}(\boldsymbol{t})$, tidak tergantung pada $\boldsymbol{\theta}$. Dalam kasus satu dimensi, kita hanya sebut bahwa *S* adalah **statistik cukup** bagi θ.

Ide bahwa jika $\boldsymbol{S}$ teramati, maka informasi tambahan tentang $\boldsymbol{\theta}$ tak dapat diperoleh dari $\boldsymbol{T}$ jika sebaran bersyarat $\boldsymbol{T}$ jika $\boldsymbol{S} = \boldsymbol{s}$ bebas dari $\boldsymbol{\theta}$. Biasanya kita asumsikan bahwa $X_1, \ldots, X_n$ adalah contoh acak dari suatu fungsi kepekatan peluang populasi $f(\mathbf{x};\boldsymbol{\theta})$, dan untuk kenyamanan, sekalilagi, kita akan gunakan $\boldsymbol{X} = (X_1, \ldots, X_n)$ sebagai contoh acak. Namun demikian, secara umum $\boldsymbol{X}$ dapat merepresentasikan beberapa vektor peubah acak yang diamati lainnya, seperti contoh tersensor, atau barangkali beberapa gugus statistik tataan lainnya.

Tujuan utamanya adalah mengurangi contoh hingga mencapai gugus statistik cukup terkecil, yang sering disebut dengan "gugus minimal" statistik cukup. Jika ada sebanyak *k* parameter yang diketahui di dalam model, seringkali akan ada sebanyak *k* statistik cukup. Dalam beberapa kasus banyaknya statistik cukup akan melebihi banyaknya parameter, dan sudah tentu tak akan ada reduksi banyaknya statistik. Keseluruhan contoh sendiri adalah gugus statistik contoh, tetapi bila kita bicara statistik cukup, biasanya kita akan memikirkan gugus statistik cukup yang lebih kecil.

Definisi 10. 2.

Suatu gugus statistik dikatakan gugus **cukup minimal** jika anggota gugusnya adalah cukup bersama bagi parameter-parameternya dan jika merupakan fungsi dari setiap gugus statistik cukup bersama lainnya.

Statistik tataan akan diperlihatkan sebagai cukup bersama. Hal ini merepresentasikan reduksi contoh, meskipun banyaknya statistik dalam hal ini tidak tereduksi. Dalam beberapa kasus, statistik tataan dapat merupakan gugus cukup minimal, tetapi tentunya kita berharap dapat mereduksi contoh hingga menjadi beberapa statistik cukup bersama saja.

Terlihat jelas bahwa, sesungguhnya kita tak dapat menggunakan seluruh statistik yang mungkin, ***T***, sebagaimana tersebut dalam definisi statistik cukup bersama untuk memverifikasi bahwa ***S*** adalah statistik cukup. Namun demikian, karena ***T*** dapat dituliskan sebagai fungsi suatu contoh acak $\boldsymbol{X} = (X_1, \ldots, X_n)$ salah satu pendekatan yang dapat dilakukan adalah menunjukkan bahwa $f_{\boldsymbol{X}|\boldsymbol{S}}(\boldsymbol{x})$ bebas dari $\boldsymbol{\theta}$.

Teladan 10. 2.

Suatu contoh acak berasal dari sebaran Eksponensial, $X_i \sim Eks(\theta)$. Dengan demikian kita peroleh

$$f(x_1,\ldots,x_n;\theta) = \frac{1}{\theta^n} e^{-\frac{\sum x_i}{\theta}} \qquad x_i > 0$$

yang mengisyaratkan untuk melakukan *checking* statistik $S = \sum X_i$. Kita juga dapat tunjukkan bahwa $S \sim Gam(\theta,n)$, sehingga

$$f_S(s;\theta) = \frac{1}{\theta^n \Gamma(n)} s^{n-1} e^{-\frac{s}{\theta}} \qquad s > 0$$

Jika $\sum x_i = s$, maka

$$\frac{f(x_1,\ldots,x_n;\theta)}{f_S(s;\theta)} = \frac{\Gamma(n)}{s^{n-1}}$$

yang bebas dari θ, dan dengan demikian berdasarkan definisi, maka

S adalah statistik cukup bagi θ.

Suatu kriteria yang sedikit lebih sederhana juga dapat diturunkan. Secara khusus, jika *S₁, …, Sₖ* adalah statistik cukup bersama untuk $\boldsymbol{\theta}$, maka

$$f(x_1,...,x_n;\theta) = f_S(s;\theta) f_{X|s}(x_1,...,x_n)$$
$$= g(s;\theta)h(x_1,...,x_n)$$

Dengan demikian, fungsi kepekatan peluang bersama contoh dapat difaktorkan menjadi suatu fungsi dari ***s*** dan $\boldsymbol{\theta}$ dikalikan dengan suatu fungsi dari ***x*** = $(x_1, \ldots, x_n)$ yang tidak tergantung dari $\boldsymbol{\theta}$. Sebaliknya, misalkan bahwa $f(x_1,...,x_n;\theta) = g(s;\theta)h(x_1,...,x_n)$ dimana diasumsikan bahwa untuk ***s*** yang tetap, $h(x_1,...,x_n)$ tidak tergantung pada $\boldsymbol{\theta}$. Jika $f(x_1,...,x_n;\theta) = g(s;\theta)h(x_1,...,x_n)$ berlaku untuk beberapa fungsi *g* dan *h*, fungsi kepekatan peluang marjinal ***S*** harus dalam bentuk

$$f_S(s;\theta) = g(s;\theta)c(s)$$

karena untuk ***s*** yang tetap, pengintegralan atau penjumlahan untuk seluruh peubah tersisa tak dapat membawa $\boldsymbol{\theta}$ ke dalam fungsi. Sehingga

$$f(x_1,...,x_n;\theta) = f_S(s;\theta)h(x_1,...,x_n)/c(s)$$

dan

$$\frac{f(x_1,...,x_n;\theta)}{f_S(s;\theta)} = \frac{h(x_1,...,x_n)}{c(s)}$$

yang bebas dari $\boldsymbol{\theta}$.

Teorema 10. 1.. Kriteria Faktorisasi.

Jika $X_1, \ldots, X_n$ memiliki fungsi kepekatan peluang bersama $f(x_1,...,x_n;\theta)$ dan jika ***S***=$(S_1, \ldots, S_k)$, maka $S_1, \ldots, S_k$ adalah statistik cukup bersama untuk $\boldsymbol{\theta}$ jika dan hanya jika $f(x_1,...,x_n;\theta) = g(s;\theta)h(x_1,...,x_n)$ dimana $g(s;\theta)$ tidak tergantung pada $x_1, \ldots, x_n$ kecuali melalui $\boldsymbol{s}$, dan $h(x_1, \ldots, x_n)$ tidak mengandung $\boldsymbol{\theta}$.

Teladan 10. 3.

Suatu contoh acak dari sebaran Bernoulli, $X_i \sim Bin(1,\theta)$. Kita ingin tunjukkan bahwa $S = \sum X_i$ merupakan statistik cukup. Kita tahun bahwa $f(x_i;\theta) = \theta^x (1-\theta)^{1-x}$, sehingga

$$\begin{aligned} f(x_1,\ldots,x_n;\theta) &= \theta^{\sum x_i}(1-\theta)^{n-\sum x_i} \\ &= \theta^s (1-\theta)^{n-s} \\ &= g(s;\theta)h(x_1,\ldots.,x_n) \end{aligned}$$

dimana $s = \sum x_i$, dan dalam kasus ini, kita definisikan $h(x_1, \ldots, x_n)$ = 1 jika semua x_i = 0 atau 1, dan nol selainnya. Perlu dicatat bahwa proporsi contoh, $\hat{\theta} = S/n$ juga merupakan statistik cukup bagi θ. Secara umum, jika suatu statistik S merupakan statistik cukup bagi θ, maka sembarang fungsi satu-satu dari S juga merupakan statistik cukup bagi θ.

Teladan 10. 4.

Suatu contoh acak dari sebaran Seragam, $X_i \sim SK(0,\theta)$, dimana θ tidak diketahui. Fungsi kepekatan peluang bersama $X_1, \ldots, X_n$ adalah

$$f(x_1,\ldots,x_n;\theta) = \prod_{i=1}^{n} \frac{1}{\theta} I_{(0,\theta)}(x_i)$$

atau dapat dituliskan dalam bentuk minimum, $x_{(1)}$, dan maksimum, $x_{(n)}$, dari $x_1, \ldots, x_n$ seperti berikut

$$f(x_1,\ldots,x_n;\theta) = \frac{1}{\theta^n} I_{(0,\infty)}(x_{(1)}) I_{(0,\theta)}(x_{(n)})$$

yang berarti bahwa

$$f(x_1,\ldots,x_n;\theta) = g(x_{(n)};\theta)h(x_1,\ldots.,x_n)$$

dimana $g(s;\theta) = \frac{1}{\theta^n} I_{(0,\theta)}(s)$ dan $h(x_1, \ldots, x_n) = I_{(0,\infty)}(x_{(1)})$.
Berdasarkan teorema Faktorisasi, maka $S = X_{(n)}$ merupakan statistik cukup bagi θ.

Teladan 10. 5.

Contoh acak dari sebaran Normal, $X_i \sim N(\mu,\sigma^2)$, dimana μ dan σ^2 tak diketahui. Dengan demikian

$$f(x_1,\ldots,x_n;\mu,\sigma^2) = \frac{1}{(2\pi\sigma^2)^{n/2}} \exp\left[-\frac{1}{2\sigma^2}\sum(x_i-\mu)^2\right]$$

atau

$$f(x_1,\ldots,x_n;\mu,\sigma^2) = \frac{1}{(2\pi\sigma^2)^{n/2}} \exp\left[-\frac{1}{2\sigma^2}\left(\sum x_i^2 - 2\mu\sum x_i + n\mu^2\right)\right]$$

Dengan $s_1 = \sum x_i$ dan $s_2 = \sum x_i^2$, kita bisa pilih

$$g(s_1,s_2;\mu,\sigma^2) = \frac{1}{(2\pi\sigma^2)^{n/2}} \exp\left[-\frac{1}{2\sigma^2}\left(s_2 - 2\mu s_1 + n\mu^2\right)\right]$$

dan $h(x_1, \ldots, x_n) = 1$.

Dengan demikian, berdasarkan kriteria faktorisasi, $S_1 = \sum X_i$ dan $S_2 = \sum X_i^2$ adalah statistik cukup bersama bagi $\boldsymbol{\theta} = (\mu,\sigma^2)$. Selanjutnya dapat ditunjukkan bahwa $\overline{X}$ dan $\frac{1}{n}\sum(X_i - \overline{X})^2$ adalah juga statistik cukup bersama bagi $\boldsymbol{\theta} = (\mu,\sigma^2)$.

Bila terdapat gugus statistik cukup minimal, kita berharap bahwa banyaknya statistik cukup akan sama dengan banyaknya parameter yang tak diketahui.

Teladan 10. 6.

Suatu contoh acak dari sebaran Seragam, X_i ~ $SK(\theta,\theta+1)$, dimana θ tak diketahui. Fungsi kepekatan peluang bersamanya adalah

$$f(x_1,...,x_n;\theta)=\prod_{i=1}^{n} I_{(\theta,\theta+1)}(x_i)$$

Fungsi ini mengasumsikan bernilai 1 jika dan hanya jika $\theta < x_i$ dan $x_i < \theta+1$ untuk semua x_i, sehingga

$$f(x_1,...,x_n;\theta)=\prod_{i=1}^{n} I_{(\theta,\infty)}(x_i)\prod_{i=1}^{n} I_{(-\infty,\theta+1)}(x_i)$$

atau

$$f(x_1,...,x_n;\theta)=I_{(\theta,\infty)}(x_{(1)})I_{(-\infty,\theta+1)}(x_{(n)})$$

yang menunjukkan, dengan kriteria faktorisasi, bahwa statistik tataan terkecil $S_1 = X_{(1)}$ dan statistik terbesar $S_2 = X_{(n)}$ merupakan statistik cukup bersama bagi θ. Juga dapat ditunjukkan bahwa $S_1 = X_{(1)}$ dan $S_2 = X_{(n)}$ adalah statistik cukup minimal.

Teorema 10. 2.

Jika $S_1, \ldots, S_k$ adalah statistik cukup bersama untuk $\boldsymbol{\theta}$ dan jika $\hat{\theta}$ merupakan penduga kemungkinan maksimum yang khas bagi $\boldsymbol{\theta}$, maka $\hat{\theta}$ merupakan fungsi dari $\mathbf{S} = (S_1, \ldots, S_k)$.

Bukti

Dengan kriteria faktorisasi,

$$L(\theta)=f(x_1,...,x_n;\theta)=g(s;\theta)h(x_1,...,x_n)$$

yang berarti bahwanilai yang memaksimumkan fungsi likelihood harus tergantung pada $\boldsymbol{s}$, misalkan $\hat{\theta}=t(s)$. Jika penduga kemungkinan maksimum tersebut khas atau unik, akan mendefinisikan fungsi dari $\boldsymbol{s}$.

Sesungguhnya, hasilnya dapat dinyatakan lebih umum lagi:

Jika terdapat statistik cukup bersama, dan jika terdapat sebuah penduga kemungkinan maksimum, maka akan ada sebuah penduga kemungkinan maksimum yang merupakan fungsi dari statistik cukup.

Hal tersebut juga berarti bahwa, jika terdapat penduga kemungkinan penduga kemungkinan $\hat{\theta}_1, \ldots, \hat{\theta}_k$ yang khas dan merupakan statistik cukup bersama, maka mereka juga merupakan gugus statistik cukup minimal, karena kriteria faktorisasi juga berlaku untuk setiap gugus statistik cukup bersama.

Berikut ini adalah teladan yang memperlihatkan bahwa dimungkinakan mendapatkan statistik cukup *S* dan penduga kemungkinan maksimum yang tidak merupakan fungsi dari *S*.

Teladan 10. 7.

Misalkan *X* adalah peubah acak diskrit dengan fungsi kepekatan peluang *f(x;θ)* dan Ω = *{0, 1}*, dimana dengan menggunakan notasi fungsi indikator

$$f(x;\theta) = \frac{1}{4} I_{\{1,4\}}(x) + \left(\frac{1-\theta}{4} + \frac{\theta}{6} \right) I_{\{2\}}(x) + \left(\frac{1-\theta}{4} + \frac{\theta}{3} \right) I_{\{3\}}(x)$$

Jika $s(x) = 3I_{\{1,4\}}(x) + 2I_{\{2\}}(x) + 4I_{\{3\}}(x)$, maka *S = s(X)* adalah statistik cukup bagi *θ,* yang dapat dilihat dari kriteria faktorisasi dengan

$$h(x) = I_{\{1,2,3,4\}}(x) \text{ dan } g(s;\theta) = \frac{1-\theta}{4} + \frac{\theta}{12} s$$

Lebih jauh lagi, terdapat lebih dari satu penduga kemungkinan maksimum. Misalnya, $t_1(x) = I_{\{1,3\}}(x)$ dan $t_2(x) = I_{\{1,3,4\}}(x)$ keduanya menghasilkan penduga kemungkinan maksimum $\hat{\theta}_1 = t_1(x)$ dan $\hat{\theta}_2 = t_2(x)$ karena nilai-nilai dugaannya akan memaksimumkan *f(x;θ)* untuk setiap nilai *x* yang tetap. Terlihat jelas bahwa $\hat{\theta}_1 = t_1(x)$ bukan merupakan fungsi dari *S*, karena *s(1) =*

s(4) = 3, tetapi $t_1(1) = 1$ sementara $t_1(4) = 0$. Namun demikian, $\hat{\theta}_2 = t_2(x) = t(s) = I_{\{3,4\}}(s)$.

Hal ini menunjukkan bahwa kita harus berhati-hati dalam membuat pernyataan mengenai hubungan antara statistik cukup dan penduga kemungkinan maksimumnya. Namun demikian, jika penduga kemungkinan maksimum tersebut khas, maka situasinya lebih jelas dan langsung.

Teorema 10. 3.

Jika *S* merupakan statistik cukup bagi θ, maka sembarang penduga Bayes akan merupakan fungsi dari *S*.

Bukti

Karena fungsi $h(x_1, \ldots, x_n)$ dalam kriteria faktorisasi tidak tergantung dari θ, hal ini dapat dieliminasi berdasarkan definisi kepekatan peluang posterior, dan $f_{\theta|x}(\theta)$ dapat diganti dengan

$$f_{\theta|x}(\theta) = \frac{g(s;\theta)p(\theta)}{\int g(s;\theta)p(\theta)d\theta}$$

Teorema 10. 4.

Jika $X_1, \ldots, X_n$ merupakan contoh acak dari sebaran kontinu dengan fungsi kepekatan peluang $f(x;\theta)$, maka statistik tataan akan membentuk suatu gugus statistik cukup bersama bagi $\boldsymbol{\theta}$.

Bukti

Untuk $x_{(1)}, \ldots, x_{(n)}$ yang tetap

$$\frac{f(x_1;\theta)\cdots f(x_n;\theta)}{n!\,f(x_{(1)};\theta)\cdots f(x_{(n)};\theta)} = \frac{1}{n!} \qquad x_1 < \cdots < x_n$$

dan nol selainnya.

Pada umumnya, statistik-statistik cukup digunakan dalam pembuatan penduga tak bias ragam minimum seragam (UMVUE).

Teorema 10. 5. Rao-Blackwell.

Misalkan $X_1, \ldots, X_n$ memiliki fungsi kepekatan peluang bersama $f(x_1, \ldots, x_n;\theta)$, dan misalkan $\boldsymbol{S} = (S_1, \ldots, S_k)$ adalah vektor statistik cukup bersama bagi $\boldsymbol{\theta}$. Jika T adalah sembarang penduga tak bias bagi $\tau(\boldsymbol{\theta})$, dan jika $T^* = E(T|\boldsymbol{S})$, maka

1. T^* merupakan penduga tak bias bagi $\tau(\boldsymbol{\theta})$,
2. T^* merupakan fungsi dari $\boldsymbol{S}$, dan
3. Var(T^*) $\leq$ Var(T) untuk setiap $\boldsymbol{\theta}$, dan Var(T^*) < Var(T) untuk beberapa $\boldsymbol{\theta}$ kecuali bila $T^* = T$ dengan peluang 1.

Bukti

Berdasarkan teori kecukupan, $f_{T|s}(t)$ tidak mengandung $\boldsymbol{\theta}$, dan dengan demikian maka fungsi $t^*(\boldsymbol{s}) = E(T|\boldsymbol{s})$ tidak tergantung pada $\boldsymbol{\theta}$. Sehingga, $T^* = t^*(\boldsymbol{S}) = E(T|\boldsymbol{S})$ merupakan penduga yang merupakan fungsi dari $\boldsymbol{S}$, dan selanjutnya

$$E(T^*) = E_S(T^*) = E_S[E(T \mid S)] = E(T) = \tau(\theta)$$

Juga dapat diperlihatkan bahwa

$$\begin{aligned} \text{Var}(T) &= \text{Var}[\text{E}(T|\boldsymbol{S})] + \text{E}[\text{Var}(T|\boldsymbol{S})] \\ &\geq \text{Var}[\text{E}(T|\boldsymbol{S})] \\ &= \text{Var}(T^*) \end{aligned}$$

dengan Var(T) = Var(T^*) jika dan hanya jika E[Var($T|\boldsymbol{S}$)] = 0, yang hanya terjadi jika dan hanya jika Var($T|\boldsymbol{S}$) = 0 dengan peluang 1, atau setara dengan pernyataan $T = E(T|\boldsymbol{S}) = T^*$.

Berdasarkan teorema diatas, jika kita menginginkan penduga tak bias dengan ragam yang kecil, kita dapat membatasi perhatian kita pada fungsi statistik cukup. Jika terdapat penduga yang tak bias, maka akan ada satu yang merupakan fungsi dari statistik cukup, sebut saja, $E(T|\boldsymbol{S})$, yang juga tak bias dan memiliki ragam yang kecil. Teorema diatas sudah mengarahkan kita untuk mendapatkan UMVUE untuk suatu parameter. Kita hanya perhitungkan fungsi tak bias dari S dalam pencarian sebuah UMVUE. Dalam beberapa kasus dimungkinkan untuk memperlihatkan bahwa hanya ada satu fungsi *S* yang tak bias, dan UMVUE. Konsep “lengkap” sangat membantu dalam penentuan penduga tak bias yang khas (unik).

Statistik Lengkap dan Kelas Eksponensial

Definisi 10. 3. Statistik Lengkap.

Suatu famili fungsi kepekatan peluang, $\{f_T(\mathbf{t};\boldsymbol{\theta});\boldsymbol{\theta}\in\boldsymbol{\Omega}\}$ dikatakan **lengkap** jika $E[u(\mathbf{T})] = 0$ untuk semua $\boldsymbol{\theta}\in\boldsymbol{\Omega}$ yang juga berarti $u(\mathbf{T}) = 0$ dengan peluang 1 untuk semua $\boldsymbol{\theta}\in\boldsymbol{\Omega}$.

Hal ini kadang diekspresikan dengan mengatakan bahwa tak ada penduga tak bias *nontrivial* dari nol. Secara khsus, ini berarti bahwa dua fungsi $\mathbf{T}$ yang berbeda tak dapat memiliki nilai harapan yang sama. Sebagai contoh, apabila $E[u_1(\mathbf{T})] = \tau(\boldsymbol{\theta})$ dan $E[u_2(\mathbf{T})] = \tau(\boldsymbol{\theta})$, maka $E[u_1(\mathbf{T})- u_2(\mathbf{T})] = 0$, yang berarti bahwa $u_1(\mathbf{T})- u_2(\mathbf{T}) = 0$, atau $u_1(\mathbf{T}) = u_2(\mathbf{T})$ dengan peluang 1, jika famili fungsi kepekatan peluang tersebut lengkap. Dengan demikian, sembarang penduga tak bias bersifat unik dalam hal ini. Kita ingin mengetahui bahwa famili fungsi kepekatakan peluang statistik cukup juga lengkap, karena pada kasus ini, suatu fungsi statistik cukup yang tak bias akan khas (unik) dan haruslah merupakan UMVUE berdasarkan teorema Rao-Blackwell.

Statistik cukup yang fungsi kepekatan peluang nya adalah anggota dari famili fungsi kepekatan yang lengkap selanjutnya disebut dengan **statistik cukup dan lengkap** atau untuk singkatnya **statistik cukup lengkap** saja.

Teorema 10. 6. Lehmann-Scheffe

Misalkan $X_1, \ldots, X_n$ memiliki fungsi kepekatan peluang bersama

$f(x_1, \ldots, x_n; \boldsymbol{\theta})$, dan misalkan ***S*** merupakan vektor statistik cukup lengkap bersama untuk $\boldsymbol{\theta}$. Jika $T^* = t^*(\mathbf{S})$ adalah statistik yang tak bias bagi $\tau(\boldsymbol{\theta})$ dan merupakan fungsi dari ***S***, maka T^* adalah UMVUE bagi $\tau(\boldsymbol{\theta})$.

Bukti

Berdasarkan definisi kelengkapan bahwa sembarang statistik yang merupakan fungsi dari ***S*** dan merupakan penduga tak bias bagi $\tau(\boldsymbol{\theta})$ haruslah sama dengan T^* dengan peluang 1. Jika T adalah sembarang statistik lainnya dan juga tak bias bagi $\tau(\boldsymbol{\theta})$, maka berdasarkan teorema Rao-Blackwell, $E(T|\boldsymbol{S})$ juga tak bias bagi $\tau(\boldsymbol{\theta})$ dan merupakan fungsi dari ***S***. Dengan kekhasannya, $T^* = E(T|\boldsymbol{S})$ dengan peluang 1. Lebih jauh lagi dapat ditunjukkan bahwa $\text{Var}(T^*) \leq \text{Var}(T)$ untuk semua $\boldsymbol{\theta}$. Dengan demikian T^* adalah UMVUE bagi $\tau(\boldsymbol{\theta})$.

Teladan 10. 8.

Misalkan $X_1, \ldots, X_n$ merupakan contoh acak dari sebaran Poisson, $X_i \sim Poi(\mu)$, sehingga

$$f(x_1, \ldots, x_n; \mu) = \frac{e^{-n\mu} \mu^{\sum x_i}}{\prod x_i!}$$

Dengan kriteria faktorisasi, $S = \sum X_i$ adalah statistik cukup. Kita juga dapat perlihatkan bahwa $S = \sum X_i \sim Poi(n\mu)$, dan kita ingin tunjukkan bahwa famili Poisson juga lengkap. Untuk kenyamanan, dimisalkan $\theta = n\mu$, dan misalkan juga sembarang fungsi $u(s)$. Selanjutnya

$$E[u(S)] = \sum_{s=0}^{\infty} u(s) \frac{e^{-\theta}\theta^s}{s!} = e^{-\theta}\left(\sum_{s=0}^{\infty} u(s) \frac{\theta^s}{s!}\right)$$

atau

$$E[u(S)] = e^{-\theta}\left(u(0)\frac{\theta^0}{0!} + u(1)\frac{\theta^1}{1!} + u(2)\frac{\theta^2}{2!} + u(3)\frac{\theta^3}{3!} + \ldots\right)$$

Karena $e^{-\theta} \neq 0$, bila dibuat $E[u(S)] = 0$ hanya bisa dipenuhi apabila semua koeffisien θ^s harus nol. Jadi untuk semua s, berlaku $u(s)/s! = 0$ yang berimplikasi bahwa $u(s) = 0$. Dengan menggunakan terminologi kelengkapan, $\overline{X} = S/n$ adalah fungsi dari S yang khas dan tak bias bagi $E(\overline{X}) = \mu$. Berdasarkan teorema Lehmann-Scheffe maka $\overline{X} = S/n$ adalah UMVUE bagi μ.

Hasil ini dapat diverifikasi dengan membandingkan $Var(\overline{X})$ terhadap CRLB; namun, pendekatan dengan CRLB tidak akan sesuai untuk fungsi nonlier dari S. Pendekatan yang sedang kita bahas dapat digunakan untuk mencari UMVUE dari $\tau(\theta) = E[u(S)]$, untuk setiap fungsi $u(s)$ yang memiliki nilai harapan.

Jika wilayah peubah acak tidak tergantung parameter, sepertinya kita perlu membatasi perhatian pada famili kepekatan peluang yang memiliki bentuk "kelas eksponensial" bilamana kita membahas statistik cukup lengkap, sehingga kita tidak perlu mempertimbangkan secara detil famili-famili ini secara terpisah.

Definisi 10. 4. Kelas Eksponensial

Suatu fungsi kepekatan peluang dikatakan sebagai anggota dari **kelas eksponensial reguler (KER)** atau ***Regular Exponential Class (REC)*** jika dapat diekspresikan dalam bentuk

$$f(x;\theta) = c(\theta)h(x)\exp\left[\sum_{j=1}^{k} q_j(\theta)t_j(x)\right] \qquad x \in A$$

dan nol selainnya, dimana $\boldsymbol{\theta} = (\theta_1, \ldots, \theta_k)$ adalah vektor k-parameter yang tak diketahui, jika ruang parametr memiliki bentuk

$$\Omega = \{\theta \mid a_i \leq \theta_i \leq b_i, i = 1,\ldots,k\}$$

dan jika kondisi dibawah ini (1, 2, dan 3.a atau 3.b) dipenuhi:

1. Gugus $A = \{x; f(x;\theta) > 0\}$ tak tergantung pada θ.
2. Fungsi-fungsi $q_j(\boldsymbol{\theta})$ nontrivial, secara fungsi bebas dari

fungsi-fungsi kontinu θ_i.

3.a. Untuk peubah acak kontinu, turunan-turunan $t_j^{'}(x)$ merupakan fugsi kontinu dari *x* yang bebas linier pada *A*.

3.b. Untuk peubah acak diskrit, $t_j(x)$ fungsi dari *x* yang nontrivial pada *A*, dan tak satupun merupakan fungsi linier dari yang lainnya.

Untuk kenyamanan, kita akan tuliskan *f(x;*$\boldsymbol{\theta}$*)* adalah anggota dari KER($q_1, \ldots, q_k$) atau REC($q_1, \ldots, q_k$) atau mungkin hanya KER atau REC saja.

Teladan 10. 9.

Pertimbangkan sebuah sebaran Bernoulli, *X ~ Bin(1,p)*. Hal ini berarti bahwa

$$f(x;p) = p^x(1-p)^{1-x}$$

$$= (1-p)\exp\left\{x\ln\left[\frac{p}{1-p}\right]\right\} \qquad x \in A = \{0,1\}$$

adalah KER(q_1), dimana $q_1(p) = \ln\left(\frac{p}{1-p}\right)$ dan $t_1(x) = x$.

Dengan sedikit modifikasi kondisi reguler, KER dapat dikembangkan pada kasus dimana ***X*** adalah vektor.

Dapat ditunjukkan bahwa KER merupakan famili yang lengkap, dan banyak fungsi kepekatan seperti Binomial, Poisson, Eksponensial, Gamma dan Normal memiliki bentuk KER, sehingga mereka tergolong famili yang lengkap. Kita ingin mengetahui bahwa fungsi kepekatan peluang statistik cukup statistik cukup dari model ini lengkap.

Teorema 10. 7.

Jika $X_1, \ldots, X_n$ adalah contoh acak dari anggota KER($q_1, \ldots, q_k$), maka statistik-statistik

$$S_1 = \sum_{i=1}^{n} t_1(X_i), \ldots, S_k = \sum_{i=1}^{n} t_k(X_i)$$

adalah gugus minimal dari statistik cukup lengkap

.

Teladan 10. 10.

Dari teladan sebelum ini, $X \sim Bin(1,p)$. Untuk sebuah contoh berukuran n, maka $t(x_i) = x_i$ dan $S = \sum X_i$ adalah statistik cukup lengkap bagi p. Jika kita menginginkan UMVUE bagi Var(X) = $p(1-p)$, maka perlu dicoba penduganya adalah $\overline{X}(1-\overline{X})$. Selanjutnya,

$$\begin{aligned} E\left[\overline{X}(1-\overline{X})\right] &= E(\overline{X}) - E(\overline{X}^2) \\ &= p - \left[p^2 + Var(\overline{X})\right] \\ &= p - p^2 - \frac{p(1-p)}{n} \\ &= p(1-p)\left(1 - \frac{1}{n}\right) \end{aligned}$$

dan dengan demikian UMVUE bagi $p(1-p)$ adalah $\dfrac{n}{n-1}\overline{X}(1-\overline{X})$.

Kita telah melihat bahwa terdapat hubungan yang dekat antara KER, statistik cukup lengkap, dan UMVUE. Juga, penduga kemungkinan maksimum merupakan fungsi dari statistik cukup minimal, dan penduga kemungkinan maksimum efisien asimtotik dengan ragam asimtotik sama dengan CRLB. Jika suatu penduga yang ragamnya sama dengan CRLB, maka penduga tersebut disebut juga sebagai penduga CRLB.

Teorema 10. 8.

Jika ada penduga CRLB, T untuk $\tau(\theta)$, maka akan ada statistik cukup tunggal, dan T merupakan fungsi dari statistik cukup tersebut. Sebaliknya, jika terdapat statistik cukup tunggal dan CRLBnya ada, maka penduga CRLB ada untuk beberapa $\tau(\theta)$.

Teorema 10. 9.

Jika CRLB ada, maka penduga CRLB akan ada untuk beberapa fungsi dari $\tau(\theta)$ jika dan hanya jika fungsi kepekatannya adalah anggota dari KER. Lebih jauh lagi, penduga CRLB bagi $\tau(\theta)$ adalah $\tau(\hat{\theta})$, dimana $\hat{\theta}$ adalah penduga kemungkinan maksimum bagi θ.

Kebanyakan fungsi kepekatan peluang yang dipelajari yang tidak termasuk dalam KER, termasuk atau tergolong pad akelas lainnya, yang membolehkan wilayah X yang dinotasikan dengan $A = \{x; f(x;\boldsymbol{\theta}) > 0\}$ tergantung pada $\boldsymbol{\theta}$.

Definisi 10. 5.

Suatu fungsi kepekatan peluang dikatakan menjadi anggota **kelas eksponensial wilayah terkait** atau **range dependent exponential class** yang dinotasikan dengan KEWT($q_1, \ldots, q_k$) atau RDEC($q_1, \ldots, q_k$), jika untuk $j = 3, \ldots, k$ memenuhi kondisi

1. Fungsi-fungsi $q_j(\boldsymbol{\theta})$ nontrivial, secara fungsi bebas dari fungsi-fungsi kontinu θ_i.

2.a. Untuk peubah acak kontinu, turunan-turunan $t_j'(x)$ merupakan fugsi kontinu dari x yang bebas linier pada A.

2.b. Untuk peubah acak diskrit, $t_j(x)$ fungsi dari x yang nontrivial pada A, dan tak satupun merupakan fungsi linier dari yang lainnya.

dan memiliki bentuk

$$f(x;\theta) = c(\theta)h(x)\exp\left[\sum_{j=3}^{k} q_j(\theta_3,...,\theta_k)t_j(x)\right]$$

dimana $A = \{x|\ q_1(\theta_1,\theta_2) < x < q_2(\theta_1,\theta_2)\}$ dan $\boldsymbol{\theta} \in \boldsymbol{\Omega}$

Hal tersebut akan mencakup kasus khusus seperti kasus satu parameter dengan $f(x;\theta) = c(\theta)h(x)$ dengan $A = \{x|\ q_1(\theta) < x < q_2(\theta)\}$; dan dua parameter dengan fungsi kepekatan peluang $f(x;\theta_1,\theta_2) = c(\theta_1,\theta_2)h(x)$ dengan $A=\{x|q_1(\theta_1,\theta_2) < x < q_2(\theta_1,\theta_2)\}$

Teorema berikut sangat berguna dalam identifikasi statistik cukup dalam kasus wilayah terkait.

Teorema 10. 10.

Misalkan $X_1, \ldots, X_n$ merupakan contoh acak dari anggota RDEC($q_1, \ldots, q_k$).

1. Jika $k > 2$, maka $S_1 = X_{(1)}$, $S_2 = X_{(n)}$ dan $S_3, \ldots, S_k$ dimana $S_j = \sum_{i=1}^{n} t_j(X_i)$ merupakan statistik cukup bersama bagi $\boldsymbol{\theta} = (\theta_1, \ldots, \theta_k)$.
2. Dalam kasus dua parameter, $S_1 = X_{(1)}$ dan $S_2 = X_{(n)}$ merupakan statistik cukup bersama bagi $\boldsymbol{\theta} = (\theta_1, \theta_2)$.
3. Dalam kasus satu parameter, $S_1 = X_{(1)}$ dan $S_2 = X_{(n)}$ merupakan statistik cukup bersama bagi θ. Jika $q_1(\theta)$ menaik dan $q_2(\theta)$ menurun, maka $T_1 = \min\left(q_1^{-1}(X_{(1)}), q_2^{-1}(X_{(n)})\right)$ merupakan statistik cukup tunggal bagi θ. Jika $q_1(\theta)$ menurun dan $q_2(\theta)$ menaik, maka $T_1 = \max\left(q_1^{-1}(X_{(1)}), q_2^{-1}(X_{(n)})\right)$ merupakan statistik cukup tunggal bagi θ.

Jika satu dari batasnya konstanta dan batas lainnya tergantung pada satu parameter, θ misalnya, maka teorema berikut berlaku.

Teorema 10. 11.

Misalkan $X_1, \ldots, X_n$ merupakan contoh acak dari anggota RDEC.

1. Jika $k > 2$ dan batas bawah konstanta, $q_1(\theta) = a$, maka $X_{(n)}$ dan statistik $\sum_{i=1}^{n} t_j(X_i)$ adalah statistik cukup bersama bagi θ dan θ_j; $j = 3, \ldots, k$. Jika batas atas konstanta, $q_2(\theta) = b$, maka $X_{(1)}$ dan statistik $\sum_{i=1}^{n} t_j(X_i)$ adalah statistik cukup bersama bagi θ dan θ_j; $j = 3, \ldots, k$.
2. Dalam kasus satu parameter, jika $q_1(\theta)$ tidak tergantung pada θ, maka $S_2 = X_{(n)}$ merupakan statistik cukup bagi θ, dan jika $q_2(\theta)$ tidak tergantung pada θ, maka $S_1 = X_{(1)}$ merupakan statistik cukup bagi θ.

Teladan 10. 11.

Misalkan kita memiliki fungsi kepekatan peluang dari sebuah peubah acak seperti berikut

$$f(x;\theta) = \frac{1}{2\theta} I_{(-\theta,\theta)}(x)$$

Disini kita punya $q_1(\theta) = -\theta$, suatu fungsi yang menurun pada θ, dan $q_2(\theta) = \theta$, suatu fungsi yang menaik pada θ. Dengan demikian, berdasarkan teorema diatas, maka $T_2 = \max\left[-X_{(1)}, X_{(n)}\right]$ merupakan statistik cukup tunggal bagi θ.

Teladan 10. 12.

Untuk sebaran Eksponensial 2-parameter, $X \sim Eks(\theta,\eta)$.

$$f(x;\theta,\eta) = \frac{1}{\theta} \exp\left[-\frac{(x-\eta)}{\theta}\right] I_{(\eta,\infty)}(x)$$

$$= \frac{1}{\theta}\exp\left(\frac{\eta}{\theta}\right)\exp\left(\frac{-x}{\theta}\right)I_{(\eta,\infty)}(x)$$

Jika $X_1, \ldots, X_n$ adalah contoh acak, maka berdasarkan teorema terakhir bahwa $X_{(1)}$ dan $\sum X_i$ adalah statistik cukup bersama bagi (θ, η). Karena $q_2(\eta) = \infty$ bukan merupakan fungsi parameter, maka $X_{(n)}$ tidak termasuk.

Misalkan θ diketahui, sebut saja $\theta = 1$. Maka

$$f(x;\eta) = \exp[-(x-\eta)]I_{(\eta,\infty)}(x)$$
$$= \exp(\eta)\exp(-x)I_{(\eta,\infty)}(x)$$

Disini kita lihat bahwa $X_{(1)}$ merupakan statistik cukup bagi η. Hal ini konsisten dengan hasil sebelumnya, dimana penduga η berdasarkan $X_{(1)}$ merupakan penduga yang lebih baik dari pada penduga lainnya, seperti $\bar{X}$.

Teladan 10. 13.

Sebuah contoh acak berukuran *n* dari sebaran Seragam, $X_i \sim SK(\theta_1, \theta_2)$. Karena

$$f(x;\theta_1,\theta_2) = \frac{1}{\theta_2 - \theta_1} I_{(\theta_1,\theta_2)}(x)$$

Berdasarkan teorema kedua terakhir diperoleh bahwa $X_{(1)}$ dan $X_{(n)}$ merupakan statistik cukup bersama bagi (θ_1, θ_2).

Tiga teladan terakhir hanya berkenaan dengan bagaimana mencari statistik cukup bagi anggota RDEC. Statistik-statistik tersebut mungkin saja merupakan statistik lengkap, namun diperlukan verifikasi dengan argumen yang terpisah.

Teladan 10. 14.

Sebuah contoh acak berukuran *n* dari sebaran Seragam, $X_i \sim SK(0,\theta)$. Dari beberapa teladan diperoleh bahwa $X_{(n)}$ merupakan statistik cukup bagi θ. Fungsi kepekatan peluang bagi $S = X_{(n)}$ adalah

$$f(s;\theta) = n\frac{s^{n-1}}{\theta^n} I_{(0,\theta)}(s)$$

Untuk memverifikasi bahwa *S* juga lengkap, asumsikan bahwa $E[u(S)] = 0$ untuk semua $\theta > 0$, yang berarti bahwa

$$\int_0^\theta \left[u(s)ns^{n-1}/\theta^n\right]ds = 0$$

Jika kedua sisi dikalikan dengan θ^n dan dideferensialkan terhadap θ, maka $u(\theta)n\theta^{n-1} = 0$ untuk semua $\theta > 0$, yang berimplikasi bahwa $u(s) = 0$ untuk semua $s > 0$, dan dengan demikian kita katakan bahwa *S* merupakan statistik cukup lengkap bagi θ.

Teorema 10. 12. Basu.

Misalkan $X_1, \ldots, X_n$ memiliki fungsi kepekatan peluang bersama $f(x_1,\ldots,x_n;\boldsymbol{\theta})$; $\boldsymbol{\theta} \in \boldsymbol{\Omega}$.
Misalkan bahwa $\boldsymbol{S} = (S_1, \ldots, S_k)$ dimana $S_1, \ldots, S_k$ adalah statistik cukup dan lengkap bersama bagi $\boldsymbol{\theta}$, dan misalkan bahwa *T* adalah statistik lainnya. Jika sebaran dari *T* tidak mencakup θ, maka $\boldsymbol{S}$ dan $\boldsymbol{T}$ saling bebas stokastik.

Bukti

Dalam kasus diskrit, notasikan $f(\mathbf{t})$, $f(s;\boldsymbol{\theta})$, dan $f(\mathbf{t}|\mathbf{s})$ berturut-turut adalah fungsi kepekatan peluang $\boldsymbol{T}$, $\boldsymbol{S}$, dan sebaran bersyarat $\boldsymbol{T}$ bilamana $\boldsymbol{S} = \boldsymbol{s}$. Selanjutnya

$$\begin{aligned} E_S\left[f(t) - f(t \mid S)\right] &= f(t) - \sum_s f(t \mid s) f(s;\theta) \\ &= f(t) - \sum_s f(s,t;\theta) \end{aligned}$$

= $f(t) - f(t) = 0$

Karena ***S*** adalah statistik cukup lengkap, *f(**t**|s)* = *f(**t**)* yang berarti bahwa ***S*** dan ***T*** saling bebas stokastik.

Teladan 10. 15.

Suatu contoh acak berukuran *n* dari sebaran Normal, $X_i \sim N(\mu,\sigma^2)$, μ dan σ^2 tidak diketahui. Telah ditunjukkan bahwa penduga kemungkinan maksimum bagi μ dan σ^2 adalah $\hat{\mu} = \overline{X}$ dan $\hat{\sigma}^2 = \dfrac{\sum (X_i - \overline{X})^2}{n}$. Sangatlah mudah untuk memverifikasi bahwa $\hat{\mu} = \overline{X}$ adalah statistik cukup lengkap bagi μ, untuk nilai σ^2 yang tetap. Juga dapat ditunjukkan bahwa

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

yang tidak tergantung pada μ. Dengan demikian $\hat{\mu} = \overline{X}$ dan $\hat{\sigma}^2 = \dfrac{\sum (X_i - \overline{X})^2}{n}$ merupakan peubah acak yang saling bebas. Juga, $\hat{\mu} = \overline{X}$ dan $\hat{\sigma}^2 = \dfrac{\sum (X_i - \overline{X})^2}{n}$ merupakan statistik cukup lengkap bersama bagi μ dan σ^2, dan kuantitas dalam bentuk $\dfrac{X_i - \overline{X}}{\hat{\sigma}}$ menyebar bebas dari μ dan σ, sehingga $\dfrac{X_i - \overline{X}}{\hat{\sigma}}$ juga bebas stokastik terhadap $\hat{\mu} = \overline{X}$ dan $\hat{\sigma}^2 = \dfrac{\sum (X_i - \overline{X})^2}{n}$.

Secara ringkas, tujuan dari bab ini adalah mengenalkan konsep cukup dan lengkap. Sebuah statistik, secara umum, menyediakan reduksi gugus data dari beberapa sebaran kedalam bentuk yang lebih ringkas. Jika statistik dikatakan cukup, maka statistik tersebut mengandung, dalam artian tertentu, seluruh informasi dalam data yang mencakup suatu parameter sebaran yang tak diketahui. Meskipun istilah cukup dapat diverifikasi, setidaknya secara teori, langsung dari definisi, hal ini biasanya dapat dilakukan dengan lebih mudah dengan menggunakan kriteria faktorisasi.

Jika statistik dikatakan cukup dan jika terdapat satu penduga kemungkinan maksimum yang unik, maka penduga kemungkinan maksimum tersebut merupakan fungsi dari statistik cukupnya. Statistik-statistik cukup juga berguna dalam pembentukan penduga tak bias ragam minimum seragam (UMVUE). Jika statistik dikatakan lengkap dan juga cukup bagi suatu parameter, dan jika terdapat penduga tak bias bagi suatu parameter (atau fungsi dari suatu parameter), maka terdapat sebuah UMVUE dan ini merupakan fungsi dari statistik cukup lengkap. Kadangkala, sangat sulit untuk memverifikasi statistik lengkap dengan menggunakan definisi secara langsung, namun kelas fungsi kepekatan peluang yang dikenal dengan kelas eksponensial memberikan cara yang nyaman untuk mengidentifikasi statistik cukup lengkap.

Latihan

1. Misalkan $X_1, \ldots, X_n$ adalah contoh acak dari sebaran Poisson, $X_i \sim Poi(\mu)$. Tunjukkan bahwa $S = \sum X_i$ adalah statistik cukup bagi μ berdasarkan definisi.
2. Misalkan $X_1, \ldots, X_n$ adalah contoh acak dari sebaran Geometrik, $X_i \sim Geo(p)$. Tunjukkan bahwa $S = \sum X_i$ adalah statistik cukup bagi p berdasarkan definisi.
3. Misalkan $X_1, \ldots, X_n$ adalah contoh acak dari sebaran Gamma, $X_i \sim Gam(\theta,2)$. Tunjukkan bahwa $S = \sum X_i$ adalah statistik cukup bagi θ berdasarkan definisi dan menggunakan kriteria faktorisasi.

4. Misalkan $X_1, \ldots, X_n$ adalah contoh yang saling bebas dari sebaran Binomial, $X_i \sim Bin(m_i,p)$. Tunjukkan bahwa $S = \sum X_i$ adalah statistik cukup bagi *p* dengan menggunakan kriteria faktorisasi.
5. Misalkan $X_1, \ldots, X_n$ adalah contoh yang saling bebas dari sebaran Beta, $X_i \sim Beta(\theta_1,\theta_2)$. Carilah statistik cukup bersama bagi θ_1 dan θ_2.
6. Suatu contoh acak berukuran *n* dari sebaran Eksponensial 2-parameter, $X_i \sim Eks(1,\eta)$.
   a. Tunjukkan bahwa $S = X_{(1)}$ adalah statistik cukup bagi η dengan menggunakan definisi.
   b. Tunjukkan bahwa *S* juga statistik lengkap.
   c. Tunjukkan bahwa $X_{(1)} - 1/n$ adalah UMVUE bagi η
   d. Carilah UMVUE bagi persentil ke-*p*.
7. Tunjukkan bahwa famili sebaran berikut termasuk anggota KER, dan untuk tiap kasus, carilah statistik cukupnya berdasarkan contoh acak $X_1, \ldots, X_n$.
   a. $Bin(1,p);\ 0 < p < 1$
   b. $Poi(\mu);\ \mu > 0$
   c. $Gam(\theta,\kappa); \theta > 0,\ \kappa > 0$
   d. $NB(r,p)$; *r* diketahui, $0 , p < 1$
   e. $Wei(\theta,\beta)$; β diketahui, $\theta > 0$
8. Misalkan $X_1, \ldots, X_n$ adalah contoh acak dari sebaran Bernoulli, $X_i \sim Bin(1,p);\ 0 < p < 1$.
   a. Carilah UMVUE untuk $Var(X) = p(1-p)$
   b. Carilah UMVUE untuk p^2.
9. Misalkan $X_1, \ldots, X_n$ adalah contoh acak dari sebaran dengan fungsi kepekatan peluang $f(x;\theta) = \theta x^{\theta-1} I_{(0,1)}(x)$ dengan $\theta > 0$.
   a. Carilah UMVUE untuk $1/\theta$.
   b. Carilah UMVUE untuk θ.
10. Misalkan $X_1, \ldots, X_n$ adalah contoh acak berukuran *n* dari sebaran dengan fungsi kepekatan peluang

$$f(x;\theta) = \theta(1+x)^{-(1+\theta)} I_{(0,\infty)}(x)$$

a. Carilah penduga kemungkinan maksimum $\hat{\theta}$ bagi θ.
b. Carilah statistik cukup dan lengkap bagi θi.
c. Carilah CRLB untuk *1/*θ.
d. Carilah UMVUE untuk *1/*θ.
e. Carilah sebaran normal asimtotik untuk $\hat{\theta}$ dan untuk 1/$\hat{\theta}$.
f. Carilah UMVUE untuk θ.

Pendugaan Interval

Pendahuluan

Kita telah bahas pendugaan titik pada beberapa bab terdahulu. Dengan nilai dugaan titik bagi suatu parameter, kita ingin mengetahui seberapa dekat nilai dugaan tersebut dengan nilai parameter yang sebenarnya. Pertanyaan ini dapat dijawab apabila kita mengetahui informasi yang berkenaan dengan varian (ragam) atau kuadrat tengah galat dari penduga tersebut. Pendekatan lain yang dapat dilakukan adalah dengan mencari penduga interval, dan dengan demikian kita perlu perhitungkan peluang interval tersebut akan mencakup nilai parameter yang sesungguhnya. Kita dapat mengatur interval sesuai dengan peluang yang kita harapkan, sehingga ukuran keakuratan secara otomatis telah termasuk dalam dugaan interval (selang) tersebut.

Teladan 11. 1.

Misalkan umur sebanyak empat puluh komponen elektronik diamati (dalam bulan), dan kita dapat berargumentasi bahwa umur komponen memiliki sebaran atau distribusi eksponensial. Sebagai konsenuensinya, kita asumsikan bahwa data diamati dari suatu sebaran Eksponensial, $X_i \sim Eks(\theta)$, dimana θ adalah rata-rata umur lampu yang sesungguhnya.

Perlu kita ingat kembali bahwa rata-rata contoh yang dinotasikan dengan $\bar{X}$ adalah merupakan penduga tak bias ragam minimum seragam (UMVUE) bagi θ. Misalkan dari data amatan kita peroleh nilai dugaan θ adalah $\bar{x}$ = 93,1 bulan. Meskipun kita tahu bahwa nilai dugaan ini berdasarkan suatu penduga dengan sifat-sifat yang optimal, suatu penduga titik tidak akan memberikan informasi mengenai **keakuratan**.

Jawaban kita dalam permasalahan ini adalah menurunkan suatu interval, dimana pada ujung-ujung interval adalah peubah-

peubah acak yang mencakup nilai parameter θ diantaranya, dengan peluang mendekati satu. Misalnya 0,99 atau 0,95.

Kita telah ketahui bahwa $\dfrac{2n\overline{X}}{\theta} \sim \chi^2_{(2n)}$. Dengan menggunakan $n = 40$ maka $\nu = 80$, dengan menggunakan tabel persentil Kai-kuadrat, akan didapatkan $\chi^2_{0,025;80} = 57,15$ dan $\chi^2_{0,975;80} = 106,63$. Yang berarti bahwa P[57,15 < 80$\overline{X}/\theta$ < 106,63] = 0,975 – 0,025 = 0,950 dan sebagai akibatnya bahwa

P[80$\overline{X}$/106,63 < θ < 80$\overline{X}$/57,15] = 0,950

Secara umum, interval dengan ujung-ujungnya yang bersifat acak disebut dengan **interval acak**. Interval (80$\overline{X}$/106,63 ; 80$\overline{X}$/57,15) adalah interval acak yang mengandung nilai θ yang sebenarnya dengan peluang 0,950.

Jika kita ganti nilai $\overline{X}$ dengan nilai dugaan contoh misalnya $\overline{x}$ = 93,1 maka hasil intervalnya menjadi (69,9 ; 130,3). Kita sebut selanjutnya interval ini dengan interval kepercayaan 95% bagi θ atau selang kepercayaan 95% bagi θ. Karena nilai pada ujung-ujung ini diketahui atau tetap, tidaklah tepat untuk mengatakan bahwa interval ini mengandung nilai yang sebenarnya bagi θ dengan peluang 0,95. Parameter θ meskipun tidak diketahui berupa suatu konstanta, dan interval tersebut dapat saja atau mungkin tidak mengandung θ. Namun demikian, fakta bahwa interval acak yang sesuai memiliki peluang 0,95 sebelum pendugaan, memberikan kepercayaan bahwa kita percaya 95% bahwa 69,9 < θ < 130,3.

Interval Kepercayaan

Misalkan $X_1, \ldots, X_n$ memiliki fungsi kepekatan peluang bersama $f(x_1, \ldots, x_n; \theta)$, $\theta \in \Omega$, dimana Ω berupa interval. Misalkan bahwa L dan U adalah statistik, sebut saja $L = l(X_1, \ldots, X_n)$ dan $U = u(X_1, \ldots, X_n)$. Jika suatu percobaan menghasilkan data $x_1, \ldots x_n$ maka kita memiliki nilai amatan $l(x_1, \ldots x_n)$ dan $u(x_1, \ldots, x_n)$.

Definisi 11. 1. Selang Kepercayaan atau Interval Kepercayaan

Suatu interval $(l(x_1, \ldots x_n); u(x_1, \ldots, x_n))$ disebut dengan selang kepercayaan 100γ% bagi θ jika $P[l(X_1, \ldots, X_n) < \theta < u(X_1, \ldots, X_n)] = \gamma$ dimana $0 < \gamma < 1$. Nilai dugaan $l(x_1, \ldots x_n)$ dan $u(x_1, \ldots, x_n)$ masing-masing berturut-turut adalah batas bawah dan batas atas kepercayaan.

Notasi lain yang sering dijumpai pada berbagai literatur statistika adalah θ_L dan θ_U untuk batas bawah dan atas kepercayaan. Kadang juga digunakan notasi $l(\mathbf{x}) = l(x_1, \ldots x_n)$ dan $u(\mathbf{x}) = u(x_1, \ldots, x_n)$ untuk melambangkan batasan nilai amatan.

Pembedaan perlu dilakukan antara interval acak $(L ; U)$ dan nilai amatan dugaan interval $(l(\mathbf{x}) ; u(\mathbf{x}))$ seperti yang telah kita sebutkan. Situasi ini analog dengan penduga dan nilai dugaannya. Terminologi lain yang berguna dalam pembedaan ini adalah dengan menyebut $(L ; U)$ sebagai penduga interval dan $(l(\mathbf{x}) ; u(\mathbf{x}))$ sebagai nilai dugaan interval. Tingkatan peluang, γ, juga disebut dengan koeffisien kepercayaan atau tingkat kepercayaan.

Interpretasi yang umum mengenai selang kepercayaan berdasarkan sifat frekuensi relatif dari peluang. Jika dugaan interval tersebut dihitung dari sekian banyak contoh, maka dalam jangka panjang, seandainya dilakukan dengan cara serupa, kita dapat berharap bahwa 100γ% interval tersebut akan mencakup nilai parameter θ yang sesungguhnya.

Definisi 11. 2. Batas Kepercayaan Satu Arah

1. Jika $P[l(X_1, \ldots, X_n) < \theta\,] = \gamma$ maka $l(\mathbf{x}) = l(x_1, \ldots x_n)$ disebut dengan batas bawah kepercayaan 100γ% bagi θ.
2. Jika $P[\theta < u(X_1, \ldots, X_n)] = \gamma$ maka $u(\mathbf{x}) = u(x_1, \ldots x_n)$ disebut dengan batas atas kepercayaan 100γ% bagi θ.

Tidaklah selalu mudah bagaimana mendapatkan batas-batas kepercayaan yang memenuhi definisi-definisi diatas. Kosep kecukupan kadang memberikan bantuan dalam permasalahan ini. Jika terdapat satu statistik cukup tunggal *S*, kita dapat mencari batas kepercayaan berdasarkan suatu fungsi dari statistik *S* ini. Jika tidak, statistik lain seperti penduga kemungkinan maksimum mungkin saja dapat dipakai.

Teladan 11. 2.

Misalkan suatu contoh acak berukuran *n* diambil dari sebaran eksponensial, $X_i \sim Eks(\theta)$, dan kita ingin menurunkan batas bawah kepercayaan 100γ% bagi θ. Kita tahu bahwa $\overline{X}$ merupakan statistik cukup bagi θ dan juga $\dfrac{2n\overline{X}}{\theta} \sim \chi^2_{(2n)}$. Dengan demikian,

$$\gamma = P[\frac{2n\overline{X}}{\theta} < \chi^2_{(2n)}] = P[2n\overline{X} / \chi^2_{2n;\gamma} < \theta]$$

Jika $\overline{x}$ diperoleh nilainya, maka batas bawah kepercayaannya adalah $l(\mathbf{x}) = 2n\overline{x} / \chi^2_{2n;\gamma}$. Dan dengan cara yang sama, maka batas atas kepercayaannya adalah $u(\mathbf{x}) = 2n\overline{x} / \chi^2_{2n;1-\gamma}$.

Namun seandainya yang ingin kita cari adalah selang kepercayaan 100γ% bagi θ. Jika kita pilih $\alpha_1 > 0$ dan $\alpha_2 > 0$ sedemikian rupa sehingga $\alpha_1+\alpha_2=\alpha = 1-\gamma$, maka kita peroleh

$$P\left[\chi^2_{\alpha_1;2n} < \frac{2n\overline{X}}{\theta} < \chi^2_{1-\alpha_2;2n}\right] = 1-\alpha_1-\alpha_2$$

Dan dengan demikian

$$P\left[\frac{2n\overline{X}}{\chi^2_{1-\alpha_2;2n}} < \theta < \frac{2n\overline{X}}{\chi^2_{\alpha_1;2n}}\right] = \gamma$$

Dalam prakteknya, biasanya kita gunakan $\alpha_1 = \alpha_2$, yang dikenal dengan pilihan ekor sama, artinya luas ekor kiri sama dengan luas ekor kanan, yang berimplikasi bahwa α_1 = α_2 = $\alpha/2$. Dengan demikian interval kepercayaan tersebut memiliki bentuk

$$\left(\frac{2n\bar{x}}{\chi^2_{1-\frac{\alpha}{2};2n}};\frac{2n\bar{x}}{\chi^2_{\frac{\alpha}{2};2n}}\right)$$

Secara umum tentunya, berdasarkan taraf kepercayaan tertentu, kita ingin mencari interval dengan jarak yang minimum. Oleh karenanya, dalam beberapa permasalahan, pemilihan luas ekor sama akan memberikan jarak minimum dimaksud, namun tidak selamanya bahwa pemilihan luas ekor yang sama akan menghasilkan jarak minimum. Hanya sebaran yang simetris tampaknya yang menghasilkan jarak minimum pada pemilihan luas ekor yang sama.

Teladan 11. 3.

Pertimbangkan sebuah contoh acak yang diambil dari sebaran Normal, $X_i \sim N(\mu,\sigma^2)$, dimana σ^2 diasumsikan diketahui. Dalam hal ini $\bar{X}$ merupakan statistik cukup bagi μ, dan kita juga tahu bahwa $Z = \frac{\sqrt{n}(\bar{X}-\mu)}{\sigma}$ ~ $N(0,1)$. Karena sebaran normal baku simetris terhadap titik nol, maka $z_{\alpha/2} = -z_{1-\alpha/2}$, sehingga

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

$$= P\left[\bar{X} - z_{1-\frac{\alpha}{2}}\frac{\sigma}{\sqrt{n}} < \mu < \bar{X} + z_{1-\frac{\alpha}{2}}\frac{\sigma}{\sqrt{n}}\right]$$

Dengan demikian selang kepercayaan 100(1-α)% bagi μ diberikan oleh

$$(\bar{x} - z_{1-\frac{\alpha}{2}}\frac{\sigma}{\sqrt{n}}\ ;\ \bar{x} + z_{1-\frac{\alpha}{2}}\frac{\sigma}{\sqrt{n}}).$$

Sebagai teladan, untuk selang kepercayaan 95%, kita peroleh 1-$\alpha/2$ = 0,975 dan dengan menggunakan tabel kita peroleh $z_{0,975}$ = 1,96. Sehingga batas bawah dan atas kepercayaan tersebut adalah $\bar{x} \pm 1{,}96 \dfrac{\sigma}{\sqrt{n}}$.

Perlu diketahui bahwa jawaban terakhir tidak dapat diterima bilamana σ^2 tidak diketahui, karena batas-batas kepercayaan tergantung pada parameter yang tak diketahui dan dengan demikian tak dapat dihitung. Dengan sedikit modifikasi, dimungkinkan untuk memperoleh selang kepercayaan untuk μ, meskipun σ^2 merupakan parameter pengganggu yang tak diketahui. Memang benar, kesulitan utama dalam penentuan selang kepercayaan muncul dalam parameter ganda dimana banyak terdapat parameter pengganggu. Metode yang lebih umum yang kadang memberikan cara penanganan masalah ini disajikan pada bagian berikut ini.

Dalam kasus parameter ganda dimungkinkan juga memiliki **wilayah kepercayaan** bersama yang berlaku pada seluruh parameter secara bersama. Juga, wilayah kepercayaan untuk parameter tunggal, dalam kasus berdimensi satu dapat berupa beberapa gugus selain interval. Secara umum, jika $\theta \in \Omega$, maka sembarang wilayah $A_\theta(x_1, \ldots, x_n)$ dalam Ω adalah memrupakan wilayah kepercayaan 100γ% jika peluang $A_\theta(X_1, \ldots, X_n)$ terdapat nilai parameter $\boldsymbol{\theta}$ yang sebenarnya adalah sebesar γ.

Metode Kuantitas Pivot

Misalkan $X_1, \ldots X_n$ memiliki fungsi kepekatan peluang bersama $f(x_1, \ldots, x_n; \theta)$ dan kita ingin mendapatkan batas-batas kepercayaan untuk θ dimana parameter pengganggu lain yang tak diketahui mungkin ada.

Definisi 11. 3. Kuantitas Pivot.

Jika $Q = q(X_1, \ldots, X_n; \theta)$ adalah peubah acak yang merupakan fungsi dari peubah acak $X_1, \ldots, X_n$ dan θ saja, maka Q disebut sebagai **kuantitas pivot** jika distribusinya tidak tergantung pada θ atau sembarang parameter lain yang tak diketahui.

Jika Q adalah kuantitas pivot untuk suatu parameter θ dan jika persentil dari Q, sebut saja q_1 dan q_2 ada sedemikian rupa sehingga

$$P[q_1 < q(X_1, \ldots, X_n;\theta) < q_2] = \gamma$$

Maka untuk contoh amatan $x_1, \ldots, x_n$, wilayah kepercayaan 100γ% bagi θ adalah gugus $\theta \in \Omega$ yang memenuhi pertidaksamaan berikut

$$q_1 < q(x_1, \ldots, x_n; \theta) < q_2$$

Wilayah kepercayaan seperti itu tidaklah selalu merupakan interval, dan secara umum dapat rumit bentuknya. Namun demikian, terdapat situasi yang lebih penting dimana selang kepercayaan dapat diperoleh. Salah satu jawaban umum yang akan selalu menghasilkan suatu interval bilamana untuk setiap gugus nilai $x_1, \ldots, x_n$ fungsi $q(x_1, \ldots, x_n; \theta)$ merupakan fungsi monoton menaik (atau menurun) dari θ. Juga dimungkinkan untuk mengidentifikasi tipe sebaran tertentu yang akan dihasilkan dari kuantitas pivot.

Teorema 11. 1.

Misalkan $X_1, \ldots, X_n$ merupakan contoh acak dari suatu sebaran dengan fungsi kepekatan $f(x;\theta)$ untuk $\theta \in \Omega$ dan asumsikan bahwa penduga kemungkinan maksimum $\hat{\theta}$ ada.

1. Jika θ adalah parameter lokasi, maka kuantitas pivotnya adalah $Q = \hat{\theta} - \theta$.

2. Jika θ adalah parameter skala, maka kuantitas pivotnya adalah Q = $\hat{\theta}/\theta$.

Kuantitas pivot yang didapat dengan cara seperti pada teorema diatas, kemudian sebaiknya dimodifikasi sehingga kita peroleh kuantitas pivot yang memiliki sebaran yang kita ketahui.

Teorema 11. 2.

Misalkan $X_1, \ldots, X_n$ merupakan contoh acak dari suatu sebaran dengan parameter lokasi dan skala yaitu yang fungsi kepekatan peluangnya memiliki bentuk seperti berikut

$$f(x;\theta_1,\theta_2) = \frac{1}{\theta_2} f_0\left(\frac{x-\theta_1}{\theta_2}\right)$$

Jika penduga kemungkinan maksimum $\hat{\theta}_1$ dan $\hat{\theta}_2$ ada, maka $(\hat{\theta}_1-\boldsymbol{\theta_1})/\hat{\theta}_2$ dan $\hat{\theta}_2/\boldsymbol{\theta_2}$ secara berturut-turut adalah kuantitas pivot bagi θ_1 dan θ_2.

Perlu dicatat bahwa $(\hat{\theta}_1-\boldsymbol{\theta_1})/\ \boldsymbol{\theta_2}$ memiliki sebaran yang bebas terhadap parameter yang tak diketahui, tetapi bukan merupakan kuantitas pivot jika $\boldsymbol{\theta_2}$ tak diketahui.

Jika terdapat statistik cukup, maka penduga kemungkinan maksimum dapat merupakan fungsi darinya, dan metode ini akan memberikan hasil yang bagus.

Teladan 11. 4.

Misalkan sebuah contoh acak berasal dari sebaran Normal, $X_i \sim N(\mu,\sigma^2)$ dimana kedua parameter μ dan σ^2 tak diketahui. Jika $\hat{\mu}$ dan $\hat{\sigma}$ berturut-turut adalah penduga kemungkinan maksimum bagi μ dan σ maka $(\hat{\mu}-\mu)/\hat{\sigma}$ dan $\hat{\sigma}/\sigma$ berturut-turut adalah kuantitas pivot bagi μ dan σ, yang dapat digunakan untuk menurunkan selang kepercayaan untuk tiap parameter dimana yang lainnya adalah parameter pengganggu. Akan lebih nyaman mengekspresikan hasil dalam bentuk penduga tak bias $S^2 = n\hat{\sigma}^2/(n-1)$ sehingga kita peroleh hasil seperti berikut

$$\frac{\overline{X}-\mu}{S/\sqrt{n}} \sim t_{(n-1)}$$

dan

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

Seandainya $t_{1-\alpha/2;n-1}$ adalah persentil ke $1-\alpha/2$ dari sebaran t dengan derajat bebas $n-1$, maka

$$\begin{aligned} 1-\alpha &= P\left[-t_{1-\frac{\alpha}{2};n-1} < \frac{\overline{X}-\mu}{S\sqrt{n}} < t_{1-\frac{\alpha}{2};n-1}\right] \\ &= P\left[\overline{X}-t_{1-\frac{\alpha}{2};n-1}\frac{S}{\sqrt{n}} < \mu < \overline{X}+t_{1-\frac{\alpha}{2};n-1}\frac{S}{\sqrt{n}}\right] \end{aligned}$$

yang berarti bahwa interval kepercayaan 100(1-α)% bagi μ adalah

$$\left(\overline{x}-t_{1-\frac{\alpha}{2};n-1}\frac{s}{\sqrt{n}};\overline{x}+t_{1-\frac{\alpha}{2};n-1}\frac{s}{\sqrt{n}}\right)$$

dengan $\overline{x}$ dan s adalah rata-rata dan standar deviasi contoh.

Dengan menggunakan cara yang sama, selang kepercayaan 100(1-α)% bagi σ^2 adalah

$$\left(\frac{(n-1)s^2}{\chi_{1-\frac{\alpha}{2};n-1}};\frac{(n-1)s^2}{\chi_{\frac{\alpha}{2};n-1}}\right)$$

Batas-batas kepercayaan untuk σ dapat diperoleh langsung dari akar pangkat dua dari masing-masing batas diatas.

Secara umum, jika $(\theta_L\ ;\ \theta_U)$ adalah merupakan selang kepercayaan 100γ% bagi parameter θ, dan jika $\tau(\theta)$ merupakan fungsi monoton menaik bagi $\theta \in \Omega$, maka $(\tau(\theta_L)\ ;\ \tau(\theta_U))$ merupakan selang kepercayaan 100γ% bagi parameter $\tau(\theta)$.

Teladan 11. 5.

Meskipun peubah acak *X* memiliki sebaran Weibull, $X_i \sim Wei(\theta,\beta)$. Sebaran ini bukanlah model sebaran dengan parameter lokasi-skala. Namun tidaklah sulit untuk menunjukkan bahwa sebaran $Y_i = \ln X_i$ adalah sebaran nilai ekstrem yang tergolong model sebaran dengan parameter lokasi-skala. Secara khusus

$$f(y;\theta_1,\theta_2) = \frac{1}{\theta_2} f_0\left[\frac{y-\theta_1}{\theta_2}\right]$$

dimana $f_0(z) = exp(z-e^z)$. Hubungannya terhadap parameter adalah $\theta_1 = \ln\theta$ dan $\theta_2 = 1/\beta$, dan dengan demikian

$$Q_1 = \hat{\beta}\ln\left(\frac{\hat{\theta}}{\theta}\right) = \frac{\hat{\theta}_1 - \theta_1}{\hat{\theta}_2}$$

$$Q_2 = \frac{\hat{\beta}}{\beta} = \left(\frac{\hat{\theta}_2}{\theta_2}\right)^{-1}$$

merupakan kuantitas pivot bagi θ dan β. Karena penduga kemungkinan maksimum bagi kedua parameter ini dihitung dengan menggunakan metode iterasi, maka tak diketahui sebaran pasti dari kedua pivot ini, tetapi nilai persentilnya dapat dicari dengan metode simulasi.

Ada kemungkinan bahwa mencari kuantitas pivot berdasarkan penduga kemungkinan maksimum tidak selalu berhasil, namun untuk contoh acak dari sebaran kontinu dengan parameter tunggal yang tak diketahui, sedikitnya ada satu kuantitas pivot yang dapat diturunkan dengan menggunakan transformasi integral peluang (***Probability Integral Transform***).

Jika X_i ~ $f(x;\theta)$ dan jika $F(x;\theta)$ adalah fungsi sebaran kumulatif dari X_i, maka berdasarkan Teorema Transformasi Integral Peluang bahwa $F(X_i;\theta)$ ~ *Seragam(0,1)*. Dan sebagai konsekuensinya $Y_i = -\ln F(X_i;\theta)$ ~ *Eks(1).*
Untuk contoh acak $X_1, \ldots, X_n$ dengan demikian berlaku

$$-2\sum_{i=1}^{n} \ln F(X_i;\theta) \sim \chi^2_{2n}$$

sehingga

$$P\left[\chi^2_{\frac{\alpha}{2};2n} < -2\sum_{i=1}^{n} \ln F(X_i;\theta) < \chi^2_{1-\frac{\alpha}{2};2n}\right] = 1-\alpha$$

dan dengan mengatur pernyataan matematis diatas dapat diperoleh wilayah kepercayaan bagi parameter θ. Jika fungsi sebaran kumulatifnya tidak dalam bentuk sederhana, inversinya dapat dilakukan dengan cara numerik. Jika $F(x;\theta)$ merupakan fungsi monoton menaik (menurun) dari θ, maka hasil dari wilayah kepercayaannya akan berupa interval. Juga perlu diperhatikan bahwa 1- $F(X_i;\theta)$ ~ *Seragam(0,1)* dan

$$-2\sum_{i=1}^{n} \ln\left[1 - F(X_i;\theta)\right] \sim \chi^2_{2n}$$

Secara umum, ekspresi transformasi yang digunakan akan menghasilkan interval yang berbeda, dan atas pertimbangan kenyamanan perhitunganlah sebagai kriteria yang beralasan dalam pemilihan transformasi mana yang dipakai.

Teladan 11. 6.

Suatu contoh acak diambil dari sebaran Pareto, X_i ~ *Par(1,*κ*)*. Fungsi sebaran kumulatifnya adalah

$$F(x;\kappa) = 1-(1+x)^{-\kappa} \quad x > 0$$

Jika kita gunakan transformasi *1- F(Xi;κ)*, maka – *ln* [*1 – F(x;κ)*] = κ *ln(1+x)*, sehingga

$$2\kappa\sum_{i=1}^{n}\ln(1+X_i)\ \sim\ \chi^2_{2n}$$

dan Selang Kepercayaan 100(1-α)% bagi κ memiliki bentuk

$$\left(\frac{\chi^2_{\frac{\alpha}{2};2n}}{2\sum_{i=1}^{n}\ln(1+x_i)};\frac{\chi^2_{1-\frac{\alpha}{2};2n}}{2\sum_{i=1}^{n}\ln(1+x_i)}\right)$$

Apabila digunakan trnasformasi *F(Xi;κ)* jawabannya akan lebih sulit karena pertidaksamaan mengharuskan jawaban diselesaikan secara numerik.

Untuk sebaran-sebaran diskret, dan beberapa permasalahan parameter ganda, kuantitas pivot mungkin tidak ada. Namun demikian, pendekatan kuantitas pivot kadang dapat diperoleh berdasarkan hasil asimtotis. Pendekatan normal terhadap sebaran binomial, merupakan suatu teladan.

Pendekatan Interval Kepercayaan

Misalkan $X_1, \ldots, X_n$ merupakan contoh acak dari suatu distribusi dengan fungsi kepekatan peluang $f(x;\theta)$. Sebagaimana telah kita ketahui dalam beberapa bab terdahulu bahwa penduga kemungkinan maksimum memiliki sifat mendekati normal atau asimtotik normal pada kondisi tertentu.

Teladan 11. 7.

Perhatikan adanya suatu contoh acak dari sebaran Bernoulli, $X_i \sim Bin(1,p)$. Penduga kemungkinan maksimum bagi *p* adalah

$\hat{p} = \dfrac{\sum_{i=1}^{n} X_i}{n}$. Kita juga tahu bahwa $\hat{p}$ merupakan statistik cukup dan $\sum_{i=1}^{n} X_i$ ~ *Bin(n,p)* namun tidak ada kuantitas pivot untuk *p.* Namun demikian, dengan menggunakan Dalil Limit Pusat,

$$\frac{\hat{p} - p}{\sqrt{p(1-p)/n}} \xrightarrow{d} Z \sim N(0,1)$$

dan sebagai konsekuensinya, untuk *n* yang besar,

$$P\left[-z_{1-\frac{\alpha}{2}} < \frac{\hat{p} - p}{\sqrt{p(1-p)/n}} < z_{1-\frac{\alpha}{2}}\right] \cong 1 - \alpha$$

Dengan menggunakan hasil limit sebaran, kita juga dapat peroleh bahwa

$$P\left[-z_{1-\frac{\alpha}{2}} < \frac{\hat{p} - p}{\sqrt{\hat{p}(1-\hat{p})/n}} < z_{1-\frac{\alpha}{2}}\right] \cong 1 - \alpha$$

sehingga pernyataan matematis ini lebih mudah untuk diinvers sehingga menghasilkan pendekatan batas-batas kepercayaan bagi *p* yaitu

$$\hat{p} \pm z_{1-\frac{\alpha}{2}} \sqrt{\hat{p}(1-\hat{p})/n}$$

Teladan 11. 8.

Contoh acak berukuran *n* diambil dari sebaran Poisson, $X_i \sim Poi(\mu)$. Dengan menggunakan Teorema Limit Pusat, kita tahu bahwa

$$\frac{\overline{X} - \mu}{\sqrt{\mu/n}} \xrightarrow{d} Z \sim N(0,1)$$

dan dengan menggunakan limit sebaran, untuk *n* yang sangat besar,

$$\frac{\overline{X} - \mu}{\sqrt{\overline{X} / n}} \xrightarrow{d} Z \sim N(0,1)$$

Kedua peubah acak diatas dapat digunakan untuk mencari pendekatan interval kepercayaan, namun ekspresi yang terakhir lebih mudah digunakan. Sebenarnya, dimungkinkan untuk membuat generalisasi pendekatan ini bilamana penduga kemungkinan maksimum memiliki sifat asimtotik normal.

Metode Umum

Jika kuantitas pivot tidak tersedia, masih dimungkinkan untuk menentukan wilayah kepercayaan suatu parameter θ, jika terdapat suatu statistik yang sebarannya tergantung θ namun tidak tergantung pada parameter pengganggu lainnya. Khususnya, misalkan $X_1, \ldots X_n$ memiliki fungsi kepekatan peluang bersama $f(x_1, \ldots, x_n; \theta)$, dan $S = s(X_1, \ldots, X_n) \sim g(s;\theta)$. Lebih baik apabila S merupakan statistik cukup bagi θ, atau mungkin beberapa penduga beralasan seperti penduga kemungkinan maksimum, tapi ini tidak diperlukan.

Untuk setiap kemungkinan nilai θ, asumsikan bahwa kita dapat memperoleh nilai $h_1(\theta)$ dan $h_2(\theta)$ sedemikian rupa sehingga

$$\boldsymbol{P[h_1(\theta) < S < h_2(\theta)] = 1-\alpha}$$

Jika kita amati nilai $S = s$, maka gugus nilai $\theta \in \Omega$ yang memenuhi $h_1(\theta) < s < h_2(\theta)$ membentuk wilayah kepercayaan 100(1-α)% bagi θ. Dengan perkataan lain bahwa, jika θ_0 merupakan nilai θ yang sebenarnya, maka θ_0 akan berada dalam wilayah kepercayaan jika dan hanya jika $h_1(\theta_0) < s < h_2(\theta_0)$ yang memiliki taraf kepercayaan 100(1-α)% bagi θ karena $P[h_1(\theta) < S < h_2(\theta)] = 1-\alpha$ benar bilamana $\theta = \theta_0$. Seringkali $h_1(\theta)$ dan $h_2(\theta)$ merupakan fungsi menaik (menurun) terhadap θ, yang mengakibatkan wilayah kemungkinannya berupa interval.

Teladan 11. 9.

Sebuah contoh acak berukuran *n* dari sebaran kontinu dengan fungsi kepekatan peluang

$$f(x;\theta) = \frac{1}{\theta^2}\exp\left[\frac{-(x-\theta)}{\theta^2}\right] I_{[\theta,\infty)}(x)$$

dimana $\theta > 0$. Tak ada statistik cukup tunggal, namun $X_{(1)}$ dan $\sum_{i=1}^{n} X_i$ merupakan statistik cukup bersama bagi θ. Kita ingin menurunkan selang kepercayaan 90% bagi θ berdasarkan statistik $S = X_{(1)}$. Fungsi sebaran kumulatif dari S adalah

$$G(s;\theta) = \left(1 - \exp\left[-n(s-\theta)/\theta^2\right]\right) I_{[\theta,\infty)}(s)$$

Salah satu pilihan fungsi $h_1(\theta)$ dan $h_2(\theta)$ yang memenuhi $P[h_1(\theta) < S < h_2(\theta)] = 1-\alpha$ adalah

$$G(h_1(\theta);\theta) = 0{,}05 \text{ dan } G(h_2(\theta);\theta) = 0{,}95$$

yang akan menghasilkan fungsi

$$h_1(\theta) = \theta - \ln(0{,}95)\ \theta^2/n \cong \theta + 0{,}0513\ \theta^2/n$$

dan

$$h_2(\theta) = \theta - \ln(0{,}05)\ \theta^2/n \cong \theta + 0{,}0513\ \theta^2/n$$

Latihan

1. Suatu contoh acak berukuran *n* dari sebaran Eksponensial dengan parameter θ.
   a. Jika nilai rata-rata contoh 17.9 dari $n = 50$, carilah batas bawah kepercayaan 95% bagi θ.
   b. Carilah batas bawah kepercayaan 95% bagi $P(X > t) = e^{-t/\theta}$ dimana *t* adalah sembarang nilai yang diketahui ?
2. Misalkan $X_1, X_2, ..., X_n$ adalah contoh acak berukuran *n* dari sebaran Weibull, $X_i \sim Weibull(\theta,2)$

a. Tunjukkan bahwa $Q = 2\sum_{i=1}^{n} X_i^2 / \theta^2 \sim \chi^2_{2n}$
b. Gunakan Q untuk menurunkan interval kepercayaan (selang kepercayaan) 100γ% dua arah sama ekor untuk θ.
c. Carilah batas bawah interval kepercayaan 100γ% untuk $P(X > t) = \exp[-(t/\theta)^2]$
d. Carilah batas atas interval kepercayaan 100γ% bagi persentil ke-p dari sebaran ini ?

3. Contoh acak berukuran *n* dari sebaran Seragam, $X_i \sim Seragam(0,\theta)$, $\theta > 0$, dan $X_{(n)}$ adalah statistik maksimum.
   a. Carilah peluang bahwa interval acak $(X_{(n)}, 2X_{(n)})$ mengandung θ.
   b. Carilah konstanta *c* sedemikian rupa sehingga $(x_{(n)}, cx_{(n)})$ adalah interval kepecayaan 100(1-α)% bagi θ.
4. Misalkan $X_1, X_2, ..., X_n$ adalah contoh acak berukuran *n* dari sebaran Weibull, $X_i \sim Weibull(\theta,\beta)$, β diketahui.
   a. Gunakan Metode Umum untuk menurunkan interval kepercayaan 100(1-α)% bagi θ berdasarkan statistik $S_1 = X_{(1)}$.
   b. Gunakan Metode Umum untuk menurunkan interval kepercayaan 100(1-α)% bagi θ berdasarkan statistik $S_2 = \sum X_i^\beta$.

Pengujian Hipotesis

Pendahuluan

Dalam kegiatan ilmiah, perhatian tercurahkan pada bagaimana menjawab pertanyaan mengenai validitas teori atau hipotesis yang berkenaan dengan fenomena fisik. Pada umumnya, informasi mengenai fenomena ini dapat diperoleh dengan melakukan percobaan. Terminologi **pengujian hipotesis** mengacu pada proses mencoba memutuskan berdasarkan bukti percobaan akan kebenaran atau ketidakbenaran hipotesis.

Sebagai misal, kita mencurigai hipotesis tertentu, barangkali suatu teori yang sudah diterima, adalah salah, dan kemudian suatu percobaan dilakukan. Luaran dari percobaan yang tak konsisten tentu akan meragukan validitasnya. Secara umum, pengukuran percobaan mengandung kesalahan acak, dan dengan demikian setiap keputusan tentang kebenaran atau ketidakbenaran suatu hipotesis, yang berdasarkan hasil percobaan, juga mengandung kesalahan. Tak akan mungkin menghindari kesalahan keputusan, namun dimungkinkan untuk membangun uji-uji sedemikian rupa sehingga kesalahan sangat jarang terjadi dan pada taraf yang ditentukan, bila ada.

Definisi 12. 1. Hipotesis Statistika.

Jika $X_i \sim f(x;\boldsymbol{\theta})$, suatu hipotesis statistika adalah suatu pernyataan tentang sebaran dari X. Jika hipotesis secara lengkap menyebut $f(x;\boldsymbol{\theta})$, maka hipotesis dikatakan **sederhana**; jika tidak hipotesis dikatakan **majemuk**.

Seringkali sebaran yang ditanyakan memiliki bentuk parametrik dengan parameter tunggal yang tak diketahui θ, dan hipotesis mencakup pertanyaan mengenai

θ. Dalam kerangka ini, hipotesis statistika berkenaan dengan suatu anak gugus ruang parameter, dan tujuan pengujian adalah untuk memutuskan apakah nilai parameter yang sebenarnya berada dalam anak gugus ini. Dengan demikian, hipotesis nol berkenaan dengan anak gugus $\Omega_0 \subset \Omega$, dan hipotesis tandingan berkenaan dengan komplemennya, Ω-Ω_0. Dalam kasus hipotesis sederhana, gugus ini hanya memiliki satu anggota $\Omega_0 = \{\theta_0\}$ dan Ω-$\Omega_0 = \{\theta_1\}$, dimana $\theta_0 \neq \theta_1$.

Dalam banyak percobaan terdapat beberapa hipotesis penelitian dimana kita berharap mendukung bukti secara statistik, dan hipotesis ini harus diletakkan sebagai hipotesis tandingan. Berdasarkan data hasil penelitian, apakah data mendukung pernyataan hipotesis nol atau hipotesis tandingan. Secara filosofi kita bagi ruang contoh menjadi dua daerah, yaitu **daerah kritis** atau **daerah penolakan** hipotesis nol, dan **daerah penerimaan** hipotesis nol.

Definisi 12. 2. Daerah Kritis.

Daerah kritis untuk suatu uji hipotesis adalah anak gugus dari ruang contoh yang berkenaan dengan penolakan hipotesis nol.

Dalam teladan kita, $\overline{X}$ merupakan statistik cukup bagi μ, sehingga kita dapat mengekspresikan daerah kritis secara langsung menggunakan peubah tunggal $\overline{X}$, dan kita akan gunakan $\overline{X}$ sebagai **statistik uji**. Karena $\mu_1 > \mu_0$, bentuk alami wilayah kritis permasalahan ini adalah dengan memisalkan $C = \{(x_1, \ldots, x_n) | \bar{x} \geq c\}$, untuk beberapa nilai c yang sesuai. Yang dimaksudkan dengan itu adalah, kita akan tolak hipotesis nol, jika $\bar{x} \geq c$ dan kita tidak akan menolak hipotesis nol jika $\bar{x} < c$. Terdapat dua kemungkinan kesalahan yang mungkin kita buat dengan prosedur ini. Kemungkinan kita akan menolak H_0 apabila H_0 benar, atau kemungkinan kita akan gagal menolak H_0 bilamana H_0 salah. Kesalahan-kesalahan tersebut dinotasikan sebagai berikut:

1. Kesalahan Tipe I : Menolak Hipotesis Nol yang benar.
2. Kesalahan Tipe II : Gagal Menolak Hipotesis Nol yang salah.

Gagal memiliki cukup bukti statistik untuk menolak Hipotesis Nol tidak sama artinya dengan memiliki cukup kuat bukti untuk mendukung Hipotesis Nol.

Kita berharap dapat memilih uji statistik dan daerah kritis sehingga peluang kita membuat kesalahan ini kecil. Notasi ringkas untuk menyatakan peluang diatas adalah:

1. P[Kesalahan Tipe I] = P(Tipe I) = α
2. P[Kesalahan Tipe II] = P(Tipe II) = β.

Definisi 12. 3. Taraf Nyata dan Ukuran Uji.

Untuk hipotesis nol sederhana, H_0, peluang menolak H_0 yang benar, α = *P(Tipe I)* sering disebut dengan **taraf nyata** pengujian. Untuk hipotesis nol majemuk, H_0, **ukuran** pengujian (atau ukuran daerah kritis) merupakan peluang maksimum penolakan H_0 bilamana H_0 benar.

Khusus untuk hipotesis nol sederhana, taraf nyata pengujian juga merupakan ukuran pengujian.

Pendekatan baku menggunakan taraf kesalahan yang masih dapat diterima, seperti $\alpha = 0,05$ atau $\alpha = 0,01$ sebagai taraf nyata pengujian. Kemudian tentukan daerah kritis tersebut dengan memperhatikan taraf nyata yang telah ditetapkan. Diantara semua daerah kritis berukuran α tersebut kita pilih satu titik c yang memiliki P{Tipe II) paling kecil. Jika misalkan $n = 25$, $\mu_0 = 10$, $\sigma = 4$, maka $\alpha = 0,05$ akan memberikan nilai

$$c = \mu_0 + z_{1-\alpha}\frac{\sigma}{\sqrt{n}} = 10 + 1,645\frac{4}{5} = 11,316$$

Hal ini mudah diverifikasi, karena

$$P[\overline{X} \geq c \mid \mu = \mu_0 = 10] \quad = P\left[\frac{\overline{X} - \mu_0}{\sigma / \sqrt{n}} \geq \frac{c - \mu_0}{\sigma / \sqrt{n}}\right]$$

$$= P\left[Z \geq \frac{11{,}316 - 10}{4 / \sqrt{25}}\right]$$

$$= P[Z \geq 1{,}645]$$

$$= 0{,}05$$

Dengan demikian, pengujian dengan menggunakan taraf uji 0,05 untuk $H_0 : \mu = 10$ lawan $H_1 : \mu = 11$ adalah Menolak hipotesis nol jika nilai amatan $\bar{x} \geq 11{,}316$. Daerah kritis ini memberikan ukuran uji 0,05 untuk setiap tandingan $\mu = \mu_1$, tapi fakta bahwa $\mu_1 > \mu_0$ berarti kita akan mendapatkan nilai P(Tipe II) semakin kecil dengan mengambil daerah kritis disisi kanan sebaran $\overline{X}$

Peluang kesalahan tipe II yang berkenaan dengan permasalahan ini dapat dihitung sebagai berikut :

$$\beta = P(\text{Tipe II}) \quad = P[\overline{X} < 11{,}316 \mid \mu = \mu_1 = 11]$$

$$= P\left[\frac{\overline{X} - 11}{4/5} < \frac{11{,}316 - 11}{4/5} \mid \mu = 11\right]$$

$$= P[Z < 0{,}35] = 0{,}654$$

Sejauh ini, tak ada alasan secara teori untuk memilih suatu daerah kritis dari pada daerah kritis yang lainnya. Sebagai misal, daerah kritis $C_1 = \{(x_1, \ldots, x_n) \mid 10 < \bar{x} < 10{,}0006\}$ juga memiliki $\alpha = 0{,}05$ karena

$$P[10 < \overline{X} < 10{,}0006] = P\left[0 < \frac{\overline{X} - 10}{4/5} < 0{,}1257\right] = 0{,}05$$

akan tetapi

$$P(\text{Tipe II}) \quad = 1 - P[10 < \overline{X} < 10{,}0006 \mid \mu = 11]$$

$$= 1 - P\left[\frac{10 - 11}{4/5} < Z < \frac{10{,}0006 - 11}{4/5}\right]$$

$$= 1 - P[-1{,}25 < Z < -1{,}12425]$$

= 0,9752

Secara lebih umum, kita dapat menguji H_0: $\mu = \mu_0$ lawan H_1: $\mu = \mu_1$ (dimana $\mu_1 > \mu_0$) pada taraf nyata α, dengan menggunakan statistik uji

$$Z_0 = \frac{\overline{X} - \mu_0}{\sigma / \sqrt{n}}$$

setara dengan penggunaan statistik $\overline{X}$, sehingga kita lebih nyaman menggunakan uji dengan menolak H_0 jika $z_0 \geq z_{1-\alpha}$, dimana z_0 dihitung dengan menggunakan nilai Z_0. Jelas bahwa, pada kondisi H_0, $P[Z_0 \geq z_{1-\alpha}] = \alpha$ dan kita memiliki daerah kritis sebesar α. Peluang tipe kesalahan II nya dapat dihitung dengan menggunakan

$$\Phi\left(z_{1-\alpha} + \frac{\mu_0 - \mu_1}{\sigma / \sqrt{n}} \right)$$

dimana $\Phi(x)$ adalah fungsi sebaran kumulatif Normal Baku.
Ukuran contoh n yang membuat P(Tipe II) = β merupakan jawaban dari

$$z_{1-\alpha} + \frac{\mu_0 - \mu_1}{\sigma / \sqrt{n}} = z_\beta = -z_{1-\beta}$$

adalah

$$n = \frac{\left(z_{1-\alpha} + z_{1-\beta}\right)^2 \sigma^2}{\left(\mu_0 - \mu_1\right)^2}$$

Definisi 12. 4. Fungsi Kuasa.

Fungsi kuasa, $\pi(\theta)$, dari suatu uji H_0 adalah peluang menolak H_0 bilamana nilai sebenarnya dari parameter adalah θ.

Untuk hipotesis sederhana H_0: $\theta = \theta_0$ lawan H_1: $\theta = \theta_1$, kita peroleh $\pi(\theta_0) = P(Tipe\ I) = \alpha$ dan $\pi(\theta_1) = 1 - P(Tipe\ II) = 1 - \beta$. Untuk hipotesis majemuk, misalkan H_0: $\theta \in \Omega_0$ lawan H_1: $\theta \in \Omega - \Omega_0$, maka ukuran ujinya (atau daerah kritisnya) adalah

$$\alpha = \max_{\theta \in \Omega} \pi(\theta)$$

dan jika nilai θ yang sebenarnya berada di dalam Ω-Ω_0, maka $\pi(\theta)$ = *1-P(Tipe II)*, dengan demikian P(Tipe II) tergantung pada θ.

Hipotesis Majemuk

Bila kita asumsikan bahwa $X \sim N(\mu, \sigma^2)$, dimana σ^2 diketahui, dan kita ingin melakukan pengujian H_0: $\mu = \mu_0$ lawan hipotesis tandingan majemuk H_1: $\mu > \mu_0$. Telah disarankan pada teladan terdahulu bahwa daerah kritis berada di sebelah ekor kanan untuk $\mu_1 > \mu_0$, tetapi nilai kritis *c* tidak tergantung pada besarnya μ_1. Dengan demikian jelas bahwa pengujian hipotesis sederhana yang telah kita punya pada teladan terdahulu juga dapat digunakan untuk hipotesis tandingan majemuk. Pengujian pada taraf nyata α masih menolak H_0 jika

$$z_0 = \frac{\bar{x} - \mu_0}{\sigma / \sqrt{n}} \geq z_{1-\alpha}$$

Kuasa uji pada sembarang nilai μ nya dapat dituliskan sebagai berikut

$$\pi(\mu) = P\left[\frac{\bar{X} - \mu_0}{\sigma / \sqrt{n}} \geq z_{1-\alpha} \mid \mu\right] = P\left[\frac{\bar{X} - \mu}{\sigma / \sqrt{n}} \geq z_{1-\alpha} + \frac{\mu_0 - \mu}{\sigma / \sqrt{n}} \mid \mu\right]$$

atau

$$\pi(\mu) = 1 - \Phi\left(z_{1-\alpha} + \frac{\mu_0 - \mu}{\sigma / \sqrt{n}}\right)$$

Kita dapat juga melakukan pengujian hipotesis nol majemuk H_0: $\mu \leq \mu_0$ lawan hipotesis tandingan majemuk H_1: $\mu > \mu_0$, dan kita tolak H_0 jika $z_0 = \frac{\bar{x} - \mu_0}{\sigma / \sqrt{n}} \geq z_{1-\alpha}$. Pengujian hipotesis nol majemuk ini masih menggunakan taraf uji sebesar α. Peluang menolak H_0 untuk sembarang $\mu \leq \mu_0$ adalah sebesar $\pi(\mu)$, dan $\pi(\mu) \leq \pi(\mu_0)$= α untuk $\mu \leq \mu_0$, dan dengan demikian $\max_{\mu \leq \mu_0} \pi(\mu) = \alpha$ Dengan demikian, jika daerah kritis dipilih memiliki luas α pada $\mu 0$, maka besarnya kesalahan Tipe I akan lebih kecil dari α untuk setiap μ < μ_0, sehingga daerah kritis aslinya masih berlaku. Oleh karenanya, pengujian pada taraf α yang dikembangkan untuk hipotesis nol sederhana kadang masih dapat dipakai pada hipotesis majemuk yang lebih realistik, dan peluang Tipe I tak akan melebihi α.

Catatan : gagal menolak hipotesis nol tidak harus selalu diinterpretasikan dengan penerimaan hipotesis alternatif. Kita dapat selalu mencari nilai tandingan μ yang cukup dekat dengan μ_0 sehingga kuasa uji, $\pi(\mu)$ mendekati α.

Ukuran amatan atau **nilai peluang** yang sering dinotasikan dengan **nilai-p** didefinisikan sebagai ukuran α terkecil dimana H_0 dapat ditolak, berdasarkan nilai amatan statistik uji.

Uji Paling Kuasa

Dalam beberapa bab terdahulu, terminologi pengujian hipotesis telah dikembangkan, dan beberapa uji secara intuitif telah dikembangkan berdasarkan kuantitas pivot atau statistik cukup yang sesuai. Berikut ini akan disajikan konsep uji paling kuasa.

Misalkan $X_1, \ldots, X_n$ memiliki fungsi kepekatan peluang $f(x_1, \ldots, x_n;\theta)$, dan daerah kritis C. Notasi untuk fungsi kuasa yang berkaitan dengan C adalah

$$\pi_C(\theta) = P[(X_1, \ldots, X_n) \in C \mid \theta]$$

Definisi 12. 5. Uji Paling Kuasa.

Suatu uji H_0: $\theta = \theta_0$ lawan H_1: $\theta = \theta_1$ berdasarkan daerah kritis C^* dikatakan sebagai **uji paling kuasa** pada taraf nyata α jika

1. $\pi_{C^*}(\theta_0) = \alpha$, dan
2. $\pi_{C^*}(\theta_1) \geq \pi_C(\theta_1)$ untuk setiap daerah kritis C pada taraf nyata α

Daerah kritis C^* disebut dengan **daerah kritis paling kuasa** pada taraf uji α.

Teorema 12. 1. Lemma Neyman-Pearson

Misalkan bahwa $X_1, \ldots, X_n$ memiliki fungsi kepekatan peluang bersama $f(x_1, \ldots, x_n; \theta)$. Misalkan

$$\lambda(x_1, \ldots, x_n; \theta_0, \theta_1) = \frac{f(x_1, \ldots, x_n; \theta_0)}{f(x_1, \ldots, x_n; \theta_1)}$$

dan misalkan C^* merupakan gugus

$$C^* = \{(x_1, \ldots, x_n) \mid \lambda(x_1, \ldots, x_n; \theta_0, \theta_1) \leq k\}$$

dimana k adalah konstanta sedemikian rupa sehingga

$$P[(X_1, \ldots, X_n) \in C^* \mid \theta_0] = \alpha$$

Dengan demikian, C^* merupakan daerah kritis paling kuasa untuk pengujian H_0: $\theta = \theta_0$ lawan H_1: $\theta = \theta_1$ pada taraf nyata pengujian α.

Teladan 12. 1.

Sebuah contoh acak berukuran n dari sebaran Eksponensial, $X_i \sim Eks(\theta)$. Kita ingin melakukan pengujian H_0: $\theta = \theta_0$ lawan H_1: $\theta = \theta_1$ pada taraf nyata pengujian α. [$\theta_1 > \theta_0$]. Lemma Neyman-Pearson mengatakan bahwa: Tolak H_0 jika

$$\lambda(x_1,...,x_n;\theta_0,\theta_1)=\frac{f(x_1,...,x_n;\theta_0)}{f(x_1,...,x_n;\theta_1)}=\frac{\theta_0^{-n}\exp\left(-\sum_{i=1}^{n}\frac{x_i}{\theta_0}\right)}{\theta_1^{-n}\exp\left(-\sum_{i=1}^{n}\frac{x_i}{\theta_1}\right)}\leq k$$

dimana *h* dipilih sedemikian rupa sehingga $P[\lambda(\mathbf{X};\theta_0,\theta_1)\leq k]=\alpha$ bilamana $\theta=\theta_0$.
Sekarang kita punya

$$P[\lambda(X_1,...,X_n;\theta_0,\theta_1)\leq k\mid\theta_0]=P\left[\sum_{i=1}^{n}X_i\left(\frac{1}{\theta_1}-\frac{1}{\theta_0}\right)\leq\ln\left(\left(\frac{\theta_0}{\theta_1}\right)^n k\right)\Bigg|\theta_0\right]$$

sehingga

$$P[\mathbf{X}\in C^*|\theta_0]=P\left[\sum_{i=1}^{n}X_i\geq k^*\mid\theta_0\right]$$

dimana $k^*=\ln\left(\left(\frac{\theta_0}{\theta_1}\right)^n k\right)/\left(\frac{1}{\theta_1}-\frac{1}{\theta_0}\right)$. Dengan demikian, daerah kritis paling kuasa untuk uji tersebut memiliki bentuk $C^*=\{(x_1,...,x_n)\mid\sum_{i=1}^{n}x_i\geq k^*\}$. Perlu kita ingat bahwa berdasarkan H_0: $\theta=\theta_0$, peubah acak $\frac{2\sum_{i=1}^{n}X_i}{\theta_0}$ memiliki sebaran χ^2_{2n}, sehingga pemilihan $k^*=\theta_0\frac{\chi^2_{1-\alpha;2n}}{2}$ akan memberikan daerah kritis pada taraf nyata pengujian α.. Dan pengujian yang ekuivalen dengan itu adalah dengan kriteria: Menolak hipotesis nol jika $\frac{2\sum_{i=1}^{n}x_i}{\theta_0}\geq\chi^2_{1-\alpha;2n}$.

Teladan 12. 2.

Sebuah contoh acak berukuran *n* berasal dari sebaran Normal dengan rata-rata nol dan varian σ^2, $X_i \sim N(0, \sigma^2)$. Kita ingin melakukan pengujian $H_0 : \sigma^2 = \sigma_0^2$ lawan $H_1 : \sigma^2 = \sigma_1^2$ dimana $\sigma_1 > \sigma_0$. Dalam kasus ini, kita punya

$$\lambda(x_1,...,x_n;\sigma_0^2,\sigma_1^2) = \frac{\dfrac{1}{\left(\sqrt{2\pi}\sigma_0\right)^n}\exp\left[-\dfrac{\sum_{i=1}^{n} x_i^2}{2\sigma_0^2}\right]}{\dfrac{1}{\left(\sqrt{2\pi}\sigma_1\right)^n}\exp\left[-\dfrac{\sum_{i=1}^{n} x_i^2}{2\sigma_1^2}\right]} \le k$$

yang setara dengan $\left(\dfrac{1}{\sigma_1^2} - \dfrac{1}{\sigma_0^2}\right)\sum_{i=1}^{n} x_i^2 \le 2\ln\left[k^*\left(\dfrac{\sigma_0}{\sigma_1}\right)^n\right]$. Karena $\sigma_1 > \sigma_0$. maka $\left(\dfrac{1}{\sigma_1^2} - \dfrac{1}{\sigma_0^2}\right) < 0$, dan daerah kritis paling kuasanya memiliki bentuk $C^* = \{(x_1,...,x_n) \mid \sum_{i=1}^{n} x_i^2 \ge k^{**}\}$. Perhatikan juga bahwa berdasarkan H_0 atau apabila H_0 benar, maka $\sum_{i=1}^{n} X_i^2 / \sigma_0^2 \sim \chi_n^2$. Dengan demikian uji paling kuasa dengan taraf nyata α adalah dengan kriteria Menolak H_0 jika $\sum_{i=1}^{n} x_i^2 / \sigma_0^2 \ge \chi_{1-\alpha;n}^2$.

Teladan 12. 3.

Kita ingin menentukan bentuk uji paling kuasa H_0: $p = p_0$ lawan H_1: $p = p_1 > p_0$ berdasarkan statistik $S \sim Bin(n,p)$. Dengan demikian kita punya

$$\lambda = \frac{C_s^n p_0^s (1-p_0)^{n-s}}{C_s^n p_1^s (1-p_1)^{n-s}} \le k$$

sehingga

$$\left\{ \frac{p_0(1-p_1)}{p_1(1-p_0)} \right\}^s \le k_1$$

atau setara dengan

$$s \ln\left(\frac{p_0(1-p_1)}{p_1(1-p_0)} \right) \le \ln k_1$$

Karena $\left(\frac{p_0(1-p_1)}{p_1(1-p_0)} \right) \le 1$ maka nilai ln nya negatif, dan dengan demikian, kriteria pengujiannya adalah Tolak H_0 jika $S \ge k_2$.

Kita perlu catat bahwa Lemma Neyman-Pearson berlaku untuk pengujian sembarang hipotesis nol yang secara lengkap terspesifikasi, H_0: $f_0(x;\theta_0)$, lawan sembarang hipotesis alternatif yang secara lengkap juga terspesifikasi, H_1: $f_1(x;\theta_1)$.

Dalam kebanyakan aplikasi ***x*** sebagai hasil dari suatu contoh acak dari fungsi kepekatan peluang yang kemungkinan dari nilai parameter yang berbeda, tetapi ***x*** juga dapat berasal dari gugus statistik tataan atau peubah ganda. Fungsi kepekatan peluang pada hipotesis nol juga tak perlu sama dengan fungsi kepekatan peluang pada hipotesis tandingan.

Teladan 12. 4.

Misalkan kita punya sebuah contoh acak berukuran *n*, dan ingin menguji H_0: *X ~ Seragam(0,1)* lawan H_1: *X ~ Eksponensial(1).* Dengan demikian kita punya

$$\lambda(x_1,...,x_n) = \frac{f_0(x_1,...,x_n)}{f_1(x_1,...,x_n)} = \frac{1}{e^{-\sum_{i=1}^{n} x_i}} \leq k$$

sehingga kita dapatkan kriteria penolakan : Tolak H_0 jika $\sum_{i=1}^{n} x_i \leq k_1 = \ln k$. Sebaran dari jumlah peubah acak Seragam tidak mudah diekspresikan, namun teorema limit pusat dapat digunakan untuk mendapatkan pendekatan nilai kritis. Karena *X ~ Ser(0,1)*, maka *E(X) = ½* dan *Var(X) = 1/12.* Sehingga

$$z_n = \frac{\overline{X} - 0{,}5}{\sqrt{1/(12n)}} \xrightarrow{d} N(0,1)$$

Sehingga, pendekatan uji paling kuasa pada taraf α nya adalah Tolak *H0* jika

$$z_n = \sqrt{12n}(\overline{x} - 0{,}5) \leq -z_{1-\alpha}$$

Uji Paling Kuasa Seragam

Definisi 12. 6. Uji Paling Kuasa Seragam.

Misalkan $X_1, \ldots, X_n$ memiliki fungsi kepekatan peluang bersama $f(x_1, \ldots, x_n;\theta)$ untuk $\theta \in \Omega$, dan pertimbangkan hipotesis dalam bentuk H_0: $\theta \in \Omega_0$ lawan H_1: $\theta \in \Omega - \Omega_0$, dimana Ω_0 merupakan anak gugus dari Ω. Suatu daerah kritis *C**, dan ujinya, dikatakan paling kuasa seragam pada taraf α jika

$$\max_{\theta \in \Omega_0} \pi_{C*}(\theta) = \alpha$$

dan

$$\pi_{C*}(\theta) \geq \pi_C(\theta)$$

untuk semua $\theta \in \Omega - \Omega_0$ dan semua daerah kritis C pada taraf nyata α.

Teladan 12. 5.

Sebuah contoh acak berukuran n dari sebaran Eksponensial, $X_i \sim Eks(\theta)$. Kita ingin melakukan pengujian H_0: $\theta = \theta_0$ lawan H_1: $\theta = \theta_1$ pada taraf nyata pengujian α., menghasilkan kriteria penolakan :

Menolak hipotesis nol jika $\dfrac{2\sum_{i=1}^{n} x_i}{\theta_0} \geq \chi^2_{1-\alpha;2n}$.

Kriteria ini tidak tergantung pada pemilihan nilai θ_1 yang dipakai, namun hanya berdasarkan kriteria bahwa $\theta_1 > \theta_0$. Dengan demikian uji ini merupakan uji paling kuasa seragam untuk pengujian H_0: $\theta = \theta_0$ lawan H_1: $\theta > \theta_0$.

Fungsi kuasa dari uji ini juga dapar diekspresikan dalam fungsi sebaran kumulatif kai-kuadrat dengan derajat bebas 2n, $H(c;\nu)$.

$$\pi(\theta) = 1 - H\left[\frac{\theta_0}{\theta}\chi^2_{1-\alpha;2n};2n\right]$$

Karena $\pi(\theta)$ merupakan fungsi menaik dari θ, maka $\max_{\theta \leq \theta_0} \pi(\theta) = \pi(\theta_0) = \alpha$, maka uji ini juga uji paling kuasa seragam untuk hipotesis majemuk H_0: $\theta \leq \theta_0$ lawan H_1: $\theta > \theta_0$.

Definisi 12. 7.

Suatu fungsi kepekatan peluang bersama $f(\mathbf{x};\theta)$ dikatakan memiliki rasio kemungkinan monoton *(monotone likelihood ratio)* dalam statistik $T=t(\mathbf{X})$ jika untuk sembarang dua nilai parameter $\theta_1<\theta_2$, rasio $f(\mathbf{x};\theta_2)/ f(\mathbf{x};\theta_1)$ tergantung pada $\mathbf{x}$ hanya melalui fungsi $t(\mathbf{x})$, dan rasio ini merupakan suatu fungsi tak menurun dari $t(\mathbf{x})$.

Teladan 12. 6.

Suatu contoh acak berukuran *n* dari sebaran eksponensial, *Xi~Eksponensial(θ).* Karena $f(\underline{x};\theta) = \frac{1}{\theta^n}\exp(-\sum x_i/\theta)$, maka $\frac{f(\underline{x};\theta_2)}{f(\underline{x};\theta_1)} = (\theta_1/\theta_2)^2 \exp\left[-\left(\frac{1}{\theta_2}-\frac{1}{\theta_1}\right)\sum x_i\right]$ merupakan fungsi tak menurun dari $t(\underline{x}) = \sum x_i$ jika θ2 > θ1. Dengan demikian $t(\underline{x}) = \sum x_i$ dikatakan memiliki sifat rasio kemungkinan monoton dalam statistik $T = \sum X_i$. Perhatikan bahwa sifat ini juga berlaku untuk statistik $\bar{X}$, karena statistik ini merupakan fungsi menaik dari *T*.

Sifat rasio kemungkinan monoton berguna untuk menurunkan uji paling kuasa seragam.

Teorema 12. 2.

Jika fungsi kepekatan peluang bersama *f(**x**;θ)* memiliki sifat rasio kemungkinan monoton dalam statistik *T=t(**X**)*, maka uji paling kuasa seragam berukuran α untuk H_0: $\theta \le \theta_0$ lawan H_1: $\theta > \theta_0$ adalah dengan menolak H_0 jika $t(\mathbf{x}) \ge k$ sedemikian rupa sehingga $P[t(\mathbf{X}) \ge k | \theta_0] = \alpha$.

Permasalahan dual dari pengujian H_0: $\theta \ge \theta_0$ lawan H_1: $\theta < \theta_0$ dapat dilakukan dengan pendekatan ini juga, hanya dengan membalik pertidaksamaan menjadi $\pi_{C*}(\theta) \le \pi_C(\theta)$. Juga seandainya rasio merupakan fungsi tak menaik dari *t(**x**)*, maka H_0 dari torema tersebut dapat ditolak dengan membalik pertidaksamaan dalam *t(**x**)*.

Teladan 12. 7.

Contoh acak berukuran *n* dari sebaran Eksponensial 2 parameter, X_i*~Eksponensial(1,*η*).* Fungsi kepekatan peluang bersamanya adalah $f(\underline{x};\eta) = \exp[-\sum(x_i - \eta)]I_{(\eta,\infty)}(x_{(1)})$. Dengan demikian, jika $\eta_1 < \eta_2$, maka

$$\frac{f(\underline{x};\eta_2)}{f(\underline{x};\eta_1)} = \begin{cases} 0 & , \quad \eta_1 < x_{(1)} \leq \eta_2 \\ \exp[n(\eta_2 - \eta_1)] & , \quad \eta_2 < x_{(1)} \end{cases}.$$

Fakta bahwa fungsi ini tak terdefinisi untuk $x_{(1)} \leq 0$ tidak jadi masalah, karena $P\left[X_{(1)} \leq \eta_1\right] = 0$ bilamana η_1 merupakan nilai η yang sebenarnya. Rasio diatas merupakan fungsi tak menurun dari nilai minimumnya dan sifat rasio kemungkinan maksimum berlaku untuk $T=X_{(1)}$. Dengan demikian, uji paling kuasa seragam berukuran α untuk H_0: $\eta\leq\eta_0$ lawan H_1:$\eta\leq\eta_0$ adalah dengan membuat keputusan Tolak H_0 jika $x_{(1)}\geq k$, sedemikian rupa sehingga $\alpha = P[X_{(1)} \geq k \mid \eta_0] = \exp[-n(k - \eta_0)]$ atau nilai $k = \eta_0 - (\ln\alpha)/n$.

Teorema 12. 3.

Misalkan bahwa X_1, …, X_n memiliki fungsi kepekatan peluang bersama dalam bentuk $f(\underline{x};\theta) = c(\theta)h(\underline{x})\exp[q(\theta)t(\underline{x})]$ dimana $q(\theta)$ merupakan fungsi menaik pada θ.

2. Uji Paling Kuasa Seragam berukuran α untuk H_0:$\theta\leq\theta$ lawan H_1:$\theta>\theta$ adalah menolak H_0 jika $t(\mathbf{x})\geq k$, dimana $P[t(\mathbf{X})\geq k|\theta_0]=\alpha$.
3. Uji Paling Kuasa Seragam berukuran α untuk H_0:$\theta\geq\theta$ lawan H_1:$\theta<\theta$ adalah menolak H_0 jika $t(\mathbf{x})\geq k$, dimana $P[t(\mathbf{X})\leq k|\theta_0]=\alpha$.

Bukti

Sebagai latihan.

Teladan 12. 8.

Sebuah contoh acak berukuran *n* dari sebaran Poisson, X_i~*Poisson*(μ). Fungsi kepekatan peluang bersamanya adalah $f(\underline{x};\mu)=\dfrac{e^{-n\mu}\mu^{\sum x_i}}{x_1!\cdots x_n!}$ untuk semua x_i = *0, 1, 2, ...* yang juga dapat dituliskan $f(\underline{x};\mu)=e^{-n\mu}(x_1!\cdots x_n!)^{-1}\exp[(\ln\mu)\sum x_i]$. Dengan menggunakan teorema diatas, $q(\mu)=\ln\mu$ dan $t(\mathbf{x})=\Sigma x_i$. Uji paling kuasa berukuran α untuk $H_0:\mu\le\mu_0$ lawan $H_1:\mu>\mu_0$ adalah menolak H_0 jika $t(\mathbf{x})\ge k$, dimana $P[t(\mathbf{X})\ge k|\mu_0]=\alpha$. Karena *T*~*Poisson*($n\mu$), maka

$$\alpha=\sum_{t=k}^{\infty}\frac{e^{-n\mu_0}(n\mu_0)^t}{t!}.$$

Dalam beberapa kasus, uji paling kuasa seragam mungkin tak akan dapat diperoleh, khususnya untuk hipotesis alternatif dua arah, mungkin ada uji paling kuasa diantara kelas uji tak bias terbatas.

Definisi 12. 8.

Suatu uji $H_0:\theta\in\Omega_0$ lawan $H_a:\theta\in\Omega-\Omega_0$ dikatkan tak bias jika

$$\min_{\theta\in\Omega-\Omega_0}\pi(\theta)\ge\max_{\theta\in\Omega}\pi(\theta)$$

Definisi 12. 9

Misalkan $\mathbf{X}$ = (X_1, …, X_n) dimana X_1, …, X_n memiliki fungsi kepekatan peluang bersama $f(\mathbf{x};\theta)$ untuk $\theta\in\Omega$, dan perhatikan hipotesis H_0: $\theta\in\Omega_0$ lawan H_1: $\theta\in\Omega-\Omega_0$. rasio kemungkinan umum (*generalized likelihood ratio*) didefinisikan sebagai

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

dimana $\underline{\hat{\theta}}$ adalah penduga kemungkinan maksimum bagi $\underline{\theta}$ dan $\underline{\theta}_0$ adalah penduga kemungkinan maksimum pada kondisi hipotesis nol benar.

Pada dasarnya, prinsip rasio kemungkinan umum menentukan daerah kritis dengan memilih titik-titik mana yang harus dimasukkan kedalamnya berdasarkan rasio kemungkinan hasil estimasi dari data amatan, dimana pembilangnya diestimasi pada kondisi hipotesis nol benar. Hal ini sama dengan prinsip Neyman-Pearson dimana kemungkinan tesebut disebutkan dengan lengkap, tetapi ini bukan generalisasi prinsip Neyman-Pearson, sebab nilai estimasi tak terbatasi, $\underline{\hat{\theta}}$, dapat saja berada dalam $\underline{\Omega}_0$.

$\lambda(\underline{X})$ merupakan statistik uji yang sahih yang tidak merupakan fungsi dari parameter yang tak diketahui; dalam banyak kasus sebaran $\lambda(\underline{X})$ bebas parameter, dan nilai kritis pastinya, *k*, dapat ditentukan. Dalam beberapa kasus, sebaran $\lambda(\underline{X})$ bilamana hipotesis nol benar tergantung pada parameter yang tak diketahui, dan daerah kritis berukuran pasti α tak dapat ditentukan. Jika kondisi reguler berlaku, yang menjamin penduga kemungkinan maksimumnya berdistribusi asimtotik Normal, dapat ditunjukkan bahwa sebaran asimtotik dari $\lambda(\underline{X})$ bebas parameter, dan uji berukuran mendekati α akan diperoleh bilamana ukuran contoh *n* menjadi besar.

Jika $\underline{X} \sim f(\underline{x};\theta_1,...,\theta_k)$, maka bila hipotesis nol $H_0:(\theta_1,...,\theta_r)$ $=(\theta_{10},...,\theta_{r0})$ $r<k$ benar, untuk *n* yang besar, maka $-2\ln\lambda(\underline{X}) \sim \chi^2_{(r)}$

Teladan 12. 9

Misalkan $X_i \sim N(\mu,\sigma^2)$ dimana σ^2 diketahui. Ingin diuji H_0: $\mu = \mu_0$ lawan H_1: $\mu \neq \mu_0$. Maka penduga kemungkinan maksimum dari μ adalah $\hat{\mu} = \overline{x}$ dan rasio kemungkinan maksimumnya adalah :

$$\lambda(\underline{x}) = \frac{f(\underline{x};\mu_0)}{f(\underline{x};\hat{\mu})}$$

$$= \frac{(2\pi\sigma^2)^{-n/2} \exp\left[-\sum_{i=1}^{n}(x_i-\mu_0)^2/(2\sigma^2)\right]}{(2\pi\sigma^2)^{-n/2} \exp\left[-\sum_{i=1}^{n}(x_i-\overline{x})^2/(2\sigma^2)\right]}$$

Bila disederhanakan, maka akan diperoleh

$$\lambda(\underline{x}) = \exp\left[\frac{-n(\overline{x}-\mu_0)^2}{2\sigma^2}\right]$$

Dengan demikian, kita tolak hipotesis nol jika $\lambda(\underline{x}) \leq k$ setara dengan menyatakan menolak hipotesis nol jika

$$z^2 = \left[\frac{\overline{x}-\mu_0}{\sigma/\sqrt{n}}\right]^2 \geq k^*$$

dimana $Z \sim N(0,1)$ dan $Z^2 \sim \chi^2(1)$. Jadi, uji berukuran α adalah menolak hipotesis nol jika $z^2 \geq \chi^2_{1-\alpha;1}$ yang ekuivalen dengan menolak hipotesis nol jika $z \leq -z_{1-\frac{\alpha}{2}}$ atau $z \geq z_{1-\frac{\alpha}{2}}$. Ini berarti bahwa uji rasio kemungkinan tersebut dapat direduksi menjadi uji normal dua arah berukuran sama. Perlu juga dicatat bahwa sebaran asimtotik $-2\ln\lambda(\underline{X}) \sim \chi^2_1$ merupakan sebaran pasti.

Latihan

1. Misalkan $X_1, \ldots, X_n$ adalah contoh acak berukuran *n* dari sebaran dengan fungsi kepekatan peluang $f(x;\theta) = \theta x^{\theta-1} I_{(0,1)}(x)$, $\theta > 0$
   a. Carilah penduga kemungkinan maksimum $\hat{\theta}$ bagi θ.
   b. Carilah statistik cukup dan lengkap bagi θ.
   c. Carilah CRLB untuk $1/\theta$.
   d. Carilah UMVUE untuk $1/\theta$.
   e. Carilah sebaran normal asimtotik untuk $\hat{\theta}$ dan untuk $1/\hat{\theta}$.
   f. Carilah UMVUE untuk θ.
   g. Carilah selang kepercayaan dua arah 100γ % bagi θ.
   h. Berdasarkan ukuran contoh sebesar n, carilah Uji Paling Kuasa Seragam pada taraf α untuk H_0: $\theta = \theta_0$ vs H_1: $\theta = \theta_1$ dimana $\theta_1 > \theta_0$.
2. Misalkan $X_1, \ldots, X_n$ adalah contoh acak berukuran *n* dari sebaran dengan fungsi kepekatan peluang $f(x;\theta) = \theta(1+x)^{-(1+\theta)} I_{(0,\infty)}(x)$
   a. Carilah penduga kemungkinan maksimum $\hat{\theta}$ bagi θ.
   b. Carilah statistik cukup dan lengkap bagi θ.
   c. Carilah CRLB untuk $1/\theta$.
   d. Carilah UMVUE untuk $1/\theta$.
   e. Carilah sebaran normal asimtotik untuk $\hat{\theta}$ dan untuk $1/\hat{\theta}$.
   f. Carilah UMVUE untuk θ.
   g. Carilah selang kepercayaan dua arah 100γ % bagi θ.

h. Berdasarkan ukuran contoh sebesar n, carilah Uji Paling Kuasa Seragam pada taraf α untuk $H_0: \theta = \theta_0$ vs $H_1: \theta = \theta_1$ dimana $\theta_1 > \theta_0$.

LAMPIRAN

Seluruh isi tabel dibangkitkan dengan menggunakan
Fungsi-fungsi yang tersedia pada
Microsoft Excel ®

Oleh : Sigit Nugroho

Tabel Lampiran 1. Sebaran Kumulatif Normal Baku.

	0,00	0,01	0,02	0,03	0,04	0,05	0,06	0,07	0,08	0,09
-3,0	0,0013	0,0013	0,0013	0,0012	0,0012	0,0011	0,0011	0,0011	0,0010	0,0010
-2,9	0,0019	0,0018	0,0018	0,0017	0,0016	0,0016	0,0015	0,0015	0,0014	0,0014
-2,8	0,0026	0,0025	0,0024	0,0023	0,0023	0,0022	0,0021	0,0021	0,0020	0,0019
-2,7	0,0035	0,0034	0,0033	0,0032	0,0031	0,0030	0,0029	0,0028	0,0027	0,0026
-2,6	0,0047	0,0045	0,0044	0,0043	0,0041	0,0040	0,0039	0,0038	0,0037	0,0036
-2,5	0,0062	0,0060	0,0059	0,0057	0,0055	0,0054	0,0052	0,0051	0,0049	0,0048
-2,4	0,0082	0,0080	0,0078	0,0075	0,0073	0,0071	0,0069	0,0068	0,0066	0,0064
-2,3	0,0107	0,0104	0,0102	0,0099	0,0096	0,0094	0,0091	0,0089	0,0087	0,0084
-2,2	0,0139	0,0136	0,0132	0,0129	0,0125	0,0122	0,0119	0,0116	0,0113	0,0110
-2,1	0,0179	0,0174	0,0170	0,0166	0,0162	0,0158	0,0154	0,0150	0,0146	0,0143
-2,0	0,0228	0,0222	0,0217	0,0212	0,0207	0,0202	0,0197	0,0192	0,0188	0,0183
-1,9	0,0287	0,0281	0,0274	0,0268	0,0262	0,0256	0,0250	0,0244	0,0239	0,0233
-1,8	0,0359	0,0351	0,0344	0,0336	0,0329	0,0322	0,0314	0,0307	0,0301	0,0294
-1,7	0,0446	0,0436	0,0427	0,0418	0,0409	0,0401	0,0392	0,0384	0,0375	0,0367
-1,6	0,0548	0,0537	0,0526	0,0516	0,0505	0,0495	0,0485	0,0475	0,0465	0,0455
-1,5	0,0668	0,0655	0,0643	0,0630	0,0618	0,0606	0,0594	0,0582	0,0571	0,0559
-1,4	0,0808	0,0793	0,0778	0,0764	0,0749	0,0735	0,0721	0,0708	0,0694	0,0681
-1,3	0,0968	0,0951	0,0934	0,0918	0,0901	0,0885	0,0869	0,0853	0,0838	0,0823
-1,2	0,1151	0,1131	0,1112	0,1093	0,1075	0,1056	0,1038	0,1020	0,1003	0,0985
-1,1	0,1357	0,1335	0,1314	0,1292	0,1271	0,1251	0,1230	0,1210	0,1190	0,1170
-1,0	0,1587	0,1562	0,1539	0,1515	0,1492	0,1469	0,1446	0,1423	0,1401	0,1379
-0,9	0,1841	0,1814	0,1788	0,1762	0,1736	0,1711	0,1685	0,1660	0,1635	0,1611
-0,8	0,2119	0,2090	0,2061	0,2033	0,2005	0,1977	0,1949	0,1922	0,1894	0,1867
-0,7	0,2420	0,2389	0,2358	0,2327	0,2296	0,2266	0,2236	0,2206	0,2177	0,2148
-0,6	0,2743	0,2709	0,2676	0,2643	0,2611	0,2578	0,2546	0,2514	0,2483	0,2451
-0,5	0,3085	0,3050	0,3015	0,2981	0,2946	0,2912	0,2877	0,2843	0,2810	0,2776
-0,4	0,3446	0,3409	0,3372	0,3336	0,3300	0,3264	0,3228	0,3192	0,3156	0,3121
-0,3	0,3821	0,3783	0,3745	0,3707	0,3669	0,3632	0,3594	0,3557	0,3520	0,3483
-0,2	0,4207	0,4168	0,4129	0,4090	0,4052	0,4013	0,3974	0,3936	0,3897	0,3859
-0,1	0,4602	0,4562	0,4522	0,4483	0,4443	0,4404	0,4364	0,4325	0,4286	0,4247
0,0	0,5000	0,5040	0,5080	0,5120	0,5160	0,5199	0,5239	0,5279	0,5319	0,5359

... Lanjutan

	0,00	0,01	0,02	0,03	0,04	0,05	0,06	0,07	0,08	0,09
0,1	0,5398	0,5438	0,5478	0,5517	0,5557	0,5596	0,5636	0,5675	0,5714	0,5753
0,2	0,5793	0,5832	0,5871	0,5910	0,5948	0,5987	0,6026	0,6064	0,6103	0,6141
0,3	0,6179	0,6217	0,6255	0,6293	0,6331	0,6368	0,6406	0,6443	0,6480	0,6517
0,4	0,6554	0,6591	0,6628	0,6664	0,6700	0,6736	0,6772	0,6808	0,6844	0,6879
0,5	0,6915	0,6950	0,6985	0,7019	0,7054	0,7088	0,7123	0,7157	0,7190	0,7224
0,6	0,7257	0,7291	0,7324	0,7357	0,7389	0,7422	0,7454	0,7486	0,7517	0,7549
0,7	0,7580	0,7611	0,7642	0,7673	0,7704	0,7734	0,7764	0,7794	0,7823	0,7852
0,8	0,7881	0,7910	0,7939	0,7967	0,7995	0,8023	0,8051	0,8078	0,8106	0,8133
0,9	0,8159	0,8186	0,8212	0,8238	0,8264	0,8289	0,8315	0,8340	0,8365	0,8389
1,0	0,8413	0,8438	0,8461	0,8485	0,8508	0,8531	0,8554	0,8577	0,8599	0,8621
1,1	0,8643	0,8665	0,8686	0,8708	0,8729	0,8749	0,8770	0,8790	0,8810	0,8830
1,2	0,8849	0,8869	0,8888	0,8907	0,8925	0,8944	0,8962	0,8980	0,8997	0,9015
1,3	0,9032	0,9049	0,9066	0,9082	0,9099	0,9115	0,9131	0,9147	0,9162	0,9177
1,4	0,9192	0,9207	0,9222	0,9236	0,9251	0,9265	0,9279	0,9292	0,9306	0,9319
1,5	0,9332	0,9345	0,9357	0,9370	0,9382	0,9394	0,9406	0,9418	0,9429	0,9441
1,6	0,9452	0,9463	0,9474	0,9484	0,9495	0,9505	0,9515	0,9525	0,9535	0,9545
1,7	0,9554	0,9564	0,9573	0,9582	0,9591	0,9599	0,9608	0,9616	0,9625	0,9633
1,8	0,9641	0,9649	0,9656	0,9664	0,9671	0,9678	0,9686	0,9693	0,9699	0,9706
1,9	0,9713	0,9719	0,9726	0,9732	0,9738	0,9744	0,9750	0,9756	0,9761	0,9767
2,0	0,9772	0,9778	0,9783	0,9788	0,9793	0,9798	0,9803	0,9808	0,9812	0,9817
2,1	0,9821	0,9826	0,9830	0,9834	0,9838	0,9842	0,9846	0,9850	0,9854	0,9857
2,2	0,9861	0,9864	0,9868	0,9871	0,9875	0,9878	0,9881	0,9884	0,9887	0,9890
2,3	0,9893	0,9896	0,9898	0,9901	0,9904	0,9906	0,9909	0,9911	0,9913	0,9916
2,4	0,9918	0,9920	0,9922	0,9925	0,9927	0,9929	0,9931	0,9932	0,9934	0,9936
2,5	0,9938	0,9940	0,9941	0,9943	0,9945	0,9946	0,9948	0,9949	0,9951	0,9952
2,6	0,9953	0,9955	0,9956	0,9957	0,9959	0,9960	0,9961	0,9962	0,9963	0,9964
2,7	0,9965	0,9966	0,9967	0,9968	0,9969	0,9970	0,9971	0,9972	0,9973	0,9974
2,8	0,9974	0,9975	0,9976	0,9977	0,9977	0,9978	0,9979	0,9979	0,9980	0,9981
2,9	0,9981	0,9982	0,9982	0,9983	0,9984	0,9984	0,9985	0,9985	0,9986	0,9986
3,0	0,9987	0,9987	0,9987	0,9988	0,9988	0,9989	0,9989	0,9989	0,9990	0,9990

Tabel Lampiran 2. Tabel Sebaran Kai-kuadrat

db	0,001	0,005	0,010	0,025	0,050	0,100
1	10,828	7,879	6,635	5,024	3,841	2,706
2	13,816	10,597	9,210	7,378	5,991	4,605
3	16,266	12,838	11,345	9,348	7,815	6,251
4	18,467	14,860	13,277	11,143	9,488	7,779
5	20,515	16,750	15,086	12,833	11,070	9,236
6	22,458	18,548	16,812	14,449	12,592	10,645
7	24,322	20,278	18,475	16,013	14,067	12,017
8	26,124	21,955	20,090	17,535	15,507	13,362
9	27,877	23,589	21,666	19,023	16,919	14,684
10	29,588	25,188	23,209	20,483	18,307	15,987
11	31,264	26,757	24,725	21,920	19,675	17,275
12	32,909	28,300	26,217	23,337	21,026	18,549
13	34,528	29,819	27,688	24,736	22,362	19,812
14	36,123	31,319	29,141	26,119	23,685	21,064
15	37,697	32,801	30,578	27,488	24,996	22,307
16	39,252	34,267	32,000	28,845	26,296	23,542
17	40,790	35,718	33,409	30,191	27,587	24,769
18	42,312	37,156	34,805	31,526	28,869	25,989
19	43,820	38,582	36,191	32,852	30,144	27,204
20	45,315	39,997	37,566	34,170	31,410	28,412
21	46,797	41,401	38,932	35,479	32,671	29,615
22	48,268	42,796	40,289	36,781	33,924	30,813
23	49,728	44,181	41,638	38,076	35,172	32,007
24	51,179	45,559	42,980	39,364	36,415	33,196
25	52,620	46,928	44,314	40,646	37,652	34,382
26	54,052	48,290	45,642	41,923	38,885	35,563
27	55,476	49,645	46,963	43,195	40,113	36,741
28	56,892	50,993	48,278	44,461	41,337	37,916
29	58,301	52,336	49,588	45,722	42,557	39,087
30	59,703	53,672	50,892	46,979	43,773	40,256
40	73,402	66,766	63,691	59,342	55,758	51,805
50	86,661	79,490	76,154	71,420	67,505	63,167

... Lanjutan

db	0,900	0,950	0,975	0,990	0,995	0,999
1	0,016	0,004	0,001	0,000	0,000	0,000
2	0,211	0,103	0,051	0,020	0,010	0,002
3	0,584	0,352	0,216	0,115	0,072	0,024
4	1,064	0,711	0,484	0,297	0,207	0,091
5	1,610	1,145	0,831	0,554	0,412	0,210
6	2,204	1,635	1,237	0,872	0,676	0,381
7	2,833	2,167	1,690	1,239	0,989	0,598
8	3,490	2,733	2,180	1,646	1,344	0,857
9	4,168	3,325	2,700	2,088	1,735	1,152
10	4,865	3,940	3,247	2,558	2,156	1,479
11	5,578	4,575	3,816	3,053	2,603	1,834
12	6,304	5,226	4,404	3,571	3,074	2,214
13	7,042	5,892	5,009	4,107	3,565	2,617
14	7,790	6,571	5,629	4,660	4,075	3,041
15	8,547	7,261	6,262	5,229	4,601	3,483
16	9,312	7,962	6,908	5,812	5,142	3,942
17	10,085	8,672	7,564	6,408	5,697	4,416
18	10,865	9,390	8,231	7,015	6,265	4,905
19	11,651	10,117	8,907	7,633	6,844	5,407
20	12,443	10,851	9,591	8,260	7,434	5,921
21	13,240	11,591	10,283	8,897	8,034	6,447
22	14,041	12,338	10,982	9,542	8,643	6,983
23	14,848	13,091	11,689	10,196	9,260	7,529
24	15,659	13,848	12,401	10,856	9,886	8,085
25	16,473	14,611	13,120	11,524	10,520	8,649
26	17,292	15,379	13,844	12,198	11,160	9,222
27	18,114	16,151	14,573	12,879	11,808	9,803
28	18,939	16,928	15,308	13,565	12,461	10,391
29	19,768	17,708	16,047	14,256	13,121	10,986
30	20,599	18,493	16,791	14,953	13,787	11,588
40	29,051	26,509	24,433	22,164	20,707	17,916
50	37,689	34,764	32,357	29,707	27,991	24,674

Tabel Lampiran 3. Sebaran *t-student*

	0,100	0,050	0,025	0,010	0,005	0,001
1	3,078	6,314	12,706	31,821	63,657	318,309
2	1,886	2,920	4,303	6,965	9,925	22,327
3	1,638	2,353	3,182	4,541	5,841	10,215
4	1,533	2,132	2,776	3,747	4,604	7,173
5	1,476	2,015	2,571	3,365	4,032	5,893
6	1,440	1,943	2,447	3,143	3,707	5,208
7	1,415	1,895	2,365	2,998	3,499	4,785
8	1,397	1,860	2,306	2,896	3,355	4,501
9	1,383	1,833	2,262	2,821	3,250	4,297
10	1,372	1,812	2,228	2,764	3,169	4,144
11	1,363	1,796	2,201	2,718	3,106	4,025
12	1,356	1,782	2,179	2,681	3,055	3,930
13	1,350	1,771	2,160	2,650	3,012	3,852
14	1,345	1,761	2,145	2,624	2,977	3,787
15	1,341	1,753	2,131	2,602	2,947	3,733
16	1,337	1,746	2,120	2,583	2,921	3,686
17	1,333	1,740	2,110	2,567	2,898	3,646
18	1,330	1,734	2,101	2,552	2,878	3,610
19	1,328	1,729	2,093	2,539	2,861	3,579
20	1,325	1,725	2,086	2,528	2,845	3,552
21	1,323	1,721	2,080	2,518	2,831	3,527
22	1,321	1,717	2,074	2,508	2,819	3,505
23	1,319	1,714	2,069	2,500	2,807	3,485
24	1,318	1,711	2,064	2,492	2,797	3,467
25	1,316	1,708	2,060	2,485	2,787	3,450
26	1,315	1,706	2,056	2,479	2,779	3,435
27	1,314	1,703	2,052	2,473	2,771	3,421
28	1,313	1,701	2,048	2,467	2,763	3,408
29	1,311	1,699	2,045	2,462	2,756	3,396
30	1,310	1,697	2,042	2,457	2,750	3,385
40	1,303	1,684	2,021	2,423	2,704	3,307
60	1,296	1,671	2,000	2,390	2,660	3,232
inf	1,282	1,645	1,960	2,326	2,576	3,090

Tabel Lampiran 4. Sebaran *F*

db-2	alpha	Db-1 1	2	3	4	5	6	7	8
1	**0,100**	39,86	49,50	53,59	55,83	57,24	58,20	58,91	59,44
	0,050	161,45	199,50	215,71	224,58	230,16	233,99	236,77	238,88
	0,010	4052,18	4999,50	5403,35	5624,58	5763,65	5858,99	5928,36	5981,07
2	**0,100**	8,53	9,00	9,16	9,24	9,29	9,33	9,35	9,37
	0,050	18,51	19,00	19,16	19,25	19,30	19,33	19,35	19,37
	0,010	98,50	99,00	99,17	99,25	99,30	99,33	99,36	99,37
3	**0,100**	5,54	5,46	5,39	5,34	5,31	5,28	5,27	5,25
	0,050	10,13	9,55	9,28	9,12	9,01	8,94	8,89	8,85
	0,010	34,12	30,82	29,46	28,71	28,24	27,91	27,67	27,49
4	**0,100**	4,54	4,32	4,19	4,11	4,05	4,01	3,98	3,95
	0,050	7,71	6,94	6,59	6,39	6,26	6,16	6,09	6,04
	0,010	21,20	18,00	16,69	15,98	15,52	15,21	14,98	14,80
5	**0,100**	4,06	3,78	3,62	3,52	3,45	3,40	3,37	3,34
	0,050	6,61	5,79	5,41	5,19	5,05	4,95	4,88	4,82
	0,010	16,26	13,27	12,06	11,39	10,97	10,67	10,46	10,29
6	**0,100**	3,78	3,46	3,29	3,18	3,11	3,05	3,01	2,98
	0,050	5,99	5,14	4,76	4,53	4,39	4,28	4,21	4,15
	0,010	13,75	10,92	9,78	9,15	8,75	8,47	8,26	8,10
7	**0,100**	3,59	3,26	3,07	2,96	2,88	2,83	2,78	2,75
	0,050	5,59	4,74	4,35	4,12	3,97	3,87	3,79	3,73
	0,010	12,25	9,55	8,45	7,85	7,46	7,19	6,99	6,84
8	**0,100**	3,46	3,11	2,92	2,81	2,73	2,67	2,62	2,59
	0,050	5,32	4,46	4,07	3,84	3,69	3,58	3,50	3,44
	0,010	11,26	8,65	7,59	7,01	6,63	6,37	6,18	6,03

... Lanjutan

db-2	alpha	db-1: 1	2	3	4	5	6	7	8
9	**0,100**	3,36	3,01	2,81	2,69	2,61	2,55	2,51	2,47
	0,050	5,12	4,26	3,86	3,63	3,48	3,37	3,29	3,23
	0,010	10,56	8,02	6,99	6,42	6,06	5,80	5,61	5,47
10	**0,100**	3,29	2,92	2,73	2,61	2,52	2,46	2,41	2,38
	0,050	4,96	4,10	3,71	3,48	3,33	3,22	3,14	3,07
	0,010	10,04	7,56	6,55	5,99	5,64	5,39	5,20	5,06
11	**0,100**	3,23	2,86	2,66	2,54	2,45	2,39	2,34	2,30
	0,050	4,84	3,98	3,59	3,36	3,20	3,09	3,01	2,95
	0,010	9,65	7,21	6,22	5,67	5,32	5,07	4,89	4,74
12	**0,100**	3,18	2,81	2,61	2,48	2,39	2,33	2,28	2,24
	0,050	4,75	3,89	3,49	3,26	3,11	3,00	2,91	2,85
	0,010	9,33	6,93	5,95	5,41	5,06	4,82	4,64	4,50
13	**0,100**	3,14	2,76	2,56	2,43	2,35	2,28	2,23	2,20
	0,050	4,67	3,81	3,41	3,18	3,03	2,92	2,83	2,77
	0,010	9,07	6,70	5,74	5,21	4,86	4,62	4,44	4,30
14	**0,100**	3,10	2,73	2,52	2,39	2,31	2,24	2,19	2,15
	0,050	4,60	3,74	3,34	3,11	2,96	2,85	2,76	2,70
	0,010	8,86	6,51	5,56	5,04	4,69	4,46	4,28	4,14
15	**0,100**	3,07	2,70	2,49	2,36	2,27	2,21	2,16	2,12
	0,050	4,54	3,68	3,29	3,06	2,90	2,79	2,71	2,64
	0,010	8,68	6,36	5,42	4,89	4,56	4,32	4,14	4,00
16	**0,100**	3,05	2,67	2,46	2,33	2,24	2,18	2,13	2,09
	0,050	4,49	3,63	3,24	3,01	2,85	2,74	2,66	2,59
	0,010	8,53	6,23	5,29	4,77	4,44	4,20	4,03	3,89

... Lanjutan

db-2	alpha	db-1: 1	2	3	4	5	6	7	8
17	**0,100**	3,03	2,64	2,44	2,31	2,22	2,15	2,10	2,06
	0,050	4,45	3,59	3,20	2,96	2,81	2,70	2,61	2,55
	0,010	8,40	6,11	5,18	4,67	4,34	4,10	3,93	3,79
18	**0,100**	3,01	2,62	2,42	2,29	2,20	2,13	2,08	2,04
	0,050	4,41	3,55	3,16	2,93	2,77	2,66	2,58	2,51
	0,010	8,29	6,01	5,09	4,58	4,25	4,01	3,84	3,71
19	**0,100**	2,99	2,61	2,40	2,27	2,18	2,11	2,06	2,02
	0,050	4,38	3,52	3,13	2,90	2,74	2,63	2,54	2,48
	0,010	8,18	5,93	5,01	4,50	4,17	3,94	3,77	3,63
20	**0,100**	2,97	2,59	2,38	2,25	2,16	2,09	2,04	2,00
	0,050	4,35	3,49	3,10	2,87	2,71	2,60	2,51	2,45
	0,010	8,10	5,85	4,94	4,43	4,10	3,87	3,70	3,56
21	**0,100**	2,96	2,57	2,36	2,23	2,14	2,08	2,02	1,98
	0,050	4,32	3,47	3,07	2,84	2,68	2,57	2,49	2,42
	0,010	8,02	5,78	4,87	4,37	4,04	3,81	3,64	3,51
22	**0,100**	2,95	2,56	2,35	2,22	2,13	2,06	2,01	1,97
	0,050	4,30	3,44	3,05	2,82	2,66	2,55	2,46	2,40
	0,010	7,95	5,72	4,82	4,31	3,99	3,76	3,59	3,45
23	**0,100**	2,94	2,55	2,34	2,21	2,11	2,05	1,99	1,95
	0,050	4,28	3,42	3,03	2,80	2,64	2,53	2,44	2,37
	0,010	7,88	5,66	4,76	4,26	3,94	3,71	3,54	3,41
24	**0,100**	2,93	2,54	2,33	2,19	2,10	2,04	1,98	1,94
	0,050	4,26	3,40	3,01	2,78	2,62	2,51	2,42	2,36
	0,010	7,82	5,61	4,72	4,22	3,90	3,67	3,50	3,36

... Lanjutan

db-2	alpha	db-1: 9	10	11	12	13	14	15	16
1	**0,100**	59,86	60,19	60,47	60,71	60,90	61,07	61,22	61,35
	0,050	240,54	241,88	242,98	243,91	244,69	245,36	245,95	246,46
	0,010	6022,47	6055,85	6083,32	6106,32	6125,86	6142,67	6157,28	6170,10
2	**0,100**	9,38	9,39	9,40	9,41	9,41	9,42	9,42	9,43
	0,050	19,38	19,40	19,40	19,41	19,42	19,42	19,43	19,43
	0,010	99,39	99,40	99,41	99,42	99,42	99,43	99,43	99,44
3	**0,100**	5,24	5,23	5,22	5,22	5,21	5,20	5,20	5,20
	0,050	8,81	8,79	8,76	8,74	8,73	8,71	8,70	8,69
	0,010	27,35	27,23	27,13	27,05	26,98	26,92	26,87	26,83
4	**0,100**	3,94	3,92	3,91	3,90	3,89	3,88	3,87	3,86
	0,050	6,00	5,96	5,94	5,91	5,89	5,87	5,86	5,84
	0,010	14,66	14,55	14,45	14,37	14,31	14,25	14,20	14,15
5	**0,100**	3,32	3,30	3,28	3,27	3,26	3,25	3,24	3,23
	0,050	4,77	4,74	4,70	4,68	4,66	4,64	4,62	4,60
	0,010	10,16	10,05	9,96	9,89	9,82	9,77	9,72	9,68
6	**0,100**	2,96	2,94	2,92	2,90	2,89	2,88	2,87	2,86
	0,050	4,10	4,06	4,03	4,00	3,98	3,96	3,94	3,92
	0,010	7,98	7,87	7,79	7,72	7,66	7,60	7,56	7,52
7	**0,100**	2,72	2,70	2,68	2,67	2,65	2,64	2,63	2,62
	0,050	3,68	3,64	3,60	3,57	3,55	3,53	3,51	3,49
	0,010	6,72	6,62	6,54	6,47	6,41	6,36	6,31	6,28
8	**0,100**	2,56	2,54	2,52	2,50	2,49	2,48	2,46	2,45
	0,050	3,39	3,35	3,31	3,28	3,26	3,24	3,22	3,20
	0,010	5,91	5,81	5,73	5,67	5,61	5,56	5,52	5,48

... Lanjutan

db-2	alpha	9	10	11	12	13	14	15	16
		db-1							
9	**0,100**	2,44	2,42	2,40	2,38	2,36	2,35	2,34	2,33
	0,050	3,18	3,14	3,10	3,07	3,05	3,03	3,01	2,99
	0,010	5,35	5,26	5,18	5,11	5,05	5,01	4,96	4,92
10	**0,100**	2,35	2,32	2,30	2,28	2,27	2,26	2,24	2,23
	0,050	3,02	2,98	2,94	2,91	2,89	2,86	2,85	2,83
	0,010	4,94	4,85	4,77	4,71	4,65	4,60	4,56	4,52
11	**0,100**	2,27	2,25	2,23	2,21	2,19	2,18	2,17	2,16
	0,050	2,90	2,85	2,82	2,79	2,76	2,74	2,72	2,70
	0,010	4,63	4,54	4,46	4,40	4,34	4,29	4,25	4,21
12	**0,100**	2,21	2,19	2,17	2,15	2,13	2,12	2,10	2,09
	0,050	2,80	2,75	2,72	2,69	2,66	2,64	2,62	2,60
	0,010	4,39	4,30	4,22	4,16	4,10	4,05	4,01	3,97
13	**0,100**	2,16	2,14	2,12	2,10	2,08	2,07	2,05	2,04
	0,050	2,71	2,67	2,63	2,60	2,58	2,55	2,53	2,51
	0,010	4,19	4,10	4,02	3,96	3,91	3,86	3,82	3,78
14	**0,100**	2,12	2,10	2,07	2,05	2,04	2,02	2,01	2,00
	0,050	2,65	2,60	2,57	2,53	2,51	2,48	2,46	2,44
	0,010	4,03	3,94	3,86	3,80	3,75	3,70	3,66	3,62
15	**0,100**	2,09	2,06	2,04	2,02	2,00	1,99	1,97	1,96
	0,050	2,59	2,54	2,51	2,48	2,45	2,42	2,40	2,38
	0,010	3,89	3,80	3,73	3,67	3,61	3,56	3,52	3,49
16	**0,100**	2,06	2,03	2,01	1,99	1,97	1,95	1,94	1,93
	0,050	2,54	2,49	2,46	2,42	2,40	2,37	2,35	2,33
	0,010	3,78	3,69	3,62	3,55	3,50	3,45	3,41	3,37

... Lanjutan

db-2	alpha	db-1 9	10	11	12	13	14	15	16
17	**0,100**	2,03	2,00	1,98	1,96	1,94	1,93	1,91	1,90
	0,050	2,49	2,45	2,41	2,38	2,35	2,33	2,31	2,29
	0,010	3,68	3,59	3,52	3,46	3,40	3,35	3,31	3,27
18	**0,100**	2,00	1,98	1,95	1,93	1,92	1,90	1,89	1,87
	0,050	2,46	2,41	2,37	2,34	2,31	2,29	2,27	2,25
	0,010	3,60	3,51	3,43	3,37	3,32	3,27	3,23	3,19
19	**0,100**	1,98	1,96	1,93	1,91	1,89	1,88	1,86	1,85
	0,050	2,42	2,38	2,34	2,31	2,28	2,26	2,23	2,21
	0,010	3,52	3,43	3,36	3,30	3,24	3,19	3,15	3,12
20	**0,100**	1,96	1,94	1,91	1,89	1,87	1,86	1,84	1,83
	0,050	2,39	2,35	2,31	2,28	2,25	2,22	2,20	2,18
	0,010	3,46	3,37	3,29	3,23	3,18	3,13	3,09	3,05
21	**0,100**	1,95	1,92	1,90	1,87	1,86	1,84	1,83	1,81
	0,050	2,37	2,32	2,28	2,25	2,22	2,20	2,18	2,16
	0,010	3,40	3,31	3,24	3,17	3,12	3,07	3,03	2,99
22	**0,100**	1,93	1,90	1,88	1,86	1,84	1,83	1,81	1,80
	0,050	2,34	2,30	2,26	2,23	2,20	2,17	2,15	2,13
	0,010	3,35	3,26	3,18	3,12	3,07	3,02	2,98	2,94
23	**0,100**	1,92	1,89	1,87	1,84	1,83	1,81	1,80	1,78
	0,050	2,32	2,27	2,24	2,20	2,18	2,15	2,13	2,11
	0,010	3,30	3,21	3,14	3,07	3,02	2,97	2,93	2,89
24	**0,100**	1,91	1,88	1,85	1,83	1,81	1,80	1,78	1,77
	0,050	2,30	2,25	2,22	2,18	2,15	2,13	2,11	2,09
	0,010	3,26	3,17	3,09	3,03	2,98	2,93	2,89	2,85

... Lanjutan

db-2	alpha	db-1 17	18	19	20	21	22	23	24
1	**0,100**	61,46	61,57	61,66	61,74	61,81	61,88	61,95	62,00
	0,050	246,92	247,32	247,69	248,01	248,31	248,58	248,83	249,05
	0,010	6181,43	6191,53	6200,58	6208,73	6216,12	6222,84	6228,99	6234,63
2	**0,100**	9,43	9,44	9,44	9,44	9,44	9,45	9,45	9,45
	0,050	19,44	19,44	19,44	19,45	19,45	19,45	19,45	19,45
	0,010	99,44	99,44	99,45	99,45	99,45	99,45	99,46	99,46
3	**0,100**	5,19	5,19	5,19	5,18	5,18	5,18	5,18	5,18
	0,050	8,68	8,67	8,67	8,66	8,65	8,65	8,64	8,64
	0,010	26,79	26,75	26,72	26,69	26,66	26,64	26,62	26,60
4	**0,100**	3,86	3,85	3,85	3,84	3,84	3,84	3,83	3,83
	0,050	5,83	5,82	5,81	5,80	5,79	5,79	5,78	5,77
	0,010	14,11	14,08	14,05	14,02	13,99	13,97	13,95	13,93
5	**0,100**	3,22	3,22	3,21	3,21	3,20	3,20	3,19	3,19
	0,050	4,59	4,58	4,57	4,56	4,55	4,54	4,53	4,53
	0,010	9,64	9,61	9,58	9,55	9,53	9,51	9,49	9,47
6	**0,100**	2,85	2,85	2,84	2,84	2,83	2,83	2,82	2,82
	0,050	3,91	3,90	3,88	3,87	3,86	3,86	3,85	3,84
	0,010	7,48	7,45	7,42	7,40	7,37	7,35	7,33	7,31
7	**0,100**	2,61	2,61	2,60	2,59	2,59	2,58	2,58	2,58
	0,050	3,48	3,47	3,46	3,44	3,43	3,43	3,42	3,41
	0,010	6,24	6,21	6,18	6,16	6,13	6,11	6,09	6,07
8	**0,100**	2,45	2,44	2,43	2,42	2,42	2,41	2,41	2,40
	0,050	3,19	3,17	3,16	3,15	3,14	3,13	3,12	3,12
	0,010	5,44	5,41	5,38	5,36	5,34	5,32	5,30	5,28

... Lanjutan

db-2	alpha	db-1 17	18	19	20	21	22	23	24
9	**0,100**	2,32	2,31	2,30	2,30	2,29	2,29	2,28	2,28
	0,050	2,97	2,96	2,95	2,94	2,93	2,92	2,91	2,90
	0,010	4,89	4,86	4,83	4,81	4,79	4,77	4,75	4,73
10	**0,100**	2,22	2,22	2,21	2,20	2,19	2,19	2,18	2,18
	0,050	2,81	2,80	2,79	2,77	2,76	2,75	2,75	2,74
	0,010	4,49	4,46	4,43	4,41	4,38	4,36	4,34	4,33
11	**0,100**	2,15	2,14	2,13	2,12	2,12	2,11	2,11	2,10
	0,050	2,69	2,67	2,66	2,65	2,64	2,63	2,62	2,61
	0,010	4,18	4,15	4,12	4,10	4,08	4,06	4,04	4,02
12	**0,100**	2,08	2,08	2,07	2,06	2,05	2,05	2,04	2,04
	0,050	2,58	2,57	2,56	2,54	2,53	2,52	2,51	2,51
	0,010	3,94	3,91	3,88	3,86	3,84	3,82	3,80	3,78
13	**0,100**	2,03	2,02	2,01	2,01	2,00	1,99	1,99	1,98
	0,050	2,50	2,48	2,47	2,46	2,45	2,44	2,43	2,42
	0,010	3,75	3,72	3,69	3,66	3,64	3,62	3,60	3,59
14	**0,100**	1,99	1,98	1,97	1,96	1,96	1,95	1,94	1,94
	0,050	2,43	2,41	2,40	2,39	2,38	2,37	2,36	2,35
	0,010	3,59	3,56	3,53	3,51	3,48	3,46	3,44	3,43
15	**0,100**	1,95	1,94	1,93	1,92	1,92	1,91	1,90	1,90
	0,050	2,37	2,35	2,34	2,33	2,32	2,31	2,30	2,29
	0,010	3,45	3,42	3,40	3,37	3,35	3,33	3,31	3,29
16	**0,100**	1,92	1,91	1,90	1,89	1,88	1,88	1,87	1,87
	0,050	2,32	2,30	2,29	2,28	2,26	2,25	2,24	2,24
	0,010	3,34	3,31	3,28	3,26	3,24	3,22	3,20	3,18

... Lanjutan

db-2	alpha	db-1: 17	18	19	20	21	22	23	24
17	**0,100**	1,89	1,88	1,87	1,86	1,86	1,85	1,84	1,84
	0,050	2,27	2,26	2,24	2,23	2,22	2,21	2,20	2,19
	0,010	3,24	3,21	3,19	3,16	3,14	3,12	3,10	3,08
18	**0,100**	1,86	1,85	1,84	1,84	1,83	1,82	1,82	1,81
	0,050	2,23	2,22	2,20	2,19	2,18	2,17	2,16	2,15
	0,010	3,16	3,13	3,10	3,08	3,05	3,03	3,02	3,00
19	**0,100**	1,84	1,83	1,82	1,81	1,81	1,80	1,79	1,79
	0,050	2,20	2,18	2,17	2,16	2,14	2,13	2,12	2,11
	0,010	3,08	3,05	3,03	3,00	2,98	2,96	2,94	2,92
20	**0,100**	1,82	1,81	1,80	1,79	1,79	1,78	1,77	1,77
	0,050	2,17	2,15	2,14	2,12	2,11	2,10	2,09	2,08
	0,010	3,02	2,99	2,96	2,94	2,92	2,90	2,88	2,86
21	**0,100**	1,80	1,79	1,78	1,78	1,77	1,76	1,75	1,75
	0,050	2,14	2,12	2,11	2,10	2,08	2,07	2,06	2,05
	0,010	2,96	2,93	2,90	2,88	2,86	2,84	2,82	2,80
22	**0,100**	1,79	1,78	1,77	1,76	1,75	1,74	1,74	1,73
	0,050	2,11	2,10	2,08	2,07	2,06	2,05	2,04	2,03
	0,010	2,91	2,88	2,85	2,83	2,81	2,78	2,77	2,75
23	**0,100**	1,77	1,76	1,75	1,74	1,74	1,73	1,72	1,72
	0,050	2,09	2,08	2,06	2,05	2,04	2,02	2,01	2,01
	0,010	2,86	2,83	2,80	2,78	2,76	2,74	2,72	2,70
24	**0,100**	1,76	1,75	1,74	1,73	1,72	1,71	1,71	1,70
	0,050	2,07	2,05	2,04	2,03	2,01	2,00	1,99	1,98
	0,010	2,82	2,79	2,76	2,74	2,72	2,70	2,68	2,66

DAFTAR PUSTAKA

Bain, L.J. and M. Engelhardt. 1987. *Introduction to Probability and Mathematical Statistics*. Duxbury Press, Boston, MA. USA.

Bickel, P.J. and K.A. Doksum. 1977. *Mathematical Statistics: Basic Ideas and Selected Topics.* Holden-Day, Inc. Oakland, CA. USA.

Brunk, H.D. 1965. *An Introduction to Mathematical Statistics.* Blaisdell Publishing Company. Waltham, MA. USA

Heathcote, C.R. 1971. *Probability: Elements of the Mathematical Theory*. John Wiley & Sons, Inc. New York, USA.

Kiefer, J.C. 1987. *Introduction to Statistical Inference*. Springer Texts in Statistics. Springer-Verlag, New York. USA.

Larsen, R.J. and M.L. Marx. 1985. *An Introduction to Probability and Its Applications.* Prentice-Hall, Inc. Englewood Cliffs, NJ. USA.

Lehmann, E.L. 1986. *Testing Statistical Hypotheses*. John Wiley & Sons. USA.

Lehmann, E.L. 1991. *Theory of Point Estimation*. Wadsworth & Brooks/ Cole Advanced Books & Software. Pacific Groove, CA. USA.

Mendenhall, W., R.L. Scheaffer, and D.D. Wackerly. 1986. *Mathematical Statistics with Applications.* Duxbury Press, Boston. USA.

Mood, A.M., F.A. Graybill, and D.C. Boes. 1974. *Introduction to the Theory of Statistics.* 3rd ed. International Student Edition. McGraw-Hill International Book Company. Singapore.

Rohatgi, V.K. 1976. *An Introduction to Probability Theory and Mathematical Statistics*. John Wiley and Sons. Singapore.

Indeks

SIGIT NUGROHO, Ph.D. (University of Kentucky-USA, 1994) dilahirkan di Surakarta pada tanggal 30 Nopember 1960. Ia menyelesaikan Pendidikan Dasar dan Menengahnya di Yogyakarta. Setelah tamat **SMA Negeri III 'Padmanaba' Yogyakarta**, ia meneruskan studinya di Institut Pertanian Bogor pada tahun 1980 melalui jalur Proyek Perintis II. Lulus sebagai **Sarjana Statistika (Ir.)** tahun 1984 dari Jurusan Statistika – **Fakultas Matematika dan Ilmu Pengetahuan Alam – Institut Pertanian Bogor (FMIPA-IPB)**. Sejak awal 1986 ia bekerja sebagai staf pengajar pada Fakultas Pertanian Universitas Bengkulu (Faperta UNIB), yang selanjutnya pada tahun 2000 pindah ke Jurusan Matematika, Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Bengkulu. Sampai buku ini ditulis, jabatan akademiknya adalah **Lektor Kepala** dalam bidang Statistika.

Pada tahun 1987 ia melanjutkan studinya di Department of Statistics, University of Kentucky, U.S.A. dan meraih gelar **Master of Science (M.Sc.**) dalam bidang Statistika pada tahun 1989. Setelah dua tahun kembali ke Universitas Bengkulu mengasuh mata kuliah Matematika I dan II, Metode Statistika I dan II, serta Rancangan Percobaan di Faperta UNIB ia kembali meneruskan studinya pada tahun 1991 ke jenjang yang lebih tinggi di tempat yang sama (Department of Statistics, University of Kentucky, U.S.A). Dibawah bimbingan **Zakkula Govindarajulu, Ph.D.** (Minnesota, 1961), ia menyelesaikan disertasinya yang berjudul "*On the Locally Most Powerful Rank Test of the Two-way Experiment*" dan dinyatakan lulus pada tanggal 15 April 1994 dihadapan tim penguji yang terdiri dari: William S. Griffith, Ph.D., William S. Rayens, Ph.D., Mokhtar Ali, Ph.D., Mai Zhou, Ph.D. dan mendapatkan gelar **Doctor of Philosophy (Ph.D.)** dalam bidang Statistika.

Pada tahun 1988 penulis mengikuti Kursus "*Analysis of Messy Data* di Washington, D.C. yang diberikan langsung oleh penulis buku tentang analisis tersebut, yaitu: George M. Milliken, Ph.D. dan Dallas T. Johnson, Ph.D.

Selain sebagai staf pengajar (**Lektor Kepala** dalam bidang Statistika) Universitas Bengkulu, ia juga sebagai dosen tamu pada program doktor di Jurusan Statistika IPB (2003) dan beberapa pendidikan tinggi lainnya. Sebagai tambahan, ia juga sebagai konsultan *Data Analysis*. Pada tahun 2003-2006 penulis juga menjadi **Senior Instruktur** pada Divisi Pendidikan dan Pelatihan **PT. Bank Rakyat Indonesia (Persero) Tbk.** Sampai dengan tahun 1997 penulis juga menjadi anggota *American Statistical Association.* Berbagai kegiatan seminar dalam bidang statistika telah diikutinya baik lokal, nasional, regional, ataupun internasional.

Beberapa Publikasi Jurnal yang berhubungan dengan bidang ilmunya:

1. Uji Nonparametrik Perlakuan Acak dalam Rancangan Acak Kelompok Lengkap. *Forum Statistika dan Komputasi IPB* (1997) **2** (1), 10-14 ISSN 0853-8115 Tests
2. Tests for Random Effects in Two-way Experiment with One Observation per Cell. *Indian Journal of Mathematics* vol **41** No 1 January 1999. B.N. Prasad Birth Centenary Commemoration Volume. (with Dr. Z. Govindarajulu, Univ of Kentucky)
3. Nonparametric tests for random effects in the balanced incomplete block design. *Statistics & Probability Letters* **56**, 431-437 2002. (with Dr. Z. Govindarajulu, Univ of Kentucky)
4. Some Notes on Nonparametric Test of Random Treatment Effects in One-way and Two-way Experiments. *Journal of Quantitative Methods* nol **3** no 2. 2007 (with Dr. Z. Govindarajulu, Univ of Kentucky)

SIGIT NUGROHO, Ph.D.
Website: http://www.stasignug.cjb.net/
Email : snugroho@unib.ac.id, snsw1960@telkom.net dan sinugsta@yahoo.com

Istilah lain Statistika Matematika adalah Teori Statistika. Sesuai dengan namanya, materi ini membahas berbagai hal yang berkenaan dengan statistik atau formula secara simbolis matematis; atau bahkan mempelajari bagaimana penurunan suatu dalil atau teorema berdasarkan suatu definisi ataupun tambahan asumsi.

Materi ini wajib diberikan untuk mahasiswa dari program studi matematika dan statistika pada fakultas matematika dan ilmu pengetahuan alam; bahkan dalam beberapa kondisi juga diberikan untuk mahasiswa pendidikan matematika pada fakultas keguruan dan ilmu pendidikan.

Diawali dengan membahas peluang, kemudian peubah acak (*random variable*), sebaran bersama, sebaran fungsi peubah acak, sifat-sifat peubah acak, dan sifat beberapa sebaran kontinu. Materi tersebut diberikan untuk satu semester.

Materi berikut ini kemudian diberikan untuk semester berikutnya : Limit Sebaran, Sebaran Nilai Ekstrim, Teori Pendugaan Titik, Statistik Cukup dan Lengkap, Pendugaan Interval dan Pengujian Hipotesis.

Dengan penyajian yang sistematis, disertai dengan beberapa teladan, diharapkan buku ini dapat dipakai seba gai rujukan atau referensi bagi siapa saja yang membutuh kan.

UNIB Press
Jalan WR Supratman
Bengkulu 38371

ISBN 978-979-9431-33-2
9 789799 431332
Pengntr Stat Matematika