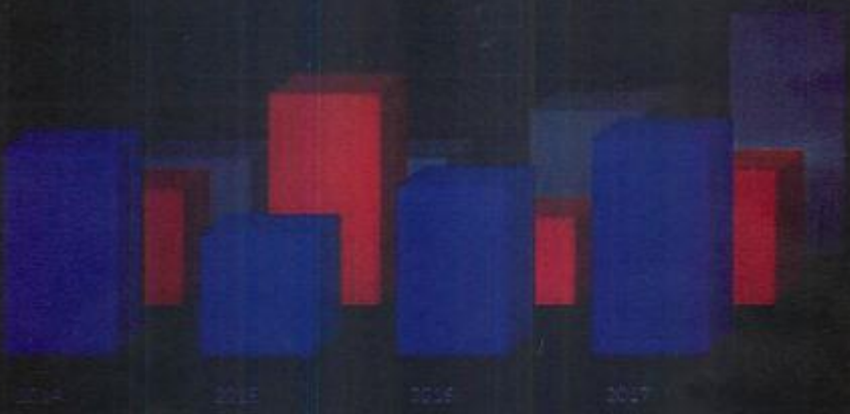
Pengantar Statistika Matematis

Book · February 2018

1 author:

Susiswo Susiswo
State University of Malang

Pengantar Statistika Matematis



Dr. Susiswo, M.Si

Penerbit & Percetakan

PENGANTAR STATISTIKA MATEMATIS

Oleh
Dr. SUSISWO, M.Si

Penerbit & Percetakan

Universitas Negeri Malang
Anggota IKAPI No. 059 / JTI / 89
Jl. Semarang 5 (Jl. Gombong 1) Malang, Kode Pos 65145
Kotak Pos 13, MLG /IKIP Telp. (0341) 562391, 551312
psw 453

Susiswo

Pengantar Statistika Matematis – Oleh: Susiswo.–Cet. I– Universitas Negeri Malang, 2017.

xvi, 365 hlm; 15,5 x 23 cm

ISBN: 978.979.495.948.0

Pengantar Statistika Matematis
Dr. Susiswo, M.Si

• Universitas Negeri Malang (UM PRESS)
d/a Penerbit IKIP Malang, Anggota IKAPI No. 059/JTI/89
Jl. Semarang 5 (Jl. Gombong 1) Malang, Kode Pos 65145
Kotak Pos 13. MLG/IKIP Telp. (0341) 562391, 551312 psw. 453

• Cetakan I : 2017

PRAKATA

Puji syukur kami panjatkan ke hadirat Tuhan Yang Maha Esa, karena hanya dengan kemurahan dan kasih-Nya buku Pengantar Statistika Matematis ini dapat diselesaikan.

Buku Pengantar Statistika Matematis ini ditulis untuk memenuhi kebutuhan pada perkuliahan Statistika Matematis Perguruan Tinggi jenjang S1. Dengan tersedianya buku ini diharapkan mahasiswa tidak menemui kesulitan dalam menempuh matakuliah Statistika Matematis. Lebih lanjut diharapkan mahasiswa dapat mengembangkan diri lebih optimal dengan mengaji pula materi-materi yang bersesuaian pada referensi-referensi lain.

Buku Pengantar Statistika Matematis ini merupakan gabungan dan perbaikan dari dua buku sebelumnya, yaitu buku Teori Peluang dan Statistika Matematis. Materi-materi pada buku Pengantar Statistika Matematis ini disusun menjadi 8 bab, dengan rincian: bab 1 berisi tentang konsep dasar peluang dan sifat-sifatnya, bab 2 berisi tentang peubah acak dan sebarannya, bab 3 berisi tentang beberapa sebaran khusus, dari sebaran diskret dan sebaran kontinu, bab 4 berisi tentang sebaran bersama dan sifat-sifat peubah acak, bab 5 berisi tentang statistic dan sebarang sampel, bab 6 berisi tentang sebaran batas, bab 7 berisi tentang teori pendugaan, dan bab 8 berisi tentang uji hipotesis. Dalam mempelajari buku ini kepada pengguna atau mahasiswa khususnya disarankan untuk mempelajarinya secara terurut dari bab pertama sampai bab terakhir.

Materi pada buku ini telah diujicobakan beberapa kali di Jurusan Matematika FMIPA Universitas Negeri Malang dan perguruan tinggi sekitarnya pada perkuliahan Statistika Matematis sebanyak dua kali. Uji coba pertama dilakukan pada tahun 2006 diikuti uji coba pada tahun-tahun berikutnya sampai dengan tahun 2016. Banyak perbaikan dilakukan baik berdasarkan masukan dari mahasiswa peserta perkuliahan, maupun masukan dari teman dosen sesama pembina matakuliah Statistika Matematis. Oleh karena itu pada kesempatan ini kami menyampaikan ucapan terima kasih kepada mereka semua atas masukannya. Secara khusus kami menyampaikan ucapan terima kasih kepada Bapak Prof. Ir. Siswadi, Ph.D dosen jurusan Statistika IPB sebagai pendamping penulisan buku ini pada proyek hibah penulisan buku Dikti tahun 2008, atas kesabaran dan kesungguhannya dalam memberikan koreksi dan masukan.

Prasyarat yang diperlukan dalam mempelajari buku ini adalah materi himpunan dan kalkulus, diantaranya materi-materi: himpunan bagian, operasi pada himpunan, fungsi titik, fungsi himpunan, limit, teknik integral, dan integral lipat. Mahasiswa diharapkan ”membuka-buka” lagi materi-materi tersebut sehingga tidak menemui masalah dalam menerapkannya pada buku ini.

Malang, April 2017

GLOSARIUM

Barisan cdf adalah suatu barisan yang terdiri dari cdf. Barisan ini digunakan untuk menentukan konvergen dalam sebaran barisan peubah acak

Barisan peubah acak adalah barisan yang terdiri dari peubah acak.

Batas atas kepercayaan merupakan nilai ujung kanan suatu selang kepercayaan.

Batas bawah kepercayaan merupakan nilai ujung kiri suatu selang kepercayaan.

Batas atas kepercayaan satu sisi merupakan batas atas selang kepercayaan di mana selang kepercayaan tersebut hanya mempunyai batas atas saja.

Batas bawah Cramer-Rao adalah batas bawah dari suatu varians yang ditentukan menggunakan rumus Cramer-Rao

Batas bawah kepercayaan satu sisi merupakan batas bawah selang kepercayaan di mana selang kepercayaan tersebut hanya mempunyai batas bawah saja.

Bebas stokastik adalah kebebasan beberapa peubah acak berdasarkan konsep peluang.

Bergantung stokastik sama artinya dengan tidak bebas stokastik.

Daerah kritis adalah suatu daerah jika sampel jatuh pada daerah kritis maka hipotesis nol ditolak. Daerah kritis ini disebut juga daerah penolakan.

Daerah kritis paling kuasa seragam adalah daerah kritis paling kuasa untuk pengujian hipotesis sederhana lawan hipotesis komposit atau hipotesis komposit lawan dengan hipotesis komposit.

Daerah kritis terbaik adalah daerah kritis yang lebih kuasa dibanding daerah kritis yang lain dengan ukuran yang sama.

Ekspektasi matematis biasa ditulis sebagai ekspektasi, merupakan nilai dari jumlah nilai peubah acak dengan peluangnya dalam hal peubah acak diskret, mengganti jumlah dengan integral dalam hal peubah acak kontinu.

Frekuensi relatif adalah perbandingan frekuensi dengan banyak percobaan.

Fungsi kepadatan peluang atau disingkat pdf merupakan suatu fungsi peluang dari suatu peubah acak.

Fungsi sebaran atau fungsi sebaran kumulatif, diberi simbol cdf, adalah jumlah kumulatif peluang mulai $-\infty$ sampai pada nilai peubah acak yang ditentukan untuk peubah acak diskret. Dan mengganti jumlah dengan integral dalam hal peubah acak kontinu.

Fungsi sebaran bersama adalah fungsi sebaran bersama dari beberapa peubah acak.

Fungsi sebaran marginal merupakan fungsi sebaran yang diperoleh dari fungsi sebaran bersama.

Fungsi kepadatan peluang bersama atau ditulis sebagai pdf bersama adalah fungsi kepadatan peluang dari beberapa peubah acak.

Fungsi pembangkit momen adalah ekspektasi eksponen dari perkalian peubah acak dengan konstanta t.

Galat jenis I adalah galat yang diakibatkan karena menolak H_0 padahal H_0 benar.

Galat jenis II adalah galat yang diakibatkan karena tidak menolak H_0 padahal H_0 salah.

Hipotesis statistika adalah pernyataan tentang sebaran dari satu atau lebih peubah acak.

Kaidah Bayes merupakan suatu kaidah yang digunakan untuk menentukan peluang bersyarat kejadian C_1 setelah munculnya kejadian C_2 ditulis $P(C_1|C_2)$.

Kejadian adalah himpunan bagian dari ruang sampel.

Kejadian saling asing adalah kejadian-kejadian yang irisannya adalah kosong.

Kejadian saling bebas adalah kejadian-kejadian yang peluang irisan kejadian tersebut sama dengan perkalian masing-masing peluang kejadian.

Kejadian sederhana adalah kejadian yang hanya mempunyai satu hasil atau unsur.

Ketaksaman Boole merupakan ketaksamaan dari peluang gabungan beberapa kejadian dengan jumlah masing-masing peluang kejadian tersebut.

Koefisien korelasi adalah koefisien dari hubungan dua peubah acak.

Konvergen dalam sebaran lihat barisan cdf.

Konvergen dalam peluang adalah suatu kenvergenan dari barisan peubah acak jika ditinjau dari peluangnya.

Koreksi kekontinuan digunakan pada perubahan dari peubah acak diskret ke peubah acak kontinu.

Lemma Neyman-Pearson adalah suatu lemma yang digunakan untuk menentukan daerah kritis terbaik.

Matriks acak adalah matriks yang unsur-unsurnya peubah acak.

Median adalah nilai tengah dari nilai-nilai peubah acak.

Modus adalah nilai yang mempunyai frekuensi tertinggi.

Model peluang klasik adalah model peluang yang digunakan untuk menentukan peluang suatu kejadian, jika masing-masing anggota kejadian tersebut mempunyai kesempatan yang sama untuk muncul, atau *equally likely*.

Nilai p adalah nilai yang digunakan untuk menentukan ditolak atau diterimanya suatu hipotesis.

Parameter adalah besaran penentu dari suatu populasi.

Parameter bentuk adalah parameter yang menentukan bentuk kurva suatu sebaran.

Peluang adalah suatu fungsi yang nilainya ditentukan oleh suatu kejadian dan memenuhi tiga asumsi peluang. Peluang ini dapat dipandang sebagai jumlah bobot masing-masing anggota kejadian yang dimaksud, jika kejadian dari ruang sampel diskret.

Peluang bersyarat adalah peluang pada suatu kejadian dengan syarat kejadian yang lainnya.

Peluang klasik lihat model klasik.

Peluang terinduksi adalah peluang yang ditimbulkan oleh suatu peubah acak.

Peluang total adalah jumlah peluang dari suatu kejadian bila kejadian tersebut bergantung pada beberapa kejadian.

Penduga adalah suatu statistik yang digunakan untuk menduga parameter.

Penduga terkonsentrasi adalah suatu penduga yang variansnya lebih kecil dibanding dengan penduga yang lain

Pendugaan metode kemungkinan maksimum adalah suatu pendugaan yang diperoleh dari memaksimumkan fungsi kemungkinan.

Pendugaan n metode momen adalah suatu pendugaan berdasarkan momen.

Penduga tak bias adalah penaksir yang ekspektasinya sama dengan parameter yang diduga.

Percobaan acak adalah suatu percobaan yang memenuhi tiga syarat, yaitu: dapat diulang di bawah kondisi yang sama, hasil belum dapat diketahui sebelum dilakukan percobaan**,** semua hasil yang mungkin dapat ditentukan.

Persentil adalah suatu nilai yang membagi data menjadi seratus bagian.

Peubah acak adalah suatu fungsi yang domainnya ruang sampel dan kodomainnya adalah himpunan bilangan real.

Peubah acak diskret adalah peubah acak yang nilainya pada himpunan diskret.

Peubah acak kontinu adalah peubah acak yang nilainya pada suatu selang.

Populasi homogen adalah populasi-populasi yang mempunyai varians sama.

Proses Poisson adalah proses yang memenuhi asumsi Poisson.

Purata adalah istilah lain dari suatu ekspektasi peubah acak. Purata ini berkaitan dengan purata populasi.

Purata sampel adalah purata dari sampel, nilainya diperoleh dari jumlah nilai anggota sampel dibagi dengan ukuran sampel.

Ruang parameter adalah himpunan semua nilai yang mungkin dari suatu parameter.

Ruang sampel adalah himpunan semua hasil yang mungkin dari suatu percobaan acak.

Sampel acak terdiri dari beberapa peubah acak yang saling bebas dan mempunyai sebaran identik.

Sebaran bersama adalah sebaran bersama dari beberapa peubah acak.

Sebaran khusus merupakan sebaran dari suatu peubah acak yang mempunyai sifat-sifat khusus di antaranya pada: pdf, cdf, ekspektasi, varians dan mgf.

Sebaran merosot merupakan suatu sebaran dari peubah acak jika fungsi sebarannya terpusat pada satu titik.

Sebaran turunan adalah suatu sebaran dari sampel acak.

Selang kepercayaan adalah suatu selang dari selang acak suatu parameter, di mana parameter itu berada dalam selang acak dengan peluang tertentu.

Sifat tanpa memori adalah suatu sifat dari peluang, di mana peluang suatu kejadian tidak tergantung pada kejadian sebelumnya.

Simpangan baku adalah ukuran simpangan suatu peubah acak dari ekspektasinya.

Statistik adalah suatu fungsi dari peubah acak yang tidak bergantung pada parameter.

Statistik urutan adalah suatu statistik yang diperoleh dari mengurutkan peubah-peubah acak yang membentuk statistik terlebih dahulu.

Taraf signifikan adalah nilai maksimum dari peluang menolak hipotesis nol padahal hipotesis nol benar.

Teorema limit pusat adalah suatu teorema untuk mendapatkan pendekatan menggunakan sebaran normal baku.

Varians adalah ekspektasi kuadrat dari selisih peubah acak dengan ekepektasinya.

Vektor acak adalah vector yang unsur-unsurnya peubah acak.

DAFTAR ISI

BAB 1
PELUANG

1.1 Pendahuluan

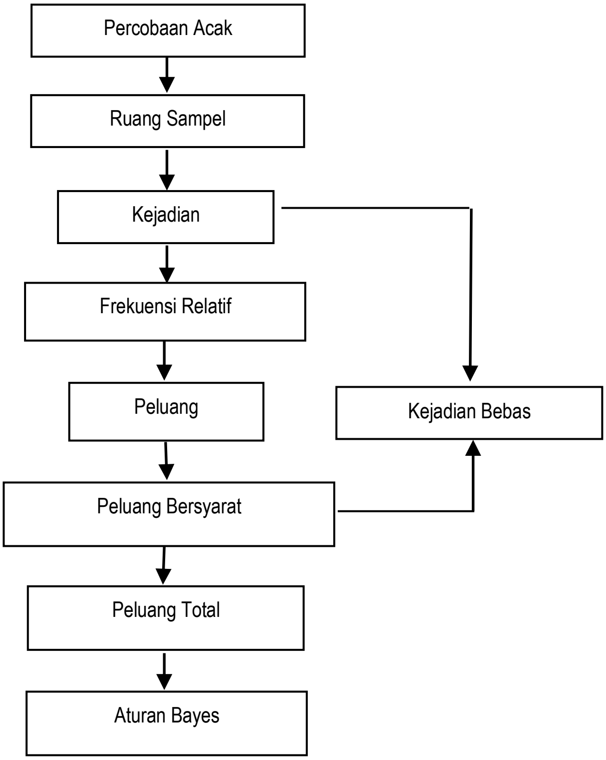
Dalam beberapa studi, ilmu pengetahuan diperoleh dari fenomena alam. Dari fenomena itu kita berusaha untuk menyederhanakannya menjadi model matematika. Dengan model tersebut memungkinkan kita untuk dapat menentukan atau memprediksi nilai pengamatan dari suatu karakteristik tertentu. Sebagai contoh, kita perhatikan kecepatan benda jatuh pada ruang hampa udara setelah t detik, yang modelnya adalah $v = gt$, di mana $g = 31.17$ kaki per detik. Model tersebut tentunya diperoleh dari percobaan yang berulang-ulang dengan kondisi yang sama. Selanjutnya dengan kondisi tadi kita dapat memprediksi kecepatan dalam sebarang t detik. Di pihak lain model tersebut tidak akan berlaku untuk kondisi yang berbeda.

Motivasi dari belajar peluang adalah menjadikan model matematika dari situasi yang belum ditentukan. Kegiatan ini disebut sebagai model peluang.

Tujuan instruksional umum dari mempelajari bab ini adalah Anda diharapkan memahami konsep tentang peluang.

Tujuan instruksional khusus dari mempelajari bab ini adalah Anda diharapkan dapat menentukan: ruang sampel suatu percobaan, kejadian suatu percobaan, frekuensi relatif suatu kejadian, peluang suatu kejadian, kekontinuan dalam peluang, peluang bersyarat suatu kejadian, peluang total suatu kejadian yang bergantung dengan beberapa kejadian yang lainnya, peluang bersyarat suatu kejadian menggunakan aturan Bayes, dan kejadian-kejadian bebas.

Sebelum membahas konsep peluang, terlebih dahulu kita lihat hubungan beberapa konsep yang akan kita bahas dalam Bab 1, yang diberikan pada skema berikut.



1.2 Ruang Sampel, Kejadian dan Frekuensi Relatif

Suatu kegiatan atau proses untuk memperoleh hasil pengamatan dari suatu fenomena tertentu disebut sebagai ***percobaan (experiment*)**. Dalam Statistika Matematis percobaan diharuskan memenuhi tiga syarat, yaitu:

1) dapat diulang di bawah kondisi yang sama,
2) ***hasil (outcome)*** belum dapat diketahui sebelum dilakukan percobaan ***(trial),***
3) semua hasil yang mungkin dapat ditentukan.

Percobaan yang demikian ini disebut sebagai ***percobaan acak.*** Untuk menyederhanakan penulisan, selanjutnya yang dimaksud percobaan adalah percobaan acak.

Definisi 1.2.1 Ruang sampel adalah himpunan semua hasil yang mungkin dari suatu percobaan, diberi notasi S.

Contoh 1.2.1 Suatu percobaan pelemparan dua keping uang logam. Jika kita perhatikan hasil muka atau belakang yang muncul, maka ruang sampelnya dapat dinyatakan sebagai pasangan berurutan berikut ini,

$$S = \{(\text{muka,muka}), (\text{muka,belakang}), (\text{belakang,muka}), (\text{belakang,belakang})\},$$

berturut-turut suku pertama dan suku kedua pasangan berurutan tersebut menunjukkan hasil dari uang logam pertama dan uang logam kedua. Secara sederhana penulisan ruang sampel di atas adalah

$$S = \{mm, mb, bm, bb\},$$

m menyatakan muka dan b menyatakan belakang uang logam. Penulisan muka atau belakang bergantung pada kesepakatan. Dapat pula dengan kesepakatan bersama kita menulis 0 jika muka pertama yang muncul dan 1 jika kedua yang muncul. Dengan demikian kita akan mendapatkan ruang sampel

$$S = \{(0,0), (0,1), (1,0), (1,1)\}.$$

Banyak anggota dari ruang sampel pada contoh di atas adalah empat, atau dapat kita katakan bahwa banyak anggota ruang sampel adalah terhingga. Tentu hasil ini tidak dapat kita perumum untuk sebarang ruang sampel. Kita lihat contoh berikut ini.

Contoh 1.2.2 Suatu percobaan pelemparan sekeping uang logam akan berhenti jika telah mendapatkan hasil muka. Maka ruang sampelnya adalah

$$S = \{m, bm, bbm, ...\}.$$

Untuk keperluan pembahasan berikutnya ada baiknya kita lihat definisi tentang himpunan terbilang berikut ini.

Definisi 1.2.2 Suatu himpunan T dikatakan ***himpunan terbilang*** jika ada fungsi satu-satu f dengan domain T dan kisaran dalam himpunan bilangan asli N.

Dari kedua contoh di atas kita mendapat kesan bahwa ruang sampel S selalu merupakan himpunan terbilang. Kesan tersebut akan hilang jika kita perhatikan Contoh 1.2.3.

Contoh 1.2.3 Seseorang menunggu angkutan di suatu pemberhentian, yang lewat tiap 10 menit sekali. Jika lama waktu menunggu dinyatakan sebagai t, maka ruang sampel dari lama waktu menunggu tersebut adalah

$$S = \{t \mid 0 \le t \le 10\}.$$

Berdasarkan contoh-contoh di atas kita mendapatkan hasil bahwa ada dua jenis ruang sampel jika dilihat dari keanggotaannya.

Definisi 1.2.3 Ruang sampel S yang merupakan himpunan terbilang dikatakan ***ruang sampel diskret***.

Definisi 1.2.4 Ruang sampel S yang anggotanya terdefinisi pada suatu selang himpunan bilangan real dikatakan ***ruang sampel kontinu***.

Selanjutnya kita akan mendefinsikan satu hal lagi yang sangat penting dalam peluang, yaitu suatu kejadian.

Definisi 1.2.5 Sebarang himpunan bagian dari ruang sampel disebut ***kejadian***.

Kejadian diberi notasi huruf kapital, sebagai contoh A, B, dan C. Pada bab ini kita akan lebih sering menggunakan simbol C daripada yang lainnya. Perlu kita sepakati pula bahwa untuk membedakan anggota ruang sampel dengan kejadian, maka anggota ruang sampel kita tulis sebagai huruf kecil.

Contoh 1.2.4 Diberikan suatu percobaan seperti pada Contoh 1.2.1. Jika kita menginginkan kejadian paling sedikit muncul satu muka, maka kita peroleh kejadian

$$C = \{mb, bm, mm\}.$$

Kita perhatikan bahwa kejadian merupakan himpunan. Oleh karena itu sifat-sifat yang berlaku pada himpunan, berlaku pula pada kejadian. Secara khusus jika suatu kejadian hanya mempunyai satu anggota, maka kejadian itu disebut kejadian sederhana, seperti yang diberikan oleh Definisi 1.2.6.

Definisi 1.2.6 Suatu kejadian dikatakan ***kejadian sederhana***, jika dia memuat tepat satu hasil dari suatu percobaan.

Ruang sampel diskret dapat kita tulis sebagai gabungan dari kejadian-kejadian sederhana. Sebagai contoh kita lihat kembali Contoh 1.2.1 dan Contoh 1.2.2. Ruang sampel pada Contoh 1.2.1 dapat kita tulis sebagai

$$S = C_1 \cup C_2 \cup C_3 \cup C_4,$$

di mana $C_1 = \{mm\}, C_2 = \{mb\}, C_3 = \{bm\},$ dan $C_4 = \{bb\},$ dan ruang sampel pada Contoh 1.2.2 dapat kita tulis sebagai

$$S = C_1 \cup C_2 \cup C_3 \cup ...$$

di mana $C_1 = \{m\}, C_2 = \{bm\}, C_3 = \{bbm\}, ...$

Konsep penting lainnya dalam teori peluang adalah konsep tentang dua himpunan yang saling asing atau saling lepas.

Definisi 1.2.7 Dua kejadian C_1 dan C_2 dikatakan ***saling asing*** jika

$$C_1 \cap C_2 = \varnothing.$$

Pada pembahasan tentang kejadian sederhana sebelumnya, jika kita mengambil sebarang dua kejadian, maka dua kejadian tersebut saling asing. Kejadian-kejadian yang demikian dikatakan kejadian saling asing, yang merupakan bentuk umum dari Definisi 1.2.7.

Definisi 1.2.8 Kejadian-kejadian $C_1, C_2, C_3, ...$ dikatakan ***saling asing*** jika sebarang pasangan dari kejadian-kejadian tersebut adalah saling asing, yaitu jika

$$C_i \cap C_j = \varnothing,$$

untuk sebarang $i \neq j.$

Satu lagi konsep penting dalam pembahasan ini, yang juga merupakan salah satu konsep munculnya gagasan peluang, yaitu konsep tentang frekuensi relatif.

Definisi 1.2.9 Misal $n(C)$ banyak muncul kejadian C dari suatu percobaan yang diulang sebanyak N kali. Maka ***frekuensi relatif*** dari kejadian $C,$ ditulis $f_R(C)$ adalah

$$f_R(C) = \frac{n(C)}{N}.$$

Contoh 1.2.5 Sekeping uang logam dilempar 50 kali, muncul muka sebanyak 18 kali. Misal C kejadian muncul muka. Maka frekuensi relatif dari kejadian C adalah

$$f_R(C) = \frac{n(C)}{N} = \frac{18}{50}.$$

Selanjutnya jika uang logam dilempar beberapa kali sehingga mendekati tak hingga, maka frekuensi relatif ini akan mendekati nilai tertentu. Nilai tertentu inilah yang disebut sebagai ***peluang***, diberi simbol p. Dalam contoh ini bila uang logam yang digunakan seimbang, maka $p = \frac{1}{2}$. Hal ini dapat dikatakan bahwa peluang kejadian C ditulis $P(C)$ adalah $p = \frac{1}{2}$. Secara matematis ditulis

$$P(C) = \lim_{N\to\infty} f_R(C) = \lim_{N\to\infty} \frac{n(C)}{N} = p = \frac{1}{2}.$$

Teorema berikut ini merupakan sifat-sifat dari frekuensi relatif yang menjadi dasar dari konsep peluang.

Teorema 1.2.1 Misal S menyatakan ruang sampel dari suatu percobaan dan $C, C_1, C_2, \ldots$ merupakan kejadian-kejadian yang mungkin dari S. Jika $f_R(C)$ menyatakan frekuensi relatif dari sebarang kejadian C, maka berlaku:

(i) $f_R(C) \geq 0$,

(ii) $f_R(S) = 1$

(iii) $f_R(C_1 \cup C_2 \cup \ldots) = f_R(C_1) + f_R(C_2) + \ldots$,
asalkan $C_1, C_2, \ldots$ merupakan kejadian-kejadian saling asing.

Bukti dari teorema di atas ditinggalkan sebagai latihan.

Latihan 1.2

1.2.1. Suatu mesin memproduksi bola yang terdiri dari bola-bola: merah, hijau dan ungu. Satu bola diambil dari mesin tersebut.
(a) Tentukan ruang sampelnya.
(b) Daftarlah semua kejadian yang mungkin.

1.2.2. Satu bola diambil secara acak dari mesin pada Latihan 1, kemudian diambil lagi satu bola. Anggap mesin menghasilkan paling sedikit dua bola dari masing-masing warna:

(a) Tentukan ruang sampelnya.

(b) Misal C_1 merupakan kejadian mendapatkan bola merah pada pengambilan pertama, dan C_2 kejadian mendapatkan paling sedikit satu bola merah pada kedua pengambilan. Tentukan C_1, C_2, $C_1 \cap C_2$, dan $C_1^C \cap C_2$.

1.2.3. Bola-bola diambil secara acak dari mesin pada Latihan 1.2.1., sampai mendapatkan warna merah. Tentukan ruang sampelnya.

1.2.4. Terdapat empat golongan darah dasar: O, A, B, dan AB. Seseorang dapat menerima darah dari golongannya sendiri, dan menerima dari golongan darah O. Sebaliknya orang yang golongan darahnya AB dapat menerima sebarang golongan darah. Suatu percobaan pengambilan darah untuk menentukan tipenya dari masing-masing dua donor yang akan dimasukkan ke bank darah.

(a) Daftarlah semua hasil yang mungkin sebagai pasangan berurutan.

(b) Daftarlah hasil-hasil yang berkorespondensi dengan kejadian bahwa pendonor kedua dapat menerima darah dari pendonor pertama.

(c) Daftarlah hasil-hasil yang berkorespondensi dengan kejadian masing-masing pendonor dapat menerima darah dari yang lainnya.

1.2.5. Suatu percobaan memilih satu titik dalam:

(a) selang $(0,1)$,

(b) daerah persegi dengan titik sudut-titik sudut yang diagonalnya $(-1,-1)$ dan $(1,1)$,

(c) daerah lingkaran dengan pusat $(1,2)$ dan jari-jari 3.

Tentukan masing-masing ruang sampelnya.

1.2.6. Buktikan Teorema 1.2.1.

1.3 Peluang

Pada Subpokok Bahasan sebelumnya kita telah menyinggung tentang pendefinisian peluang melalui frekuensi relatif. Pendefinisian peluang ini jelas tidak praktis, karena untuk mengetahui peluang suatu kejadian kita harus mengulang percobaan berkali-kali sampai mendapatkan suatu nilai tertentu. Oleh karena itu perlu adanya definisi yang memungkinkan kita bekerja dalam peluang tanpa harus mengulang percobaan berkali-kali. Karena kita telah mempunyai definisi peluang dari frekuensi relatif, maka definisi peluang kita nanti haruslah memenuhi sifat-sifat serupa dengan sifat-sifat pada frekuensi relatif. Dengan kata lain sifat-sifat pada frekuensi relatif kita adopsi untuk pendefinisian peluang. Definisi peluang disajikan setelah defnisi σ - *field* berikut ini.

Definisi 1.3.1 Misal **B** merupakan koleksi dari himpunan-impunan bagian dari ruang sampel S. Kita katakan **B** adalah ***sigma field*** ditulis $\sigma - field$jika

(i) $\emptyset \in \mathbf{B}$,

(ii) Jika $C \in \mathbf{B}$ maka $C^C \in \mathbf{B}$

(iii) Jika barisan dari himpunan-himpunan $\{C_1, C_2, C_2, ...\}$ dalam **B**
,
maka $\bigcup_{i=1}^{\infty} C_i \in \mathbf{B}$.

Catatan oleh (i) dan (ii), σ - *field* selalu memuat $\emptyset$ dan S. Oleh (ii) dan (iii), dan hokum DeMorgan,s tertutup di bawah irisan terhitung.

Dua contoh dari σ - *field* diberikan di bawah ini.

1. Misal C sebarang himpunan dan missal $C \subset S$. Maka $\mathbf{B} = \{C, C^C, \emptyset, S\}$ adalah σ - *field.*
2. Misal S sebarang himpunan dan missal **B** adalah himpunan kuasan dari S, yaitu, koleksi dari semua himpunan bagian dari S. Maka **B** adalah σ - *field.*

Definisi 1.3.2 Misal S menyatakan ruang sampel dan misal **B** adalah σ - *field* pada S. Misal P suatu fungsi bernilai real yang terdefinisi pada **B**. Maka P disebut sebagai ***fungsi himpunan peluang***, jika memenuhi sifat-sifat berikut ini:

(i) $P(C) \geq 0,$ untuk semua $C \in \mathbf{B}$,

(ii) $P(S) = 1,$

(iii) $P(C_1 \cup C_2 \cup C_3 \cup ...) = P(C_1) + P(C_2) + P(C_3) + ...,$
asalkan $C_1, C_2, C_3 ...$ himpunan-himpunan dalam **B** dan saling asing.

Untuk menyederhanakan penulisan fungsi himpunan peluang ditulis sebagai ***peluang***, dan bilangan real $P(C)$ kita sebut sebagai peluang dari kejadian C.

Akibat 1.3.1

(i) $P(\varnothing) = 0.$

(ii) Jika $C_1, C_2, ..., C_n$ kejadian-kejadian saling asing, maka
$P(C_1 \cup C_2 \cup ... \cup C_n) = P(C_1) + P(C_2) + ... + P(C_n).$

Bukti diberikan sebagai Latihan 1.3.8.

Contoh 1.3.1 Keberhasilan penyelesaian rancangan konstruksi memerlukan bekerjanya suatu peralatan dengan sempurna. Anggap bahwa salah satu kejadian rancangan sukses dinyatakan (C_1) atau kejadian rancangan gagal karena satu dari berikut ini: mekanik gagal (C_2) atau listrik gagal (C_3). Misal kesempatan mekanik gagal adalah tiga kali dari kesempatan listrik gagal, dan kesempatan rancangan sukses adalah dua kali dari kesempatan mekanik gagal. Hitung peluang kejadian rancangan gagal.

Penyelesaian

Karena kesempatan mekanik gagal adalah tiga kali dari kesempatan listrik gagal, dan kesempatan rancangan sukses adalah dua kali dari kesempatan mekanik gagal, maka kita peroleh model peluang

$$P(C_2) = 3P(C_3), P(C_1) = 2P(C_2).$$

Menggunakan Definisi 1.3.2 bagian (ii) dan Akibat 1.3.1 bagian (ii) untuk $n = 3$ kita peroleh

$$P(C_1) + P(C_2) + P(C_3) = 1.$$

Pemecahan dari kedua persamaan di atas adalah

$$P(C_1)=0.6,\ P(C_2)=0.3,\ \text{dan}\ P(C_3)=0.1.$$

Sehingga peluang kejadian rancangan gagal adalah

$$P(C_2\cup C_3)=P(C_2)+P(C_3)=0.3+0.1=0.4.$$

Dalam banyak permasalahan, masing-masing anggota sampel mempunyai kesempatan yang sama untuk muncul ***(equally likely)***. Dengan kata lain peluang masing-masing kejadian sederhananya adalah sama. Model seperti ini sering disebut sebagai ***model peluang klasik.***

Contoh 1.3.2 Misal dua keping uang logam dilempar sekali dan misal uang logam tersebut seimbang. Jika C kejadian muncul sedikitnya satu muka, maka hitunglah $P(C)$.

Penyelesaian

Ruang sampel dari percobaan ini adalah $S=\{mm,mb,bm,bb\}$. Dengan menganggap uang logam tersebut seimbang, maka kita dapat menentukan bahwa kesempatan masing-masing anggota sampel untuk muncul adalah sama. Menggunakan Definisi 1.3.2 bagian (ii) dan Akibat 1.3.1 bagian (ii) kita peroleh

$$P(\{mm\})=P(\{mb\})=P(\{bm\})=P(\{bb\})=\frac{1}{4}.$$

Karena C kejadian muncul sedikitnya satu muka, maka $C=\{mm,mb,bm\}$. Oleh karena itu

$$\begin{aligned}P(C)&=P(\{mm,mb,bm\})=P(\{mm\})+P(\{mb\})+P(\{bm\})\\&=\frac{1}{4}+\frac{1}{4}+\frac{1}{4}=\frac{3}{4}.\end{aligned}$$

Kebanyakan aplikasi dari model peluang klasikal muncul pada pemilihan objek atau himpunan dari objek-objek secara acak.

Definisi 1.3.3 Jika suatu objek dipilih dari sekumpulan objek di mana masing-masing objek mempunyai kesempatan yang sama untuk dipilih, maka dikatakan bahwa objek telah dipilih ***secara acak***.

Contoh 1.3.3 Misal satu kartu diambil dari satu pak kartu *bridge* secara acak. Maka kesempatan masing-masing kartu untuk muncul adalah sama, yaitu $\frac{1}{52}$. Misal C kejadian muncul kartu hati. Karena banyak kartu hati ada 13, maka $P(C) = \frac{13}{52}$.

Pembahasan berikutnya adalah sifat-sifat dari peluang, sebagian buktinya dijadikan sebagai latihan.

Teorema 1.3.1 Misal C adalah sebarang kejadian pada ruang sampel S, dan C^C adalah komplemen dari C. Maka

$$P(C^C) = 1 - P(C).$$

Bukti

Karena C^C komplemen dari C maka $C^C \cup C = S$. Dan karena $C^C \cap C = \varnothing$, maka C^C dan C saling asing. Oleh karena itu menggunakan Definisi 1.3.2 bagian (ii) dan Akibat 1.3.1 bagian (ii) untuk $n = 2$ kita peroleh

$$1 = P(S) = P(C^C \cup C) = P(C^C) + P(C).$$

Jadi $P(C^C) = 1 - P(C)$.

Catatan: Teorema di atas dapat pula dinyatakan sebagai

$P(C) = 1 - P(C^C)$.

Kembali ke Contoh 1.3.1, karena kejadian rancangan gagal dan kejadian rancangan sukses saling berkomplemen, maka

$$P(C_2 \cup C_3) = 1 - P(C_1) = 1 - 0.6 = 0.4.$$

Hasil ini persis sama dengan hasil pada Contoh 1.3.1.

Contoh 1.3.4 Misal sekeping uang logam yang seimbang dilempar empat kali, dan misal C adalah kejadian muncul paling sedikit satu muka. Hitunglah $P(C)$.

Penyelesaian

Untuk menghitung $P(C)$ akan lebih mudah jika kita hitung terlebih dahulu $P(C^C)$, di mana C^C adalah kejadian muncul tidak satupun muka, yaitu $C^C = \{bbbb\}$. Kita akan dengan mudah mendapatkan bahwa $n(S) = 16.$ Jadi

$$P(C^C) = \frac{1}{16.}$$

Oleh karena itu

$$P(C) = 1 - P(C^C) = 1 - \frac{1}{16} = \frac{15}{16}.$$

Teorema 1.3.2 Untuk sebarang kejadian $C,$ berlaku

$$P(C) \leq 1.$$

Bukti dari teorema di atas ditinggalkan sebagai latihan.

Teorema 1.3.3 Untuk sebarang kejadian-kejadian C_1 dan C_2 berlaku

$$P(C_1 \cup C_2) = P(C_1) + P(C_2) - P(C_1 \cap C_2).$$

Bukti

Untuk membuktikan teorema di atas kita ubah terlebih dahulu suatu kejadian menjadi gabungan dari dua kejadian yang saling asing, yaitu

$$C_1 \cup C_2 = (C_1 \cap C_2^C) \cup C_2,$$

dan

$$C_1 = (C_1 \cap C_2) \cup (C_1 \cap C_2^C).$$

Oleh karena itu

$$P(C_1 \cup C_2) = P(C_1 \cap C_2^C) + P(C_2),$$

dan

$$P(C_1) = P(C_1 \cap C_2) + P(C_1 \cap C_2^C).$$

Substitusi persamaan terakhir ke persamaan sebelumnya kita peroleh

$$P(C_1 \cup C_2) = P(C_1) + P(C_2) - P(C_1 \cap C_2).$$

Contoh 1.3.5 Misal satu kartu diambil dari satu pak kartu *bridge* secara acak. Hitung peluang kejadian mendapatkan kartu hati atau kartu as.

Penyelesaian

Misal C_1 kejadian mendapatkan kartu hati dan C_2 kejadian mendapatkan kartu as. Maka $P(C_1)=\frac{13}{52}$, $P(C_2)=\frac{4}{52}$, dan $P(C_1 \cap C_2)=\frac{1}{52}$. Oleh karena itu peluang kejadian mendapatkan kartu hati atau kartu as adalah

$$P(C_1 \cup C_2)=P(C_1)+P(C_2)-P(C_1 \cap C_2)=\frac{13}{52}+\frac{4}{52}-\frac{1}{52}=\frac{16}{52}.$$

Akibat 1.3.2 Untuk sebarang dua kejadian C_1 dan C_2 berlaku

$$P(C_1 \cup C_2) \le P(C_1)+P(C_2).$$

Bukti

Dari Teorema 1.3.3 kita peroleh

$$P(C_1 \cup C_2)=P(C_1)+P(C_2)-P(C_1 \cap C_2).$$

Karena peluang suatu kejadian paling kecil nol, atau

$$P(C_1 \cap C_2) \ge 0,$$

maka

$$P(C_1 \cup C_2) \le P(C_1)+P(C_2).$$

Teorema berikut ini merupakan perluasan dari teorema sebelumnya, dan buktinya ditinggalkan sebagai latihan.

Teorema 1.3.4 Untuk sebarang kejadian-kejadian $C_1, C_2,$ dan C_3 berlaku

$$\begin{aligned}P(C_1 \cup C_2 \cup C_3) = P(C_1)+P(C_2)+P(C_3)\\ -P(C_1 \cap C_2)-P(C_1 \cap C_3)-P(C_2 \cap C_3)\\ +P(C_1 \cap C_2 \cap C_3).\end{aligned}$$

Secara intuisi kita dapat menerima bahwa jika setiap hasil dari kejadian C_1 juga hasil dari kejadian C_2, maka peluang C_1 muncul tidak akan lebih besar dari pada peluang C_2 untuk muncul. Secara formal diberikan oleh teorema berikut ini, dan buktinya dijadikan sebagai latihan.

Teorema 1.3.5 Misal C_1 dan C_2 dua kejadian sehingga $C_1 \subset C_2$. Maka

$$P(C_1) \le P(C_2).$$

Secara umum, jika $C_1, C_2, C_3, \ldots$ merupakan barisan kejadian, di mana $C_n \subset C_{n+1}$, maka barisan yang demikian dikatakan barisan kejadian tidak turun, dan $\lim_{n\to\infty} C_n$ didefinisikan sebagai $\bigcup_{n=1}^{\infty} C_n$, atau ditulis sebagai

$$\lim_{n\to\infty} C_n = \bigcup_{n=1}^{\infty} C_n.$$

Di pihak lain, jika $C_n \supset C_{n+1}$, maka barisan kejadian yang demikian dikatakan barisan kejadian tidak naik, dan $\lim_{n\to\infty} C_n$ didefinisikan sebagai $\bigcap_{n=1}^{\infty} C_n$, atau ditulis sebagai

$$\lim_{n\to\infty} C_n = \bigcap_{n=1}^{\infty} C_n.$$

Kita perhatikan $\lim_{n\to\infty} P(C_n)$. Pertanyaannya adalah: dapatkah limit dan P kita tukar? Jawaban dari pertanyaan ini disajikan pada teorema berikut, yaitu teorema tentang kekontinuan dalam peluang, yang akan kita gunakan pada pembuktian ketaksamaan Boole. Selain itu teorema ini juga akan digunakan pada pembuktian sifat-sifat fungsi sebaran yang akan dibahas pada Bab 2.

Teorema 1.3.6 Kekontinuan dalam Peluang.

(i) Misal $\{C_n\}$ merupakan barisan kejadian tidak turun. Maka

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

(ii) Misal $\{C_n\}$ merupakan barisan kejadian tidak naik. Maka

$$\lim_{n\to\infty} P(C_n) = P(\lim_{n\to\infty} C_n) = P\left(\bigcap_{n=1}^{\infty} C_n\right).$$

Bukti

Kita buktikan bagian (i) dan ditinggalkan sebagai Latihan bagian (ii). Untuk membuktikannya terlebih dahulu $\bigcup_{n=1}^{\infty} C_n$ kita bentuk menjadi gabungan tak hingga kejadian yang saling asing, misal $\bigcup_{n=1}^{\infty} B_n$, di mana $B_1, B_2, ...$ adalah kejadian-kejadian yang saling asing. Kita ambil $B_1 = C_1$, dan $B_2 = C_2 \cap C_1^C$. Maka

$$C_1 \cup C_2 = B_1 \cup B_2$$

dan

$$B_1 \cap B_2 = \varnothing.$$

Kita peroleh juga hasil

$$B_1 \subset C_1, B_2 \subset C_2,$$

sehingga

$$P(B_1) \leq P(C_1),\ \ P(B_2) \leq P(C_2).$$

Oleh karena itu

$$P(C_1 \cup C_2) = P(B_1 \cup B_2) = P(B_1) + P(B_2) \leq P(C_1) + P(C_2).$$

Secara umum jika kita ambil

$$B_n = C_n \cap \left(\bigcup_{j=1}^{n-1} C_j \right)^C,$$

maka

$$B_n \subset C_n, n = 1, 2, 3, ...,$$

$$\bigcup_{n=1}^{\infty} C_n = \bigcup_{n=1}^{\infty} B_n$$

dan

$B_1, B_2, B_3, ...$ kejadian-kejadian saling asing. Oleh karena itu

$$P\left[\lim_{n \to \infty} C_n \right] = P\left(\bigcup_{n=1}^{\infty} C_n \right) = P\left(\bigcup_{n=1}^{\infty} B_n \right) = \sum_{n=1}^{\infty} P(B_n) = \lim_{n \to \infty} \sum_{j=1}^{n} P(B_j)$$

$$= \lim_{n \to \infty} \left\{ P(C_1) + \sum_{j=2}^{n} \left[P(C_j) - P(C_{j-1}) \right] \right\} = \lim_{n \to \infty} P(C_n).$$

Teorema 1.3.7 Ketaksamaan Boole. Misal $C_1, C_2, C_3, ...$ sebarang barisan kejadian. Maka

$$P\left(\bigcup_{n=1}^{\infty} C_n \right) \leq \sum_{n=1}^{\infty} P(C_n).$$

Bukti

Misal $D_n = \bigcup_{i=1}^{n} C_i$. Maka $D_1, D_2, D_3, ...$ merupakan barisan kejadian tidak turun. Untuk semua j, $D_j = D_{j-1} \cup C_j$. Oleh karena itu menggunakan Akibat 1.3.2 kita peroleh

$$P(D_j) \le P(D_{j-1}) + P(C_j),$$

atau

$$P(D_j) - P(D_{j-1}) \le P(C_j).$$

Jadi

$$P\left(\bigcup_{n=1}^{\infty} C_n\right) = P\left(\bigcup_{n=1}^{\infty} D_n\right) = \lim_{n\to\infty}\left\{P(D_1) + \sum_{j=2}^{n}\left[P(D_j) - P(D_{j-1})\right]\right\}$$

$$\le \lim_{n\to\infty}\sum_{j=1}^{n} P(C_j) = \sum_{n=1}^{\infty} P(C_n).$$

Ruas kanan Ketaksamaan Boole disebut sebagai batas peluang dari peluang gabungan sebarang kejadian C_i.

Latihan 1.3

1.3.1. Misal suatu mesin menghasilkan 100 bola karet, yang terdiri dari 20 merah (M), 30 putih (T) dan sisanya adalah hijau (H).

(a) Dapatkah kita gunakan model peluang untuk $P(M) = 0.2$, $P(T) = 0.3$, dan $P(H) = 0.5$?

(b) Dapatkah kita gunakan model peluang jika mesin juga menghasilkan bola kuning (K) sehingga $P(M) = 0.2$, $P(T) = 0.3$, $P(H) = 0.5$, dan $P(K) = 0.1$?

1.3.2 Misal suatu percobaan hanya menghasilkan tiga kejadian C_1, C_2 dan C_3. Tentukan masing-masing peluangnya jika:

(a) $P(C_1) = P(C_2) = P(C_3)$.

(b) $P(C_1) = P(C_2)$, dan $P(C_3) = \frac{1}{2}$.

(c) $P(C_1)=2P(C_2)=3P(C_3).$

1.3.3. Misal suatu ruang sampel $S=\{c:0<c<\infty\}$ dan untuk sebarang $C\subset S$ fungsi P didefinisikan sebagai $P(C)=\int_C e^{-x}dx.$

(a) Tunjukkan bahwa fungsi P mendefinisikan suatu peluang.
(b) Misal $C=\{c:4<c<\infty\}$. Tentukan $P(C)$, $P(C^C)$, dan $P(C\cup C^C)$.
1.3.4. Misal suatu ruang sampel $S=\{c:-\infty<c<\infty\}$ dan untuk sebarang $C\subset S$ fungsi Q didefinisikan sebagai $Q(C)=\int_C e^{-|x|}dx.$ Tunjukkan bahwa fungsi Q tidak mendefinisikan sebagai peluang.

1.3.5. Misal suatu percobaan melempar sekeping uang logam sebanyak empat kali. Hitunglah peluang untuk mendapatkan:
(a) tepat tiga muka,
(b) paling sedikit satu muka,
(c) banyak muka dan banyak belakang sama,
(d) banyak muka kurang dari banyak belakang,
(e) bergantian muka dan belakang atau belakang dan muka.

1.3.6. Ulangilah Latihan 1.3.5 untuk:
(a) muka muncul pada pelemparan pertama,
(b) muka muncul pada pelemparan terakhir,
(c) muka pertama pada pelemparan terakhir,
(d) muka kedua pada pelemparan terakhir,
(e) muka ketiga pada pelemparan terakhir.

1.3.7. Ulangilah Latihan 1.3.5 sampai pada n pelemparan, dan hitung peluang untuk mendapatkan:
(a) muka pertama pada pelemparan terakhir,
(b) muka ke k, di mana $k<n$, pada pelemparan terakhir.

1.3.8. Buktikan Akibat 1.3.1, menggunakan Definisi 1.3.2.

1.3.9. Buktikan Teorema 1.3.2 menggunakan :
(a) Teorema 1.3.1 dan Definisi 1.3.2 bagian (i).

(b) Teorema 1.3.5.

1.3.10. Untuk sebarang kejadian C buktikan bahwa $0 \le P(C) \le 1$.

1.3.11 Buktikan Teorema 1.3.4.

1.3.12. Buktikan Teorema 1.3.5.

1.3.13. Misal C_1 dan C_2 sebarang dua kejadian. Tunjukkan bahwa:

(a) $P(C_1 \cap C_2^C) = P(C_1) - P(C_1 \cap C_2)$.

(b) $P(C_1 \cup C_2) = 1 - P(C_1^C \cap C_2^C)$.

1.3.14. Buktikan Teorema 1.3.6 bagian (ii).

1.3.15. Anggap $P(C_i) = \dfrac{1}{3+i}$, $i = 1, 2, 3, 4$. Tentukan batas atas dari $P(C_1 \cup C_2 \cup C_3 \cup C_4)$.

1.3.16. Suatu perkumpulan terdiri dari 8 pasang suami istri. Dipilih 3 pengurus yang terdiri dari: ketua, sekretaris dan bendahara. Hitung peluang bahwa tidak ada pengurus dari pasangan yang sama.

1.3.17. Perhatikan Contoh 1.2.2.

(a) Jika C kejadian percobaan berakhir pada pelemparan ke 4, hitung $P(C)$.

(b) Ulangi (a) untuk pelemparan ke 5.

(c) Ulangi (a) untuk pelemparan ke 6.

(d) Buktikan bahwa $P(S) = 1$.

1.4 Peluang Bersyarat

Kebanyakan tujuan model peluang adalah menentukan bagaimana kejadian C_2 terjadi jika percobaan dilakukan. Meskipun demikian banyak juga kasus kejadian C_2 baru akan terjadi jika kejadian C_1 terjadi lebih dahulu. Dalam hal ini sama saja dengan menentukan peluang bersyarat kejadian C_2 setelah munculnya kejadian C_1, ditulis sebagai

$P(C_2 \mid C_1)$. Permasalahan terakhir yang akan kita bahas pada subpokok bahasan ini. Di samping itu pada subpokok bahasan ini kita juga akan membahas bagaimana menentukan peluang bersyarat kejadian C_1 setelah munculnya kejadian C_2 ditulis $P(C_1 \mid C_2)$ yang disebut sebagai Kaidah Bayes.

Sekarang marilah kita lihat kembali peluang dari sebarang kejadian C, yaitu $P(C)$. Peluang ini dapat kita pandang sebagai peluang bersyarat dari kejadian C relatif terhadap ruang sampel S, yaitu

$$P(C) = P(C \mid S).$$

Di lain pihak

$$P(C) = \frac{P(C \cap S)}{P(S)}.$$

Oleh karena itu kita memperoleh hubungan

$$P(C \mid S) = \frac{P(C \cap S)}{P(S)}.$$

Dari hasil ini kita dapat menduga bahwa peluang bersayarat dari kejadian C_2 relatif terhadap kejadian C_1 dapat kita rumuskan sebagai

$$P(C_2 \mid C_1) = \frac{P(C_2 \cap C_1)}{P(C_1)}.$$

Perumusan di atas tentunya perlu kita periksa apakah memenuhi asumsi-asumsi pada definisi peluang atau tidak, yaitu apakah:

(i) $P(C_2 \mid C_1) \geq 0$?

(ii) $P(C_1 \mid C_1) = 1$?

(iii) $P(C_2 \cup C_3 \cup \ldots \mid C_1) = P(C_2 \mid C_1) + P(C_3 \mid C_1) + \ldots$, jika $C_2, C_3, \ldots$ kejadian-kejadian yang saling asing?

Definisi 1.4.1 Peluang bersyarat sebarang kejadian C_2 setelah munculnya kejadian C_1 adalah

$$P(C_2 \mid C_1) = \frac{P(C_2 \cap C_1)}{P(C_1)},$$

asalkan $P(C_1) > 0$.

Kita ingat pula bahwa jika masing-masing anggota sampel mempunyai kesempatan yang sama untuk muncul, maka kita mempunyai hubungan

$$P(C) = \frac{n(C)}{n(S)}.$$

Hasil di atas dapat pula ditulis sebagai

$$P(C) = P(C \mid S) = \frac{n(C \cap S)}{n(S)}.$$

Hal ini tentunya berlaku pula untuk peluang bersyarat, yaitu

$$P(C_2 \mid C_1) = \frac{n(C_2 \cap C_1)}{n(C_1)}.$$

Lebih lanjut

$$\frac{n(C_2 \cap C_1)}{n(C_1)} = \frac{n(C_2 \cap C_1)/n(S)}{n(C_1)/n(S)} = \frac{P(C_2 \cap C_1)}{P(C_1)}$$

Hasil ini persis sama dengan Definisi 1.4.1.

Contoh 1.4.1 Suatu kotak berisi 100 mikrocip rusak dan bagus yang dihasilkan oleh pabrik 1 dan pabrik 2. Misal C_1 kejadian untuk memperoleh mikrocip rusak, konsekuensinya C_1^C kejadian untuk memperoleh mikrocip bagus. Misal C_2 kejadian mikrocip diperoleh dari pabrik 1 dan C_2^C kejadian mikrocip diperoleh dari pabrik 2. Daftar mikrocip selengkapnya seperti diberikan oleh Tabel 1.4.1 berikut ini.

TABEL 1.4.1

	C_2	C_2^C	Total
C_1	15	5	20
C_1^C	45	35	80
Total	60	40	100

Suatu percobaan mengambil satu mikrocip secara acak dari kotak tersebut. Tentukan:

(i) $P(C_1)$

(ii) $P(C_1 \mid C_2)$

Penyelesaian

Karena masing-masing anggota sampel mempunyai kesempatan sama untuk muncul, maka

$$P(C_1) = \frac{n(C_1)}{n(S)} = \frac{20}{100} = 0.2.$$

Terdapat beberapa cara untuk meyelesaikan permasalahan kita ini. Pertama seperti berikut ini.

$$P(C_1 \mid C_2) = \frac{n(C_1 \cap C_2)}{n(C_2)} = \frac{15}{60} = 0.25.$$

Kedua kita pandang kejadian C_2 ruang sampel yang baru. Sehingga menentukan $P(C_1 \mid C_2)$ sama saja dengan menentukan peluang dari kejadian C_1 relatif terhadap ruang sampel C_2. Karena banyak anggota kejadian C_1 pada ruang sampel C_2 adalah 15 dan banyak anggota ruang sampel C_2 adalah 60, maka

$$P(C_1 \mid C_2) = \frac{15}{60} = 0.25.$$

Sedangkan cara yang ketiga dapat kita kerjakan menggunakan rumus pada Definisi 1.4.1. Dalam contoh ini dengan cara ketiga tentunya diperlukan prosedur yang lebih panjang dan kita akan mendapatkan hasil seperti pada cara pertama.

Contoh 1.4.2 Misal suatu kantong berisi 5 bola merah dan 4 bola putih. Dua bola diambil dari kantong tersebut secara berurutan tanpa pengembalian. Hitung peluang bola kedua putih dengan syarat bola pertama merah.

Penyelesaian

Misal C_1 kejadian pengambilan bola pertama merah, dan C_2 kejadian pengambilan bola kedua putih. Menghitung peluang bola kedua putih dengan syarat bola pertama merah sama saja dengan menghitung $P(C_2 \mid C_1).$ Bola kedua putih dengan syarat bola pertama merah sama artinya dengan kejadian C_1 sudah dilakukan. Oleh karena itu dalam kantong tinggal 4 bola merah dan 4 bola putih. Jadi

$$P(C_2 \mid C_1) = \frac{4}{8} = 0.5.$$

Teorema berikut ini merupakan peluang dari sebarang dua kejadian, yang merupakan hasil kali peluang kejadian tertentu dan peluang kejadian bersyaratnya.

Teorema 1.4.1 Untuk sebarang dua kejadian C_1 dan $C_2,$ berlaku

$$P(C_1 \cap C_2) = P(C_1)P(C_2 \mid C_1).$$

Bukti dari teorema ini kita peroleh dari Definisi 1.4.1, dengan mengalikan kedua ruas dengan $P(C_1).$

Kembali ke Contoh 1.4.2, jika kita ingin menentukan peluang mendapatkan bola merah pada pengambilan pertama dan bola putih pada pengambilan kedua, maka menggunakan rumus pada Teorema 1.4.1 kita peroleh

$$P(C_1 \cap C_2) = P(C_1)P(C_2 \mid C_1) = \frac{5}{9}.\frac{4}{8} = \frac{5}{18}.$$

Teorema berikut ini merupakan perluasan dari Teorema 1.4.1, dan bukti ditinggalkan sebagai latihan.

Teorema 1.4.2 Untuk sebarang kejadian-kejadian C_1, C_2 dan C_3 berlaku

$$P(C_1 \cap C_2 \cap C_3) = P(C_1)P(C_2 \mid C_1)P(C_3 \mid C_1 \cap C_2).$$

Kembali lagi ke Contoh 1.4.2, jika kita ingin mengambil tiga bola berturutan secara acak tanpa pengembalian dan kita ingin menentukan peluang mendapatkan bola merah pada pengambilan pertama, bola putih pada pengambilan kedua, dan bola putih pada pengambilan ketiga maka menggunakan rumus pada Teorema 1.4.2 kita peroleh

$$P(C_1 \cap C_2 \cap C_3) = P(C_1)P(C_2 \mid C_1)P(C_3 \mid C_1 \cap C_2)$$

$$= \frac{5}{9} \cdot \frac{4}{8} \cdot \frac{3}{7} = \frac{5}{56}.$$

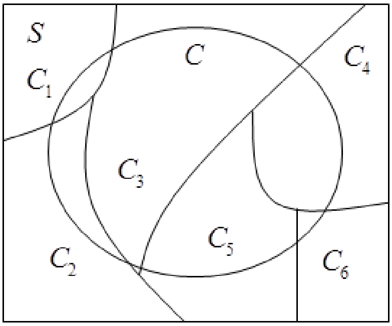
Kita telah membahas peluang bersyarat dalam hal menentukan peluang kejadian kedua yang terjadi kemudian dengan syarat kejadian pertama muncul terlebih dahulu. Sekarang bagaimana jika urutan kejadian bersyaratnya dibalik, yaitu menentukan peluang bersyarat dari kejadian pertama dengan syarat kejadian kedua muncul terlebih dahulu, padahal kejadian kedua ini belum terjadi. Seperti telah kita kemukakan pada pendahuluan subpokok bahasan ini bahwa aturan yang berkaitan dengan peluang bersyarat ini disebut sebagai Kaidah Bayes. Untuk hal ini kita kaji terlebih dahulu konsep tentang peluang total, karena konsep ini akan kita gunakan untuk keperluan yang dimaksud.

Teorema 1.4.3 Peluang Total. Misal $C_1, C_2, \ldots, C_n$ koleksi dari kejadian-kejadian yang saling asing dan menyeluruh (*exhaustive*). Maka untuk sebarang kejadian C berlaku

$$P(C) = \sum_{i=1}^{n} P(C_i)P(C \mid C_i).$$

Bukti

Kejadian-kejadian pada Teorema 1.4.3 jika kita gambar dengan diagram Venn untuk beberapa kejadian akan terlihat seperti Gambar 1.4.1 berikut ini.



Gambar 1.4.1

Karena $C_1, C_2, ..., C_n$ koleksi dari kejadian-kejadian yang saling asing dan menyeluruh, maka

$$C_i \cap C_j = \varnothing, i \neq j,$$

dan

$$S = C_1 \cup C_2 \cup ... \cup C_n.$$

Selanjutnya kita pandang kejadian C sebagai irisan dari dia sendiri dengan ruang sampel S, dan dengan mensubstitusi persamaan terakhir di atas maka akan kita peroleh

$$\begin{aligned} C &= C \cap S = C \cap (C_1 \cup C_2 \cup ... \cup C_n) \\ &= (C \cap C_1) \cup (C \cap C_1) \cup ... \cup (C \cap C_n). \end{aligned}$$

Selain kita peroleh hasil persamaan di atas kita peroleh pula hasil bahwa

$$\begin{aligned} (C \cap C_i) \cap (C \cap C_j) &= C \cap (C_i \cap C_j) \\ &= C \cap \varnothing, i \neq j \\ &= \varnothing. \end{aligned}$$

Oleh karena itu

$$P(C) = \sum_{i=1}^{n} P(C \cap C_i) = \sum_{i=1}^{n} P(C_i) P(C \mid C_i).$$

Contoh 1.4.3 Kita perluas Contoh 1.4.1 seperti berikut ini. Suatu kotak berisi 130 mikrocip rusak dan bagus yang dihasilkan oleh pabrik 1, pabrik 2 dan pabrik 3. Misal C kejadian untuk memperoleh mikrocip rusak, konsekuensinya C^C kejadian untuk memperoleh mikrocip bagus. Misal C_1, C_2 dan C_3 berturut-turut kejadian mikrocip yang diperoleh dari pabrik 1, 2 dan 3. Daftar mikrocip selengkapnya seperti diberikan oleh Tabel 1.4.2 berikut ini.

TABEL 1.4.2

	C_1	C_2	C_3	Total
C	15	5	10	30
C^C	45	35	20	100
Total	60	40	30	130

Suatu percobaan mengambil satu mikrocip secara acak dari kotak tersebut. Menggunakan peluang total hitung $P(C)$.

Penyelesaian

$$P(C)=\sum_{i=1}^{3}P(C_i)P(C\,|\,C_i)=P(C_1)P(C\,|\,C_1)+P(C_2)P(C\,|\,C_2)+P(C_3)P(C\,|\,C_3)$$

$$=\frac{60}{130}.\frac{15}{60}+\frac{40}{130}.\frac{5}{40}+\frac{30}{130}.\frac{10}{30}=\frac{15}{130}+\frac{5}{130}+\frac{10}{130}=\frac{30}{130}.$$

Hasil ini persis sama jika kita menggunakan rumus

$$P(C)=\frac{n(C)}{n(S)}.$$

Teorema 1.4.4 Kaidah Bayes. Misal $C_1, C_2, ..., C_n$ koleksi dari kejadian-kejadian yang saling asing dan menyeluruh. Maka untuk sebarang kejadian C berlaku

$$P(C_j\,|\,C)=\frac{P(C_j)P(C\,|\,C_j)}{\sum_{i=1}^{n}P(C_i)P(C\,|\,C_i)}.$$

Bukti

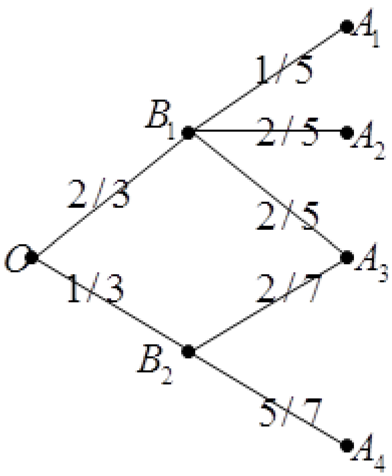
Menggunakan rumus pada Definisi 1.4.1, kita peroleh

$$P(C_j\,|\,C)=\frac{P(C_j\cap C)}{P(C)}.$$

Selanjutnya jika kita gunakan Teorema 1.4.1 untuk pembilang dan Teorema 1.4.3 untuk penyebut, maka akan kita peroleh hasil teorema ini, yaitu

$$P(C_j \mid C) = \frac{P(C_j)P(C \mid C_j)}{\sum_{i=1}^{n} P(C_i)P(C \mid C_i)}.$$

Contoh 1.4.4 Seseorang berjalan dimulai pada titik O, kemudian melanjutkan perjalanannya ke titik-titik B_1 atau B_2. Dari salah satu titik yang sudah dipilihnya dia melanjutkan perjalanannya lagi ke titik-titik A_1, A_2, A_3 atau A_4. Lebih jelas tentang lintasan yang dapat dipilih oleh dalam perjalanan orang tersebut dan peluang yang berkorespondensi dengan titik-titiknya diberikan oleh Gambar 1.4.2 berikut ini. Hitung $P(B_1 \mid A_3)$.



Gambar 1.4.2

Penyelesaian

$$P(B_1 \mid A_3) = \frac{P(B_1)P(A_3 \mid B_1)}{P(B_1)P(A_3 \mid B_1) + P(B_2)P(A_3 \mid B_2)} = \frac{(2/3)(2/5)}{(2/3)(2/5) + (1/3)(2/7)}$$

$$= \frac{4/15}{4/15 + 2/21} = \frac{4}{15} . \frac{105}{38} = \frac{14}{19}.$$

Latihan 1.4

1.4.1 Buktikan bahwa pendefinisian peluang bersyarat pada Definisi 1.4.1 memenuhi ketiga asumsi definisi peluang.

1.4.2 Buktikan Teorema 1.4.2.

1.4.3. Suatu kantong berisi 5 bola merah dan 3 bola putih. Seseorang mengambil satu bola secara acak tanpa pengembalian, kemudian mengambil satu bola lagi secara acak dari kantong tersebut. Hitunglah peluang mendapatkan:

(a) Dua bola merah.
(b) Bola merah pada pengambilan kedua dengan syarat bola putih pada pengambilan pertama.
(c) Satu bola merah dan satu bola putih.
(d) Paling sedikit satu bola merah.
(e) Bola putih pada pengambilan kedua.
(f) Bola merah pada pengambilan pertama dengan syarat bola putih pada pengambilan kedua.
(g) Bola merah pada pengambilan pertama dengan syarat paling sedikit satu bola putih pada kedua pengambilan.

1.4.4 Suatu kotak berisi tiga kartu merah dan dua kartu putih. Pemain A mengambil kartu dan kemudian pemain B mengambil kartu. Hitung peluang:

(a) Pemain A mendapatkan kartu merah.
(b) Pemain B mendapatkan kartu merah dengan syarat pemain A mendapatkan kartu merah.
(c) Pemain A mendapatkan kartu merah dan pemain B mendapatkan kartu merah.
(d) Pemain B mendapatkan kartu merah.
(e) Pemain A mendapatkan kartu merah dengan syarat pemain B mendapatkan kartu merah.
(f) Tulislah ruang sampelnya sebagai pasangan berurutan, kemudian hitung kembali Latihan (a) sampai (e) berdasarkan ruang sampel tersebut.

1.4.5. Ulangi Contoh 1.4.4 untuk menghitung:

(a) $P(B_2 \mid A_2)$.

(b) $P(B_2 \mid A_4)$.

1.4.6. Buktikan bahwa pendefinisian peluang bersyarat pada Definisi 1.4.5 memenuhi ketiga asumsi Definisi 1.3.2.

1.4.7. Misal kantong I berisi 3 kelereng merah dan 7 kelereng putih, kantong II berisi 6 kelereng merah dan 4 kelereng putih. Satu kantong dipilih secara acak kemudian 1 kelereng diambil dari kantong tersebut.

(a) hitung peluang bahwa kelereng yang terambil merah,

(b) relativ terhadap hipotesis bahwa kelereng merah, tentukan peluang bersyarat dia terambil dari kantong II.

1.4.8. Misal kantung I berisi enam merah kelereng dan empat kelereng putih. Misal lima kelereng diambil dari kantong pertama dan dimasukkan ke kantong II yang masih kosong. Hitunglah peluang bersyarat:

a. satu kelereng yang terambil dari kantong II adalah kelereng putih dengan syarat dua kelereng merah dan tiga kelereng putih dipindah dari kantong I ke kantong II,
b. dua kelereng merah dan tiga kelereng putih dipindah dari kantong I ke kantong II dengan syarat satu kelereng yang terambil dari kantong II adalah kelereng putih.

1.5 Kejadian Bebas

Pada Subpokok Bahasan 1.4 kita telah membahas peluang bersyarat, secara khusus kita telah mempunyai

$$P(C_2 \mid C_1) = \frac{P(C_2 \cap C_1)}{P(C_1)}.$$

Dari rumus di atas terlihat bahwa peluang bersyarat dari kejadian C_2 setelah munculnya kejadian C_1 bergantung pada peluang dari C_1, atau $P(C_1)$. Meskipun demikian kenyataannya tidak selalu demikian, yaitu ada peluang bersyarat dari kejadian C_2 setelah munculnya kejadian C_1 yang tidak bergantung pada peluang dari C_1. Dalam hal demikian, maka kita katakan bahwa kejadian C_2 bebas dari kejadian C_1. Hasil ini berlaku pula sebaliknya, yaitu kejadian C_1 bebas dari kejadian C_2. Oleh karena itu kita katakan bahwa kejadian C_1 dan C_2 bebas. Jika kita kembali lagi ke rumus peluang bersyarat tepatnya rumus perkalian pada Teorema 1.4.1, yaitu

$$P(C_1 \cap C_2) = P(C_1)P(C_2 \mid C_1),$$

maka jika kejadian C_1 dan C_2 bebas kita peroleh rumus

$$P(C_1 \cap C_2) = P(C_1)P(C_2).$$

Definisi formal tentang dua kejadian bebas diberikan oleh definisi berikut ini.

Definisi 1.5.1 Misal C_1 dan C_2 sebarang dua kejadian. Maka C_1 dan C_2 dikatakan ***bebas*** jika

$$P(C_1 \cap C_2) = P(C_1)P(C_2).$$

Dalam hal lain dikatakan tidak bebas.

Contoh 1.5.1 Kita kembali ke Contoh 1.4.1, untuk melihat apakah kejadian-kejadian C_1 dan C_2 bebas. Dari Tabel 1.4.1 kita peroleh

$$P(C_1) = 0.2, \quad P(C_2) = 0.6,$$

dan

$$P(C_1 \cap C_2) = 0.15.$$

Di pihak lain

$$P(C_1)P(C_2) = (0.2)(0.6) = 0.12 \neq P(C_1 \cap C_2).$$

Jadi C_1 dan C_2 dua kejadian yang tidak bebas.

Rumus lain dari kejadian bebas diberikan oleh teorema berikut ini.

Teorema 1.5.1 Misal C_1 dan C_2 sebarang dua kejadian. Maka C_1 dan C_2 dikatakan bebas jika

$$P(C_1 \mid C_2) = P(C_1).$$

Dalam hal lain dikatakan tidak bebas.

Dari Contoh 1.4.1 kita mempunyai

$$P(C_1) = 0.2,$$

dan

$$P(C_1 \mid C_2) = 0.25.$$

Menggunakan rumus pada Teorema 1.5.1 kita langsung dapat mengatakan bahwa C_1 dan C_2 dua kejadian yang tidak bebas.

Contoh 1.5.2 Selidiki apakah dua kejadian C_1 dan C_2 pada Contoh 1.4.2 bebas?

Penyelesaian

Dari hasil pada Contoh 1.4.2 kita peroleh

$$P(C_2 \mid C_1) = \frac{1}{2}.$$

Selanjutnya kita lihat peluang dari kejadian C_2. Menggunakan peluang total kita peroleh

$$P(C_2) = P(C_1)P(C_2 \mid C_1) + P(C_1^C)P(C_2 \mid C_1^C) = \frac{5}{9}.\frac{4}{8} + \frac{4}{9}.\frac{3}{8} = \frac{4}{9}.$$

Oleh karena itu

$$P(C_2 \mid C_1) \neq P(C_2).$$

Jadi dua kejadian C_1 dan C_2 tidak bebas.

Contoh 1.5.3 Ulangi Contoh 1.4.2 untuk pengambilan ***dengan pengembalian***. Apakah dua kejadian C_1 dan C_2 bebas?

Penyelesaian

Secara intuitif kita dapat mengatakan bahwa dua kejadian C_1 dan C_2 bebas, karena pengambilan dengan pengembalian mengandung pengertian pula bahwa apakah bola sudah terambil atau belum terambil tidak ada pengaruhnya. Hal ini tentunya akan berakibat bahwa

$$P(C_2 \mid C_1) = P(C_2).$$

Ada baiknya pula jika kita mencoba menyelesaikannya seperti prosedur pada Contoh 1.5.2. Karena pengambilan dengan pengembalian, maka

$$P(C_2 \mid C_1) = \frac{4}{9}.$$

Di pihak lain

$$P(C_2) = P(C_1)P(C_2 \mid C_1) + P(C_1^C)P(C_2 \mid C_1^C)$$
$$= \frac{5}{9}.\frac{4}{9} + \frac{4}{9}.\frac{4}{9}$$
$$= \frac{4}{9}.$$

Oleh karena itu

$$P(C_2 \mid C_1) = P(C_2).$$

Jadi dua kejadian C_1 dan C_2 bebas.

Pada Definisi 1.2.6 kita telah mempunyai dua kejadian yang saling asing, yaitu jika irisan dua kejadian tersebut adalah kosong. Hal ini sering kali timbul kerancuan pada mahasiswa untuk membedakan dua konsep antara dua kejadian saling asing dan bebas. Perlu kita tekankan di sini bahwa dua kejadian saling asing titik perhatiannya pada irisan kedua kejadian tersebut, yaitu lebih kepada himpunannya. Sedangkan dua kejadian bebas titik perhatiannya pada peluangnya. Jadi sebenarnya dua konsep ini merupakan dua konsep yang sangat berbeda.

Sering pula mahasiswa berkeinginan untuk mengetahui bagaimana hubungan kedua konsep tersebut. Kedua konsep yang sangat berbeda ini tentunya tidak dapat dikaitkan, artinya yang satu tidak berakibat ke yang lain demikian pula sebaliknya. Meskipun demikian memang tidak dapat kita pungkiri bahwa ada kecenderungan kalau dua kejadian itu saling asing, maka dua kejadian itu tidak bebas. Argumentasinya adalah dua kejadian saling asing menghendaki irisannya kosong. Oleh karena itu peluang irisan kedua kejadian tersebut adalah nol. Jika kita kaitkan dengan dua kejadian bebas, maka peluang nol tersebut akan terjadi jika paling sedikit peluang satu kejadiannya adalah nol. Oleh karena itu jika dua kejadian masing-masing peluangnya tidak nol, tetapi dua kejadian tersebut irisannya kosong atau saling asing, maka dapat dipastikan bahwa dua kejadian tersebut tidak bebas. Untuk lebih jelasnya kita lihat contoh berikut ini.

Contoh 1.5.4 Kita kembali ke Contoh 1.4.4. Sekarang kita perhatikan dua kejadian B_1 dan A_4. Kita melihat bahwa

$$P(A_4 \mid B_1) = 0.$$

Karena

$$P(A_4 \mid B_1) = \frac{P(A_4 \cap B_1)}{P(B_1)},$$

maka

$$P(A_4 \cap B_1) = 0.$$

Jadi dua kejadian B_1 dan A_4 saling asing. Di pihak lain

$$P(B_1) = \frac{2}{3},$$

dan menggunakan peluang total kita peroleh

$$\begin{aligned} P(A_4) &= P(B_1)P(A_4 \mid B_1) + P(B_2)P(A_4 \mid B_2) \\ &= \frac{2}{3}.0 + \frac{1}{3}.\frac{5}{7} \\ &= \frac{5}{21} \\ &\neq 0. \end{aligned}$$

Jadi dua kejadian B_1 dan A_4 tidak bebas.

Teorema berikut ini merupakan bentuk lain untuk melihat bahwa dua kejadian bebas.

Teorema 1.5.2 Misal C_1 dan C_2 sebarang dari dua kejadian. Maka C_1 dan C_2 bebas jika dan hanya jika pasangan dua kejadian berikut ini bebas

1. C_1 dan C_2^C.
2. C_1^C dan C_2.
3. C_1^C dan C_2^C.

Bukti dari Teorema di atas ditinggalkan sebagai latihan.

Konsep mengenai tentang dua kejadian bebas dapat pula diperumum untuk lebih dari dua kejadian, seperti definisi berikut ini.

Definisi 1.5.2 Misal $C_1, C_2, ..., C_n$ sebarang n kejadian. Maka $C_1, C_2, ..., C_n$ dikatakan bebas jika untuk setiap $j = 2, ..., n$ dan setiap himpunan bagian dari indeks yang berbeda $i_1, i_2, ..., i_j$,

$$P(C_{i_1} \cap C_{i_2} \cap ... \cap C_{i_j}) = P(C_{i_1})P(C_{i_2})...P(C_{i_j}).$$

Secara khusus, kejadian-kejadian $C_1, C_2,$ dan C_3 dikatakan bebas jika dan hanya jika

$$P(C_1 \cap C_2) = P(C_1)P(C_2),$$
$$P(C_1 \cap C_3) = P(C_1)P(C_3),$$
$$P(C_2 \cap C_3) = P(C_2)P(C_3)$$

dan

$$P(C_1 \cap C_2 \cap C_3) = P(C_1)P(C_2)P(C_3).$$

Contoh 1.5.5 Suatu percobaan mengambil satu kartu secara acak dari satu pak kartu *bridge*. Misal C_1 kejadian terambilnya kartu merah; C_2 kejadian terambilnya kartu gambar, yaitu kartu-kartu J, Q, dan K; C_3 kejadian terambilnya kartu hati atau skop. Maka

$$P(C_1 \cap C_2) = P(C_1)P(C_2) = \frac{3}{26},$$
$$P(C_1 \cap C_3) = P(C_1)P(C_3) = \frac{1}{4},$$
$$P(C_2 \cap C_3) = P(C_2)P(C_3) = \frac{3}{26},$$

dan

$$P(C_1 \cap C_2 \cap C_3) = P(C_1)P(C_2)P(C_3) = \frac{3}{52}.$$

Jadi kejadian-kejadian $C_{1,}$ $C_2,$ dan C_3 bebas.

Tidak selalu berlaku bahwa jika

$$P(C_1 \cap C_2) = P(C_1)P(C_2),$$
$$P(C_1 \cap C_3) = P(C_1)P(C_3),$$
$$P(C_2 \cap C_3) = P(C_2)P(C_3),$$

berakibat

$$P(C_1 \cap C_2 \cap C_3) = P(C_1)P(C_2)P(C_3),$$

dan sebaliknya. Hal ini dapat dilihat pada contoh-contoh berikut ini.

Contoh 1.5.6 Misal suatu ruang sampel

$$S=\{(1,0,0),(0,1,0),(0,0,1),(1,1,1)\}$$

yang bobot masing-masing anggota sampelnya sama dan misal $C_1=\{(1,0,0),(1,1,1)\}$, $C_2=\{(0,1,0),(1,1,1)\}$ dan $C_3=\{(0,0,1),(1,1,1)\}$. Maka

$$P(C_1\cap C_2)=P(C_1)P(C_2)=\frac{1}{4},$$

$$P(C_1\cap C_3)=P(C_1)P(C_3)=\frac{1}{4},$$

$$P(C_2\cap C_3)=P(C_2)P(C_3)=\frac{1}{4},$$

tetapi

$$P(C_1\cap C_2\cap C_3)=\frac{1}{4}\neq\frac{1}{8}=P(C_1)P(C_2)P(C_3).$$

Jadi kejadian-kejadian $C_{1,}$ C_2, dan C_3 tidak bebas.

Contoh 1.5.7 Misal suatu ruang sampel $S=\{c_1,c_2,c_3,c_4,c_5\}$ yang bobot masing-masing anggota sampelnya berturut-turut adalah $\frac{2}{16},\frac{3}{16},\frac{3}{16},\frac{3}{16},\frac{5}{16}$ dan misal $C_1=\{c_1,c_2,c_3\}$, $C_2=\{c_1,c_2,c_4\}$ dan $C_3=\{c_1,c_3,c_4\}$. Maka

$$P(C_1\cap C_2)=\frac{5}{16}\neq\frac{1}{4}=P(C_1)P(C_2)$$

$$P(C_1\cap C_3)=\frac{5}{16}\neq\frac{1}{4}=P(C_1)P(C_3)=\frac{1}{4},$$

$$P(C_2\cap C_3)=\frac{5}{16}\neq\frac{1}{4}=P(C_2)P(C_3)=\frac{1}{4}$$

tetapi

$$P(C_1\cap C_2\cap C_3)=P(C_1)P(C_2)P(C_3).$$

Jadi kejadian-kejadian $C_{1,}$ C_2 dan C_3 tidak bebas.

Latihan 1.5

1.5.1. Buktikan Teorema 1.5.2.

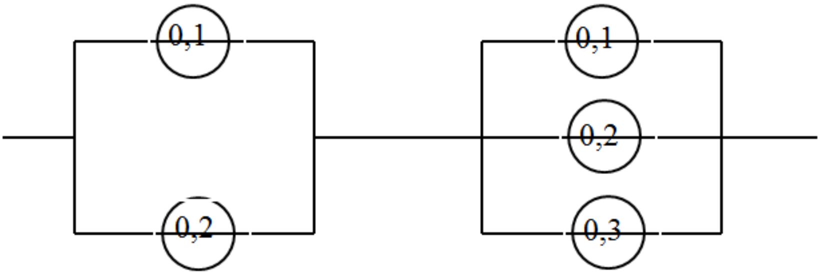
1.5.2. Ulangi Contoh 1.4.4, untuk menyelidiki bebas tidaknya dua kejadian-kejadian B_1 dan A_3.

1.5.3. Misal $P(C_1)=0.4$ dan $P(C_1\cup C_2)=0.6.$

(a) Berapakah nilai $P(C_2)$, jika dua kejadian C_1 dan C_2 saling asing?

(b) Berapakah nilai $P(C_2)$, jika dua kejadian C_1 dan C_2 bebas?

1.5.4. Perhatikan suatu sistem berikut ini lengkap dengan peluang tidak berfungsi komponennya. Anggap kejadian-kejadian tersebut bebas.



Hitunglah peluang sistem tersebut berfungsi.

1.5.5. Beri contoh dua kejadian C_1 dan C_2 sehingga $P(C_1\cap C_2)<P(C_1)P(C_2).$

1.5.6. Misal C_1 dan C_2 sebarang dua kejadian.

(a) Tunjukkan bahwa jika $P(C_1)=1,$ maka $P(C_1\cap C_2)=P(C_2).$

(b) Buktikan bahwa untuk sebarang kejadian C_1 dengan $P(C_1)=0$ atau $P(C_1)=1$ adalah bebas dengan sebarang kejadian C_2.

1.5.7. Tunjukkan bahwa jika kejadian C bebas dengan dirinya sendiri, maka $P(C)=0$ atau $P(C)=1.$

1.5.8. Misal C_1, C_2, dan C_3 kejadan-kejadian bebas.

(a) Tunjukkan bahwa C_1 dan $C_2\cup C_3$ bebas.

(b) Tunjukkan bahwa $C_1 \cap C_2^C$ dan C_3 bebas.

1.5.9. Suatu percobaan melempar dadu dua kali. Misal C_1 kejadian dadu kedua muncul 1, 2 atau 5; C_2 kejadian dadu kedua muncul 4, 5 atau 6; dan C_3 kejadian jumlah mata dadu adalah 9. Selidiki apakah kejadian-kejadian $C_{1,}$ C_2, dan C_3 bebas?

1.5.10. Misal kejadian-kejadian $C_{1,}$ C_2, dan C_3 bebas dengan peluang berturut-turut adalah 1/2, 1/3 , 1/4. Hitung $P(C_1 \cup C_2 \cup C_3)$.

1.5.11. Misal kejadian-kejadian C dan C_3 bebas sehingga $P(C_1) = P(C_2) = P(C_3) = \dfrac{1}{4}$. Tentukan $P\left[(C_1^C \cap C_2^C) \cup C_3\right]$.

BAB 2
PEUBAH ACAK

2.1 Pendahuluan

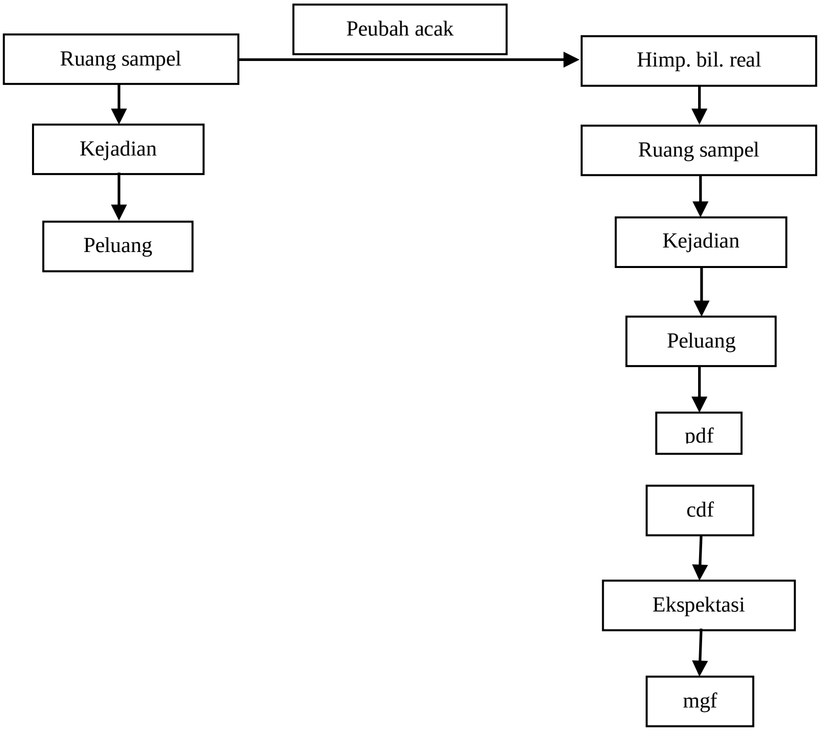
Pada Bab 1 kita telah membahas ruang sampel, kejadian, dan peluang. Ruang sampel yang telah kita hadapi sering kali anggotanya tidak berupa bilangan, walaupun memang adakalanya anggotanya berupa bilangan. Dalam hal ruang sampel yang anggotanya tidak berupa bilangan tentu kita tidak dapat berbuat banyak, karena tidak dapat mengoperasikannya baik secara aljabar maupun secara kalkulus.

Pada ruang sampel ini paling tinggi kita hanya bisa mengembangkannya sampai pada konsep peluang. Oleh karena itu perlu adanya suatu fungsi yang mengubah ruang sampel yang anggotanya bukan berupa bilangan menjadi himpunan yang anggotanya berupa bilangan khususnya bilangan real. Fungsi yang demikian ini disebut sebagai peubah acak**.**

Tujuan instruksional umum dari mempelajari bab ini adalah Anda diharapkan memahami konsep tentang peubah acak.

Tujuan instruksional khusus mempelajari bab ini adalah Anda diharapkan dapat menentukan: peubah acak suatu ruang sampel, peluang terinduksi, fungsi kepadatan peluang, modus, median, persentil, fungsi sebaran, ekspektasi, varians, dan fungsi pembangkit momen.

Selanjutnya kita lihat bagan berikut untuk memperoleh gambaran tentang keberadaan peubah acak dan konsekuensinya.



2.2 Peubah Acak

Sekarang tibalah saatnya kita kaji lebih dalam tentang peubah acak. Kita mulai dari definisi peubah acak berikut ini.

Definisi 2.2.1 Misal X suatu fungsi yang terdefinisi atas ruang sampel S, pada himpunan bilangan real R, yaitu

$$S \xrightarrow{X} R,$$

yang mengaitkan untuk setiap $c \in S$ ke $x \in R$, ditulis $X(c) = x, \forall c \in S$. Maka X disebut ***peubah acak***. Dan ***daerah jelajah*** dari X, ditulis sebagai $\mathfrak{A} = \{x : X(c) = x, c \in S\}$ disebut ***ruang peubah acak***.

Dalam buku ini peubah acak ditulis sebagai huruf besar $X, Y,$ atau $Z,$ sedangkan nilainya kita tulis sebagai huruf kecil $x, y,$ atau $z.$ Jadi jika tertulis huruf besar itu berarti peubah acak yang berupa suatu fungsi. Di pihak lain jika tertulis huruf kecil berarti nilai fungsi yang berupa bilangan real. Cara penulisan ini perlu sekali kita perhatikan mulai sekarang, karena kesalahan dalam penulisan akan berakibat kerancuan dalam memahami konsep peubah acak.

Contoh 2.2.1 Misal suatu percobaan melempar sekeping uang logam dua kali. Maka ruang sampelnya adalah

$$S = \{mm, mb, bm, bb\}.$$

Misal X peubah acak yang menyatakan banyak muncul muka. Maka

$$X(mm) = 2, \quad X(mb) = X(bm) = 1, \; X(bb) = 0.$$

Sehingga kita peroleh ruang peubah acak

$$\mathfrak{A} = X(S) = \{0, 1, 2\}.$$

Sekarang kita pandang ruang peubah acak di atas sebagai ruang sampel. Oleh karena itu sebarang himpunan bagiannya juga dapat kita pandang sebagai kejadian. Permasalahannya sekarang bagaimana menentukan peluang kejadian ini? Sebagai contoh kita ambil kejadian

$$A_1 = \{0\}.$$

Kejadian ini ekivalen dengan kejadian

$$C_1 = \{bb\}.$$

Dan dengan mudah kita dapat melihat bahwa

$$P(C_1) = \frac{1}{4}.$$

Oleh karena itu cukup beralasan jika kita tentukan bahwa

$$P(A_1) = P(C_1) = \frac{1}{4}.$$

Dengan cara yang sama jika

$$A_2 = \{1\},$$

maka kejadian yang ekivalen adalah

$$C_2 = \{mb, bm\}.$$

Oleh karena itu

$$P(A_2) = P(C_2) = \frac{2}{4}.$$

Demikian pula untuk

$$A_3 = \{0,1\},$$

ekivalen dengan

$$C_3 = \{bb, mb, bm\},$$

dan

$$P(A_3) = P(C_3) = \frac{3}{4}.$$

Pekerjaan kita pada Contoh 2.2.1 adalah menentukan peluang kejadian pada ruang peubah acak menggunakan peluang kejadian pada ruang sampel. Konsep peluang seperti ini disebut sebagai ***peluang terinduksi*** oleh peubah acak X, didefinisikan sebagai

$$P(A) = P_X(A) = P(X \in A) = P(C),$$

di mana $C = \{c : X(c) \in A, c \in S\}$.

Untuk memeriksa apakah pendefinisian di atas memenuhi ketiga asumsi peluang ditinggalkan sebagai latihan. Sifat-sifat peluang pada Subpokok Bahasan 1.3, berlaku pula pada peluang terinduksi ini.

Kita perhatikan cara-cara penulisan berikut ini:

1. Penulisan $X \in A$. Penulisan ini merupakan penyederhanaan dari penulisan $\{c : X(c) \in A, c \in S\}$. Atau
$$X \in A \Leftrightarrow C = \{c : X(c) \in A, c \in S\}.$$
2. Dalam hal A kejadian sederhana, seperti $A_1 = \{0\}$, penulisan $X \in A_1$ biasanya ditulis sebagai $X = 0$, yang merupakan penyederhanaan dari $\{c : X(c) = 0, c \in S\}$. Atau
$$X \in A_1 \Leftrightarrow X = 0 \Leftrightarrow C_1 = \{c : X(c) = 0, c \in S\}.$$
3. Dimungkinkan juga muncul penulisan $X \le x$. Penulisan ini dimaksudkan penyederhanaan dari penulisan $\{c : X(c) \le x, c \in S\}$. Atau
$$X \le x \Leftrightarrow \{c : X(c) \le x, c \in S\}.$$

Selanjutnya penulisan tentang peluang terinduksi dapat dipilih salah satu dari beberapa penulisan di atas disesuaikan dengan keperluan. Oleh karena itu hasil pada Contoh 2.2.1 dapat kita buat dalam bentuk Tabel 2.2.1 berikut ini.

TABEL 2.2.1

x	0	1	2
$P(X=x)$	$\frac{1}{4}$	$\frac{2}{4}$	$\frac{1}{4}$

Tabel 2.2.1 sering disebut sebagai tabel sebaran peluang dari peubah acak X.

Contoh 2.2.2 Lihat kembali Contoh 2.2.1, kemudian hitung $P(X \le 1)$.

Penyelesaian

Menghitung $P(X \le 1)$ sama saja dengan menghitung $P(X = 0 \text{ atau } 1)$, atau sama juga menghitung $P(A)$, di mana $A = \{0,1\}$. Pada Contoh 2.2.1 kita telah memperoleh hasil ini, yaitu

$$P(A) = \frac{3}{4}.$$

Jadi

$$P(X \le 1) = \frac{3}{4}.$$

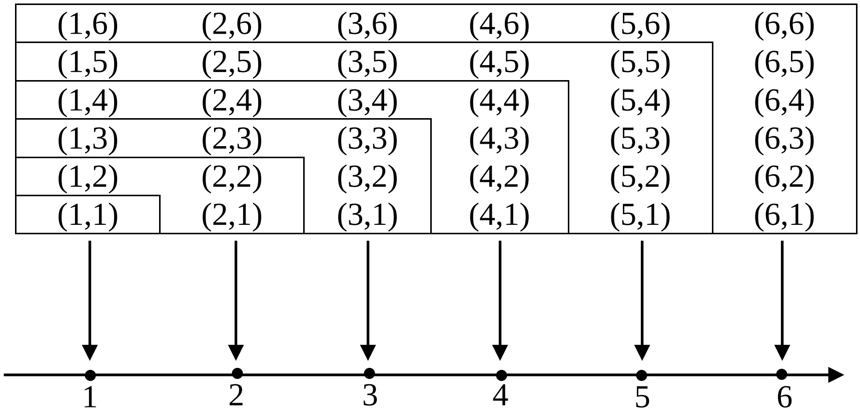
Hasil di atas dapat pula diperoleh langsung dari Tabel 2.2.1.

Contoh 2.2.3 Misal suatu dadu bermata $1,2,3,4,5,$ dan 6 dilempar sebanyak dua kali, dan misal X suatu peubah acak yang menyatakan nilai maksimum untuk mata dadu dari kedua lemparan tersebut, yaitu $X = maks\{i, j\}$, di mana $i, j \in \{1,2,3,4,5,6\}$.

a. Buatlah tabel sebaran peluangnya.
b. Hitung $P(X \le 2)$ langsung dari tabel pada hasil a.

Penyelesaian

Untuk memudahkan pekerjaan kita akan lebih baik jika kita buat gambar dari ruang sampel pelemparan dadu sebanyak dua kali seperti pada Gambar 2.2.1 berikut ini.



Gambar 2.2.1

a. Kita peroleh tabel sebaran peluang berikut ini.

TABEL 2.2.2

x	1	2	3	4	5	6
$P(X=x)$	$\frac{1}{36}$	$\frac{3}{36}$	$\frac{5}{36}$	$\frac{7}{36}$	$\frac{9}{36}$	$\frac{11}{36}$

b. Dari tabel di atas kita memperoleh hasil

$$P(X \leq 2) = \frac{1}{36} + \frac{3}{36} = \frac{4}{36}.$$

Semua peubah acak yang telah kita bahas nilai-nilainya dari himpunan terbilang, atau ruang peubah acaknya berupa himpunan terbilang. Peubah yang demikian disebut ***peubah acak diskret***. Sekarang kita akan melihat contoh-contoh peubah acak yang nilai-nilainya dari suatu selang, yang disebut sebagai ***peubah acak kontinu***, yaitu peubah acak yang ruangnya berupa selang.

Contoh 2.2.4 Kita lihat kembali Latihan 1.2.5, yang ruang sampelnya adalah

$$S = \{c : 0 < c < 1\}.$$

Misal peluang dari sebarang kejadian C dari ruang sampel ini didefinisikan oleh

$$P(C) = \int_C dz,$$

bukti bahwa pendefinisian ini memenuhi ketiga asumsi peluang ditinggalkan sebagai latihan.

Sebagai contoh, jika $C = \left\{c : \frac{1}{3} < c < \frac{1}{2}\right\},$ maka

$$P(C) = \int_{\frac{1}{3}}^{\frac{1}{2}} dz = [z]_{\frac{1}{3}}^{\frac{1}{2}} = \frac{1}{6}.$$

Misal peubah acak X didefinisikan sebagai

$$x = X(c) = 3c + 2.$$

Maka ruang peubah acak X adalah

$$\mathfrak{A} = \{x : 2 < x < 5\}.$$

Kita ingin menentukan peluang terinduksi ini, kita sebut

$$P(A), A \subset \mathfrak{A}.$$

Misal

$$A = \{x : 2 < x < b\},$$

di mana $2 < b < 5.$ Mudah kita lihat bahwa himpunan A ini ekivalen dengan

$$C = \left\{c : 0 < c < \frac{b-2}{3}\right\}.$$

Oleh karena itu berdasarkan pada definisi peluang terinduksi kita peroleh

$$P(A) = P(C) = \int_0^{\frac{b-2}{3}} dz = \frac{b-2}{3}.$$

Kita dapat juga mengubah bentuk integral ini dalam peubah $x = 3z + 2,$ dan kita peroleh

$$P(A) = \int_2^b \frac{1}{3} dx = \int_A \frac{1}{3} dx,$$

di mana $A = \{x : 2 < x < b\}.$ Bentuk terakhir ini merupakan bentuk peluang yang terdefinisi pada kejadian yang ditimbulkan oleh peubah acak $X,$ yang nilainya ditentukan langsung oleh kejadian $A.$

Contoh 2.2.5 Misal P menyatakan fungsi peluang dari peubah acak X, sehingga untuk sebarang $A \subset \mathfrak{A}$ berlaku

$$P(A) = \int_A f(x)dx,$$

di mana

$$f(x) = \frac{3x^2}{8}, x \in \mathfrak{A} = \{x : 0 < x < 2\}.$$

Misal $A_1 = \left\{x : 0 < x < \frac{3}{2}\right\}$, dan $A_2 = \{x : 1 < x < 2\}$. Hitung: a. $P(A_1)$, b. $P(A_2)$ dan c. $P(A_1 \cup A_2)$.

Penyelesaian

a.

$$P(A_1) = \int_{A_1} f(x)dx = \int_0^{\frac{3}{2}} \frac{3x^2}{8} dx = \left[\frac{x^3}{8}\right]_0^{\frac{3}{2}} = \frac{27}{64}.$$

b.

$$P(A_2) = \int_{A_2} f(x)dx = \int_1^2 \frac{3x^2}{8} dx = \left[\frac{x^3}{8}\right]_1^2 = \frac{7}{8}.$$

c. Karena $A_1 \cap A_2 = \left\{1 < x < \frac{3}{2}\right\} \neq \varnothing$, maka untuk menghitung $P(A_1 \cup A_2)$ kita gunakan rumus pada Teorema 1.3.3, yaitu

$$P(A_1 \cup A_2) = P(A_1) + P(A_2) - P(A_1 \cap A_2).$$

Kita hitung terlebih dahulu $P(A_1 \cap A_2)$ seperti berikut ini:

$$P(A_1 \cap A_2) = \int_{A_1 \cap A_2} f(x)dx = \int_1^{\frac{3}{2}} \frac{3x^2}{8} dx = \left[\frac{x^3}{8}\right]_1^{\frac{3}{2}} = \frac{19}{64}.$$

Jadi

$$P(A_1 \cup A_2) = \frac{27}{64} + \frac{56}{64} - \frac{19}{64} = 1.$$

Hasil ini dapat pula kita peroleh dari kenyataan bahwa

$$\begin{aligned} A_1 \cup A_2 &= \left\{x : 0 < x \leq \frac{3}{2}\right\} \cup \{x : 1 \leq x < 2\} \\ &= \{x : 0 < x < 2\} \\ &= \mathfrak{A}. \end{aligned}$$

Sehingga

$$P(A_1 \cup A_2) = P(\mathfrak{A}) = 1.$$

Lebih lanjut mengenai peubah acak diskret dan peubah acak kontinu akan dibahas pada subpokok bahasan berikutnya.

Latihan 2.2

2.2.1 Misal $c = (i, j),$ menyatakan sebarang hasil dari pelemparan satu dadu sebanyak dua kali, dan misal $X(c) = i + j.$ Buatlah tabel sebaran peluang dari peubah acak X.

2.2.2. Ulangilah Latihan 2.2.1, untuk $Y(c) = i - j.$

2.2.3. Ulangilah Latihan 2.2.1, untuk $Z(c) = (i - j)^2.$

2.2.4. Buktikan bahwa pendefinisian peluang terinduksi oleh peubah acak X , yaitu

$$P(A) = P(C),$$

di mana $C = \{c : X(c) \in A, c \in S\}$, memenuhi ketiga asumsi definisi peluang.

2.2.5. Misal satu kartu diambil secara acak dari satu pak kartu *bridge*. Misal c salah satu kartu dari satu pak kartu tersebut, dan misal peubah acak X didefinisikan sebagai: $X(c) = 4,$ jika c kartu as, $X(c) = 3,$ jika c kartu raja, $X(c) = 2,$ jika c kartu ratu, $X(c) = 1,$ jika c kartu *jack*, dan $X(c) = 0,$ jika yang lainnya. Tentukan peluang terinduksi $P(A),$ dari peubah acak X, di mana $A \subset \mathfrak{A} = \{x : x = 0, 1, 2, 3, 4\},$ dan $A = \{0, 1, 2\}.$

2.2.6. Misal suatu percobaan mempunyai ruang sampel

$$S = \{c : 0 < c < 1\},$$

dan peluang dari sebarang kejadian C dari ruang sampel ini didefinisikan oleh

$$P(C) = \int_C dz.$$

Buktikan bahwa pendefinisian ini memenuhi ketiga asumsi peluang.

2.2.7. Misal P menyatakan fungsi peluang dari peubah acak X, sehingga untuk sebarang $A \subset \mathfrak{A}$ berlaku

$$P(A) = \int_A f(x)dx,$$

di mana $f(x) = \dfrac{2x}{9}, x \in \mathfrak{A} = \{x : 0 < x < 3\}$. Misal $A_1 = \{x : 0 < x < 1\}$, dan $A_2 = \{x : 2 < x < 3\}$. Hitung: a. $P(A_1)$, b. $P(A_2)$, dan c. $P(A_1 \cup A_2)$.

2.3 Fungsi Kepadatan Peluang

Kita lihat kembali Tabel 2.2.1 tentang sebaran peluang dari peubah acak X pada Contoh 2.2.1. Tabel tesebut dapat kita ubah menjadi bentuk sederhana

$$f(x) = \frac{\binom{2}{x}}{4}, x = 0,1,2,$$

atau

$$f(x) = \frac{\binom{2}{x}}{4}, x \in \mathfrak{A}.$$

Sedangkan dari Contoh 2.2.4 untuk sebarang $A \subset \mathfrak{A}$ kita peroleh

$$P(A) = \int_A \frac{1}{3} dx,$$

yang dapat kita tulis sebagai

$$P(A) = \int_A f(x)dx,$$

di mana $f(x) = \dfrac{1}{3}, x \in A,$ dan $A = \{x : 2 < x < 5\}$.

Kedua fungsi di atas, yaitu f sehingga

$$f(x)=\frac{\binom{2}{x}}{4},\quad x\in\mathfrak{A}=\{x:x=0,1,2\}$$

dan f sehingga

$$f(x)=\frac{1}{3}, x\in\mathfrak{A}=\{x:2<x<5\}$$

disebut sebagai ***fungsi kepadatan peluang*** (***probability density function***) dari peubah acak X. Untuk selanjutnya fungsi kepadatan peluang ini ditulis sebagai ***pdf*** yang merupakan kependekan dari ***probability density function***.

Dua pdf di atas merupakan pdf dari dua jenis yang berlainan, yang pertama pdf yang terdefinisi pada himpunan terbilang dan yang kedua pdf yang terdefinisi pada suatu selang. Beberapa referensi membedakan nama dari fungsi ini. Untuk ruang sampel diskret fungsi ini disebut sebagai fungsi massa peluang, sedangkan untuk ruang sampel kontinu fungsi ini disebut sebagai fungsi kepadatan peluang. Dalam buku ini penulis menyebut fungsi kepadatan peluang baik untuk ruang sampel diskret maupun ruang sampel kontinu.

Definisi 2.3.1 Misal X peubah acak dengan ruang $\mathfrak{A}$, yang merupakan himpunan dari titik-titik diskret. Maka X disebut ***peubah acak diskret***, dan fungsi f sehingga

(i) $f(x)>0, x\in\mathfrak{A}$,

(ii) $\sum_{\mathfrak{A}} f(x)=1$,

(iii) untuk sebarang $A\subset\mathfrak{A}$, $P(A)=\sum_{A} f(x)$,

disebut sebagai ***fungsi kepadatan peluang***.

Contoh 2.3.1 Misal X peubah acak diskret dengan ruang $\mathfrak{A}=\{x:x=0,1,2,3,4\}$ dengan pdf

$$f(x)=\binom{4}{x}\left(\frac{1}{2}\right)^4, x\in\mathfrak{A}.$$

Jika $A_1=\{x:x=0,1\}$, dan $A_2=\{x:x=1,3\}$, maka hitung: a. $P(A_1)$, b. $P(A_2)$, dan $P(A_1\cup A_2)$.

Penyelesaian

a.

$$P(A_1) = \sum_{A_1} f(x) = f(0) + f(1) = \binom{4}{0}\left(\frac{1}{2}\right)^4 + \binom{4}{1}\left(\frac{1}{2}\right)^4 = \frac{5}{16}.$$

b.

$$P(A_2) = \sum_{A_2} f(x) = f(1) + f(3) = \binom{4}{1}\left(\frac{1}{2}\right)^4 + \binom{4}{3}\left(\frac{1}{2}\right)^4 = \frac{1}{2}.$$

c.

$$P(A_1 \cup A_2) = \sum_{A_1 \cup A_2} f(x) = f(0) + f(1) + f(3) = \binom{4}{0}\left(\frac{1}{2}\right)^4 + \binom{4}{1}\left(\frac{1}{2}\right)^4 + \binom{4}{3}\left(\frac{1}{2}\right)^4 = \frac{9}{16}.$$

Hasil c dapat pula diperoleh dari rumus

$$P(A_1 \cup A_2) = P(A_1) + P(A_2) - P(A_1 \cap A_2).$$

Definisi 2.3.2 Misal X peubah acak dengan ruang $\mathfrak{A}$, yang merupakan selang dari bilangan real. Maka X disebut ***peubah acak kontinu***, dan fungsi f sehingga

(ii) $f(x) > 0, x \in \mathfrak{A}$,

(ii) $\int_{\mathfrak{A}} f(x)dx = 1$,

(iii) untuk sebarang $A \subset \mathfrak{A}$, $P(A) = \int_A f(x)dx$,

disebut sebagai ***fungsi kepadatan peluang***.

Contoh 2.3.2 Misal X peubah acak kontinu dengan ruang $\mathfrak{A} = \{x : 0 < x < 1\}$ dan pdf

$$f(x) = 3x^2, x \in \mathfrak{A}.$$

Jika $A_1 = \{x : 0 < x < \frac{1}{2}\}$, dan $A_2 = \{x : \frac{1}{4} < x < \frac{3}{4}\}$, maka hitung: a. $P(A_1)$, b. $P(A_2)$, dan $P(A_1 \cup A_2)$.

Penyelesaian

a.

$$P(A_1) = \int_{A_1} f(x)dx = \int_0^{\frac{1}{2}} 3x^2 dx = \left[x^3\right]_0^{\frac{1}{2}} = \frac{1}{8}.$$

b.

$$P(A_2) = \int_{A_2} f(x)dx = \int_{\frac{1}{4}}^{\frac{3}{4}} 3x^2 dx = \left[x^3 \right]_{\frac{1}{4}}^{\frac{3}{4}} = \frac{13}{32}.$$

c.

$$P(A_1 \cup A_2) = \int_{A_1 \cup A_2} f(x)dx = \int_0^{\frac{3}{4}} 3x^2 dx = \left[x^3 \right]_0^{\frac{3}{4}} = \frac{27}{64}.$$

Definisi 2.3.1 dan Definisi 2.3.2 mensyaratkan bahwa $f(x) > 0, x \in \mathfrak{A}$. Ruang peubah acak dari definisi ini dapat kita perluas ke himpunan bilangan real. Hal ini akan berakibat perubahan pada penulisan $f(x) > 0, x \in \mathfrak{A}$. Untuk Contoh 2.3.1 kita tulis sebagai

$$f(x) = \begin{cases} \binom{4}{x} \left(\frac{1}{2} \right)^4, x = 0, 1, 2, 3, 4, \\ \text{nol untuk yang lainnya.} \end{cases}$$

Sedangkan untuk Contoh 2.3.2 kita tulis sebagai

$$f(x) = \begin{cases} 3x^2, 0 < x < 1, \\ \text{nol untuk yang lainnya.} \end{cases}$$

Untuk meyederhanakan penulisan tentang pdf, kita sajikan terlebih dahulu konsep tentang fungsi indikator berikut ini.

Definisi 2.3.3 Suatu fungsi indikator dari himpunan bagian A didefinisikan sebagai

$$I(A) = \begin{cases} 1, \text{bila } A \text{ terpenuhi,} \\ 0, \text{selainnya.} \end{cases}$$

Menggunakan definisi tersebut, penulisan masing-masing pdf pada Contoh 2.3.1 dan Contoh 2.3.2 berturut-turut menjadi

$$f(x) = \binom{4}{x} \left(\frac{1}{2} \right)^4 I(x = 0, 1, 2, 3, 4),$$

dan

$$f(x) = 3x^2 I(0 < x < 1).$$

Penulisan menggunakan fungsi indikator yang akan kita gunakan dalam buku ini.

Kita telah melihat bahwa jika X kontinu dan jika f suatu pdf dari X, maka jika $A = \{x : a < x < b\}$,

$$P(A) = \int_A f(x)dx = \int_a^b f(x)dx.$$

Lebih lanjut, jika $A = \{x : x = a\}$, maka

$$P(A) = \int_A f(x)dx = \int_a^a f(x)dx = 0.$$

Akibatnya

$$P(a < X < b) = P(a < X \le b) = P(a \le X < b) = P(a \le X \le b).$$

Latihan 2.3

2.3.1 Untuk masing-masing fungsi berikut ini, tentukan konstanta k sehingga f merupakan pdf dari peubah acak X.

(a) $f(x) = k\left(\frac{2}{3}\right)^x I(x = 1, 2, 3, \ldots)$.

(b) $f(x) = kxe^{-x} I(0 < x < \infty)$.

(c) $f(x) = kx^{-(k+1)} I(1 < x < \infty)$.

2.3.2. Misal X peubah acak diskret dengan pdf

$$f(x) = \left(\frac{1}{2}\right)^x I(x = 1, 2, 3, \ldots).$$

Misal $A_1 = \{x : x = 1, 3, 5, \ldots\}$, dan $A_2 = \{x : x = 2, 4, 6, \ldots\}$. Hitung:

(a) $P(A_1)$.

(b) $P(A_2)$.

2.3.3. Misal pdf dari peubah acak X, adalah $f(x) = \left(x/15\right) I(x = 1, 2, 3, 4, 5)$. Hitung:

(a) $P(X = 1 \text{ atau } 2)$.

(b) $P(1/2 < X < 5/2)$.

(c) $P(1 \le X \le 2)$.

2.3.4. Untuk masing-masing pdf dari peubah acak X berikut ini, hitunglah $P(|X|<1)$ dan $P(X^2>4)$.

(a) $f(x)=\left(x^2/18\right)I(-3<x<3)$.

(b) $f(x)=\left[(x+2)/18\right]I(-2<x<4)$.

2.3.5. Jika X peubah acak kontinu, maka $P(X=a)=0$. Berikan contoh untuk untuk hasil ini.

2.3.6. Suatu ***modus*** dari peubah acak X adalah suatu nilai x sehingga $f(x)$ maksimum. Tentukan modus peubah acak X untuk masing-masing pdf berikut ini.

(a) $f(x)=\left(\frac{1}{2}\right)^x I(x=1,2,3,...)$.

(b) $f(x)=12x^2(1-x)I(0<x<1)$.

2.3.7. Suatu ***median*** dari peubah acak X adalah suatu nilai x sehingga $P(X<x)\le\frac{1}{2}$ dan $P(X\le x)\ge\frac{1}{2}$. Tentukan median peubah acak X untuk masing-masing pdf berikut ini.

(a) $f(x)=\dfrac{4!}{x!(4-x)!}\left(\frac{1}{4}\right)^x\left(\frac{3}{4}\right)^{4-x} I(x=0,1,2,3,4)$.

(b) $f(x)=3x^2I(0<x<1)$.

2.3.8 Suatu ***persentil*** ke $100p$, di mana $0<p<1$, dari peubah acak X adalah suatu nilai ξ_p sehingga $P(X<\xi_p)\le p$ dan $P(X\le\xi_p)\ge p$. Ulangi Latihan 2.3.7 untuk persentil ke 60.

2.4 Fungsi Sebaran

Misal peubah acak X mempunyai fungsi peluang P dan misal A himpunan pada ruang berdimensi satu yang terbatas pada x, termasuk x sendiri, yaitu $A=\{x:-\infty\le x\}$. Maka

$$P(A)=P(X\le x).$$

Peluang ini merupakan peluang kumulatif peubah acak X dari $-\infty$ sampai dengan x. Dapat pula dikatakan sebagai sebaran peluang kumulatif dari X pada x. Selain itu peluang ini juga merupakan fungsi dari x, ditulis sebagai $F(x)$. Fungsi F yang demikian disebut sebagai ***fungsi sebaran kumulatif/cumulative distribution function*** dari peubah acak X, disingkat sebagai cdf.

Definisi 2.4.1 Misal X suatu peubah acak. Maka untuk sebarang bilangan real x fungsi F sehingga

$$F(x) = P(X \leq x),$$

disebut ***fungsi sebaran kumulatif***.

Untuk selanjutnya fungsi sebaran kumulatif ditulis sebagai fungsi sebaran.

Akibat langsung dari definisi di atas adalah

$$F(x) = \sum_{-\infty}^{x} f(t),$$

jika X peubah acak diskret, dan

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

jika X peubah acak kontinu.

Contoh 2.4.1 Berdasarkan Tabel 2.2.1 kita dapat membuat tabel cdf seperti berikut ini.

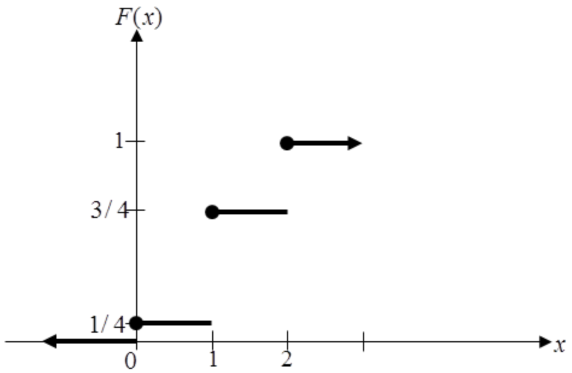
TABEL 2.4.1

x	0	1	2
$F(x) = P(X \leq x)$	$\frac{1}{4}$	$\frac{3}{4}$	1

Karena cdf berlaku untuk sebarang bilangan real, maka secara lengkap kita dapat menulisnya sebagai

$$F(x) = \frac{1}{4} I(0 \le x < 1) + \frac{3}{4} I(1 \le x < 2) + I(2 \le x).$$

Grafik dari cdf di atas adalah seperti pada gambar berikut ini.



Gambar 2.4.1

Contoh 2.4.2 Misal X peubah acak kontinu dengan pdf

$$f(x) = 3x^2 I(0 < x < 1).$$

Tentukan cdf dari X dan buatlah grafiknya.

Penyelesaian

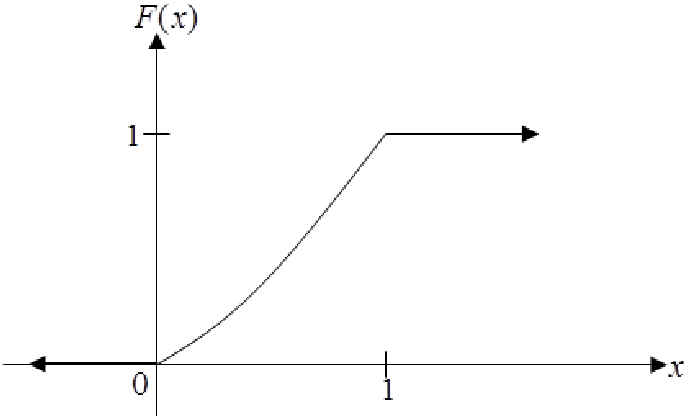
Karena X peubah acak kontinu, maka

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

Jadi

$$F(x) = x^3 I(0 \le x < 1) + I(1 \le x).$$

Grafiknya adalah seperti pada gambar berikut ini.



Gambar 2.4.2

Dua grafik pada dua contoh di atas memberikan ilustrasi tentang sifat-sifat cdf seperti pada teorema berikut ini.

Teorema 2.4.1 Suatu fungsi F adalah suatu cdf dari peubah acak X jika dan hanya jika memenuhi sifat-sifat:

(i) $0 \le F(x) \le 1$

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

(iii) $a < b \Rightarrow F(a) \le F(b)$

(iv) $\lim_{h \to 0^+} F(x+h) = F(x)$

Sifat pertama mengatakan bahwa nilai fungsi sebaran F paling kecil nol dan paling besar satu. Sifat kedua mengatakan bahwa $F(x)$ bernilai 0 atau 1 untuk sebarang bilangan yang cukup besar, berturut-turut negatif atau positif. Sifat ketiga mengatakan bahwa F tidak turun. Dan sifat terakhir mengatakan bahwa fungsi sebaran F adalah ***kontinu kanan*** di mana-mana. Bahkan untuk peubah acak kontinu, fungsi sebarannya tidak sekedar kontinu kanan, tetapi kontinu.

Berikut ini adalah bukti dari Teorema 2.4.1.

(i) Dari definisi cdf kita peroleh $F(x) = P(X \le x)$. Karena peluang nilai terkecilnya 0 dan nilai terbesarnya 1, maka

$$0 \le F(x) \le 1.$$

(ii) Kita misalkan $A = \{t : t \le x\}$. Oleh karena itu $F(x) = P(A)$. Selanjutnya karena $A = \varnothing$ jika $x \to -\infty$, maka menggunakan sifat kekontinuan dalam peluang,

$$\lim_{x\to-\infty} F(x) = \lim_{x\to-\infty} P(A) = P\left(\lim_{x\to-\infty} A\right) = P(\varnothing) = 0.$$

Dengan cara yang serupa kita peroleh

$$\lim_{x\to\infty} F(x) = P(S) = 1.$$

(iii) Kita punya $a < b$. Maka

$$\{x : x \le b\} = \{x : x \le a\} \cup \{x : a < x \le b\}$$

Oleh karena itu

$$P(X \le b) = P(X \le a) + P(a < X \le b).$$

Atau

$$F(b) = F(a) + P(a < X \le b).$$

Karena nilai peluang tidak negatif, maka

$$F(a) \le F(b).$$

(iv) Menggunakan hasil pada bagian (iii), kita peroleh

$$P(a < X \le b) = F(b) - F(a).$$

Oleh karena itu

$$\lim_{h\to0^+} P(x < X \le x+h) = \lim_{h\to0^+} \left[F(x+h) - F(x)\right].$$

Sekali lagi menggunakan kekontinuan dalam peluang kita peroleh

$$\lim_{h\to0^+} P(x < X \le x+h) = P\left[\lim_{h\to0^+}(x < X \le x+h)\right] = P(\varnothing) = 0.$$

Oleh karena itu

$$\lim_{h\to0^+} \left[F(x+h) - F(x)\right] = 0.$$

Jadi

$$\lim_{h\to0^+} F(x+h) = F(x).$$

Teorema selanjutnya menyatakan hubungan pdf dengan cdf. Pertama untuk peubah acak diskret, dan kedua untuk peubah acak kontinu.

Teorema 2.4.2 Misal X peubah acak diskret dengan pdf dan cdf berturut-turut f dan F. Jika nilai-nilai yang mungkin dari adalah $x_1, x_2, x_3, \ldots$ di mana $x_1 < x_2 < x_3 < \ldots$, maka

$$f(x_1) = F(x_1),$$

Dan untuk sebarang $i > 1,$

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

Bukti teorema di atas dapat diperoleh langsung dari definisi cdf. Demikian pula untuk teorema berikut ini.

Teorema 2.4.3 Misal X peubah acak kontinu dengan pdf dan cdf berturut-turut f dan F. Maka

$$F'(x) = f(x),$$

untuk setiap x sehingga f kontinu. Dengan kata lain pdf merupakan turunan pertama dari cdf untuk setiap titik x sehingga f kontinu.

Contoh 2.4.3 Misal f merupakan pdf dari peubah acak X, di mana $f(x) = (1/2)I(-1 < x < 1).$ Definisikan suatu peubah acak Y, sehingga $Y = X^2$. Kita akan menentukan pdf dari Y. Misal G dan g berturut-turut menyatakan cdf dan pdf dari Y. Jika $Y \geq 0,$ maka

$$G(y) = P(Y \leq y) = P(X^2 \leq y) = P(-\sqrt{y} \leq X \leq \sqrt{y}).$$

Oleh karena itu untuk $0 \leq y < 1$ kita peroleh

$$G(y) = \int_{-\sqrt{y}}^{\sqrt{y}} 1/2\, dx = \sqrt{y}\,.$$

Jadi

$$G(y) = \sqrt{y}\, I(0 \leq y < 1) + I(1 \leq y).$$

Selanjutnya karena Y peubah acak kontinu, maka $g(y) = G'(y)$ untuk setiap y sehingga g kontinu. Jadi

$$g(y) = \frac{1}{2\sqrt{y}} I(0 < y < 1).$$

Latihan 2.4

2.4.1. Misal $c = (i, j),$ menyatakan sebarang hasil dari pelemparan satu dadu sebanyak dua kali, dan misal $X(c) = i + j.$ Tentukan cdf peubah acak X dan buatlah grafik cdf-nya.

2.4.2. Ulangilah Latihan 2.4.1, untuk $Y(c) = i - j.$

2.4.3 Selidiki apakah fungsi-fungsi berikut ini merupakan cdf atau bukan. Berikan alasan jawaban Anda.

(a) $F(x) = e^{-x} I(0 \le x).$

(b) $F(x) = e^{x} I(x \le 0).$

(c) $F(x) = (1 - e^{-x}) I(-1 < x).$

(d) $F(x) = x^2 I(-1 \le x < 1) + I(1 \le x).$

2.4.4. Misal X suatu peubah acak diskret dengan cdf,

$$F(x) = \left(1 - (1/2)^{x+1}\right) I(x = 0, 1, 2, \ldots).$$

(a) Tentukan pdf dari X.
(b) Hitung $P(10 < X \le 20).$
(c) Hitung peluang bahwa X bernilai genap.

2.4.5. Misal suatu peubah acak mempunyai cdf

$$F(x) = \frac{x}{2} I(0 \le x < 1) + (x - \tfrac{1}{2}) I(1 \le x < 3/2) + I(3/2 \le x).$$

(a) Hitung $P(X \le 1/2).$
(b) Hitung $P(X \le 1.2).$
(c) Hitung $P(X = 1.2).$
(d) Tentukan pdf-nya.

2.4.6. Misal peubah acak diskret X mempunyai pdf

$$f(x) = \frac{x}{6} I(x = 1, 2, 3).$$

Tentukan cdf dan pdf dari $Y = X^2.$

2.4.7. Misal $f(x) = 2x I(0 < x < 1)$ merupakan pdf X. Tentukan cdf dan pdf dari $Y = X^2.$

2.4.8. Misal $f(x) = 4x^3 I(0 < x < 1)$ merupakan pdf X. Tentukan cdf dan pdf dari $Y = -2 \ln X^4.$

2.4.9. Misal X mempunyai pdf $f(x)=\frac{1}{\pi}I\left(-\frac{\pi}{2}<x<\frac{\pi}{2}\right)$. Tentukan pdf dari $Y=\tan X$. Peubah acak ini disebut ***sebaran Cauchy.***

2.4.10. Misal $f(x)=\frac{1}{3}I(-1<x<2)$ merupakan pdf X. Tentukan pdf dari $Y=X^2$.

2.4.11. Misal F suatu cdf dari X. Jika m adalah suatu bilangan sehingga $F(m)=1/2$, tunjukkan bahwa m adalah median dari X.

2.5 Ekspektasi Matematis

Satu dari beberapa konsep yang penting dari sebaran peluang adalah ekspektasi matematis. Kita lihat terlebih dahulu ilustrasi pada contoh berikut ini untuk peubah acak diskret.

Contoh 2.5.1 Suatu percobaan melempar sekeping uang logam yang seimbang dua kali. Dari percobaan ini tentunya kita berharap bahwa muka akan muncul sekali. Hasil ini secara matematika dapat kita tulis seperti berikut ini. Kita misalkan peubah acak X sebagai banyak muncul muka. Maka

$$f(x)=\frac{\binom{2}{x}}{4}I(x=0,1,2).$$

Selanjutnya nilai total dari X dikalikan dengan bobotnya dalam hal ini sama dengan $f(x)$ adalah

$$\sum_x xf(x)=0.\frac{1}{4}+1.\frac{2}{4}+2.\frac{1}{4}=1.$$

Nilai ini persis sama dengan nilai yang kita harapkan. Oleh karena itu nilai ini disebut sebagai nilai harapan/ekspektasi matematis dari X ditulis sebagai $E(X)$. Karena nilai total dari X dikalikan dengan bobotnya dapat pula dipandang sebagai ***purata*** dari X, maka nilai ini disebut juga sebagai purata dari X ditulis sebagai μ atau μ_X.

Definisi 2.5.1 Misal X peubah acak dengan pdf f sehingga

$$\sum_x |x|\, f(x)$$

ada untuk X diskret, dan

$$\int_{-\infty}^{\infty} |x|\, f(x)\, dx$$

ada untuk X kontinu. Maka ***ekspektasi matematis*** dari X ditulis $E(X)$ adalah

$$E(X) = \sum_x xf(x),$$

untuk X diskret, atau

$$E(X) = \int_{-\infty}^{\infty} xf(x)dx,$$

untuk X kontinu.

Contoh 2.5.2 Misal suatu peubah acak X mempunyai pdf bentuk

$$f(x) = \frac{2}{(1+x)^3} I(0 < x).$$

Tentukan $E(X)$.

Penyelesaian

Karena X peubah acak kontinu, maka untuk menentukan $E(X)$ kita menggunakan rumus yang kedua, yaitu dalam bentuk integral. Sehingga

$$E(X) = \int_{-\infty}^{\infty} xf(x)dx = \int_{-\infty}^{0} x.0dx + \int_{0}^{\infty} x \frac{2}{(1+x)^3} dx.$$

Kita substitusi $t = 1 + x,$ kita peroleh

$$E(X) = 2\int_{1}^{\infty} (t-1)t^{-3} dt = 1.$$

Ekspektasi matematis secara fisika dapat dipandang sebagai pusat massa. Untuk itu kita lihat contoh berikut ini.

Contoh 2.5.3 Misal dua benda masing-masing mempunyai massa 0.25 gram dan 0.75 gram. Benda tersebut kita letakkan pada sumbu horisontal berturut-turut pada $x = 2,$ dan $4.$ Pusat massa dari kedua benda itu adalah

$$2.(0.25) + 4.(0.75) = 3.5.$$

Pekerjaan menentukan pusat masa di atas sama saja dengan pekerjaan menentukan ekspektasi dengan memandang masa benda sebagai

$$f(x) = 0.25I(x = 2) + 0.75I(x = 4),$$

dan letak benda pada sumbu horisontal sebagai peubah acak X. Oleh karena itu kita peroleh

$$E(X) = \sum_x xf(x) = 2.(0.25) + 4.(0.75) = 3.5.$$

Misal $\mathfrak{A}$ merupakan ruang peubah acak X, dan misal $Y = u(X)$. Anggap X peubah acak kontinu dan $y = u(x)$ fungsi kontinu naik dari X dengan balikan $x = w(y)$. Misal G merupakan cdf dari Y, maka

$$G(y) = P(Y \leq y) = P\left[u(X) \leq y\right]$$

$$= P\left[X \leq w(y)\right] = \int_{-\infty}^{w(y)} f(x)\,dx,$$

di mana f menyatakan pdf dari X. Menggunakan Teorema Dasar Kalkulus kita peroleh

$$g(y) = G'(y) = f\left[w(y)\right]w'(y)\,I(y \in \mathfrak{M}),$$

di mana

$$\mathfrak{M} = \left\{y : y = u(x), x \in \mathfrak{A}\right\}.$$

Jika

$$\int_{-\infty}^{\infty} |\,y\,|\,g(y)\,dy$$

ada, maka ekspektasi matematis dari Y adalah

$$E(Y) = \int_{-\infty}^{\infty} yg(y)\,dy.$$

Selanjutnya kita bandingkan $E(Y)$ dengan integral

$$I = \int_{-\infty}^{\infty} u(x)\,f(x)\,dx.$$

Karena

$$\frac{dx}{dy} = w'(y) > 0,$$

maka

$$I = \int_{-\infty}^{\infty} yf\left[w(y)\right]w'(y)\,dy = \int_{-\infty}^{\infty} yg(y)\,dy.$$

Jadi

$$E(Y)=\int_{-\infty}^{\infty} yg(y)\,dy=\int_{-\infty}^{\infty} u(x)\,f(x)\,dx.$$

Integral tersebut diganti dengan jumlah untuk peubah acak diskret.

Teorema 2.5.1 Misal peubah acak X mempunyai pdf f, k, dan l konstanta, $g(x)$ dan $h(x)$ adalah fungsi-fungsi bernilai real yang domainnya berisi nilai-nilai yang mungkin dari X. Maka

$$E\left[kg(X)+lh(X)\right]=kE\left[g(X)\right]+lE\left[h(X)\right].$$

Bukti teorema tersebut dapat diperoleh dari sifat jumlah dan integral, ditinggalkan sebagai latihan. Teorema tersebut menyatakan bahwa E merupakan operator linear.

Contoh 2.5.4 Misal X mempunyai pdf

$$f(x)=\frac{x}{6}I(x=1,2,3).$$

Tentukan $E(X^3)$.

Penyelesaian

$$E(X^3)=\sum_x x^3 f(x)=\sum_{x=1}^{3} x^3\,\frac{x}{6}=\frac{1}{6}+\frac{16}{6}+\frac{81}{6}=\frac{98}{6}=\frac{49}{3}.$$

Contoh 2.5.5 Misal X mempunyai pdf

$$f(x)=2(1-x)I(0<x<1).$$

Tentukan: (i) $E(X)$, (ii) $E(X^2)$, dan $E(6X+3X^2)$.

Penyelesaian

(i)

$$E(X)=\int_{-\infty}^{\infty} x(f(x)dx=\int_0^1 (x)2(1-x)dx$$

$$=2\int_0^1 (x-x^2)dx=2\left[\frac{1}{2}x^2-\frac{1}{3}x^3\right]_0^1=\frac{1}{3}.$$

(ii)

$$E(X^2)=\int_{-\infty}^{\infty} x^2(f(x)dx=\int_0^1 (x^2)2(1-x)dx$$

$$=2\int_0^1 (x^2-x^3)dx=2\left[\frac{1}{3}x^3-\frac{1}{4}x^4\right]_0^1=\frac{1}{6}.$$

(iii)

$$E(6X+3X^2)=6E(X)+3E(X^2)=6\left(\frac{2}{6}\right)+3\left(\frac{1}{6}\right)=2.5.$$

Satu sifat khusus dari ekspektasi ini yang sangat penting dalam Statistika diperoleh dari fungsi $u(x)=(x-\mu)^2$. Dan ekspektasi dari fungsi ini disebut sebagai varians peubah acak X, seperti diberikan oleh definisi berikut ini.

Defnisi 2.5.2 Varians dari peubah acak X ditulis $Var(X)$, diberikan oleh

$$Var(X)=E\left[(X-\mu)^2\right].$$

Notasi lain dari varians peubah acak X adalah σ^2, σ_X^2, atau $Var(X)$. Selanjutnya akar positif dari varians ini disebut sebagai ***simpangan baku***, yaitu $\sigma=\sigma_X=\sqrt{Var(X)}$. Simpangan baku ini merupakan ukuran penyimpangan nilai-nilai peubah acak X dari puratanya. Akibatnya jika peubah acak X hanya mempunyai satu nilai, maka simpangan bakunya adalah nol.

Teorema berikut ini merupakan rumus lain dari varians, yang sering digunakan dalam menentukan varians.

Teorema 2.5.2 Misal X suatu peubah acak. Maka

$$Var(X)=E(X^2)-\mu^2.$$

Bukti

Pembuktian dari teorema ini cukup menggunakan sifat-sifat ekspektasi.

$$\begin{aligned}Var(X)&=E\left[(X-\mu)^2\right]=E\left[X^2-2X\mu+\mu^2\right]\\&=E(X^2)-E(2X\mu)+\mu^2=E(X^2)-2\mu E(X)+\mu^2\\&=E(X^2)-2\mu\mu+\mu^2=E(X^2)-\mu^2.\end{aligned}$$

Tidak seperti pada ekspektasi, varians bukan merupakan operator linear. Hal ini diberikan oleh teorema berikut ini.

Teorema 2.5.3 Misal X suatu peubah acak, a dan b suatu konstanta. Maka

$$Var(aX+b)=a^2Var(X).$$

Bukti

$$\begin{aligned}
Var(aX+b) &= E\left[(aX+b)^2\right]-\left[E(aX+b)\right]^2 \\
&= E\left[a^2X^2+2aXb+b^2\right]-\left[aE(X)+b\right]^2 \\
&= \left[a^2E(X^2)+2abE(X)+b^2\right]-\left[a^2(E(X))^2+2abE(X)+b^2\right] \\
&= a^2\left[E(X^2)-(E(X))^2\right]=a^2\left[E(X^2)-\mu^2\right]=a^2Var(X).
\end{aligned}$$

Teorema berikutnya adalah teorema tentang ketaksamaan Markov atau disebut juga teorema tentang batas peluang.

Teorema 2.5.4 Ketaksamaan Markov. Misal $u(X)$ adalah fungsi tak negatif dari peubah acak X. Jika $E\left[u(X)\right]$ ada, maka untuk setiap konstanta positif c berlaku

$$P\left[u(X)\geq c\right]\leq\frac{E\left[u(X)\right]}{c}.$$

Bukti

Kita buktikan untuk peubah acak X kontinu, sedangkan untuk peubah acak X diskret tinggal mengganti integral dengan jumlah. Gagasan penting dalam pembuktian teorema ini adalah pembentukan himpunan

$$A=\{x:u(x)\geq c\}.$$

Dari himpunan tersebut kita peroleh

$$E\left[u(X)\right]=\int_{-\infty}^{\infty}u(x)f(x)dx=\int_A u(x)f(x)dx+\int_{A^c}u(x)f(x)dx.$$

Karena $u(x)$ tidak negatif, demikian pula untuk $f(x)$, maka integral pada suku terakhir tidak negatif, atau

$$\int_{A^c}u(x)f(x)dx\geq 0.$$

Oleh karena itu

$$E\left[u(X)\right]\geq\int_A u(x)f(x)dx.$$

Di pihak lain, karena $u(x) \geq c$, untuk semua $x \in A$, maka

$$\int_A u(x) f(x) dx \geq \int_A cf(x) dx.$$

Sehingga kita peroleh

$$E[u(X)] \geq \int_A cf(x) dx = cP[X \in A] = cP[u(X) \geq c].$$

Jadi

$$P[u(X) \geq c] \leq \frac{E[u(X)]}{c}.$$

Teorema di atas merupakan bentuk umum dari ketaksamaan yang sering disebut sebagai ***Ketaksamaan Chebyshev,*** yang diberikan oleh teorema berikut ini.

Teorema 2.5.5 Ketaksamaan Chebyshev. Misal peubah acak X mempunyai purata u dan varians σ^2. Maka untuk sebarang k positif berlaku

$$P[|X - \mu| \geq k\sigma] \leq \frac{1}{k^2},$$

atau, ekivalen dengan

$$P[|X - \mu| < k\sigma] \geq 1 - \frac{1}{k^2}.$$

Bukti

Berdasarkan Teorema 2.5.4, jika kita ambil

$$u(X) = (X - \mu)^2$$

dan

$$c = k^2\sigma^2,$$

maka kita peroleh

$$P\left[(X - \mu)^2 \geq k^2\sigma^2\right] \leq \frac{E\left[(X - \mu)^2\right]}{k^2\sigma^2}.$$

Karena

$$E\left[(X - \mu)^2\right] = \sigma^2,$$

maka

$$P\left[(X - \mu)^2 \geq k^2\sigma^2\right] \leq \frac{1}{k^2}.$$

Di pihak lain

$$\{x : (x-\mu)^2 \geq k^2\sigma^2\} \Leftrightarrow \{x : |x-\mu| \geq k\sigma\}.$$

Maka

$$P\left[(X-\mu)^2 \geq k^2\sigma^2\right] = P\left[|X-\mu| \geq k\sigma\right].$$

Jadi

$$P\left[|X-\mu| \geq k\sigma\right] \leq \frac{1}{k^2}.$$

Ketaksamaan di atas sering pula dinyatakan dalam bentuk ε dengan memisalkan

$$\varepsilon = k\sigma.$$

Sehingga kita peroleh hasil

$$P\left[|X-\mu| \geq \varepsilon\right] \leq \frac{\sigma^2}{\varepsilon^2},$$

atau

$$P\left[|X-\mu| < \varepsilon\right] \geq 1 - \frac{\sigma^2}{\varepsilon^2}.$$

Terlihat bahwa bilangan $\frac{1}{k^2}$ merupakan batas atas peluang $P\left[|X-\mu| \geq k\sigma\right]$. Contoh berikut ini membandingkan peluang eksak dan batas atas peluang.

Contoh 2.5.6 Misal peubah acak X mempunyai pdf bentuk

$$f(x) = \frac{1}{2\sqrt{3}} I(-\sqrt{3} < x < \sqrt{3}).$$

Tentukan:

(i) Peluang eksak dari $P\left(|X| \geq \frac{3}{2}\right)$.

(ii) Batas peluang dari $P\left(|X| \geq \frac{3}{2}\right)$.

Penyelesian

(i)

$$P\left(|X|\geq \frac{3}{2}\right)=1-P\left(|X|<\frac{3}{2}\right)=1-\int_{-3/2}^{3/2}\frac{1}{2\sqrt{3}}dx$$

$$=1-\left[\frac{1}{2\sqrt{3}}x\right]_{-3/2}^{3/2}=1-\frac{\sqrt{3}}{2}\approx 0.134.$$

(ii) Kita gunakan ketaksamaan Chebyshev untuk menentukan batas peluang dari $P\left(|X|\geq \frac{3}{2}\right)$. Untuk itu terlebih dahulu kita hitung μ dan $\sigma,$ kemudian ditentukan k seperti berikut ini:

$$\mu=E(X)=\int_{-\infty}^{\infty}xf(x)dx=\int_{-\sqrt{3}}^{\sqrt{3}}x\frac{1}{2\sqrt{3}}dx=\left[\frac{x^2}{4\sqrt{3}}\right]_{-\sqrt{3}}^{\sqrt{3}}=0,$$

$$\sigma^2=E(X^2)-\left[E(X)\right]^2=\int_{-\infty}^{\infty}x^2f(x)dx=\int_{-\sqrt{3}}^{\sqrt{3}}x^2\frac{1}{2\sqrt{3}}dx=\left[\frac{x^3}{6\sqrt{3}}\right]_{-\sqrt{3}}^{\sqrt{3}}=1.$$

Untuk menentukan k kita lihat lagi peluang berikut

$$P\left(|X|\geq \frac{3}{2}\right)=P\left(|X-0|\geq \frac{3}{2}\right)=P\left(|X-\mu|\geq k\sigma\right).$$

Karena $\sigma=1,$ kita peroleh $k=\frac{3}{2}.$ Sehingga kita peroleh

$$P\left(|X|\geq \frac{3}{2}\right)\leq \frac{4}{9}\approx 0.444.$$

Dengan kata lain batas atas dari $P\left(|X|\geq \frac{3}{2}\right)$ adalah $\frac{4}{9}$.

Kenyataannya batas atas ini memang lebih dari nilai eksak, yaitu

$$0.134\leq 0.444.$$

Hasil ini sesuai dengan ketaksamaan Chebyshev.

Contoh 2.5.7 Misal peubah acak diskret X mempunyai peluang berturut-turut $\frac{1}{8},\frac{6}{8},$ dan $\frac{1}{8}$ pada titik-titik $x=-1,0,1.$ Tentukan:

(i) Peluang eksak dari $P\left(|X|\geq 1\right).$

(ii) Batas peluang dari $P\left(|X|\geq 1\right).$

Penyelesaian
(i)

$$P\left(|X|\geq 1\right)=P(X=-1)+P(X=1)=\frac{1}{8}+\frac{1}{8}=\frac{1}{4}.$$

(ii) Karena X peubah acak diskret maka,

$$\mu=\sum_x xf(x)=(-1).\frac{1}{8}+0.\frac{6}{8}+1.\frac{1}{8}=0,$$

dan

$$\sigma^2=E(X^2)-\mu^2=\sum_x x^2 f(x)=1.\frac{1}{8}+0.\frac{6}{8}+1.\frac{1}{8}=\frac{1}{4}.$$

Kita peroleh $\sigma=\frac{1}{2}.$ Selanjutnya kita tentukan k seperti berikut ini

$$P\left(|X|\geq 1\right)=P\left(|X-0|\geq 1\right)=P\left(|X-\mu|\geq k\sigma\right).$$

Karena $\mu=0$ dan $\sigma=\frac{1}{2},$ maka $k=2.$ Oleh karena itu

$$P\left(|X|\geq 1\right)\leq\frac{1}{k^2}=\frac{1}{4}.$$

Dengan kata lain batas atas dari $P\left(|X|\geq 1\right)$ adalah $\frac{1}{4}.$

Terlihat bahwa batas atas ini sama dengan nilai eksaknya. Meskipun demikian ketaksamaan Chebyshev masih berlaku, karena

$$\frac{1}{4}\leq\frac{1}{4}.$$

Kita tutup pembahasan pada subpokok bahasan ini dengan Ketaksamaan Jensen. Meskipun demikian kita sajikan definisi tentang fungsi cembung seperti berikut ini.

Definisi 2.5.3 Suatu fungsi ϕ yang terdefinisi pada selang terbuka $(a,b),\ -\infty\leq a<b\leq\infty,$ dikatakan fungsi cembung (konveks) jika untuk semua $x,\ y$ dalam (a,b) dan untuk semua $0<\gamma<1,$

$$\phi\left[\gamma x+(1-\gamma)y\right]\leq\gamma\phi(x)+(1-\gamma)\phi(y).$$

Kita katakana ϕ adalah cembung sempurna jika ketaksamaan di atas adalah sempurna.

Bergantung pada keberadaan dari turunan pertama atau kedua dari ϕ, teorema berikut dapat dibuktikan.

Teorema 2.5.6 Jika ϕ dapat diturunkan pada (a,b), maka:

(a) ϕ adalah cembung jika dan hanya jika $\phi'(x) \le \phi'(y)$, untuk semua $a < x < y < b$,

(b) ϕ adalah cembung sempurna jika dan hanya jika $\phi'(x) < \phi'(y)$, untuk semua $a < x < y < b$.

Jika ϕ mempunyai turunan kedua pada (a,b), maka:

(a) ϕ adalah cembung jika dan hanya jika $\phi''(x) \ge 0$, untuk semua $a < x < b$,

(b) ϕ adalah cembung sempurna jika dan hanya jika $\phi''(x) > 0$, untuk semua $a < x < b$.

Bukti dari Teorema 2.5.6 dapat ditemui pada buku-buku analisis.

Teorema 2.5.7 Ketaksamaan Jensen. Jika ϕ cembung pada selang terbuka I dan X adalah peubah acak yang terdefinisi pada selang yang termuat dalam I dan mempunyai ekspektasi terhingga, maka

$$\phi\left[E(X)\right] \le E\left[\phi(X)\right].$$

Jika ϕ cembung sempurna, maka ketaksamaan adalah sempurna, kecuali X peubah acak konstan.

Bukti

Kita anggap ϕ mempunyai turunan kedua, tetapi secara umum ini dipenuhi oleh kecembungan. Kita perluas $\phi(x)$ ke dalam deret Taylor sekitar $\mu = E(X)$ dari urutan kedua:

$$\phi(x) = \phi(u) + \phi'(u)(x-\mu) + \frac{\phi''(\zeta)(x-\mu)^2}{2},$$

di mana ζ antara μ dan x. Karena suku terakhir pada ruas kanan dari kesamaan di atas tidak negatif, maka kita punya

$$\phi(x) \ge \phi(u) + \phi'(u)(x-\mu).$$

Dengan mengambil ekspektasi dari kedua sisi, maka bukti selesai. Ketaksamaan akan sempurna asalkan X tidak konstan.

Latihan 2.5

2.5.1. Misal peubah acak X mempunyai pdf dengan rumus
$f(x) = \dfrac{x+2}{18} I(-2 < x < 4).$ Tentukan:

(a) $E(X).$

(b) $E\left[(X+2)^3\right].$

(c) $E\left[6X - 2(X+2)^3\right].$

2.5.2. Anggap bahwa $f(x) = \dfrac{1}{5} I(x = 1, 2, 3, 4, 5),$ adalah pdf dari peubah acak diskret X. Hitung $E(X)$ dan $E(X^2).$ Gunakan hasil ini untuk menghitung $E\left[(X+2)^2\right].$

2.5.3. Misal X peubah acak dengan pdf f yang positif pada $x = -1, 0, 1,$ dan nol untuk yang lainnya.

(a) Jika $f(0) = \dfrac{1}{2},$ maka tentukan $E(X^2).$

(b) Jika $f(0) = \dfrac{1}{2}$ dan jika $E(X) = \dfrac{1}{6},$ maka tentukanlah $f(-1)$ dan $f(1).$

2.5.4 Tentukan purata dan varians, jika ada, dari masing-masing sebaran berikut ini.

(a) $f(x) = \dbinom{3}{x}\left(\dfrac{1}{2}\right)^3 I(x = 0, 1, 2, 3).$

(b) $f(x) = 6x(1-x) I(0 < x < 1).$

2.5.5. Misal peubah acak kontinu X mempunyai pdf $f(x) = \frac{1}{x^2} I(1 < x < \infty)$.

(a) Apakah $E(X)$ ada?

(b) Apakah $E\left(\frac{1}{X}\right)$ ada?

(c) Berapakah nilai k agar $E\left(X^k\right)$ ada?

2.5.6. Misal varians dari peubah acak X ada. Tunjukkan bahwa $E(X^2) \geq [E(X)]^2$.

2.5.7. Misal peubah acak kontinu X mempunyai pdf $f(x)$, yang grafiknya simetrik terhadap titik $x = c$. Kita katakan bahwa X adalah bersebaran simetrik terhadap c. Tunjukkan:

(a) $E(X) = c$, jika nilai purata dari X ada. *Petunjuk*: Tunjukkan bahwa $E(X - c) = 0$ dengan menulis $E(X - c)$ sebagai jumlah dari dua integral, yaitu satu dari $-\infty$ ke c dan yang lainnya dari c ke ∞. Dalam kasus pertama misalkan $y = c - x$, dan kasus yang lainnya misalkan $z = x - c$. Akhirnya gunakan sifat simetri $f(c - y) = f(c + y)$.

(b) $Y = X - c$ bersebaran simetrik terhadap 0.

(c) $Z = -(X - c)$ bersebaran sama dengan bagian (b).

2.5.8. Buktikan Teorema 2.5.1.

2.5.9. Misal X suatu peubah acak sehingga $E\left[\left(X - b\right)^2\right]$ ada untuk semua bilangan real b.

(a) Tunjukkan bahwa $E\left[\left(X - b\right)^2\right]$ minimum jika $b = E(X)$.

(b) Beri contoh suatu sebaran dari peubah acak diskret di mana $E\left[\left(X - b\right)^2\right] = 0$. Sebaran semacam ini disebut sebagai ***sebaran merosot (degenerate).***

2.5.10. Misal X peubah acak kontinu dengan pdf f. Jika m adalah median tunggal dari X dan b sebarang konstanta real. Tunjukkan bahwa

$$E\left(|X-b|\right)=E\left(|X-m|\right)+2\int_m^b (b-x)f(x)dx,$$

asalkan kedua ekspektasi tersebut ada. Apakah b nilai minimum dari $E\left(|X-b|\right)$?

2.5.11. Misal X peubah acak kontinu dengan pdf f, yang terdefinisi pada bilangan positif $0<x<b<\infty$. Tunjukkan bahwa

$$E(X)=\int_0^b \left[1-F(x)\right]dx,$$

di mana F cdf dari X.

2.5.12. Misal X peubah acak kontinu sehingga pdf-nya adalah $f(x)=3x^2 I(0<x<1)$.

(a) Gunakan ketaksamaan Chebyshev untuk memperoleh batas bawah dari $P(5/8<X<7/8)$.

(b) Ulangi Latihan (a) untuk $P(1/2<X<1)$.

2.5.13. Misal X peubah acak positif, yaitu, $P(X\le 0)=0$. Beri argumentasi ketaksamaan berikut:

(a) $E(1/X)\ge 1/E(X)$,

(b) $E\left[-\log X\right]\ge -\log\left[E(X)\right]$,

(c) $E\left[\log(1/X)\right]\ge \log\left[1/E(X)\right]$,

(d) $E\left[X^3\right]\ge \left[E(X)\right]^3$.

2.5.14. Misal X peubah acak kontinu sehingga pdf-nya adalah $f(x)=\dfrac{1}{x^2}I(1<x)$. Tunjukkan bahwa $E(X)$ tidak ada.

2.6 Fungsi Pembangkit Momen

Nilai ekspektasi khusus lainnnya lagi yang sering digunakan dalam Statistika adalah fungsi pembangkit momen.

Definisi 2.6.1 Misal X suatu peubah acak. Maka nilai ekspektasi

$$M_X(t) = E(e^{tX})$$

disebut ***fungsi pembangkit momen*** dari X jika nilai ekspektasi tersebut ada pada selang $-h < t < h,$ untuk suatu $h > 0.$

Untuk selanjutnya fungsi pembangkit momen ini ditulis sebagai **mgf,** yang merupakan singkatan dari ***moment generating function.*** Jika tidak menimbulkan kerancuan, yaitu dalam hal kita hanya menghadapi satu peubah acak, maka mgf ini cukup ditulis sebagai $M(t).$

Seperti halnya pada ekspektasi sebelumnya, untuk fungsi pembangkit momen ini berakibat pula bahwa: jika X peubah acak diskret maka

$$M(t) = \sum_x e^{tx} f(x),$$

dan jika X peubah acak kontinu maka

$$M(t) = \int_{-\infty}^{\infty} e^{tx} f(x)dx.$$

Khususnya untuk $t = 0,$ kita mempunyai $M(0) = 1$ (mengapa?). Jelas bahwa mgf ada pada kitaran (*neighborhood*) terbuka dari $0.$ Oleh karena itu jika mgf ada, maka dia haruslah ada pada selang terbuka sekitar $0.$ Meskipun demikian mgf dari suatu sebaran tidak selalu ada.

Contoh 2.6.1 Misal X peubah acak diskret dengan pdf

$$f(x) = \frac{\binom{2}{x}}{4} I(x = 0,1,2).$$

Maka mgf dari X adalah

$$M(t) = \sum_x e^{tx} f(x) = e^{t.0}.\frac{1}{4} + e^{t.1}.\frac{2}{4} + e^{t.2}.\frac{1}{4} = \frac{1}{4} + \frac{2}{4}e^{t} + \frac{1}{4}e^{2t}.$$

Contoh 2.6.2 Misal seseorang menunggu suatu bus pada tempat tertentu. Misal bus akan lewat tempat tersebut setiap 15 menit. Jika waktu tunggu dinyatakan sebagai peubah acak X, maka pdf-nya adalah

$$f(x) = \frac{1}{15} I(0 < x < 15).$$

Oleh karena itu fungsi pembangkit momen dari X adalah

$$M(t) = \int_{-\infty}^{\infty} e^{tx} f(x)dx = \int_{0}^{15} \frac{1}{15} e^{tx} dx = \frac{1}{15}\left[\frac{1}{t}e^{tx}\right]_0^{15} = \frac{1}{15t}\left[e^{15t} - 1\right], t \neq 0.$$

Misal X dan Y dua peubah acak yang mempunyai mgf. Jika X dan Y mempunyai cdf yang sama, yaitu, $F_X(z) = F_Y(z)$ untuk semua z, maka tentu $M_X(t) = M_Y(t)$ dalam kitaran terbuka 0, dan sebaliknya berlaku benar. Ini berarti mgf dari suatu sebaran, jika ada, dia adalah tunggal. Pernyataan tersebut disajikan pada teorema berikut, tanpa bukti.

Teorema 2.6.1 Misal X dan Y dua peubah acak dengan mgf berturut-turut $M_X(t)$ dan $M_Y(t)$, ada dalam kitaran terbuka 0. Maka $F_X(z) = F_Y(z)$ untuk semua $z \in R$ jika dan hanya jika $M_X(t) = M_Y(t)$ untuk semua $t \in (-h, h)$ untuk suatu $h > 0$.

Contoh 2.6.3 Misal peubah acak X mempunyai mgf

$$M(t) = \frac{1}{10}e^{t} + \frac{2}{10}e^{2t} + \frac{3}{10}e^{3t} + \frac{4}{10}e^{4t}$$

untuk semua bilangan real t. Jika kita misalkan f merupakan pdf dari X dan misal $a, b, c, d, ...$ merupakan titik-titik diskret dalam ruang peubah acak X sehingga $f(x) > 0$, maka

$$M(t) = \sum_{x} e^{tx} f(x).$$

Oleh karena itu

$$\frac{1}{10}e^{t} + \frac{2}{10}e^{2t} + \frac{3}{10}e^{3t} + \frac{4}{10}e^{4t} = f(a)e^{at} + f(b)e^{bt} + f(c)e^{ct} + f(d)e^{dt} + ...$$

Karena kesamaan ini merupakan suatu identitas, yaitu berlaku untuk semua bilangan real t, maka kita peroleh

$$a=1,\ f(a)=\frac{1}{10},b=2,\ f(b)=\frac{2}{10},c=3,\ f(c)=\frac{3}{10},d=4,\ f(d)=\frac{4}{10.}$$

Jadi pdf dari X adalah f sehingga

$$f(x)=\frac{x}{10}I(x=1,2,3,4).$$

Contoh 2.6.4 Misal kontinu X mempunyai mgf

$$M(t)=\frac{1}{(1-t)^2},\ (t<1).$$

Oleh karena itu

$$\frac{1}{(1-t)^2}=\int_{-\infty}^{\infty}e^{tx}f(x)\,dx,\ (t<1).$$

Umumnya tidak mudah untuk menentukan $f(x)$. Meskipun demikian dalam contoh ini kita mudah melihat bahwa pdf dari X

$$f(x)=xe^{-x}I(0<x<\infty).$$

Teorema 2.6.2 Misal mgf dari peubah acak X ada. Maka

(i) $E(X^r)=M^r(0), r=1,2,3,...$

(ii) $M(t)=1+\sum_{r=1}^{\infty}\frac{E(X^r)t^r}{r!}.$

Bukti

Kita buktikan untuk kasus peubah acak X kontinu. Sedangkan untuk peubah acak X diskret tinggal mengganti integral dengan jumlah. Kita mulai dengan membuktikan bagian pertama, dengan menulis terlebih dahulu mgf dari X seperti berikut ini

$$M(t)=\int_{-\infty}^{\infty}e^{tx}f(x)dx.$$

Jika mgf tersebut ada maka turunan pertamanya adalah

$$M'(t)=\int_{-\infty}^{\infty}xe^{tx}f(x)dx.$$

Oleh karena itu

$$M'(0)=\int_{-\infty}^{\infty}xe^{0x}f(x)dx=\int_{-\infty}^{\infty}xf(x)dx=E(X).$$

Turunan keduanya adalah

$$M''(t)=\int_{-\infty}^{\infty} x^2 e^{tx} f(x)dx.$$

Sehingga kita peroleh

$$M''(0)=\int_{-\infty}^{\infty} x^2 e^{0x} f(x)dx=\int_{-\infty}^{\infty} x^2 f(x)dx=E(X^2).$$

Proses ini dapat dilanjutkan sampai ke r di mana $r=1,2,...$ Oleh karena itu kita peroleh

$$M^r(t)=\int_{-\infty}^{\infty} x^r e^{tx} f(x)dx,$$

dan

$$M^r(0)=E(X^r), r=1,2,...$$

Untuk membuktikan bagian kedua kita gunakan rumus perluasan deret di sekitar nol, yaitu

$$M(t)=M(0)+\frac{M'(0)t}{1!}+\frac{M''(0)t^2}{2!}+...=M(0)+\sum_{r=1}^{\infty}\frac{M^r(0)t^r}{r!}.$$

Karena mgf pada saat $t=0$ adalah satu, maka kita peroleh

$$M(t)=1+\sum_{r=1}^{\infty}\frac{E(X^r)t^r}{r!},$$

bukti selesai.

Bentuk

$$E(X^r)=M^r(0).$$

disebut sebagai ***momen ke*** r di sekitar pusat dari peubah acak X. Inilah alasan bahwa $M(t)$ disebut fungsi pembangkit momen.

Selanjutnya kita lihat hubungan antara mgf dengan μ dan σ^2 seperti berikut ini. Menggunakan Teorema 2.6.2 bagian (i), maka untuk $r=1,$ kita peroleh

$$\mu=E(X)=M'(0).$$

Sedangkan untuk σ^2 kita peroleh

$$\sigma^2=E(X^2)-\left[E(X)\right]^2=M''(0)-\left[M'(0)\right]^2.$$

Kembali ke Contoh 2.6.1,

$$M'(t)=\frac{2}{4}e^t+\frac{2}{4}e^{2t}.$$

Jadi

$$\mu = M'(0) = \frac{2}{4} + \frac{2}{4} = 1.$$

Lebih lanjut untuk memperoleh varians kita kerjakan seperti berikut ini:

$$M'(t) = \frac{2}{4}e^{t} + \frac{2}{4}e^{2t},$$

dan

$$M''(t) = \frac{2}{4}e^{t} + \frac{4}{4}e^{2t} = \frac{1}{2}e^{t} + e^{2t}.$$

Sehingga

$$M''(0) = \frac{1}{2} + 1.$$

Jadi

$$\sigma^2 = M''(0) - \left[M'(0)\right]^2 = \left(\frac{1}{2} + 1\right) - 1 = \frac{1}{2}.$$

Teorema 2.6.3 Misal X dan Y dua peubah acak sehingga

$$Y = aX + b.$$

Maka

$$M_Y(t) = e^{bt} M_X(at).$$

Bukti

$$\begin{aligned} M_Y(t) &= E(e^{tY}) = E(e^{t(aX+b)}) = E(e^{taX+tb}) = E(e^{taX}e^{tb}) \\ &= e^{bt}E(e^{taX}) = e^{bt}M_X(at), \end{aligned}$$

terbukti.

Latihan 2.6

2.6.1 Tunjukkan bahwa mgf dari peubah acak X yang mempunyai pdf $f(x) = \frac{1}{3}I(-1 < x < 2),$ adalah

$$M(t) = \begin{cases} \dfrac{e^{2t} - e^{-t}}{3t}, & t \neq 0, \\ 1 & , t = 0. \end{cases}$$

2.6.2. Misal peubah acak diskret X mempunyai pdf

$$f(x)=\left(\frac{1}{2}\right)^{x+1} I(x=0,1,2,\ldots).$$

Tentukan:

(a) mgf. *Petunjuk*: gunakan rumus pada deret geometrik.

(b) Purata.

(c) Varians.

2.6.3. Misal peubah acak kontinu X mempunyai pdf $f(x)=e^{-x}I(x>0).$ Tentukan

(a) mgf. (b) Purata. (c) Varians .

2.6.4. Anggap peubah acak diskret X mempunyai mgf

$$M(t)=\frac{1}{8}e^{t}+\frac{1}{4}e^{2t}+\frac{5}{8}e^{5t}.$$

Tentukan:

(a) Purata. (b) Varians. (c) pdf. (d) $P(X=2).$ (e) cdf.

2.6.5. Misal peubah acak X mempunyai purata μ, simpangan baku σ dan mgf $M(t), -h<t<h.$ Tunjukkan bahwa

$$E\left(\frac{X-\mu}{\sigma}\right)=0, \qquad E\left[\left(\frac{X-\mu}{\sigma}\right)^{2}\right]=1,$$

dan

$$E\left\{\exp\left[t\left(\frac{X-\mu}{\sigma}\right)\right]\right\}=e^{-t\frac{\mu}{\sigma}}M\left(\frac{t}{\sigma}\right), -h\sigma<t<h\sigma.$$

2.6.6. Tentukan momen-momen dari peubah acak yang mempunyai mgf

$$M(t)=(1-t)^{-1}, t<1.$$

2.6.7. Misal X peubah acak dengan mgf $M(t),\ -h<t<h.$ Buktikan bahwa

$$P(X\geq a)\leq e^{-at}M(t),\ 0<t<h,$$

dan

$$P(X \leq a) \leq e^{-at} M(t), \ -h < t < 0.$$

Petunjuk: Misalkan $u(x) = e^{tx}$ dan $c = e^{ta}$ dalam Teorema 2.5.4.

2.6.8. Misal mgf dari X ada untuk semua t dan diberikan oleh

$$M(t) = \frac{e^t - e^{-t}}{2t}, \ t \neq 0, \ M(0) = 1.$$

Gunakan Latihan 2.6.7 untuk menentukan $P(X \geq 1)$ dan $P(X \leq -1)$.

2.6.9. Misal X peubah acak dengan pdf $f(x)$ dan mgf $M(t)$. Anggap f simetri terhadap 0, yaitu $f(-x) = f(x)$. Tunjukkan bahwa $M(-t) = M(t)$.

2.6.10. Misal $\psi(t) = \ln M(t)$, di mana $M(t)$ menyatakan mgf dari X. Buktikan bahwa $\psi'(0) = \mu$ dan $\psi''(0) = \sigma^2$.

BAB 3
BEBERAPA SEBARAN KHUSUS

3.1 Pendahuluan

Dalam bab ini kita akan membahas beberapa sebaran peluang khusus. Bentuk terapannya adalah kita mengorganisasikan karakteristik tertentu, kemudian menentukan sebarannya. Sebaran yang akan kita bahas ini dikatakan sebaran khusus karena mempunyai sifat-sifat khusus, di antaranya adalah pdf, ekspektasi, varians, dan mgf mempunyai bentuk tertentu.

Tujuan instruksional umum dari mempelajari bab ini adalah Anda diharapkan memahami konsep tentang sebaran khusus.

Tujuan instruksional khusus mempelajari bab ini adalah Anda diharapkan dapat mengidentifikasi, menentukan ekspektasi, varians, dan fungsi pembangkit momen peubah acak yang bersebaran: Bernoulli, binomial, hipergeometrik, binomial negatif, geometrik, Poisson, seragam, gamma, eksponensial, Weibull, Pareto, dan normal.

3.2 Sebaran Bernoulli, Binomial, dan Hipergeometrik

Anggap suatu percobaan dilakukan sekali dan kita hanya memperhatikan dua kemungkinan kejadian, katakanlah kejadian E dan kejadian E^C. Percobaan yang demikian disebut sebagai percobaan ***Bernoulli*.** Sebagai contoh E merupakan kejadian muncul muka dan E^C kejadian muncul belakang pada percobaan pelemparan sekeping uang logam yang dilakukan sekali. Contoh lainnya adalah E menyatakan kejadian mendapatkan cip rusak dan E^C menyatakan kejadian mendapatkan cip baik pada percobaan pengambilan satu cip dari suatu pabrik. Secara umum E biasa disebut sebagai ***kejadian sukses*** dan E^C disebut kejadian tidak sukses atau disebut gagal.

Sekarang anggap bahwa peluang untuk mendapatkan kejadian sukses adalah p, dan peluang mendapatkan kejadian gagal adalah q, maka kita mendapatkan $P(E) = p$ dan $P(E^C) = q = 1 - p.$

Misal suatu peubah acak X hanya bernilai 0 atau 1, yaitu

$$X(e) = I(e \in E).$$

Maka

$$f(0) = P(X = 0) = q,$$

dan

$$f(1) = P(X = 1) = p.$$

Secara umum kita menulis pdf dari X sebagai

$$f(x) = p^x q^{1-x} I(x = 0,1).$$

Peubah acak X yang demikian dikatakan ***sebaran Bernoulli***. Mudah dibuktikan bahwa

$$\sum_{x=0}^{1} f(x) = 1,$$

bukti ditinggalkan sebagai latihan.

Ekspektasi dan varians dari peubah acak ini mudah kita tentukan, yaitu

$$E(X) = \sum_x x f(x) = \sum_{x=0}^{1} x p^x q^{1-x} = 0.p^0 q^{1-0} + 1.p^1 q^{1-1} = p.$$

Sedangkan untuk menentukan variansnya kita tentukan terlebih dahulu $E(X^2)$,

$$E(X^2) = \sum_x x^2 f(x) = \sum_{x=0}^{1} x^2 p^x q^{1-x} = 0^2.p^0 q^{1-0} + 1^2.p^1 q^{1-1} = p.$$

Sehingga

$$Var(X) = E(X^2) - \left[E(X)\right]^2 = p - p^2 = p(1-p) = pq.$$

Contoh 3.2.1 Misal suatu dadu dilempar sekali. Anggap kejadian sukses adalah kejadian jika mendapatkan mata dadu 1. Maka $E = \{1\}$, $E^C = \{2,3,4,5,6\}$, $p = \frac{1}{6}$ dan $q = \frac{5}{6}$. Oleh karena itu

$$f(x) = \left(\frac{1}{6}\right)^x \left(\frac{5}{6}\right)^{1-x} I(x = 0,1),$$

$$E(X) = \frac{1}{6},$$

dan

$$Var(X) = \frac{1}{6}.\frac{5}{6} = \frac{5}{36}.$$

Contoh 3.2.2 Misal suatu kantong berisi 10 kelereng merah dan 20 kelereng putih. Dari kantong tersebut diambil satu kelereng secara acak. Misal pengambilan dikatakan sukses jika mendapatkan kelereng merah. Maka $E = \{m\},\ E^C = \{p\},\ p = \frac{1}{3},$ dan $q = \frac{2}{3}.$ Oleh karena itu

$$f(x) = \left(\frac{1}{3}\right)^x \left(\frac{2}{3}\right)^{1-x} I(x = 0,1),$$

$$E(X) = \frac{1}{3},$$

dan

$$Var(X) = \frac{1}{3}.\frac{2}{3} = \frac{2}{9}.$$

Selain pdf, ekspektasi, dan varians, kita dapat juga menentukan mgf dari sebaran ini, yaitu

$$M(t) = E\left(e^{tX}\right) = \sum_x e^{tx} f(x) = \sum_{x=0}^{1} e^{tx} p^x q^{1-x} = q + e^t p = pe^t + q.$$

Dari mgf ini dapat pula diturunkan ekspektasi dan varians. Dapat kita periksa (sebagai latihan) bahwa hasilnya akan sama dengan hasil ekspektasi dan varians di atas. Demikian juga untuk mgf dari Contoh 3.2.1 dan Contoh 3.2.2 (ditinggalkan sebagai latihan).

Pembahasan berikutnya adalah suatu sebaran yang ditimbulkan oleh percobaan Bernoulli dengan pengulangan sebanyak n kali di mana masing-masing pengulangan bebas.

Kita kembali ke Contoh 3.2.2. Anggap dari kantong tersebut diambil satu persatu sebanyak lima kali ***dengan pengembalian***. Misal X menyatakan banyak merah. Maka $X = 2$ ekivalen dengan kejadian $C = \{mmppp, mpmpp, ..., pppmm\}.$ Banyak anggota kejadian C merupakan kombinasi 2 dari 5 unsur, atau $\binom{5}{2}$, dan masing-masing anggota/titik mempunyai kesempatan yang sama yaitu, $\left(\frac{1}{3}\right)^2 \left(\frac{2}{3}\right)^3$. Oleh karena itu kita peroleh peluang

$$P(X=2)=\binom{5}{2}\left(\frac{1}{3}\right)^2\left(\frac{2}{3}\right)^3.$$

Dari hasil ini kita dapat memperoleh pula pdf dari X, yaitu

$$f(x)=P(X=x)=\binom{5}{x}\left(\frac{1}{3}\right)^x\left(\frac{2}{3}\right)^{5-x} I(x=0,1,2,3,4,5).$$

Secara umum jika suatu percobaan Bernoulli diulang sebanyak n kali di mana masing-masing pengulangan bebas dengan peluang kejadian sukses p, peluang kejadian gagal q, dan jika X menyatakan banyak sukses, maka pdf-nya adalah

$$f(x)=\binom{n}{x}p^x q^{n-x} I(x=0,1,...,n).$$

Peubah acak X yang demikian ini dikatakan ***bersebaran binomial***, ditulis $X\sim BIN(n,p)$ pdf-nya ditulis sebagai $b(x;n,p)$. Sehingga kita peroleh

$$b(x;n,p)=\binom{n}{x}p^x q^{n-x} I(x=0,1,...,n).$$

Kita dapat memeriksa bahwa pdf ini memenuhi ketiga syarat definisi pdf. Khususnya untuk syarat kedua dapat kita lihat seperti berikut ini:

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

Selanjutnya cdf dari sebaran ini ditulis sebagai $B(x;n,p)$, di mana

$$B(x;n,p)=\sum_{k=0}^{x} b(k;n,p), x=0,1,...,n,$$

untuk n dan p tertentu nilai masing-masing x dapat diperoleh dari tabel sebaran binomial atau Komputer dengan perangkat lunak Minitab atau perangkat lunak statistik yang lainnya. Sebagai contoh jika tentukan $n=10$ dan $p=0.1$, maka

$$B(6;10,0.1)=0.99999,$$

dari Minitab.

Selanjutnya kita akan menentukan ekspektasi sebaran ini berdasarkan definisi ekspektasi seperti berikut ini:

$$
\begin{aligned}
E(X) &= \sum_{x} xf(x) = \sum_{x=0}^{n} xb(x;n,p) = \sum_{x=0}^{n} x\binom{n}{x} p^x q^{n-x} = \sum_{x=1}^{n} x\binom{n}{x} p^x q^{n-x} \\
&= \sum_{x=1}^{n} x \frac{n!}{x!(n-x)!} p^x q^{n-x} = np\sum_{x=1}^{n} \frac{(n-1)!}{(x-1)!(n-x)!} p^{x-1} q^{n-x} \\
&= np\sum_{x=1}^{n} \frac{(n-1)!}{(x-1)!\left[(n-1)-(x-1)\right]!} p^{x-1} q^{(n-1)-(x-1)} \\
&= np\sum_{y=0}^{n-1} \frac{(n-1)!}{y!\left[(n-1)-y\right]!} p^{y} q^{(n-1)-y} \\
&= np(p+q)^{n-1} = np.1^{n-1} = np.
\end{aligned}
$$

Dengan cara serupa kita peroleh

$$E(X^2) = np + (np)^2 - np^2,$$

secara rinci ditinggalkan sebagai latihan. Oleh karena itu

$$Var(X) = np + (np)^2 - np^2 - (np)^2 = np - np^2 = np(1-p) = npq.$$

Kita dapat pula memperoleh mgf-nya, yaitu

$$M(t) = \sum_{x=0}^{n} e^{tx}\binom{n}{x} p^x q^{n-x} = \sum_{x=0}^{n}\binom{n}{x}(pe^t)^x q^{n-x} = (pe^t + q)^n.$$

Hasil ini sama dengan hasil pada sebaran Bernoulli, jika $n = 1$.

Pembahasan berikutnya subpokok bahasan ini adalah tentang sebaran hipergeometrik**.** Sebaran ini diperoleh dari pengambilan sebanyak n barang ***tanpa pengembalian*** dari N barang, di mana N barang tersebut terdiri dari M barang jenis satu dan $N-M$ barang jenis kedua. Misal X menyatakan banyak barang dari jenis M yang terambil, maka pdf-nya diberikan oleh

$$h(x;n,M,N) = \frac{\binom{M}{x}\binom{N-M}{n-x}}{\binom{N}{n}} I(x = 0,1,\ldots,n).$$

Peubah acak yang demikian disebut ber***sebaran hipergeometrik***, ditulis $X \sim HIP(n,M,N)$ dan cdf-nya ditulis sebagai $H(x;n,M,N)$. Sehingga

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

Kembali ke Contoh 3.2.2, jika diambil 5 kelereng secara acak tanpa pengembalian, maka kita peroleh pdf dari X adalah

$$h(x;5,10,30)=\frac{\binom{10}{x}\binom{20}{5-x}}{\binom{30}{5}}I(x=0,1,...,n).$$

Secara khusus jika $X=2,$ maka

$$h(2;5,10,30)=\frac{\binom{10}{2}\binom{20}{3}}{\binom{30}{5}}.$$

dapat ditunjukkan (sebagai latihan) bahwa

$$E(X)=\frac{nM}{N},$$

dan

$$Var(X)=\frac{n\frac{M}{N}\left(1-\frac{M}{N}\right)(N-n)}{N-1}.$$

Pembahasan berikutnya tentang hubungan sebaran binomial dengan disribusi hipergeometrik. Hubungan ini menyatakan bahwa untuk n kecil, relatif terhadap N dan $M,$ maka cukup beralasan bahwa jika pengambilan tanpa pengembalian kita pandang sebagai pengambilan dengan pengembalian. Untuk lebih jelasnya kita lihat contoh berikut ini.

Contoh 3.2.3 Suatu kantong berisi 1000 kelereng terdiri dari 400 kelereng merah dan 600 kelereng warna lain. Dari kantong ini diambil 10 kelereng secara acak tanpa pengembalian. Maka peluang kita mendapatkan 5 kelereng merah adalah

$$\binom{10}{5}\frac{400}{1000}\cdot\frac{399}{999}\cdots\frac{396}{996}\cdot\frac{600}{995}\cdot\frac{599}{994}\cdots\frac{596}{991}.$$

Nilai ini mendekati

$$\binom{10}{5}(0.4)^5(0.6)^5,$$

yang merupakan peluang dari mendapatkan 5 kelereng merah dari 10 kelereng yang terambil dengan pengembalian. Oleh karena itu jika X menyatakan banyak kelereng merah yang terambil, maka kita mendapatkan hubungan

$$\frac{\binom{400}{x}\binom{600}{10-x}}{\binom{1000}{10}}\approx\binom{10}{x}(0,4)^x(0,6)^{10-x}, x=0,1,...,10.$$

Teorema 3.2.1 Misal $X \sim HIP(n, M, N)$ Maka untuk masing-masing nilai $x=0,1,...,n,$ dan untuk $M\to\infty$ dengan $\frac{M}{N}\to p,$ suatu konstanta, maka

$$\lim_{N\to\infty}\frac{\binom{M}{x}\binom{N-M}{n-x}}{\binom{N}{n}}=\binom{n}{x}p^x(1-p)^{n-x}.$$

Bukti dari teorema tersebut ditinggalkan sebagai latihan.

Latihan 3.2

3.2.1 Misal X peubah acak bersebaran Bernoulli dengan parameter p.

(a) Jika f merupakan pdf dari X buktikan bahwa $\sum_{x=0}^{1} f(x)=1.$

(b) Turunkan rumus untuk ekspektasi dan varians dari X menggunakan mgf.

3.2.2. Tentukan mgf untuk peubah acak dari Contoh 3.2.1 dan Contoh 3.2.2.

3.2.3. Misal $X \sim BIN(n,p)$. Buktikan bahwa $E(X^2) = np + (np)^2 - np^2$.

3.2.4. Misal $X \sim BIN(n,p)$. Turunkan rumus untuk ekspektasi dan varians melalui mgf.

3.2.5. Misal 10 bola diambil secara acak dari suatu kantong yang berisi 30 bola merah dan 20 bola putih. Tentukan peluang memperoleh tepat 4 bola merah, jika pengambilan:
(a) dengan pengembalian,
(b) tanpa pengembalian.

3.2.6. Misal 5 kartu diambil secara acak tanpa pengembalian dari satu pak kartu *bridge*, yang berisi 52 kartu. Tentukan peluang mendapatkan;
(a) tepat dua kartu as,
(b) tepat tiga kartu raja,
(c) paling sedikit dua kartu raja.

3.2.7. Misal Y menyatakan banyak sukses suatu percobaan yang diulang sebanyak n kali dengan peluang sukses $p = 1/4$. Tentukan nilai terkecil n sehingga $P(1 \le Y) \ge 0,70$.

3.2.8. Misal mgf suatu peubah acak diskret X adalah

$$M(t) = \left(\frac{2}{3} + \frac{1}{3}e^t\right)^{9.}$$

Tunjukkan bahwa

$$P\left(\mu - 2\sigma < X < \mu + 2\sigma\right) = \sum_{x=1}^{5} \binom{9}{x} \left(\frac{1}{3}\right)^x \left(\frac{2}{3}\right)^{9-x}.$$

3.2.9. Misal $X \sim HIP(n, M, N)$. Buktikan bahwa

$$E(X) = \frac{nM}{N}$$

dan

$$Var(X) = \frac{n\frac{M}{N}\left(1-\frac{M}{N}\right)(N-n)}{N-1}.$$

3.2.10. Buktikan Teorema 3.2.1.

3.2.11. Misal Y menyatakan banyak sukses dalam n pengulangan percobaan secara bebas dengan peluang sukses $p = \frac{1}{4}$. Jika $n = 3,$ hitung $P(2 \leq Y);$ jika $n = 5,$ hitung $P(3 \leq Y).$

3.3 Sebaran Geometrik dan Binomial Negatif

Sebaran binomial yang telah kita bahas berpangkal pada percobaan Bernoulli, yang diulang secara bebas sebanyak n kali. Sekarang kita akan membahas percobaan Bernoulli yang diulang berkali-kali sehingga mendapatkan sukses pertama.

Contoh 3.3.1 Misal suatu kantong berisi 3 kelereng warna merah dan 7 kelereng warna lain. Dari kantong tersebut kita ambil kelereng satu persatu secara acak dengan pengembalian sehingga mendapatkan kelereng merah. Misal X menyatakan banyak percobaan sehingga mendapatkan kelereng merah. Dengan memisalkan kelereng warna merah sebagai m dan warna lainnya sebagai b, maka kita peroleh hasil sebagai berikut:

(1) $X = 1,$ ekivalen dengan kejadian $C_1 = \{m\},$

$$f(x) = P(X = 1) = P(C_1) = \left(\frac{7}{10}\right)^{1-1}\left(\frac{3}{10}\right).$$

(2) $X = 2,$ ekivalen dengan kejadian $C_2 = \{bm\},$

$$f(x) = P(X = 2) = P(C_2) = \frac{7}{10}.\frac{3}{10} = \left(\frac{7}{10}\right)^{2-1}\left(\frac{3}{10}\right).$$

(3) $X = 3,$ ekivalen dengan kejadian $C_3 = \{bbm\},$

$$f(x) = P(X = 3) = P(C_3) = \frac{7}{10}.\frac{7}{10}.\frac{3}{10} = \left(\frac{7}{10}\right)^{3-1}\left(\frac{3}{10}\right).$$

. . . .

(3) $X = 7,$ ekivalen dengan kejadian $C_7 = \{bb...bm\},$

$$f(x) = P(X = 10) = P(C_{10}) = \frac{7}{10}.\frac{7}{10}....\frac{7}{10}.\frac{3}{10} = \left(\frac{7}{10}\right)^{7-1}\left(\frac{3}{10}\right).$$

Oleh karena itu kita peroleh

$$f(x) = \left(\frac{7}{10}\right)^{x-1}\left(\frac{3}{10}\right)I(x = 1, 2, 3, ..., 7).$$

Secara umum jika X sebagai banyak percobaan sehingga mendapatkan sukses pertama, p menyatakan peluang kejadian sukses, dan q menyatakan peluang kejadian gagal, maka kita peroleh pdf dari X adalah

$$f(x) = q^{x-1}pI(x = 1, 2, 3, ...).$$

Peubah acak semacam ini disebut ber***sebaran geometrik,*** ditulis sebagai $X \sim GEO(p)$ dan pdf-nya ditulis sebagai $g(x; p).$ Sehingga

$$g(x; p) = q^{x-1}p, x = 1, 2, 3, ...$$

Menggunakan deret geometrik kita akan memperoleh

$$\sum_{x=1}^{\infty} g(x; p) = 1,$$

demikian juga untuk cdf, ditulis sebagai $G(x; p),$ dan

$$G(x; p) = 1 - q^x,$$

bukti ditinggalkan sebagai latihan.

Contoh 3.3.2 Peluang seorang pemain basket memasukkan bola ke dalam keranjang adalah 0.7. Karena dilanggar oleh pemain lawan maka pemain tersebut mendapatkan hadiah "3 bola". Jika masing-masing kesempatan untuk memasukkan bola kita anggap bebas, maka berapakah peluang bahwa pemain tadi pertama kali memasukkan bola pada kesempatan ketiga? Berapakah peluang paling sedikit pemain tersebut memasukkan satu bola pada ketiga kesempatan?

Penyelesaian

Permasalahan ini dapat kita selesaikan menggunakan sebaran geometrik. Kita misalkan kejadian sukses sebagai kejadian memasukkan bola ke dalam keranjang, dan X peubah acak yang menyatakan banyak percobaan untuk mendapatkan sukses pertama. Maka kita peroleh

$$P(X = 3) = g(3; 0.7) = (0.3)^2(0.7) = 0.063.$$

Lebih lanjut peluang paling sedikit pemain tersebut memasukkan satu bola pada ketiga kesempatan adalah

$$P(X \leq 3) = G(3; 0.7) = 1 - (0.3)^3 = 0.973.$$

Teorema 3.3.1 Sifat tanpa Memori. Misal $X \sim GEO(p)$. Maka $P(X > j + k \mid X > j) = P(X > k)$.

Bukti

$$P(X > j + k \mid X > j) = \frac{P(X > j + k)}{P(X > j)} = \frac{1 - P(X \leq j + k)}{1 - P(X \leq j)}$$

$$= \frac{1 - \left(1 - q^{j+k}\right)}{1 - \left(1 - q^{j}\right)} = \frac{q^{j+k}}{q^{j}} = q^{k} = P(X > k).$$

Terbukti.

Purata, varians, dan mgf dari $X \sim GEO(p)$ adalah:

$$E(X) = \sum_{x=1}^{\infty} x q^{x-1} p = p \sum_{x=1}^{\infty} \frac{d}{dq} q^{x} = p \frac{d}{dq} \sum_{x=0}^{\infty} q^{x}$$

$$= p \frac{d}{dq} \frac{1}{1-q} = \frac{p}{(1-q)^2} = \frac{p}{p^2} = \frac{1}{p}.$$

Dengan cara serupa kita peroleh

$$E(X^2) = \frac{1+q}{p^2},$$

bukti ditinggalkan sebagai latihan. Oleh karena itu variansnya adalah

$$Var(X) = E(X^2) - \left[E(X)\right]^2 = \frac{1+q}{p^2} - \frac{1}{p^2} = \frac{q}{p^2}.$$

Dan fungsi pembangkit momen dari sebaran ini adalah

$$M(t) = \frac{pe^{t}}{1 - qe^{t}},$$

bukti ditinggalkan sebagai latihan. Oleh karena itu dari Contoh 3.3.1 kita peroleh

$$E(X)=\frac{10}{3},$$

dan

$$Var(X)=\left(\frac{7}{10}\right).\left(\frac{10}{3}\right)^2=\frac{70}{3},$$

sedangkan dari Contoh 3.3.2 kita peroleh

$$E(X)=\frac{10}{7},$$

dan

$$Var(X)=\left(\frac{3}{10}\right).\left(\frac{10}{7}\right)^2=\frac{30}{7}.$$

Pembahasan selanjutnya adalah tentang sebaran binomial negatif. Sebaran ini berkaitan dengan suatu percobaan yang diulang beberapa kali secara bebas sehingga mendapatkan n pertama sukses. Secara khusus, untuk $n=1$ sebaran ini sama dengan sebaran geometrik. Untuk lebih jelasnya mengenai sebaran binomial negatif kita ikuti contoh berikut ini.

Contoh 3.3.3 Misal suatu kantong berisi 7 kelereng warna putih dan 3 kelereng warna lain. Dari kantong tersebut kita ambil kelereng satu persatu secara acak dengan pengembalian sehingga mendapatkan empat kelereng putih pertama. Misal X menyatakan banyak percobaan sehingga mendapatkan empat kelereng putih. Dengan memisalkan kelereng warna putih sebagai t dan warna bukan putih sebagai b, maka kita peroleh hasil-hasil sebagai berikut:

(1) $X=4,$ ekivalen dengan kejadian $C_1=\{tttt\},$ sehingga

$$f(4)=P(X=4)=\left(\frac{7}{10}\right)^4.$$

(2) $X=5,$ ekivalen dengan kejadian $C_2=\{btttt,tbttt,ttbtt,tttbt\},$ banyak anggota kejadian ini adalah $\binom{4}{3}=\binom{5-1}{4-1}$, sehingga

$$f(5) = P(X = 5) = \binom{5-1}{4-1}\left(\frac{3}{10}\right)^{5-4}\left(\frac{7}{10}\right)^{4}.$$

(3) $X = 6,$ ekivalen dengan kejadian $C_3 = \{bbtttt, btbttt, bttbtt,$ $btttbt, ..., ttthbt\}$, banyak anggota kejadian ini adalah $\binom{6-1}{4-1}$, sehingga

$$f(6) = P(X = 6) = \binom{6-1}{4-1}\left(\frac{3}{10}\right)^{6-4}\left(\frac{7}{10}\right)^{4}.$$

(4) $X = 7,$ ekivalen dengan kejadian $C_4 = \{bbbtttt, bbtbttt, bbttbtt,$ $bbtttbt, ..., tttbbbt\}$, banyak anggota kejadian ini adalah $\binom{7-1}{4-1}$, sehingga

$$f(7) = P(X = 7) = \binom{7-1}{4-1}\left(\frac{3}{10}\right)^{7-4}\left(\frac{7}{10}\right)^{4}.$$

Oleh karena itu kita peroleh

$$f(x) = \binom{x-1}{4-1}\left(\frac{3}{10}\right)^{x-4}\left(\frac{7}{10}\right)^{4} I(x = 4, 5, 6, 7).$$

Secara umum jika X sebagai banyak percobaan sehingga mendapatkan r sukses pertama, p menyatakan peluang kejadian sukses, dan q menyatakan peluang kejadian gagal, maka kita peroleh pdf dari X adalah

$$f(x) = \binom{x-1}{r-1} q^{x-r} p^{r} I(x = r, r+1, r+2, r+3, ...)$$

Peubah acak semacam ini disebut ***sebaran binomial negatif (negative binomial),*** ditulis sebagai $X \sim NB(r, p)$, dan pdf-nya ditulis sebagai $f(x; r, p)$. Oleh karena itu

$$f(x; r, p) = \binom{x-1}{r-1} q^{x-r} p^{r}, x = r, r+1, r+2, r+3, ...$$

Contoh 3.3.4 Dalam suatu turnamen bola voli pertandingan dinyatakan berakhir jika salah satu tim sudah memperoleh tiga kali kemenangan. Misal tim A sedang berhadapan dengan tim B. Berdasarkan data yang diperoleh dari pertandingan-pertandingan sebelumnya diperoleh bahwa $P(A \text{ menang}) = 0.6,$ pada tiap pertemuan dan anggap merupakan kejadian bebas. Berapakah peluang bahwa pertandingan berakhir dalam empat pertemuan.

Penyelesaian

Permasalahan di atas dapat kita pandang sebagai permasalahan sebaran binomial negatif dengan $x = 4$ dan $r = 3.$ Pertama kita misalkan tim A menang dalam pertandingan, maka

$$P(A \text{ menang dalam pertandingan}) = f(4;3,0.6) = \binom{3}{2}(0.4)^1(0.6)^3 = 0.2592.$$

Kedua kita misalkan tim B menang dalam pertandingan, maka

$$P(B \text{ menang dalam pertandingan}) = f(4;3,0.4) = \binom{3}{2}(0.6)^1(0.4)^3 = 0.1152.$$

Oleh karena itu peluang pertandingan berakhir empat pertemuan sebesar $0.2592 + 0.1152 = 0.3744.$

Selanjutnya ditinggalkan sebagai latihan untuk menunjukkan bahwa

$$E(X) = \frac{r}{p},$$

$$Var(X) = \frac{rq}{p^2},$$

dan

$$M(t) = \left[\frac{pe^t}{1-q^{et}}\right]^r.$$

Latihan 3.3

3.3.1. Misal $X \sim GEO(p)$. Buktikan bahwa

(a) $\sum_{x=1}^{\infty} g(x;p) = 1.$

(b) $G(x;p) = 1 - q^x.$

3.3.2. Misal $X \sim GEO(p)$. Tunjukkan bahwa $E(X^2) = \dfrac{1+q}{p^2}$.

3.3.3. Misal 5 kartu diambil secara acak dengan pengembalian dari satu pak kartu *bridge*, yang berisi 52 kartu. Tentukan peluang mendapatkan:
(a) kartu as pertama pada pengambilan ketiga,
(b) kartu as ketiga pada pengambilan kelima.

3.3.4. Tiga orang melempar uang logam untuk menentukan orang yang membayar kopi pada suatu restoran, dengan ketentuan: jika ketiga-tiganya menghasilkan angka/gambar sama, maka mereka melemparnya lagi. Jika lain maka orang yang uang logamnya berbeda dengan dua orang yang lain membayar kopi. Berapakah peluang bahwa:
(a) berakhir pada pelemparan kedua,
(b) mereka memerlukan lebih dari satu pelemparan.

3.3.5. Misal $X \sim GEO(p)$. Tunjukkan bahwa

$$M(t) = \frac{pe^t}{1-qe^t}.$$

3.3.6. Misal $X \sim NB(r,p)$. Tunjukkan bahwa

$$E(X) = \frac{r}{p},\ Var(X) = \frac{rq}{p^2},$$

dan

$$M(t) = \left[\frac{pe^t}{1-qe^t}\right]^r.$$

3.4 Sebaran Poisson

Sebelum sampai pada sebaran Poisson, marilah kita ikuti uraian berikut ini. Kita perhatikan ekspansi deret Taylor dari

$$g(u) = e^u,$$

di sekitar nol, yaitu

$$e^u = 1 + \frac{u^1}{1!} + \frac{u^2}{2!} + \frac{u^3}{3!} + ... = \sum_{x=0}^{\infty} \frac{u^x}{x!}.$$

Jika kedua persamaan kita kalikan dengan e^{-u}, maka kita peroleh

$$\sum_{x=0}^{\infty} \frac{e^{-u}u^x}{x!} = 1.$$

Oleh karena itu jika kita pilih

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

maka:

(i) $f(x) > 0, x = 0,1,2,...,$

(ii) $\sum_{x=0}^{\infty} f(x) = 1.$

Jadi f mendefinisikan pdf. Peubah acak X yang mempunyai pdf seperti ini disebut ber***sebaran Poisson,*** ditulis sebagai $X \sim POI(\mu)$.

Untuk menentukan ekspektasi dan varians peubah acak X yang bersebaran Poisson terlebih dahulu kita tentukan mgf-nya seperti berikut.

$$M(t) = E(e^{tX}) = \sum_{x=0}^{\infty} e^{tx} \frac{e^{-u}u^x}{x!} = e^{-u} \sum_{x=0}^{\infty} \frac{(ue^t)^x}{x!} = e^{-u}e^{ue^t} = e^{u(e^t-1)},$$

untuk semua nilai t. Selanjutnya kita tentukan turunan pertama dan kedua

$$M'(t) = e^{u(e^t-1)}(ue^t),$$

dan

$$M''(t) = e^{u(e^t-1)}(ue^t) + e^{u(e^t-1)}(ue^t)^2.$$

Oleh karena itu

$$\mu = M'(0) = u,$$

dan

$$\sigma^2 = M''(0) - \left[M'(0)\right]^2 = u + u^2 - u^2 = u = \mu.$$

Hasil ini menyatakan bahwa jika X bersebaran Poisson, maka $\mu = \sigma^2 = u > 0$. Oleh karena itu kita dapat pula menulis sebaran ini sebagai $X \sim POI(\mu)$, pdf-nya

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

dan cdf-nya

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

Untuk nilai μ dan x tertentu cdf ini dapat diperoleh dari tabel sebaran Poisson, atau dari komputer dengan perangkat lunak statistik. Dalam hal ini kita pakai Minitab.

Contoh 3.4.1 Misal $X \sim POI(2)$. Maka

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

Selanjutnya jika kita ingin menentukan $P(1 \le X)$, maka kita dapat memperolehnya seperti berikut ini

$$\begin{aligned} P(1 \le X) &= 1-P(X \le 0) \\ &= 1-P(X=0) \\ &= 1-e^{-2} \\ &= 0.865, \end{aligned}$$

hasil di atas dapat diperoleh langsung dari kalkulator, dari tabel sebaran Poisson, atau dari Minitab.

Pembahasan selanjutnya dalam subpokok bahasan ini adalah mengenai hubungan antara sebaran binomial dengan sebaran Poisson, seperti terlihat pada teorema di bawah ini.

Teorema 3.4.1 Misal $X \sim BIN(n,p)$. Untuk masing-masing nilai $x=0,1,2,...,$ dan jika $p \to 0,$ dengan $np=\mu,$ maka

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

Bukti

Kita uraikan ruas kiri terlebih dahulu, yaitu dengan menguraikan bentuk faktorialnya, dan $np=\mu$ kita ubah menjadi bentuk $p=\dfrac{\mu}{n}$. Oleh karena itu kita peroleh seperti berikut ini

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

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

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

Karena

$$\lim_{n\to\infty}\frac{\mu^x}{x!} = \frac{\mu^x}{x!},$$

dan

$$\lim_{n\infty}\left(\frac{n}{n}\right)\left(\frac{n-1}{n}\right)...\left(\frac{n-x+1}{n}\right)\left(1-\frac{\mu}{n}\right)^{-x} = 1,$$

maka menggunakan sifat limit pada kalkulus, yaitu sifat

$$\lim_{n\to\infty}\left(1-\frac{\mu}{n}\right)^n = e^{-\mu},$$

kita peroleh

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

Contoh 3.4.2 Misal 1% dari semua komponen yang diproduksi oleh pabrik tertentu adalah rusak. Suatu komputer model baru memerlukan 100 komponen dari pabrik tersebut. Maka peluang eksak untuk memperoleh 3 komponen rusak adalah $b(3;100,0.01) = 0.0610,$ sedangkan menggunakan pendekatan Poisson kita peroleh peluang $f(3;1) = 0.0613.$

Pembahasan terakhir dari subpokok bahasan ini adalah tentang ***proses Poisson***, yaitu suatu proses untuk mendapatkan sebaran Poisson. Pada prakteknya proses Poisson ini dapat pula digunakan untuk mendeteksi apakah suatu peubah acak mengikuti sebaran Poisson atau tidak.

Kita perhatikan situasi fisik dalam kejadian tertentu, seperti dering telepon atau kerusakan kabel pada potongan tertentu. Misal $X(t)$

menyatakan banyak kejadian yang terjadi pada selang $[0,t]$, dan anggap bahwa asumsi-asumsi berikut ini benar.

(1) Peluang bahwa kejadian akan terjadi dalam selang yang pendek $[t, t+\Delta t]$ adalah mendekati perbandingan dengan panjang selang, Δt, dan tidak bergantung pada posisi selang.

(2) Kejadian pada selang yang tidak berpotongan adalah saling bebas.

(3) Peluang dua atau lebih kejadian dalam selang pendek $[t, t+\Delta t]$ diabaikan.

Jika asumsi-asumsi di atas dipenuhi maka untuk $\Delta t \to 0,\ \ X(t)$ akan bersebaran Poisson. Asumsi-asumsi di atas jika dinyatakan secara matematika seperti terlihat pada teorema berikut ini. Sebelumnya kita catat terlebih dahulu bahwa $o(\Delta t)$ menyatakan sebagai fungsi dari Δt sehingga

$$\lim_{\Delta t \to 0} \frac{o(\Delta t)}{\Delta t} = 0.$$

Sebagai contoh $o(\Delta t) = (\Delta t)^2$, maka

$$\lim_{\Delta t \to 0} \frac{o(\Delta t)}{\Delta t} = \lim_{\Delta t \to 0} \frac{(\Delta t)^2}{\Delta t} = \lim_{\Delta t \to 0} \Delta t = 0.$$

Teorema 3.4.2 Misal $X(t)$ menyatakan banyak kejadian yang terjadi pada selang $[0,t]$, dan $P_n(t) = P\left[n \text{ kejadian dalam selang } [0,\mathrm{t}]\right]$. Jika $X(t)$ memenuhi sifat-sifat berikut:

(1) $X(0) = 0,$

(2) $P\left[X(t+h) - X(t) = n \mid X(s) = m\right] = P\left[X(t+h) - X(t) = n\right],$ untuk semua $0 \le s \le t$ dan $0 < h,$

(3) $P\left[(t+\Delta t) - X(t) = 1\right] = \lambda \Delta t + o(\Delta t)$ untuk suatu konstanta $\lambda,$ dan

(4) $P\left[(t+\Delta t) - X(t) \ge 2\right] = o(\Delta t),$

maka untuk $t > 0$

$$P_n(t) = \frac{(\lambda t)^n e^{-\lambda t}}{n!} I(n = 0, 1, 2, \ldots).$$

Bukti

Terdapat dua hasil yang mungkin untuk n kejadian dalam selang $\left[0, t+\Delta t\right]$, yaitu mempunyai 1 kejadian dalam selang $\left[t, t+\Delta t\right]$ dan $n-1$ kejadian dalam selang $\left[0, t\right]$, untuk Δt kecil, atau 0 kejadian dalam $\left[t, t+\Delta t\right]$ dan n kejadian dalam selang $\left[0, t\right]$. Oleh karena itu

$$\begin{aligned}
P_n(t+\Delta t) &= P_{n-1}(t)P_1(\Delta t) + P_n(t)P_0(\Delta t) + o(\Delta t) \\
&= P_{n-1}(t)\left[\lambda\Delta t + o(\Delta t)\right] + P_n(t)\left[1-\lambda\Delta t - o(\Delta t)\right] + o(\Delta t) \\
&= P_{n-1}(t)\lambda\Delta t + P_{n-1}(t)o(\Delta t) + P_n(t) - P_n(t)\lambda\Delta t + P_n(t)o(\Delta t) + o(\Delta t).
\end{aligned}$$

Di pihak lain

$$\frac{dP_n(t)}{dt} = \lim_{\Delta t\to 0}\frac{P_n(t+\Delta t) - P_n(t)}{\Delta t}.$$

Oleh karena itu

$$\begin{aligned}
\frac{dP_n(t)}{dt} &= \lim_{\Delta t\to 0}\frac{P_{n-1}(t)\lambda\Delta t + P_{n-1}(t)o(\Delta t) + P_n(t) - P_n(t)\lambda\Delta t + P_n(t)o(\Delta t) + o(\Delta t) - P_n(t)}{\Delta t} \\
&= \lim_{\Delta t\to 0}\frac{P_{n-1}(t)\lambda\Delta t - P_n(t)\lambda\Delta t}{\Delta t} = \lambda\left[P_{n-1}(t) - P_n(t)\right].
\end{aligned}$$

Sekarang kita lihat untuk $n = 0$,

$$\begin{aligned}
P_0(t+\Delta t) &= P_0(t)P_0(\Delta t) = P_0(t)\left[1-\lambda\Delta t - o(\Delta t)\right] \\
&= P_0(t) - P_0(t)\lambda\Delta t - P_0(t)o(\Delta t),
\end{aligned}$$

sehingga

$$\begin{aligned}
\frac{dP_0(t)}{dt} &= \lim_{\Delta t\to 0}\frac{P_0(t+\Delta t) - P_0(t)}{\Delta t} = \lim_{\Delta t\to 0}\frac{P_0(t) - P_0(t)\lambda\Delta t - P_0(t)o(\Delta t) - P_0(t)}{\Delta t} \\
&= \lim_{\Delta t\to 0}\frac{-P_0(t)\lambda\Delta t - P_0(t)o(\Delta t)}{\Delta t} = -\lambda P_0(t).
\end{aligned}$$

Dalam bentuk persamaan diferensial dapat ditulis sebagai

$$\frac{1}{P_0(t)}\,dP_0(t) = -\lambda dt,$$

Penyelesaian persamaan deferensial di atas adalah

$$\ln P_0(t) = -\lambda t + C.$$

Menggunakan kenyataan bahwa

$$P_0(0) = 1,$$

yang berakibat

$$C = \ln P_0(0) = \ln 1 = 0,$$

maka kita peroleh

$$P_0(t) = e^{-\lambda t}.$$

Selanjutnya kita lihat untuk $n = 1$,

$$\frac{dP_1(t)}{dt} = \lambda\left[P_0(t) - P_1(t)\right] = \lambda\left[e^{-\lambda t} - P_1(t)\right] = \lambda e^{-\lambda t} - \lambda P_1(t)$$

yang memberikan

$$P_1(t) = \lambda t e^{-\lambda t}.$$

Menggunakan induksi matematis dapat ditunjukkan (ditinggalkan sebagai latihan) bahwa

$$P_n(t) = \frac{(\lambda t)^n e^{-\lambda t}}{n!} I(n = 0, 1, 2, ...).$$

Oleh karena itu $X(t) \sim POI(\lambda t)$, di mana $\mu = E\left[X(t)\right] = \lambda t.$

Latihan 3.4

3.4.1 Misal $X(t) \sim POI(\lambda t)$. Buktikan menggunakan induksi matematis bahwa

$$P_n(t) = \frac{(\lambda t)^n e^{-\lambda t}}{n!} I(n = 0, 1, 2, ...).$$

3.4.2. Misal X peubah acak yang bersebaran Poisson, $P(X = 1) = P(X = 2).$ Tentukan $P(X = 4).$

3.4.3. Misal X peubah acak yang bersebaran Poisson, $P(X = 0) = 0.2.$ Tentukan $P(X > 4).$

3.4.4. Misal mgf dari peubah acak X adalah

$$M(t) = e^{4(e^t - 1)}.$$

Tunjukkan bahwa

$$P\left[\mu - 2\sigma < X < \mu + 2\sigma\right] = 0,931.$$

3.4.5. Misal banyak antrian panggilan telepon pada operator tertentu dalam satu jam bersebaran Poisson dengan $\mu = 10.$ Tentukan peluang dari masing-masing kejadian berikut ini:
(a) tujuh antrian panggilan,
(b) paling banyak tujuh antrian panggilan,
(c) mulai dua sampai tujuh panggilan.

3.4.6. Misal pdf $f(x)$ positif hanya untuk bilangan bulat tidak negatif, sehingga

$$f(x) = \frac{4}{x} f(x-1) I(x = 1, 2, 3, ...).$$

Tentukan $f(x).$ *Petunjuk:* Perhatikan bahwa

$f(1) = 4f(0), f(2) = \frac{4^2}{2!} f(0), ...,$ kemudian nyatakan $f(x)$ sebagai suku-suku dari $f(0)$ dan tentukan $f(0)$ dari

$$1 = f(0) + f(1) + f(2) + ...$$

3.4.7. Misal $X \sim POI(100)$. Gunakan ketaksamaan Chebyshev untuk menentukan batas bawah dari $P(75 < X < 125).$

3.4.8. Misal X mempunyai sebaran Poisson. Jika $P(X = 1) = P(X = 3),$ tentukan:
(a) $P(X = 5),$
(b) modus dari sebaran ini.

3.4.9. Suatu pabrik elektronik memproduksi komponen tertentu, dan kejadian komponen rusak adalah bebas dengan peluang $0.1.$ Jika pabrik tersebut memproduksi 500 komponen, maka:
(a) berapakah peluang kejadian komponen rusak paling banyak dua?
(b) hitung (a) menggunakan pendekatan Poisson.

3.5 Sebaran Seragam

Terdapat dua macam sebaran seragam, yaitu sebaran seragam peubah acak diskret dan sebaran seragam peubah acak kontinu. Kasus pertama disebut sebagai sebaran seragam diskret, dan yang kedua disebut sebaran seragam kontinu. Pembahasan kita awali dengan ***sebaran seragam diskret***.

Banyak permasalahan praktis yang melibatkan peluang klasik, yaitu suatu peluang dari suatu kejadian yang masing-masing anggotanya mempunyai kesempatan sama (seragam). Untuk mempermudah permasalahan, masing-masing anggota sampel dikaitkan oleh suatu peubah acak X dengan bilangan asli $1, 2, ..., N$. Oleh karena itu kita memperoleh pdf dari peubah acak diskret X, yaitu

$$f(x) = \frac{1}{N} I(x = 1, 2, ..., N),$$

diberi simbol sebagai $X \sim DU(N)$.

Purata dan varians dari sebaran ini dapat kita peroleh seperti berikut ini:

$$E(X) = \sum_x x f(x) = \sum_{x=1}^{N} \frac{x}{N} = \frac{1}{N} \sum_{x=1}^{N} x = \frac{1}{N} . \frac{N(N+1)}{2} = \frac{N+1}{2},$$

dan

$$E(X^2) = \sum_x x^2 f(x) = \sum_{x=1}^{N} \frac{x^2}{N} = \frac{1}{N} \sum_{x=1}^{N} x^2$$

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

sehingga

$$Var(X) = E(X^2) - \left[E(X)\right]^2 = \frac{N^2 - 1}{12},$$

sedangkan mgf-nya ditinggalkan sebagai latihan.

Contoh 3.5.1 Misal X menyatakan bilangan sisi dadu yang muncul dari suatu percobaan melempar dadu sekali. Maka $X \sim DU(6)$. Oleh karena itu kita peroleh pdf

$$f(x) = \frac{1}{6} I(x = 1, 2, ..., 6),$$

purata

$$E(X) = \frac{6+1}{2} = 3.5,$$

dan varians

$$Var(X) = \frac{36-1}{12} = 2.92.$$

Pembahasan selanjutnya mengenai sebaran seragam dari peubah acak kontinu, atau sebaran seragam kontinu. Untuk itu anggap X nilai yang terdefinisi pada selang terbuka (a,b), dan anggap bahwa kesempatan nilai X jatuh pada selang tersebut adalah sama (seragam). Maka kita dapat mengatakan bahwa pdf dari X adalah konstan, yaitu

$$f(x) = cI(a < x < b).$$

Selanjutnya menggunakan sifat kedua dari Definisi 2.3.2 kita peroleh

$$\int_a^b c dx = 1.$$

Solusi dari integral ini adalah

$$c = \frac{1}{b-a}.$$

Peubah acak yang mempunyai pdf demikian disebut ber***sebaran seragam kontinu***, ditulis sebagai

$$X \sim UNIF(a,b),$$

dan pdf-nya ditulis sebagai

$$f(x;a,b) = \frac{1}{b-a} I(a < x < b).$$

Seperti halnya pada sebaran seragam diskret, purata dan varians dari sebaran seragam kontinu ini kita peroleh seperti berikut:

$$E(X) = \int xf(x)dx = \int_a^b x.\frac{1}{b-a}dx = \frac{1}{b-a}\int_a^b xdx = \frac{1}{b-a}\left[\frac{1}{2}x^2\right]_a^b$$

$$= \frac{1}{2}.\frac{1}{b-a}\left(b^2 - a^2\right) = \frac{1}{2}.\frac{1}{b-a}\left(b-a\right)\left(b+a\right) = \frac{b+a}{2},$$

dan

$$E(X^2) = \int x^2 f(x)dx = \int_a^b x^2.\frac{1}{b-a}dx = \frac{1}{b-a}\int_a^b x^2 dx = \frac{1}{b-a}\left[\frac{1}{3}x^3\right]_a^b$$

$$= \frac{1}{3}.\frac{1}{b-a}\left(b^3 - a^3\right) = \frac{1}{3}.\frac{1}{b-a}\left(b-a\right)\left(b^2 + ba + a^2\right) = \frac{(b^2 + ba + a^2)}{3},$$

sehingga

$$Var(X) = \frac{b^2+ba+a^2}{3} - \frac{(b+a)^2}{4} = \frac{4b^2+4ba+4a^2}{12} - \frac{3b^2+6ba+3a^2}{12}$$
$$= \frac{b^2-2ba+a^2}{12} = \frac{(b-a)^2}{12},$$
sedangkan mgf-nya ditinggalkan sebagai latihan.

Contoh 3.5.2 Misal satu titik dipilih dari selang $(0,1)$. Maka $X \sim UNIF(0,1)$, sehingga pdf-nya adalah
$$f(x;0,1) = I(0 < x < 1),$$
ekspektasinya adalah
$$E(X) = \frac{1}{2},$$
dan variansnya adalah
$$Var(X) = \frac{1}{12}.$$

Latihan 3.5

3.5.1. Misal $X \sim DU(N)$. Tentukanlah mgf dari X. Petunjuk: Gunakan identitas $s + s^2 + \ldots + s^N = \dfrac{s(1-s^N)}{1-s}, s \neq 1.$

3.5.2. Misal X menyatakan bilangan sisi dadu yang muncul dari suatu percobaan melempar dadu sekali. Maka $X \sim DU(6)$, dan pdf-nya adalah
$$f(x) = \frac{1}{6} I(x = 1, 2, \ldots, 6).$$
(a) Tentukanlah cdf dari X, dan buatlah grafiknya.
(b) Jika dadu dilempar dua kali, maka tentukan pdf dari Y, di mana Y menyatakan nilai terkecil dari dua kali pelemparan dadu.
(c) Jika dadu dilempar tiga kali, maka hitung peluang bahwa nilai terbesarnya kurang dari tiga.

3.5.3. Misal satu titik dipilih dari selang $(0,1)$. Maka $X \sim UNIF(0,1)$, dan pdf-nya adalah

$$f(x;0,1) = I(0 < x < 1).$$

(a) Tentukanlah cdf dari X, dan buatlah grafiknya.

(b) Tentukanlah peluang bahwa nilai dari X kurang dari 0.5.

(c) Tentukan $P(0.1 < X < 0.9)$.

3.5.4. Misal $X \sim UNIF(a, b)$. Tentukanlah mgf dari X.

3.5.5. Misal $X \sim UNIF(0,2)$. Tentukanlah peluang akar-akar persamaan $g(t) = 0$ adalah real, jika

$$g(t) = 4t^2 + 4Xt + X + 2.$$

3.6 Sebaran Gamma, Eksponensial, Weibull dan Pareto

Salah satu sebaran kontinu yang sering muncul dalam terapan adalah sebaran gamma. Sebaran ini diturunkan dari fungsi gamma.

Definisi 3.6.1 Suatu ***fungsi gamma,*** dinyatakan sebagai $\Gamma(\kappa)$ untuk semua $\kappa > 0$, adalah

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

Sebagai contoh, jika $\kappa = 1$, maka $\Gamma(1) = \int_0^\infty e^{-t} dt = 1$. Beberapa sifat dari fungsi gamma diberikan oleh teorema berikut ini.

Teorema 3.6.1 Suatu fungsi gamma memenuhi sifat-sifat berikut ini:

(1) $\Gamma(\kappa) = (\kappa - 1)\Gamma(\kappa - 1), \kappa > 1.$

(2) $\Gamma(n) = (n-1)!$, n bilangan bulat positif.

(3) $\Gamma\left(1/2\right) = \sqrt{\pi}.$

Bukti

Kita buktikan sifat (1) menggunakan integral parsial seperti berikut ini.

$$\Gamma(\kappa) = \int_0^\infty t^{\kappa-1} e^{-t} dt = -\int_0^\infty t^{\kappa-1} de^{-t} = -\left[t^{\kappa-1} e^{-t} \right]_0^\infty + \int_0^\infty e^{-t} dt^{\kappa-1}$$

$$= 0 + (\kappa - 1)\int_0^\infty t^{(\kappa-1)-1} e^{-t} dt = (\kappa - 1)\Gamma(\kappa - 1), \kappa > 1.$$

Sifat (2) dapat dibuktikan menggunakan induksi matematis (ditinggalkan sebagai latihan). Sifat (3) kita buktikan menggunakan integral lipat dengan mengubahnya terlebih dahulu menjadi koordinat kutub. Kita substitusi $x = \sqrt{t}$ ke dalam fungsi gamma kita peroleh $t = x^2$, sehingga $dt = 2xdx.$ Oleh karena itu

$$\Gamma(1/2) = \int_0^\infty t^{1/2-1} e^{-t} dt = \int_0^\infty t^{-1/2} e^{-t} dt = \int_0^\infty x^{-1} e^{-x^2} 2xdx = 2\int_0^\infty e^{-x^2} dx.$$

Selanjutnya kita tulis

$$\left[\Gamma(1/2)\right]^2 = \left[2\int_0^\infty e^{-x^2} dx\right]^2 = \left[2\int_0^\infty e^{-x^2} dx\right]\left[2\int_0^\infty e^{-x^2} dx\right]$$

$$= \left[2\int_0^\infty e^{-x^2} dx\right]\left[2\int_0^\infty e^{-y^2} dy\right] = 4\int_0^\infty \int_0^\infty e^{-(x^2+y^2)} dxdy.$$

Jika ubah integral tersebut menjadi bentuk integral kutub (polar), yaitu dengan mengambil

$$x = f(r,\theta) = r\cos\theta,$$

$$y = g(r,\theta) = r\sin\theta.$$

Maka menggunakan teknik transformasi, lihat subpokok bahasan 4.6,

$$\iint_A f_{X_1,X_2}(x_1,x_2) dx_1 dx_2 = \iint_B f_{X_1,X_2}\left[w_1(y_1,y_2), w_2(y_1,y_2)\right] |J|\, dy_1 dy_2,$$

di mana

$$J = \frac{\partial(x,y)}{\partial(r,\theta)} = \begin{vmatrix} \dfrac{\partial x}{\partial r} & \dfrac{\partial x}{\partial \theta} \\ \dfrac{\partial y}{\partial r} & \dfrac{\partial y}{\partial \theta} \end{vmatrix} = \begin{vmatrix} \cos\theta & -r\sin\theta \\ \sin\theta & r\cos\theta \end{vmatrix} = r(\cos^2\theta + \sin^2\theta) = r,$$

kita peroleh

$$\left[\Gamma(1/2)\right]^2 = 4\int_0^\infty \int_0^\infty e^{-(x^2+y^2)} dxdy = 4\int_0^{\frac{1}{2}\pi} \int_0^\infty e^{-r^2} r dr d\theta = 2\int_0^{\frac{1}{2}\pi} \int_0^\infty e^{-r^2} dr^2 d\theta$$

$$= 2\int_0^{\frac{1}{2}\pi} \left[-e^{-r^2}\right]_0^\infty d\theta = 2\int_0^{\frac{1}{2}\pi} d\theta = 2\left[\theta\right]_0^{\frac{1}{2}\pi} = \pi.$$

Karena $\left[\Gamma(1/2)\right]$ tidak negatif, maka $\left[\Gamma(1/2)\right] = \sqrt{\pi}.$

Kita kembali ke definisi fungsi gamma

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

Jika kita substitusi $t = x/\theta$ dan $\theta > 0$, maka fungsi gamma menjadi

$$\Gamma(\kappa)=\int_0^\infty \left(\frac{x}{\theta}\right)^{\kappa-1} e^{-\frac{x}{\theta}}\left(\frac{1}{\theta}\right)dx=\frac{1}{\theta^\kappa}\int_0^\infty x^{\kappa-1}e^{-\frac{x}{\theta}}dx.$$

Oleh karena itu

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

Selanjutnya jika kita pilih

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

maka:

(i) $f(x)>0, x>0$

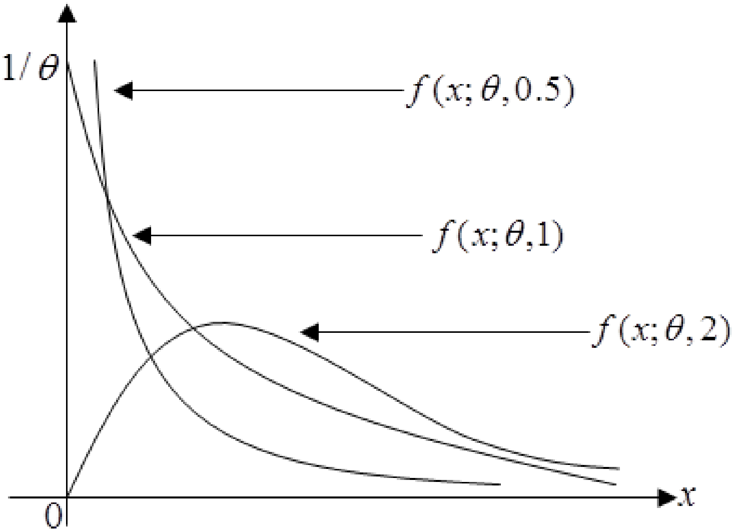
(ii) $\int_0^\infty f(x)dx=1.$

Fungsi di atas mendefinisikan suatu pdf, dan peubah acak X yang mempunyai pdf di atas disebut ber***sebaran gamma***. Secara lengkap pdf-nya ditulis sebagai

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

Peubah acak yang bersebaran demikian ditulis $X \sim GAM(\theta,\kappa)$.

Parameter κ disebut ***parameter bentuk (shape parameter*)** karena menentukan bentuk dasar dari grafik pdf. Secara khusus ada tiga bentuk dasar yang bergantung pada parameter ini, yaitu $\kappa<1, \kappa=1,$ dan $\kappa>1,$ seperti terlihat pada Gambar 3.6.1 berikut.



Gambar 3.6.1

Fungsi sebaran dari $X \sim GAM(\theta, \kappa)$ adalah

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

Substitusi $u = t/\theta$ dalam integral di atas menghasilkan

$$F(x;\theta,\kappa) = \int_0^{x/\theta} \frac{1}{\theta^\kappa \Gamma(\kappa)} (u\theta)^{\kappa-1} e^{-u} \theta du = \int_0^{x/\theta} \frac{1}{\Gamma(\kappa)} u^{\kappa-1} e^{-u} du = F(x/\theta;1,\kappa),$$

yang bergantung pada θ yang terletak pada peubah x/θ. Sehingga parameter ini biasa disebut sebagai parameter skala.

Contoh 3.6.1 Misal peubah acak X menyatakan waktu hidup suatu baterai dalam bulan dan anggap $X \sim GAM(\theta, \kappa)$, dengan $\theta = 12$. Maka peluang bahwa waktu hidup baterai kurang dari 24 bulan dapat ditulis sebagai

$$P(X \le 24) = F(24;12,\kappa) = F(2;1,\kappa) = P(Y \le 2),$$

di mana Y menyatakan waktu hidup baterai dalam tahun dan $X \sim GAM(1, \kappa)$.

Purata dari $X \sim GAM(\theta, \kappa)$ dapat kita peroleh seperti berikut ini:

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

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

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

Dengan cara yang sama (ditinggalkan sebagai latihan) kita peroleh

$$E(X^2)=\theta^2\kappa(1+\kappa).$$

Oleh karena itu

$$Var(X)=\theta^2\kappa(1+\kappa)-(\kappa\theta)^2=\kappa\theta^2.$$

Untuk mgf kita peroleh seperti berikut

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

Kita substitusi $u=-(t-1/\theta)x,$ kita peroleh

$$M(t)=\left(\frac{1}{\theta}-t\right)^{-\kappa}\frac{1}{\theta^\kappa\Gamma(\kappa)}\int_0^\infty u^{\kappa-1}e^{-u}du=\left(1-\theta t\right)^{-\kappa},t<1/\theta.$$

Dalam hal kita mempunyai parameter κ berupa bilangan bulat positif, yaitu $\kappa=n$ maka untuk parameter θ dan untuk nilai x tertentu cdf-nya dapat kita tentukan menggunakan rumus pada teorema berikut ini.

Teorema 3.6.2 Misal $X\sim GAM(\theta,n)$ di mana n bilangan bulat positif. Maka cdf dari X dapat ditulis sebagai

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

Teorema di atas dibuktikan menggunakan induksi matematis dan integral parsial secara berulang (ditinggalkan sebagai latihan). Suku dalam penjumlahan Teorema 3.6.2, serupa dengan penjumlahan pada sebaran Poisson dengan μ diganti oleh $x/\theta.$

Contoh 3.6.2 Misal kuantitas harian (dalam inchi) ukuran gelombang sungai adalah peubah acak $X\sim GAM(0.2,6)$. Misal kita ingin mengetahui peluang kuantitas gelombang lebih dari puratanya, yaitu 1.2 inchi. Maka kita peroleh

$$P\left[X > 1.2\right] = 1 - F(1.2;\ 0.2,\ 6) = 1 - \left(1 - \sum_{i=0}^{5} \frac{6^i}{i!} e^{-6}\right) = \sum_{i=0}^{5} \frac{6^i}{i!} e^{-6} = 0.4457,$$

dari Minitab atau tabel sebaran Poisson dengan $x = 5$ dan $\mu = 6$. Apa komentar Anda tentang hasil ini?

Salah satu sebaran khusus dari sebaran gamma adalah ***sebaran eksponensial***. Sebaran eksponensial diperoleh pada saat $\kappa = 1$. Oleh karena itu jika X bersebaran eksponensial, ditulis $X \sim EXP(\theta)$, maka pdf-nya adalah

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

dan cdf-nya adalah

$$F(x;\theta) = 1 - e^{-\frac{x}{\theta}} I(x > 0),$$

bukti dijadikan sebagai latihan. Untuk parameter θ dan nilai x tertentu nilai pdf dan cdf ini dapat diperoleh dari Minitab atau tabel sebaran eksponensial.

Contoh 3.6.3 Misal X menyatakan waktu hidup (dalam jam) bola lampu *merk* tertentu dan mengikuti sebaran eksponensial dengan parameter $\theta = 72$, $X \sim EXP(72)$ Misal kita ingin mengetahui peluang lampu hidup setelah 48 jam. Maka peluang tersebut dapat kita hitung seperti berikut ini

$$P\left[X > 48\right] = 1 - P\left[X \le 48\right] = 1 - F(48;72) = 1 - 0.4596 = 0.5404,$$

diperoleh dari Minitab.

Sebaran eksponensial sering digunakan untuk menentukan peluang waktu hidup suatu komponen, yang mempunyai sifat seperti diberikan oleh teorema berikut ini.

Teorema 3.6.3 Sifat tanpa Memori. Misal $X \sim EXP(\theta)$. Maka

$$P\left[X > a + t \mid X > a\right] = P\left[X > t\right]$$

untuk semua $a > 0$ dan $t > 0$.

Bukti

$$P\left[X > a+t \mid X > a\right] = \frac{P\left[X > a+t \text{ dan } X > a\right]}{P\left[X > a\right]} = \frac{P\left[X > a+t\ \right]}{P\left[X > a\right]}$$

$$= \frac{e^{-(a+t)/\theta}}{e^{-a/\theta}} = e^{-t/\theta} = P\left[X > t\right].$$

Bentuk khusus lain dari sebaran gamma adalah ***sebaran khi kuadrat***, yaitu jika $\theta = 2$ dan $\kappa = \nu/2$, dan ν disebut sebagai derajat bebas.

Pembahasan selanjutnya tentang sebaran Weibull. Sebaran ini ditemukan oleh ahli Fisika W. Weibull.

Suatu peubah acak kontinu X dikatakan ber***sebaran Weibull*** dengan parameter $\beta > 0$ dan $\theta > 0$ jika mempunyai pdf

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

dinyatakan $X \sim WEI(\theta,\beta)$. Dalam hal $\beta = 2$ disebut ber***sebaran Rayleigh***. Jika $X \sim WEI(\theta,\beta)$, maka cdf-nya adalah

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

bukti ditinggalkan sebagai latihan. Purata dari $X \sim WEI(\theta,\beta)$ diperoleh seperti berikut:

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

Substitusi $t = \left(\dfrac{x}{\theta}\right)^{\beta}$, diperoleh

$$E(X) = \theta \int_0^{\infty} t^{(1+1/\beta)-1} e^{-t}\ dt = \theta\Gamma\left(1 + \frac{1}{\beta}\right).$$

Serupa untuk $E(X^2) = \theta^2 \Gamma\left(1 + \dfrac{2}{\beta}\right)$, ditinggalkan sebagai Latihan. Jadi

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

Sebaran Weibull sering digunakan pada daya tahan suatu material.

Pembahasan akhir dari subpokok bahasan ini adalah tentang sebaran Pareto.

Suatu peubah acak X dikatakan ber***sebaran Pareto*** dengan parameter $\theta > 0$ dan $\kappa > 0$ jika mempunyai pdf

$$f(x;\theta,\kappa) = \frac{\kappa}{\theta}\left(1+\frac{x}{\theta}\right)^{-(x+1)} I(x>0),$$

dinyatakan sebagai $X \sim PAR(\theta,\kappa)$. Jika $X \sim PAR(\theta,\kappa)$ maka cdf-nya adalah

$$F(x;\theta,\kappa) = \left[1-\left(1+\frac{x}{\theta}\right)^{-\kappa}\right] I(x>0),$$

bukti ditinggalkan sebagai Latihan. Sebaran Pareto sering digunakan pada permasalahan model biomedis, seperti daya tahan hidup transplantasi jantung.

Latihan 3.6

3.6.1. Menggunakan induksi matematis, buktikan sifat kedua Teorema 3.6.1, yaitu $\Gamma(n) = (n-1)!$

3.6.2. Menggunakan induksi matematis dan integral parsial secara berulang buktikan Teorema 3.6.2.

3.6.3. Tunjukkan bahwa $E(X^2) = \theta^2\kappa(1+\kappa)$.

3.6.4. Gunakan Teorema 3.6.1 untuk menentukan:
(a) $\Gamma(6)$.
(b) $\Gamma(13/2)$.

3.6.5 Misal $X \sim GAM(5,2)$. Gunakan Teorema 3.6.2 untuk menentukan:
(a) $P[X \le 10]$

(b) $P\left[7 < X \le 10\right]$

3.6.6. Misal $X \sim GAM(\theta, 2)$. Jika $x = 2$ adalah modus tunggal dari sebaran ini, tentukan parameter θ dan $P(X < 9,49)$.

3.6.7. Misal $X \sim GAM(\theta, \kappa)$. Tunjukkan bahwa
$$P(X \ge 2\theta\kappa) \le (2/e)^{\kappa}.$$
Petunjuk: Gunakan Latihan 2.6.7.

3.6.8 Misal $X \sim EXP(\theta)$. Buktikan bahwa cdf dari X adalah
$$F(x;\theta) = 1 - e^{-\frac{x}{\theta}} I(x > 0).$$

3.6.9 Misal peubah acak X menyatakan waktu hidup suatu transistor dan mengikuti sebaran eksponensial dengan parameter $\theta = 100$. Tentukan:

(a) $P\left[X > 25\right]$

(b) $P\left[X > 125\right]$

(c) $P\left[X > 125 \mid X > 25\right]$

3.6.10. Misal $X \sim WEI(\theta, \beta)$. Buktikan:

(a) cdf-nya adalah
$$F(x;\theta,\beta) = \left(1 - e^{-\frac{x}{\theta}\beta}\right) I(x > 0).$$

(b) $E(X^2) = \theta^2 \Gamma\left(1 + \dfrac{2}{\beta}\right)$.

3.6.11. Anggap daya tahan suatu material bersebaran Weibull, $X \sim WEI\left(400, \frac{2}{3}\right)$. Tentukan $P(X > 410)$.

3.6.12. Misal waktu tunggu (dalam menit) seorang pelanggan pada suatu toko bersebaran Pareto, $X \sim PAR(100,2)$. Tentukan peluang bahwa pelanggan datang antara pukul 08.05 dan 08.10.

3.7 Sebaran Normal

Suatu sebaran yang sangat penting dalam Inferens Statistis adalah sebaran normal, karena banyak analisis data yang menggunakan asumsi sebaran ini.

Untuk mengawali pembahasan sebaran ini, mula-mula kita perhatikan integral

$$I = \int_{-\infty}^{\infty} \exp\left(-\frac{y^2}{2}\right) dy.$$

Kita tentukan nilai dari integral tersebut seperti berikut ini.

$$\begin{aligned} I^2 &= \int_{-\infty}^{\infty} \exp\left(-\frac{y^2}{2}\right) dy . \int_{-\infty}^{\infty} \exp\left(-\frac{y^2}{2}\right) dy \\ &= \int_{-\infty}^{\infty} \exp\left(-\frac{y^2}{2}\right) dy . \int_{-\infty}^{\infty} \exp\left(-\frac{z^2}{2}\right) dz \\ &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \exp\left(-\frac{y^2+z^2}{2}\right) dydz. \end{aligned}$$

Serupa dengan cara untuk mendapatkan nilai dari $\Gamma(1/2),$ jika ubah integral tersebut menjadi bentuk integral kutub, yaitu dengan mengambil $y = r\cos\theta$ dan $z = r\sin\theta,$ diperoleh

$$I^2 = \int_0^{2\pi} \int_0^{\infty} e^{-\frac{r^2}{2}} r dr d\theta = 2\pi.$$

Karena I positif (mengapa?), maka

$$I = \sqrt{2\pi},$$

atau

$$\int_{-\infty}^{\infty} \exp\left(-\frac{y^2}{2}\right) dy = \sqrt{2\pi}.$$

Jika kedua ruas dari persamaan tersebut kita bagi dengan $\sqrt{2\pi},$ maka diperoleh persamaan

$$\int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{y^2}{2}\right) dy = 1.$$

Kita ubah integral menjadi peubah baru, misal $x,$ di mana

$$y=\frac{x-a}{b}, b>0,$$

maka persamaan integral di atas menjadi

$$\int_{-\infty}^{\infty}\frac{1}{b\sqrt{2\pi}}\exp\left(-\frac{(x-a)^2}{2b^2}\right)dx=1.$$

Oleh karena itu jika diambil

$$f(x)=\frac{1}{b\sqrt{2\pi}}\exp\left(-\frac{(x-a)^2}{2b^2}\right), b>0, -\infty<x<\infty,$$

maka $f(x)$ mendefinisikan pdf dari peubah acak X, karena:

(i) $f(x)>0, -\infty<x<\infty,$

(ii) $\int_{-\infty}^{\infty}f(x)dx=1.$

Peubah acak X yang mempunyai pdf seperti ini dikatakan ber***sebaran normal.***

Sekarang kita tentukan mgf dari X seperti berikut ini.

$$M(t)=E\left(e^{tX}\right)=\int_{-\infty}^{\infty}e^{tx}\frac{1}{b\sqrt{2\pi}}\exp\left(-\frac{(x-a)^2}{2b^2}\right)dx$$

$$=\int_{-\infty}^{\infty}\frac{1}{b\sqrt{2\pi}}\exp\left(-\frac{(x-a)^2}{2b^2}+tx\right)dx$$

$$=\int_{-\infty}^{\infty}\frac{1}{b\sqrt{2\pi}}\exp\left(-\frac{-2b^2tx+x^2-2ax+a^2}{2b^2}\right)dx.$$

Dengan melengkapkan bentuk kuadrat sempurna pada eksponen di atas kita peroleh

$$M(t)=\exp\left[-\frac{a^2-(a+b^2t)^2}{2b^2}\right]\int_{-\infty}^{\infty}\frac{1}{b\sqrt{2\pi}}\exp\left(-\frac{(x-a-b^2t)^2}{2b^2}\right)dx$$

$$=\exp\left(-\frac{-2ab^2t-b^4t^2}{2b^2}\right).1=\exp\left(at+\frac{b^2t^2}{2}\right).$$

Purata dan varians dari peubah ini kita peroleh dari mgf seperti berikut:

$$M'(t)=M(t)(a+b^2t)$$

dan

$$M''(t) = M(t)(b^2) + M(t)(a+b^2t)^2.$$

Oleh karena itu kita peroleh

$$\mu = M(0)(a) = a$$

dan

$$\sigma^2 = M''(0) - \mu^2 = M(0)(b^2) + M(0)(a+b^2 0)^2 - a^2 = b^2 + a^2 - a^2 = b^2.$$

Dari hasil ini kita dapat menulis ulang pdf dari X menjadi

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

yang menyatakan secara jelas bahwa purata atau ekspektasi dari dan varians X berturut-turut adalah μ dan σ^2, sedangkan mgf-nya kita tulis sebagai

$$M(t) = \exp\left(\mu t + \frac{\sigma^2 t^2}{2}\right).$$

Peubah acak X yang bersebaran demikian ditulis sebagai

$$X \sim N(\mu, \sigma^2),$$

yang dibaca sebagai X bersebaran normal dengan purata μ dan varians σ^2.

Grafik pdf sebaran normal ini mempunyai ciri-ciri:

(1) simetrik terhadap $x = \mu$,

(2) titik puncak $\left(\mu, \frac{1}{\sigma\sqrt{2\pi}}\right)$,

(3) sumbu x sebagai asimtot,

(4) titik belok pada $x = \mu \pm \sigma$,

ditinggalkan sebagai latihan rincian bukti dari ciri-ciri di atas beserta grafiknya, untuk purata dan varians tertentu.

Dalam hal khusus, yaitu dalam hal $\mu = 0$, dan $\sigma^2 = 1$, maka pdf-nya ditulis sebagai

$$f(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}, -\infty < x < \infty,$$

dan mgf adalah

$$M(t) = e^{\frac{t^2}{2}}.$$

Sebaran normal seperti ini disebut ***sebaran normal baku***, ditulis $X \sim N(0,1)$ Jika nilai x diketahui, maka cdf dari sebaran normal baku ini dapat diperoleh dari tabel sebaran normal, atau dari Minitab. Sebagai contoh jika $x = 2$, maka $\Phi(2) = 0.9772$, diperoleh dari Minitab, Φ merupakan simbol cdf dari sebaran normal baku.

Menggunakan transformasi kita dapat mengubah sebaran normal ke sebaran normal baku. Oleh karena itu untuk sebarang nilai x yang merupakan nilai dari peubah acak bersebaran normal kita dapat menentukan nilai cdf-nya. Hal ini disajikan oleh teorema berikut ini.

Teorema 3.7.1 Misal $X \sim N(\mu, \sigma^2)$ Maka

(i) $Z = \frac{X-\mu}{\sigma} \sim N(0,1)$.

(ii) $F(x) = \Phi\left(\dfrac{x-\mu}{\sigma}\right)$.

Bukti

Kita buktikan bagian pertama seperti berikut

$$F_Z(z) = P(Z \le z) = P\left(\frac{X-\mu}{\sigma} \le z\right) = P(X \le \mu + \sigma z)$$

$$= \int_{-\infty}^{\mu+\sigma z} \frac{1}{\sigma\sqrt{2\pi}} \exp\left[-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2\right] dx.$$

Kita substitusi $w = \dfrac{x-\mu}{\sigma}$, kita peroleh

$$F_Z(z) = \int_{-\infty}^{z} \frac{1}{\sqrt{2\pi}} e^{-\frac{w^2}{2}} dw = \Phi(z).$$

Oleh karena itu $Z \sim N(0,1)$

Untuk membuktikan bagian kedua kita tempuh seperti berikut ini.

$$F_X(x) = P(X \le x) = P\left(\frac{X-\mu}{\sigma} \le \frac{x-\mu}{\sigma}\right) = \Phi\left(\frac{x-\mu}{\sigma}\right).$$

Menggunakan sifat simetris pada pdf sebaran normal baku, yaitu simetrik terhadap nol, maka kita peroleh hubungan

$$\Phi(-z) = 1 - \Phi(z).$$

Sebagai contoh kita lihat persentil ke 95. Dari tabel sebaran normal kita peroleh $z_{0.95} = 1.645.$ Oleh karena itu

$$\Phi(-1.645) = 1 - \Phi(1.645) = 1 - 0.95 = 0.05.$$

Kita peroleh hubungan bahwa $z_{0.05} = -z_{1-0.05}$.

Contoh 3.7.1 Misal peubah acak X menyatakan nilai ujian Statistika Matematis mahasiswa UM, yang dianggap dapat didekati sebagai peubah acak bersebaran normal dengan purata $\mu = 75$ dan varians $\sigma^2 = 25$. Maka peluang mahasiswa mendapatkan nilai di atas delapan puluh adalah

$$\begin{aligned} P(X \geq 80) &= 1 - P(X < 80) \\ &= 1 - \Phi\left(\frac{80-75}{5}\right) \\ &= 1 - \Phi(1) \\ &= 1 - 0.8413 \\ &= 0.1587, \end{aligned}$$

dari Minitab.

Latihan 3.7

3.7.1. Misal $X \sim N(\mu, \sigma^2)$. Tunjukkan bahwa grafik pdf dari X mempunyai ciri-ciri:

(a) simetrik terhadap $x = \mu$,

(b) titik puncak $\left(\mu, \dfrac{1}{\sigma\sqrt{2\pi}}\right)$,

(c) sumbu x sebagai asimtot,

(d) titik belok pada $x = \mu \pm \sigma$,

(e) buatlah sketsa grafiknya, untuk $\mu = 7$ dan $\sigma^2 = 4$.

3.7.2. Misal $Z \sim N(0,1)$ Tentukan peluang berikut ini:

(a) $P(Z \leq 2)$.

(b) $P(Z > -2)$.

(c) $P(0.35 < Z < 2.01)$.

3.7.3. Ulangi Latihan 3.7.2 untuk menentukan a dan b sehingga:

(a) $P(Z \leq a) = 0.85$.

(b) $P(Z \le b) = 0.01.$
(c) $P(|Z| \le b) = 0.95.$

3.7.4. Misal $X \sim N(10,16)$. Tentukan
(a) $P(X \le 5).$
(b) $P(2X - 10 \le 15).$
(c) $x_{0.95}.$
(d) Tentukan c sehingga $P(10 - c < X < 10 + c) = 0.90.$

3.7.5. Misal $X \sim N(1,4)$. Hitung $P\left(1 < X^2 < 9\right).$

3.7.6. Misal $X \sim N(5,10)$. Hitung $P\left(0.04 < (X-5)^2 < 38.4\right).$

3.7.7. Tunjukkan bahwa konstanta k dapat dipilih sehingga

$$f(x) = k2^{-x^2}, -\infty < x < \infty,$$

memenuhi kondisi pdf sebaran normal. *Petunjuk:* Tulis $2 = e^{\ln 2.}$

3.7.8. Misal peubah acak X mempunyai pdf

$$f(x) = \frac{2}{\sqrt{2\pi}} e^{-x^2/2} I(0 < x < \infty).$$

Tentukan purata dan varians dari X. *Petunjuk:* Hitung $E(X)$ langsung dan $E(X^2)$ dengan membandingkan integralnya dengan integral dari sebaran normal baku.

3.7.9. Hitunglah $\int_2^3 \exp\left[-2\left(x-3\right)^2\right] dx.$

3.7.10. Misal $X \sim N(\mu, \sigma^2)$. Tunjukkan bahwa $E(|X - \mu|) = \sigma\sqrt{2/\pi}.$

BAB 4
SEBARAN BERSAMA DAN SIFAT-SIFAT PEUBAH ACAK

4.1 Pendahuluan

Dalam banyak aplikasi kita sering berhadapan dengan lebih dari satu peubah acak, katakanlah $X_1, X_2, ..., X_n$. Secara matematis peubah-peubah acak tersebut dapat dinyatakan dalam bentuk vektor berdimensi n, yaitu $\mathbf{X} = (X_1, X_2, ..., X_n)^t$ dan nilai-nilainya dinyatakan sebagai $\mathbf{x} = (x_1, x_2, ..., x_n)^t$ dalam ruang Euclid berdimensi n, t menyatakan transpos vektor. Vektor semacam ini disebut sebagai ***vektor acak*.** Peubah-peubah acak $X_1, X_2, ..., X_n$ mungkin saja dibangkitkan dari satu peubah acak. Sebagai contoh, misal X menyatakan nilai Statistika Matematis mahasiswa, maka kita dapat memandang X_1 sebagai nilai Statistika Matematis mahasiswa pertama, X_2 sebagai nilai Statistika Matematis mahasiswa kedua, dan seterusnya sampai mahasiswa ke n. Perlu kita catat bahwa peubah acak berkaitan dengan percobaan yang belum dilakukan. Sedangkan jika kita sudah melakukan percobaan, maka yang kita peroleh adalah nilai dari peubah acak, yaitu x, atau $x_1, x_2, ..., x_n$.

Tujuan instruksional umum dari mempelajari bab ini adalah Anda diharapkan memahami konsep sebaran bersama dan sifat-sifat peubah acak.

Tujuan instruksional khusus mempelajari bab ini adalah Anda diharapkan dapat menentukan: pdf bersama, cdf bersama, ekspektasi, varians, mgf, pdf marginal, cdf marginal, pdf bersyarat, ekspektasi bersyarat, varians bersyarat, kovarians, koefisien korelasi, kebebasan peubah acak, sampel acak, dan transformasi peubah-peubah acak.

4.2 Sebaran Bersama

Kita ingat kembali materi pada Bab 2 Subpokok Bahasan 2.2, tentang konsep peluang terinduksi oleh peubah acak diskret X, yang didefinisikan sebagai

$$P(A) = P_X(A) = P(X \in A) = P(C),$$

di mana $C = \{c : X(c) \in A, c \in S\}$. Dari pendefinisian ini kita peroleh bahwa kejadian $X = x$ ekivalen dengan kejadian $C = \{e \in S \mid X(e) = x\}$, sehingga menentukan $P[X = x]$ sama artinya dengan menentukan $P(C)$. Serupa dengan hal itu kita peroleh pula kejadian $X_i = x_i$ ekivalen dengan kejadian $C_i = \{e \in S \mid X_i(e) = x_i\}, i = 1, 2, ..., n$. Oleh karena itu jika kita menentukan $P[X_1 = x_1, X_2 = x_2, ..., X_n = x_n]$ sama artinya dengan kita menentukan $P(C_1 \cap C_2 \cap ... \cap C_n)$. Kaitannya dengan pdf, untuk yang pertama kita tulis sebagai $f(x)$, yaitu

$$f(x) = P[X = x],$$

sedangkan yang kedua kita tulis sebagai $f(x_1, x_2, ..., x_n)$. Sehingga

$$f(x_1, x_2, ..., x_n) = P[X_1 = x_1, X_2 = x_2, ..., X_n = x_n].$$

Oleh karena itu $f(x_1, x_2, ..., x_n)$ disebut sebagai pdf bersama dari peubah-peubah acak $X_1, X_2, ..., X_n$. Definisi formal tentang pdf bersama diberikan oleh definisi berikut ini.

Definisi 4.2.1 Suatu ***pdf bersama*** peubah-peubah acak diskret berdimensi n, yaitu $X_1, X_2, ..., X_n$ didefinisikan sebagai

$$f(x_1, x_2, ..., x_n) = P[X_1 = x_1, X_2 = x_2, ..., X_n = x_n]$$

untuk semua nilai $x_1, x_2, ..., x_n$.

Contoh 4.2.1 Misal satu kantong berisi 15 kelereng, yang terdiri dari 5 kelereng merah, 7 kelereng putih, dan 3 kelereng kuning. Misal dari kantong tersebut diambil 5 kelereng tanpa pengembalian, dan misal X_1 banyak kelereng merah yang terambil, X_2 banyak kelereng putih yang terambil. Maka pdf bersama dari X_1 dan X_2 adalah

$$f(x_1, x_2) = \frac{\binom{5}{x_1}\binom{7}{x_2}\binom{3}{5 - x_1 - x_2}}{\binom{15}{5}}$$

$$I(x_1 = 0, 1, \ldots, 5, \ \ x_2 = 0, 1, \ldots, 5, \ x_1 + x_2 \le 5).$$

Contoh 4.2.2 Seperti pada Contoh 4.2.1, tetapi pengambilan dengan pengembalian. Maka kita peroleh pdf bersama dari X_1 dan X_2

$$f(x_1, x_2) = \frac{5!}{x_1!x_2!(5 - x_1 - x_2)!} p_1^{x_1} p_2^{x_2} (1 - p_1 - p_2)^{5 - x_1 - x_2}$$

$$I(x_1 = 0,1,...,5,\ x_2 = 0,1,...,5),$$

di mana $p_1 = \frac{5}{15}$ dan $p_2 = \frac{7}{15}$. Bentuk faktorial $\frac{5!}{x_1!x_2!(5 - x_1 - x_2)!}$ merupakan banyak susunan dari 5 kelereng yang terdiri dari x_1 kelereng merah, x_2 kelereng putih dan $5 - x_1 - x_2$ kelereng kuning. Peubah acak $\mathbf{X} = \left(X_1, X_2\right)^t$ di atas dikatakan bersebaran multinomial, secara umum ditulis sebagai $\boldsymbol{X} = (X_1, X_2)^t \sim MULT(n, p_1, p_2)$, dalam contoh ini $n = 5$.

Teorema 4.2.1 Suatu fungsi $f(x_1, x_2, ..., x_n)$ adalah pdf bersama dari peubah-peubah acak diskret $X_1, X_2, ..., X_n$ jika dan hanya jika memenuhi kedua syarat berikut:

(1) $f(x_1, x_2, ..., x_n) \geq 0$, untuk semua nilai yang mungkin dari $x_1, x_2, ..., x_n$,

(2) $\sum_{x_1} ... \sum_{x_n} f(x_1, x_2, ..., x_n) = 1.$

Tabel berikut ini merupakan contoh nilai pdf dari peubah acak $(X_1, X_2) \sim MULT(3{,}0.4{,}0.4)$.

TABEL 4.2.1
pdf DARI PEUBAH ACAK $(X_1, X_2) \sim MULT(3, 0.4, 0.4)$

		x_2				$f_1(x_1)$
		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
$f_2(x_2)$		0.216	0.432	0.288	0.064	1.000

Berdasarkan pdf bersama peubah-peubah acak X_1 dan X_2, yaitu $f(x_1, x_2)$, kita tentukan pdf peubah acak X_1, yaitu $f_1(x_1)$, seperti berikut ini

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

demikian juga untuk peubah acak X_2, kita punya

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

Fungsi semacam ini disebut sebagai pdf marginal. Untuk kasus pertama disebut pdf marginal dari X_1, dan untuk kasus kedua disebut pdf marginal dari X_2. Definisi formal tentang hal ini disajikan pada definisi berikut ini.

Definisi 4.2.2 Misal peubah-peubah acak diskret X_1 dan X_2 mempunyai pdf bersama $f(x_1, x_2)$. Maka ***pdf marginal*** dari X_1 adalah

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

dan pdf marginal dari X_2 adalah

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

Kembali ke Tabel 4.2.1, maka $f_1(0)$ diperoleh dari

$$f_1(0) = \sum_{x_2=0}^{3} f(x_1, x_2) = f(0,0) + f(0,1) + f(0,2) + f(0,3)$$
$$= 0.008 + 0.048 + 0.096 + 0.064 = 0.216,$$
demikian pula seterusnya untuk yang lain.

Secara umum jika $X_1, X_2, ..., X_n$ peubah-peubah acak diskret dari ruang berdimensi n, maka pdf marginal dari X_j adalah
$$f_j(x_j) = \sum_{x_n} ... \sum_{x_1} f(x_1, x_2, ..., x_j, ..., x_n),$$
untuk semua $x_i \neq x_j$.

Jika $\mathbf{X}$ peubah acak diskret dalam dimensi n dan $\mathbf{A}$ sebarang kejadian dalam ruang $\mathbf{X} = (X_1, X_2, ..., X_n)^t$, maka
$$P(X \in \mathbf{A}) = \sum ... \sum_{\mathbf{x} \in \mathbf{A}} f(x_1, x_2, ..., x_n).$$
Dari Tabel 4.2.1, $P[2X_1 < X_2]$ dapat kita hitung seperti berikut
$$P[2X_1 < X_2] = f(0,1) + f(0,2) + f(0,3) + f(1,3)$$
$$= 0.048 + 0.096 + 0.064 + 0.000 = 0.208.$$

Definisi tentang cdf bersama peubah-peubah acak diberikan berikut ini.

Definisi 4.2.3 Suatu ***cdf bersama*** dari peubah-peubah acak diskret berdimensi n, yaitu $X_1, X_2, ..., X_n$ didefinisikan sebagai
$$F(x_1, x_2, ..., x_n) = P[X_1 \leq x_1, X_2 \leq x_2, ..., X_n \leq x_n]$$
untuk semua nilai $x_1, x_2, ..., x_n$.

Sebaran bersama yang telah kita bahas adalah sebaran bersama peubah-acak-peubah acak tipe diskret. Pembahasan berikutnya adalah pembahasan tentang sebaran bersama peubah-peubah acak tipe kontinu.

Suatu cdf bersama dari peubah-peubah acak kontinu $\mathbf{X}^t = (X_1, X_2, ..., X_n)$ disajikan oleh definisi berikut ini.

Definisi 4.2.4 Peubah-peubah acak $X_1, X_2, ..., X_n$ dari ruang berdimensi n dikatakan kontinu, jika ada fungsi $f(x_1, x_2, ..., x_n)$ yang disebut sebagai ***pdf bersama*** dari $X_1, X_2, ..., X_n$, sehingga cdf-nya adalah

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

untuk semua nilai $x_1, ..., x_n$.

Seperti halnya pada ruang berdimensi satu, pdf bersama dapat diperoleh dari cdf dengan penurunan, yaitu

$$f(x_1, x_2, ..., x_n) = \frac{\partial^n}{\partial x_1 ... \partial x_n} F(x_1, x_2, ..., x_n),$$

asalkan turunan parsial tersebut ada.

Teorema 4.2.2 Suatu fungsi $f(x_1, x_2, ..., x_n)$ dikatakan pdf bersama peubah-peubah acak kontinu $X_1, X_2, ..., X_n$ dari ruang berdimensi n jika dan hanya jika:

(1) $f(x_1, x_2, ..., x_n) \geq 0,$ untuk semua nilai $x_1, ..., x_n$,

(2) $\int_{-\infty}^{\infty} ... \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x_1, x_2, ..., x_n) dx_1 dx_2 ... dx_n = 1.$

Contoh 4.2.3 Misal X_1 menyatakan konsentrasi dari substansi dalam percobaan pertama, dan X_2 menyatakan konsentrasi dari substansi dalam percobaan kedua. Anggap pdf bersama dari dua peubah acak tersebut adalah

$$f(x_1, x_2) = 4x_1 x_2 I(0 < x_1 < 1, 0 < x_2 < 1).$$

Maka cdf-nya adalah

$$F(x_1, x_2) = \int_{-\infty}^{x_2} \int_{-\infty}^{x_1} f(t_1, t_2) dt_1 dt_2 = \int_{-\infty}^{x_2} \int_{-\infty}^{x_1} 4t_1 t_2 dt_1 dt_2 = x_1^2 x_2^2 I(0 < x_1 < 1, 0 < x_2 < 1)$$

$$+ x_1^2 I(0 < x_1 < 1, 1 < x_2) + x_2^2 I(1 < x_1, 0 < x_2 < 1) + I(1 \leq x_1, 1 \leq x_2).$$

Seperti halnya pada peubah acak diskret dari ruang berdimensi n, pada peubah acak kontinu dalam ruang berdimensi n ini berlaku pula sifat bahwa jika $\mathbf{A}$ sebarang kejadian dalam ruang $\mathbf{X} = (X_1, ..., X_n)^t$, maka

$$P(\mathbf{X} \in \mathbf{A}) = \int ... \int_{A} f(x_1, x_{2,} ..., x_n) dx_1 ... dx_n.$$

Dari Contoh 4.2.3, $P[2X_1 < X_2]$ dapat kita hitung seperti berikut

$$P[2X_1 < X_2] = \int_0^1 \int_0^{\frac{1}{2}x_2} 4x_1x_2 dx_1 dx_2 = 2\int_0^1 \left[x_1^2\right]_0^{\frac{1}{2}x_2} x_2 dx_2$$

$$= \frac{1}{2}\int_0^1 x_2^3 dx_2 = \frac{1}{8}\left[x_2^4\right]_0^1 = \frac{1}{8}.$$

Definisi 4.2.5 Misal peubah-peubah acak kontinu X_1 dan X_2 mempunyai pdf bersama $f(x_1, x_2)$, maka ***pdf marginal*** dari X_1 adalah

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

dan pdf marginal dari X_2 adalah

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

Kembali ke Contoh 4.2.3, pdf marginal dari X_1 adalah

$$f_1(x_1) = \int_{-\infty}^{\infty} f(x_1, x_2) dx_2 = \int_0^1 4x_1x_2 dx_2 = 2x_1\left[x_2^2\right]_0^1 = 2x_1 I(0 < x_1 < 1).$$

Dengan cara serupa kita peroleh pdf bersama dari X_2

$$f_2(x_2) = 2x_2 I(0 < x_2 < 1).$$

Secara umum jika $\mathbf{X} = (X_1, X_2, ..., X_n)^t$ peubah acak dari ruang berdimensi n, maka pdf marginal dari X_j adalah

$$f_j(x_j) = \int_{-\infty}^{\infty} ... \int_{-\infty}^{\infty} f(x_1, ..., x_j, ..., x_n) dx_1 ... dx_n,$$

untuk semua $x_i \neq x_j$.

Serupa dengan sifat-sifat cdf pada Teorema 2.2.1, berikut ini merupakan sifat-sifat cdf bersama peubah-peubah acak X_1 dan X_2.

Teorema 4.2.3 Suatu fungsi $F(x_1, x_2)$ merupakan cdf bersama dari peubah-peubah acak diskrit X_1 dan X_2 jika dan hanya jika

(1) $\lim\limits_{x_1 \to -\infty, x_2 \to -\infty} F(x_1, x_2) = 0,$

(2) $\lim\limits_{x_1 \to \infty, x_2 \to \infty} F(x_1, x_2) = 1,$

(3) $F(b,d) - F(b,c) - F(a,d) + F(a,c) \geq 0,\ a < b,\ c < d,$

(4) $\lim\limits_{h \to 0^+} F(x_1 + h, x_2) = \lim\limits_{h \to 0^+} F(x_1, x_2 + h) = F(x_1, x_2).$

Contoh 4.2.4 Kita perhatikan fungsi berikut:

$$F(x_1, x_2) = I(x_1 + x_2 \geq -1).$$

Jika kita pilih $a = c = -1$ dan $b = d = 1,$ maka

$$F(1,1) - F(1,-1) - F(-1,1) + F(-1,-1) = 1 - 1 - 1 + 0 = -1.$$

Jadi $F(x_1, x_2)$ bukan cdf bersama dari X_1 dan X_2.

Pembahasan berikutnya dari subpokok bahasan ini adalah tentang ekspektasi dari dua atau lebih peubah acak.

Definisi 4.2.6 Misal peubah-peubah acak $X_1, X_2, ..., X_n$ mempunyai pdf bersama $f(x_1, x_2, ..., x_n),$ dan misal $Y = u\left(X_1, X_2, ..., X_n\right).$ Jika

$$\sum_{x_1} ... \sum_{x_n} | u\left(x_1, x_2, ..., x_n\right) | \ f(x_1, x_2, ..., x_n)$$

ada untuk Y diskret, dan

$$\int_{-\infty}^{\infty} ... \int_{-\infty}^{\infty} | u\left(x_1, x_2, ..., x_n\right) | f(x_1, x_2, ..., x_n) dx_1 ... dx_n$$

ada untuk Y kontinu. Maka ekspektasi dari Y adalah

$$E(Y) = E\left[u(X_1, X_2, ..., X_n)\right] = \sum_{x_1} ... \sum_{x_n} u(x_1, x_2, ..., x_n) f(x_1, x_2, ..., x_n),$$

jika Y diskret, dan

$$E(Y) = E\left[u(X_1, X_2, ..., X_n)\right] = \int_{-\infty}^{\infty} ... \int_{-\infty}^{\infty} u(x_1, x_2, ..., x_n) f(x_1, x_2, ..., x_n) dx_1 dx_2 ... dx_n,$$

jika Y kontinu.

Teorema 4.2.4 Misal X_1 dan X_2 dua peubah acak dengan pdf bersama $f(x_1, x_2)$. Maka

$$E\left(X_1 + X_2\right) = E\left(X_1\right) + E(X_2).$$

Bukti

Kita buktikan untuk peubah acak kontinu.

$$\begin{aligned}
E\left(X_1 + X_2\right) &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} (x_1 + x_2) f(x_1, x_2) dx_1 dx_2 \\
&= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} x_1 f(x_1, x_2) dx_1 dx_2 + \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} x_2 f(x_1, x_2) dx_1 dx_2 \\
&= \int_{-\infty}^{\infty} x_1 \left[\int_{-\infty}^{\infty} f(x_1, x_2) dx_2\right] dx_1 + \int_{-\infty}^{\infty} x_2 \left[\int_{-\infty}^{\infty} f(x_1, x_2) dx_1\right] dx_2 \\
&= \int_{-\infty}^{\infty} x_1 f_1(x_1) dx_1 + \int_{-\infty}^{\infty} x_2 f_2(x_2) dx_2 \\
&= E(X_1) + E(X_2).
\end{aligned}$$

Untuk peubah acak diskret tinggal mengganti integral dengan jumlah.

Contoh 4.2.5 Misal X dan Y dua peubah acak kontinu dengan pdf bersama

$$f(x, y) = (x + y) I(0 < x < 1, 0 < y < 1).$$

Maka

$$E(XY^2) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} xy^2 f(x, y) dx dy = \int_0^1\int_0^1 xy^2 (x + y) dx dy = \frac{17}{72}.$$

Berikutnya kita berikan definisi *MGF* bersama dari n peubah acak $X_1, X_2, ..., X_n$, untuk melengkapi pembahasan pada subpokok bahasan ini.

Definisi 4.2.7 Suatu *MGF* bersama dari vektor acak $\mathbf{X} = (X_1, X_2, ..., X_n)^t$, ditulis $M_{\mathbf{X}}(\mathbf{t})$ adalah

$$M_{\mathbf{X}}(\mathbf{t}) = E\left[\exp\left(\sum_{i=1}^{n} t_i X_i\right)\right],$$

di mana $\mathbf{t} = (t_1, t_2, ..., t_n)^t$ dan $-h < t_i < h$ untuk suatu $h > 0$.

Dari definisi di atas memungkinkan kita untuk memperoleh *MGF* marginal dari *MGF* bersama, sebagai contoh

$$M_X(t_1) = M_{X,Y}(t_1, 0),$$

dan

$$M_Y(t_2) = M_{X,Y}(0, t_2).$$

Latihan 4.2

4.2.1. Misal $c = (i, j)$, menyatakan sebarang hasil dari pelemparan satu dadu sebanyak dua kali. Misal $X(c) = i + j$, $Y(c) = i - j$, dan $Z(c) = i^2 - j^2$. Buatlah tabel pdf bersama dari:
(a) X dan Y.
(b) X dan Z.
(c) Y dan Z.
(d) X , Y , dan Z.

4.2.2. Ulangi Latihan 4.2.1, untuk menentukan:
(a) $P\left[X + Y = 5\right]$.
(b) $P\left[2X - Y < 5\right]$.
(c) $P\left[2X + Y \geq 5\right]$.
(d) $P\left[2X + Y - Z \geq 5\right]$.

4.2.3. Misal X dan Y dua peubah acak diskret dengan pdf bersama

$$f(x, y) = c(x + y)I(x = 0,1,2; y = 0,1,2).$$

Tentukan:
(a) c.
(b) pdf marginal dari X.
(c) $P\left[X + Y = 2\right]$.

4.2.4. Ulangi Latihan 4.2.3, untuk

$$f(x, y) = c\,\frac{2^{(x+y)}}{x!\,y!}\,I(x = 0,1,2,\ldots; y = 0,1,2,\ldots).$$

4.2.5. Misal X dan Y dua peubah acak diskret dengan pdf bersama

$$f(x,y)=\frac{x+2y}{18}I((x,y)=(1,1),(1,2),(2,1),(2,2)).$$

Tentukanlah

(a) pdf marginal dari X

(b) pdf marginal dari Y

Petunjuk: buatlah tabel pdf-nya terlebih dahulu.

4.2.6. Ulangi Latihan 4.2.5, untuk membuat tabel cdf bersama dari X dan Y.

4.2.7 Misal dua peubah acak kontinu X dan Y mempunyai pdf bersama

$$f(x,y)=e^{-(x+y)}I(0<x<\infty,0<y<\infty).$$

Tentukan:

(a) $P\left[X>Y\right]$.

(b) $P\left[X+Y<2\right]$.

(c) pdf marginal dari X.

(d) $P\left[X<2\right]$.

(e) cdf bersama dari X dan Y.

4.2.8. Misal X, Y, dan Z kontinu dengan pdf bersama

$$f(x,y,z)=6I(0<x<y<z<1).$$

Tentukan:

(a) masing-masing pdf marginal dari X, Y, dan Z.

(b) pdf bersama dari X dan Y.

(c) $P\left[X+Y<1\right]$.

4.2.9. Misal dua peubah acak kontinu X dan Y mempunyai pdf bersama

$$f(x,y)=k(x+y)I(0<x<y<1).$$

Tentukan:

(a) k.

(b) $P\left[X<Y<2X\right]$.

4.2.10. Misal X dan Y dua peubah acak dengan pdf bersama

$$f(x,y) = x + yI(0 < x < 1, 0 < y < 1).$$

Tentukan:

(a) $E(X)$

(b) $E(Y)$

(c) $E(X+2Y)$.

4.2.11. Misal dua peubah acak X dan Y mempunyai pdf bersama

$$f(x,y) = e^{-y} I(0 < x < y < \infty).$$

Tentukan MGF bersama dari X dan Y.

4.3 Sebaran Bersyarat

Pada definisi peluang bersyarat, jika kita mempunyai dua kejadian A dan B, maka peluang bersyarat dari kejadian B bila kejadian A muncul terlebih dahulu adalah

$$P(B \mid A) = \frac{P(A \cap B)}{P(A)}.$$

Sekarang jika kita misalkan X_1 dan X_2 dua peubah acak diskret, di mana $X_1 = x_1$ ekivalen dengan kejadian A dan $X_2 = x_2$ ekivalen dengan kejadian B, maka berdasarkan definisi peluang tersebut kita memperoleh hubungan

$$P\left[X_2 = x_2 \mid X_1 = x_1\right] = \frac{P\left[X_1 = x_1, X_2 = x_2\right]}{P\left[X_1 = x_1\right]} = \frac{f(x_1, x_2)}{f_1(x_1)}.$$

Jika $P\left[X_2 = x_2 \mid X_1 = x_1\right]$ kita tulis sebagai $f(x_2 \mid x_1)$, maka kita peroleh

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

Uraian tersebut mendasari definisi pdf bersyarat berikut ini.

Definisi 4.3.1 Misal X_1 dan X_2 dua peubah acak diskret atau kontinu, dengan pdf bersama $f(x_1, x_2)$. Maka *pdf bersyarat* dari X_2 bila $X_1 = x_1$ adalah

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

dan pdf bersyarat dari X_1 bila $X_2 = x_2$ adalah

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

Contoh 4.3.1 Misal X_1 dan X_2 dua peubah acak dengan pdf $f(x_1, x_2) = 8x_1x_2 I(0 < x_1 < x_2 < 1)$. Kita ingin menentukan $f(x_1 \mid x_2)$. Kita tentukan terlebih dahulu $f_2(x_2)$ seperti berikut

$$f_2(x_2) = \int_0^{x_2} 8x_1x_2 dx_1 = 4x_2^3 I(0 < x_2 < 1).$$

Oleh karena itu

$$f(x_1 \mid x_2) = \frac{8x_1x_2}{4x_2^3} = \frac{2x_1}{x_2^2} I(0 < x_1 < x_2 < 1).$$

Dari pdf bersyarat ini, kita dapat menghitung peluang bersyarat dari suatu kejadian $[a < X_1 < b]$ diberikan $X_2 = x_2$, yaitu

$$P[a < X_1 < b \mid X_2 = x_2] = \int_a^b f(x_1 \mid x_2) dx_1,$$

jika X_1 dan X_2 kontinu. Dalam hal X_1 dan X_2 diskret tinggal mengganti integral menjadi jumlah.

Lebih lanjut untuk ekspektasi bersyarat juga bergantung pada pdf bersyarat, lebih jelas bisa kita lihat pada definisi berikut ini.

Definisi 4.3.2 Misal X_1 dan X_2 sebarang dua peubah acak. Maka ***ekspektasi bersyarat*** dari X_2 bila $X_1 = x_1$, jika ada,

$$E(X_2 \mid x_1) = \sum_{x_2} x_2 f(x_2 \mid x_1),$$

jika X_1 dan X_2 diskret, dan

$$E(X_2 \mid x_1) = \int_{-\infty}^{\infty} x_2 f(x_2 \mid x_1) dx_2,$$

jika X_1 dan X_2 kontinu.

Serupa dengan definisi di atas, ekspektasi bersyarat dari X_1 bila $X_2 = x_2$ adalah

$$E(X_1 \mid x_2) = \sum_{x_1} x_1 f(x_1 \mid x_2),$$

jika X_1 dan X_2 diskret, dan

$$E(X_1 \mid x_2) = \int_{-\infty}^{\infty} x_1 f(x_1 \mid x_2) dx_1,$$

jika X_1 dan X_2 kontinu.

Varians bersyarat diberikan oleh definisi berikut ini.

Definisi 4.3.3 Misal X_1 dan X_2 sebarang dua peubah acak. Maka ***varians bersyarat*** dari X_2 bila $X_1 = x_1$ adalah

$$Var\left(X_2 \mid x_1\right) = E\left\{\left[X_2 - E(X_2 \mid x_1)\right]^2 \mid x_1\right\}.$$

Hal serupa untuk varians bersyarat dari X_1 bila $X_2 = x_2$, yaitu

$$Var\left(X_1 \mid x_2\right) = E\left\{\left[X_1 - E(X_1 \mid x_2)\right]^2 \mid x_2\right\}.$$

Contoh 4.3.2 Misal dua peubah acak X_1 dan X_2 mempunyai pdf bersama

$$f(x_1, x_2) = 2I(0 < x_1 < x_2 < 1).$$

Maka:

(1) pdf marginal dari X_2 adalah

$$f_2(x_2) = \int_0^{x_2} 2 dx_1 = 2x_2 I(0 < x_2 < 1).$$

(2) pdf bersyarat dari X_1 bila $X_2 = x_2$, adalah

$$f(x_1 \mid x_2) = \frac{2}{2x_2} = \frac{1}{x_2} I(0 < x_1 < x_2 < 1).$$

(3) Ekspektasi bersyarat dari X_1 bila $X_2 = x_2$, adalah

$$E(X_1 \mid x_2) = \int_{-\infty}^{\infty} x_1 f(x_1 \mid x_2) dx_1 = \int_0^{x_2} x_1 \frac{1}{x_2} dx_1 = \frac{x_2}{2}, 0 < x_2 < 1.$$

(4) Varians bersyarat dari X_1 bila $X_2 = x_2$, adalah

$$Var\left(X_1 \mid x_2\right) = E\left\{\left[X_1 - E(X_1 \mid x_2)\right]^2 \mid x_2\right\}$$

$$= \int_0^{x_2} \left(x_1 - \frac{x_2}{2}\right)^2 \left(\frac{1}{x_2}\right) dx_1 = \frac{x_2^2}{12},\ 0 < x_2 < 1.$$

(5) Peluang bersyarat dari X_1 bila $X_2 = x_2$, $P\left[0 < X_1 < \frac{1}{2} \mid x_2 = \frac{3}{4}\right]$ adalah

$$P\left[0 < X_1 < \tfrac{1}{2} \mid x_2 = \tfrac{3}{4}\right] = \int_0^{1/2} f(x_1 \mid \tfrac{3}{4}) dx_1 = \int_0^{1/2} \left(\frac{4}{3}\right) dx_1 = \frac{2}{3}.$$

Karena $E(X_2 \mid x_1)$ adalah fungsi dari x_1, maka $E(X_2 \mid X_1)$ adalah peubah acak dengan sebaran, purata, dan varians sendiri. Sebagai gambaran kita perhatikan contoh berikut ini.

Contoh 4.3.3 Misal X_1 dan X_2 peubah-peubah acak yang mempunyai pdf bersama

$$f(x_1, x_2) = 6x_2 I(0 < x_2 < x_1 < 1).$$

Maka pdf marginal dari X_1 adalah

$$f_1(x_1) = \int_0^{x_1} 6x_2 dx_2 = 3x_1^2 I(0 < x_1 < 1).$$

Oleh karena itu pdf bersyarat dari X_2 bila $X_1 = x_1$ adalah

$$f(x_2 \mid x_1) = \frac{6x_2}{3x_1^2} = \frac{2x_2}{x_1^2} I(0 < x_2 < x_1 < 1),$$

dan purata bersyarat dari X_2 bila $X_1 = x_1$ adalah

$$E\left(X_2 \mid x_1\right) = \int_0^{x_1} x_2 \left(\frac{2x_2}{x_1^2}\right) dx_2 = \frac{2}{3} x_1 I(0 < x_1 < 1).$$

Kita misalkan $Y = E\left(X_2 \mid X_1\right) = 2X_1/3.$ Maka cdf dari Y adalah

$$G(y) = P(Y \le y) = P\left(X_1 \le \frac{3y}{2} \right)$$

$$= \int_0^{3y/2} 3x_1^2 dx_1 I(0 \le y < \frac{2}{3}) + I(\frac{2}{3} \le y)$$

$$= \frac{27y^3}{8} I\left(0 \le y < \frac{2}{3} \right) + I\left(\frac{2}{3} \le y \right).$$

Jadi pdf, purata, dan varians dari Y berturut-turut adalah

$$g(y) = \frac{81y^2}{8} I\left(0 < y < \frac{2}{3} \right),$$

$$E(Y) = \int_0^{2/3} y\left(\frac{81y^2}{8} \right) dy = \frac{1}{2},$$

dan

$$Var(Y) = \int_0^{2/3} y^2\left(\frac{81y^2}{8} \right) dy - \frac{1}{4} = \frac{1}{60}.$$

Mudah ditunjukkan bahwa $E(X_2) = \frac{1}{2}$ dan $Var(X_2) = \frac{1}{20}$. Sehingga kita peroleh hubungan

$$E(Y) = E\left[E\left(X_2 \mid X_1 \right) \right] = E(X_2)$$

dan

$$Var(Y) = Var\left[E\left(X_2 \mid X_1 \right) \right] \le Var(X_2).$$

Hasil dari Contoh 4.3.3 secara umum disajikan oleh teorema berikut ini.

Teorema 4.3.1 Misal X_1 dan X_2 peubah-peubah acak sehingga varians dari X_2 terhingga. Maka

(a) $E\left[E\left(X_2 \mid X_1 \right) \right] = E(X_2)$.

(b) $Var\left[E\left(X_2 \mid X_1 \right) \right] \le Var(X_2)$.

Bukti

Kita buktikan untuk kasus kontinu. Untuk memperoleh bukti kasus diskret, tinggal mengganti integral dengan jumlah. Kita buktikan terlebih dahulu bagian (a) seperti berikut ini

$$E\left(X_2\right)=\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} x_2 f(x_1,x_2)dx_2dx_1=\int_{-\infty}^{\infty}\left[\int_{-\infty}^{\infty} x_2\frac{f(x_1,x_2)}{f_1(x_1)}dx_2\right]f_1(x_1)dx_1$$

$$=\int_{-\infty}^{\infty} E(X_2\mid x_1)f_1(x_1)dx_1=E\left[E(X_2\mid X_1)\right].$$

Selanjutnya kita buktikan bagian (b). Misal $\mu_2=E(X_2)$. Maka

$$\begin{aligned}
Var(X_2)&=E\left[X_2-\mu_2\right]^2\\
&=E\left\{\left[X_2-E(X_2\mid X_1)+E(X_2\mid X_1)-\mu_2\right]^2\right\}\\
&=E\left\{\left[X_2-E(X_2\mid X_1)\right]^2\right\}+E\left\{\left[E(X_2\mid X_1)-\mu_2\right]^2\right\}\\
&\quad+2E\left\{\left[X_2-E(X_2\mid X_1)\right]\left[E(X_2\mid X_1)-\mu_2\right]\right\}.
\end{aligned}$$

Kita akan menunjukkan bahwa suku terakhir pada ruas kanan sama dengan nol, seperti berikut ini.

$$2\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\left[x_2-E(X_2\mid X_1)\right]\left[E(X_2\mid X_1)-\mu_2\right]f(x_1,x_2)dx_2dx_1=$$

$$2\int_{-\infty}^{\infty}\left[E(X_2\mid X_1)-\mu_2\right]\left\{\int_{-\infty}^{\infty}\left[x_2-E(X_2\mid X_1)\right]\frac{f(x_1,x_2)}{f_1(x_1)}dx_2\right\}f_1(x_1)dx_1$$

$$=2\int_{-\infty}^{\infty}\left[E(X_2\mid X_1)-\mu_2\right]\left\{E(X_2\mid X_1)-E(X_2\mid X_1)\right\}f_1(x_1)dx_1=0.$$

Oleh karena itu

$$Var(X_2)=E\left\{\left[X_2-E(X_2\mid x_1)\right]^2\right\}+E\left\{\left[E(X_2\mid x_1)-\mu_2\right]^2\right\}.$$

Karena suku pertama ruas kanan tidak negatif dan suku keduanya adalah $Var\left[E(X_2\mid X_1)\right]$, maka kita peroleh

$$Var(X_2)\geq Var\left[E(X_2\mid X_1)\right].$$

Bukti selesai.

Latihan 4.3

4.3.1. Misal dua kartu diambil dari satu pak kartu yang berisi 52 kartu *bridge* tanpa pengembalian. Misal X menyatakan banyak kartu hati dan Y banyak kartu hitam yang terambil.

(a) Hitung $P\left[Y=1\mid X=1\right]$.

(b) Hitung $P\left[Y=y \mid X=1\right]$.

(c) Hitung $P\left[Y=y \mid X=x\right]$.

(d) Buat tabel $P\left[X+Y \le z\right]; z=0,1$.

4.3.2 Misal X dan Y dua peubah acak diskret dengan pdf bersama

$$f(x,y)=\frac{x+2y}{18} I\left[(x,y)=(1,1),(1,2),(2,1),(2,2)\right].$$

Tentukan ekspektasi bersyarat dan varians bersyarat dari X bila $Y=1$ atau 2.

4.3.3. Misal

$$f(x \mid y)=\frac{c_1 x}{y^2} I(0<x<y<1),$$

dan misal

$$f_2(y)=c_2 y^4 I(0<y<1),$$

Berturut-turut merupakan pdf bersyarat dari X bila $Y=y$, dan pdf marginal dari Y. Tentukan:

(a) konstanta c_1 dan c_2,

(b) pdf bersama dari X dan Y,

(c) $P\left[\frac{1}{4}<X<\frac{1}{2} \mid Y=\frac{3}{8}\right]$.

4.3.4. Misal

$$f(x,y)=21x^2y^2 I(0<x<y<1),$$

merupakan pdf bersama dari X dan Y. Tentukan ekspektasi dan varians bersyarat dari X bila $Y=y, 0<y<1$.

4.3.5. Misal suatu titik dipilih secara acak dari selang $(0,1)$ dan misal X menyatakan bilangan yang berkorespondensi dengan titik tersebut. Kemudian dipilih titik secara acak pada selang $(0,x)$, di mana x adalah hasil pada pemilihan titik pertama, dan misalkan Y menyatakan bilangan yang berkorespondensi dengan titik yang terpilih pada selang $(0,x)$.

(a) Tentukan pdf marginal $f_1(x)$.

(b) Tentukan pdf bersyarat $f(y \mid x)$.

(c) Tentukan pdf bersama $f(x, y)$.

(d) Hitung $P\left[X + Y \geq 1\right]$.

(e) Ekspektasi bersyarat $E(Y \mid x)$.

4.3.6. Misal f dan F, berturut-turut menyatakan pdf dan cdf peubah acak X. Suatu pdf bersyarat dari X bila $X > x_0$, untuk x_0 tetap, didefinisikan sebagai

$$f(x \mid X > x_0) = \frac{f(x)}{1 - F(x_0)} I(x_0 < x).$$

Jenis pdf ini digunakan dalam permasalahan umur suatu komponen, bila daya tahan hidup sampai waktu x_0.

(a) Tunjukkan bahwa $f(x \mid X > x_0)$ adalah pdf.

(b) Misal $f(x) = e^{-x} I(0 < x < \infty)$. Hitung $P\left(X > 2 \mid X > 1\right)$.

4.3.7. Misal $f(x_1, x_2) = 21 x_1^2 x_2^3 I(0 < x_1 < x_2 < 1)$ adalah pdf bersama dari X_1 dan X_2. Tentukan:

(a) Purata dan varians bersyarat dari X_1, bila $X_2 = x_2, 0 < x_2 < 1$.

(b) Sebaran dari $Y = E(X_1 \mid X_2)$.

(c) $E(Y)$ dan $Var(Y)$ dan bandingkan dengan $E(X_1)$ dan $Var(X_1)$.

4.3.8. Misal X_1 dan X_2 peubah-peubah acak sehingga sebaran dan purata bersyaratnya ada. Tunjukkan bahwa: $E\left[u(X_2) \mid X_2\right] = u(X_2)$.

4.4 Koefisien Korelasi

Pembahasan subpokok bahasan ini kita awali dengan nilai ekspektasi tentang hubungan di antara dua peubah.

Definisi 4.4.1 Kovarians dari pasangan dua peubah acak X dan Y yang masing-masing puratanya μ_X dan μ_Y, ditulis $Cov(X,Y)$, adalah

$$Cov(X,Y) = E\left[\left(X-\mu_X\right)\left(Y-\mu_Y\right)\right].$$

Notasi lain dari kovarians ini adalah σ_{XY}. Menggunakan sifat-sifat ekspektasi, kovarians tersebut dapat dinyatakan sebagai

$$\begin{aligned}\sigma_{XY} &= E\left[\left(X-\mu_X\right)\left(Y-\mu_Y\right)\right] = E(XY-\mu_Y X-\mu_X Y+\mu_X\mu_Y)\\ &= E(XY)-\mu_Y E(X)-\mu_X E(Y)+\mu_X\mu_Y\\ &= E(XY)-\mu_Y\mu_X-\mu_X\mu_Y+\mu_X\mu_Y = E(XY)-\mu_X\mu_Y,\end{aligned}$$

atau

$$E(XY) = \sigma_{XY}+\mu_X\mu_Y. \qquad (4.4.1)$$

Beberapa sifat kovarians diberikan oleh teorema berikut ini.

Teorema 4.4.1 Misal X dan Y dua peubah acak sebarang, a dan b dua konstanta sebarang. Maka

(i) $Cov(X,Y) = E(XY)-E(X)E(Y)$.

(ii) $Cov(aX,bY) = abCov(X,Y)$.

(iii) $Cov(X+a,Y+b) = Cov(X,Y)$.

(iv) $Cov(X,aX+b) = aVar(X)$.

(v) $Var(aX+bY) = a^2Var(X)+b^2Var(Y)+2abCov(X,Y)$.

Bukti ditinggalkan sebagai latihan.

Sekarang kita bahas konsep tentang koefisien korelasi, yang diberikan oleh definisi berikut ini.

Definisi 4.4.2 Misal X dan Y dua peubah acak dengan varians masing-masing σ_X^2 dan σ_Y^2 dan kovarians σ_{XY}. Maka ***koefisien korelasi*** dari X dan Y, ditulis ρ, adalah

$$\rho = \frac{\sigma_{XY}}{\sigma_X\sigma_Y}. \qquad (4.4.2)$$

Dari persamaan (4.4.1) dan (4.4.2) diperoleh persamaan

$$E(XY) = \rho\sigma_X\sigma_Y + \mu_X\mu_Y. \quad (4.4.3)$$

Contoh 4.4.1 Misal dua peubah acak X dan Y mempunyai pdf bersama

$$f(x, y) = (x + y)I(0 < x < 1, 0 < y < 1).$$

Tentukan koefisien korelasinya.

Penyelesaian

Kita hitung terlebih dahulu μ_X seperti berikut ini

$$\mu_X = E(X) = \int_0^1\int_0^1 x(x + y)dxdy = \frac{7}{12},$$

dan $E(X^2)$ adalah

$$E(X^2) = \int_0^1\int_0^1 x^2(x + y)dxdy = \frac{60}{144}.$$

Sehingga

$$\sigma_X^2 = \frac{11}{144}.$$

Dengan cara serupa kita peroleh

$$\mu_Y = E(Y) = \frac{7}{12},\ \sigma_Y^2 = \frac{11}{144}.$$

Sedangkan kovariansnya adalah

$$\begin{aligned}\sigma_{XY} &= E(XY) - E(X)E(Y) \\ &= \int_0^1\int_0^1 xy(x + y)dxdy - \left(\frac{7}{12}\right)^2 \\ &= -\frac{1}{144}.\end{aligned}$$

Jadi koefisien dari X dan Y adalah

$$\rho = \frac{-\frac{1}{144}}{\sqrt{\frac{11}{144}\cdot\frac{11}{144}}} = -\frac{1}{11}.$$

Teorema 4.4.2 Misal (X,Y) mempunyai sebaran bersama dengan varians-varians dari X dan Y terhingga dan positif. Misal purata dan varians dari X dan Y berturut-turut dinyatakan sebagai μ_X, μ_Y dan σ_X^2, σ_Y^2, dan misal ρ menyatakan koefisien korelasi dari X dan Y. Jika $E(Y \mid X)$ linear dalam X, maka

$$E(Y \mid X) = \mu_Y + \rho \frac{\sigma_Y}{\sigma_X}\left(X - \mu_X\right)$$

dan

$$E\left(Var(Y \mid X)\right) = \sigma_Y^2(1-\rho^2).$$

Bukti

Kita buktikan untuk peubah acak kontinu. Untuk peubah acak diskret tinggal mengganti integral dengan jumlah. Misal

$$E(Y \mid x) = a + bx.$$

Karena

$$E(Y \mid x) = \frac{\int_{-\infty}^{\infty} yf(x,y)dy}{f_1(x)},$$

maka

$$\int_{-\infty}^{\infty} yf(x,y)dy = (a+bx) f_1(x). \tag{4.4.4}$$

Jika kedua ruas Persamaan (4.4.4) kita integralkan atas x, maka diperoleh

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} yf(x,y)\,dy\,dx = \int_{-\infty}^{\infty}(a+bx) f_1(x)\,dx.$$

$$\int_{-\infty}^{\infty} y\left[\int_{-\infty}^{\infty} f(x,y)\,dx\right]dy = a\int_{-\infty}^{\infty} f_1(x)\,dx + b\int_{-\infty}^{\infty} xf_1(x)\,dx.$$

$$E(Y) = a + bE(X),$$

atau

$$\mu_Y = a + b\mu_X. \tag{4.4.5}$$

Serupa dengan hasil sebelumnya, jika kedua ruas Persamaan (4.4.4) kita kalikan dengan x kemudian mengintegralkannya atas x, maka kita mempunyai

$$E(XY) = aE(X) + bE(X^2).$$
(4.4.6)

Substitusi Persamaan (4.4.3) ke ruas kiri Persamaan (4.4.6) dan substitusi persamaan

$$E(X^2) = \sigma_X^2 + \mu_X^2$$

diperoleh

$$\rho\sigma_X\sigma_Y + \mu_X\mu_Y = a\mu_X + b\left(\sigma_X^2 + \mu_X^2\right).$$
(4.4.7)

Solusi simultan Persamaan (4.4.5) dan (4.4.7) adalah

$$a = \mu_Y - \rho\frac{\sigma_Y}{\sigma_X}\mu_X \text{ dan } b = \rho\frac{\sigma_Y}{\sigma_X}.$$

Jadi

$$E(Y\,|\,x) = \mu_Y - \rho\frac{\sigma_Y}{\sigma_X}\mu_X + \left(\rho\frac{\sigma_Y}{\sigma_1}\right)x = \mu_Y + \rho\frac{\sigma_Y}{\sigma_X}(x - \mu_X).$$

Hasil ini merupakan nilai dari persamaan pertama Teorema 4.4.2. Untuk membuktikan persamaan yang kedua, kita perhatikan varians bersyarat berikut ini

$$Var(Y\,|\,x) = \int_{-\infty}^{\infty}\left[y - \mu_Y - \rho\frac{\sigma_Y}{\sigma_X}(x - \mu_X)\right]^2 f(y\,|\,x)\,dy$$

$$= \frac{\int_{-\infty}^{\infty}\left[y - \mu_Y - \rho\frac{\sigma_Y}{\sigma_X}(x - \mu_X)\right]^2 f(x, y)\,dy}{f_1(x)}.$$

Varians ini tidak negatif dan hanya merupakan fungsi dari x. Jika kita kalikan dengan $f_1(x)$ dan mengintegralkannya atas x, maka diperoleh hasil tidak negatif. Oleh karena itu

$$E\left(Var(Y\,|\,X)\right) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\left[y - \mu_Y - \rho\frac{\sigma_Y}{\sigma_X}(x - \mu_X)\right]^2 f(x, y)\,dydx = \sigma_Y^2(1 - \rho^2),$$

secara rinci ditinggalkan sebagai latihan. Terbukti.

Dari persamaan terakhir diperoleh $\sigma_Y^2(1 - \rho^2) \geq 0.$ Mengapa? Oleh karena itu

$$\rho^2 \le 1, \text{ atau } -1 \le \rho \le 1.$$

Teorema 4.4.2 berlaku untuk purata bersyarat dari Y bila $X = x$ linear. Ditinggalkan sebagai latihan untuk ekspektasi bersyarat tidak linear.

Contoh 4.4.2 Misal dua peubah acak X dan Y mempunyai purata bersyarat linear, yaitu, $E(Y \mid x) = 4x + 3$ dan $E(X \mid y) = \frac{1}{16}y - 3.$ Menggunakan rumus pertama pada Teorema 4.4.2 kita mempunyai $E(Y \mid x) = \mu_Y$, jika $x = \mu_X$ dan $E(X \mid y) = \mu_X$, jika $y = \mu_Y$. Oleh karena itu $\mu_Y = 4\mu_X + 3$ dan $\mu_X = \frac{1}{16}\mu_Y - 3,$ yang memberikan solusi $\mu_X = -\frac{15}{4}$ dan $\mu_Y = -12.$ Dari rumus $b = \rho\frac{\sigma_Y}{\sigma_X},$ maka perkalian dan perbandingan koefisien kedua purata bersyarat berturut-turut adalah ρ^2 dan $\frac{\sigma_Y^2}{\sigma_X^2}$. Jadi $\rho^2 = 4.\frac{1}{16} = \frac{1}{4}$ dengan $\rho = \frac{1}{2}$ (bukan $-\frac{1}{2}$) dan $\frac{\sigma_Y^2}{\sigma_X^2} = 64.$

Kita akhiri subpokok bahasan ini dengan pembahasan varians dan kovarians dari vektor acak.

Misal $\mathbf{X} = \left(X_1, X_2, ..., X_n\right)^t$ adalah vektor acak berdimensi n. Kita definisikan

$$E(\mathbf{X}) = \left(E(X_1), E(X_2), ..., E(X_n)\right)^t,$$

yaitu, ekspektasi dari vektor acak adalah juga merupakan vektor acak ekspektasi komponen-komponennya. Sekarang anggap $\mathbf{W}$ adalah matriks $m \times n$ dari peubah-peubah acak, katakan, $\mathbf{W} = \left[W_{ij}\right]$ dari peubah acak W_{ij}, $1 \le i \le m$ dan $1 \le j \le n.$ Matriks yang demikian disebut sebagai ***matriks acak.*** Selanjutnya kita definisikan juga ekspektasi dari matriks acak sebagai

$$E(\mathbf{W}) = \left[E\left(W_{ij}\right)\right].$$

Seperti halnya pada pembahasan sebelumnya, pada matriks acak ini E juga merupakan operator linear. Hal ini ditunjukkan oleh teorema berikut.

Teorema 4.4.3 Misal $\mathbf{W}_1$ dan $\mathbf{W}_2$ adalah matriks-matriks acak $m\times n,$ $\mathbf{A}_1$ dan $\mathbf{A}_2$ adalah matriks-matriks konstanta $k\times m,$ dan $\mathbf{B}$ matriks konstanta $n\times l.$ Maka

$$E\left[\mathbf{A}_1\mathbf{W}_1+\mathbf{A}_2\mathbf{W}_2\right]=\mathbf{A}_1E\left(\mathbf{W}_1\right)+\mathbf{A}_2E\left(\mathbf{W}_2\right)$$

$$E\left[\mathbf{A}_1\mathbf{W}_1\mathbf{B}\right]=\mathbf{A}_1E\left(\mathbf{W}_1\right)\mathbf{B}.$$

Bukti

Kita buktikan persamaan pertama

$$E\left[\mathbf{A}_1\mathbf{W}_1+\mathbf{A}_2\mathbf{W}_2\right]=E\left[\sum_{s=1}^{m}a_{1is}W_{1sj}+\sum_{s=1}^{m}a_{2is}W_{2sj}\right]$$

$$=\left[\sum_{s=1}^{m}a_{1is}E\left[W_{1sj}\right]+\sum_{s=1}^{m}a_{2is}E\left[W_{2sj}\right]\right]=\mathbf{A}_1E\left(\mathbf{W}_1\right)+\mathbf{A}_2E\left(\mathbf{W}_2\right).$$

Ditinggalkan sebagai latihan bukti dari persamaan kedua.

Misal $\mathbf{X}=\left(X_1,X_2,...,X_n\right)^t$ adalah vektor acak berdimensi $n,$ sehingga $Var(X_i)=\sigma_i^2<\infty.$ Purata dari $\mathbf{X}$ adalah $\boldsymbol{\mu}=E(\mathbf{X})$ dan kita definisikan matriks kovarians sebagai

$$Cov(\mathbf{X})=E\left[\left(\mathbf{X}-\boldsymbol{\mu}\right)\left(\mathbf{X}-\boldsymbol{\mu}\right)^t\right]=\left[\sigma_{ij}\right],$$

di mana $\sigma_{ii}=\sigma_i^2.$ Ditinggalkan sebagai latihan, untuk menunjukkan unsur diagonal ke i dari $Cov(\mathbf{X})$ adalah $\sigma_i^2=Var(X_i)$ dan unsur ke (i,j) adalah $Cov(X_i,X_j).$

Contoh 4.4.3 Misal peubah-peubah acak kontinu X dan Y mempunyai pdf bersama

$$f(x,y)=e^{-y}I(0<x<y<\infty).$$

Maka

$$\mu_X=1,\mu_Y=2,\ \sigma_X^2=1,\sigma_Y^2=2,\text{ dan } E\left[\left(X-\mu_X\right)\left(Y-\mu_Y\right)\right]=1,$$

lihat Latihan 4.4.2. Misal $\mathbf{Z}=(X,Y)^t$. Maka

$$E(\mathbf{Z})=\begin{bmatrix}1\\2\end{bmatrix} \text{ dan } Cov(\mathbf{Z})=\begin{bmatrix}1 & 1\\1 & 2\end{bmatrix}.$$

Teorema 4.4.4 Misal $\mathbf{X}=\left(X_1, X_2, ..., X_n\right)^t$ adalah vektor acak berdimensi n, sehingga $Var(X_i)=\sigma_i^2=\sigma_{ii}<\infty.$ Misal $\mathbf{A}$ matriks konstanta $m\times n.$ Maka

$$Cov(\mathbf{X})=E\left[\mathbf{X}\mathbf{X}^t\right]-\boldsymbol{\mu}\boldsymbol{\mu}^t,$$

$$Cov\left(\mathbf{A}\mathbf{X}\right)=\mathbf{A}Cov(\mathbf{X})\mathbf{A}^t.$$

Bukti

Menggunakan Teorema 4.4.3 kita peroleh

$$\begin{aligned}Cov(\mathbf{X}) &= E\left[\left(\mathbf{X}-\boldsymbol{\mu}\right)\left(\mathbf{X}-\boldsymbol{\mu}\right)^t\right]\\ &= E\left[\mathbf{X}\mathbf{X}^t-\boldsymbol{\mu}\mathbf{X}^t-\mathbf{X}\boldsymbol{\mu}^t+\boldsymbol{\mu}\boldsymbol{\mu}^t\right]\\ &= E\left[\mathbf{X}\mathbf{X}^t\right]-\boldsymbol{\mu}E\left[\mathbf{X}^t\right]-E\left[\mathbf{X}\right]\boldsymbol{\mu}^t+\boldsymbol{\mu}\boldsymbol{\mu}^t\\ &= E\left[\mathbf{X}\mathbf{X}^t\right]-\boldsymbol{\mu}\boldsymbol{\mu}^t,\end{aligned}$$

terbukti bagian pertama. Ditinggalkan sebagai latihan untuk bagian kedua.

Semua matriks kovarians adalah matriks semi definit positif, yaitu, $\mathbf{a}^tCov(\mathbf{X})\mathbf{a}\geq 0,$ untuk semua vektor $\mathbf{a}\in R^n$. Untuk ini, misal $\mathbf{X}$ adalah vektor acak dan misal $\mathbf{a}$ sebarang vektor konstanta $n\times 1$. Maka $Y=\mathbf{a}^t\mathbf{X}$ adalah peubah acak. Oleh karena itu

$$0\leq Var(Y)=Var(\mathbf{a}^t\mathbf{X})=\mathbf{a}^tCov(\mathbf{X})\mathbf{a}.$$

Jadi $Cov(\mathbf{X})$ adalah semi definit positif.

Latihan 4.4

4.4.1. Buktikan Teorema 4.4.1.

4.4.2. Misal peubah-peubah acak kontinu X dan Y mempunyai pdf bersama

$$f(x,y) = e^{-y} I(0 < x < y < \infty).$$

Buktikan bahwa

$$\mu_X = 1, \mu_Y = 2, \ \sigma_X^2 = 1, \sigma_Y^2 = 2, \text{ dan } E\left[\left(X - \mu_X\right)\left(Y - \mu_Y\right)\right] = 1.$$

4.4.3. Misal dua peubah acak X dan Y mempunyai pdf bersama seperti pada tabel berikut:

(x,y)	$(1,1)$	$(1,2)$	$(1,3)$	$(2,1)$	$(2,2)$	$(2,3)$
$f(x,y)$	$\frac{2}{15}$	$\frac{4}{15}$	$\frac{3}{15}$	$\frac{1}{15}$	$\frac{1}{15}$	$\frac{4}{15}$

Hitung koefisien korelasi dari X dan Y

4.4.4. Misal dua peubah acak X dan Y mempunyai pdf bersama

$$f(x,y) = \frac{1}{20} I(0 < x < 10, x-1 < y < x+1).$$

Hitung koefisien korelasinya.

4.4.5. Buktikan persamaan

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\left[y - \mu_Y - \rho\frac{\sigma_Y}{\sigma_X}(x - \mu_1)\right]^2 f(x,y)\,dydx = \sigma_Y^2(1-\rho^2),$$

Pada bukti Teorema 4.4.2.

4.4.6. Misal $\sigma_1^2 = \sigma_2^2 = \sigma^2$ adalah varians bersama dari X_1 dan X_2. Tunjukkan bahwa

$$P\left[\left|\left(X_1 - \mu_1\right) + \left(X_2 - \mu_2\right)\right| \geq k\sigma\right] \leq \frac{2(1+\rho)}{k^2}.$$

4.4.7. Jika koefisien korelasi ρ dari X dan Y ada, tunjukkan bahwa $-1 \leq \rho \leq 1$.

Petunjuk: Perhatikan diskriminan fungsi kuadrat tidak negatif

$$h(v) = E\left\{\left[\left(X - \mu_X\right) + v\left(Y - \mu_Y\right)\right]^2\right\},$$

di mana v adalah bilangan real tidak bergantung pada X dan tidak pada Y.

4.4.8. Buktikan Teorema 4.4.3, bagian kedua.

4.4.9. Tunjukkan bahwa unsur diagonal ke i dari $Cov(\mathbf{X})$ adalah $\sigma_i^2 = Var(X_i)$ dan unsur ke (i, j) adalah $Cov(X_i, X_j)$.

4.4.10. Buktikan Teorema 4.4.4, bagian kedua.

4.5 Kebebasan Peubah-peubah Acak

Misal X_1 dan X_2 sebarang dua peubah acak dengan pdf bersama $f(x_1, x_2)$ dan pdf marginalnya berturut-turut $f_1(x_1)$ dan $f_2(x_2)$. Oleh karena itu jika $f(x_2 | x_1)$ pdf bersyarat dari X_2 bila $X_1 = x_1$, maka pdf bersama dari X_1 dan X_2 dapat ditulis sebagai

$$f(x_1, x_2) = f(x_2 | x_1) f_1(x_1).$$

Misal $f(x_2 | x_1)$ tidak bergantung pada x_1. Maka dalam hal X_2 peubah acak kontinu pdf marginalnya dapat kita peroleh seperti berikut ini:

$$f_2(x_2) = \int_{-\infty}^{\infty} f(x_1, x_2) dx_1 = \int_{-\infty}^{\infty} f(x_2 | x_1) f_1(x_1) dx_1$$

$$= f(x_2 | x_1) \int_{-\infty}^{\infty} f_1(x_1) dx_1 = f(x_2 | x_1).$$

Jadi

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

Dengan cara serupa untuk peubah acak diskret, yaitu dengan mengganti integral menjadi jumlah. Hasil di atas menjadi dasar untuk definisi berikut ini.

Definisi 4.5.1 Misal X_1 dan X_2 dua peubah acak dengan pdf bersama $f(x_1, x_2)$ dan pdf marginalnya berturut-turut adalah $f_1(x_1)$ dan $f_2(x_2)$. Maka X_1 dan X_2 dikatakan bebas jika

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

Dan jika lain dikatakan tidak bebas atau bergantung.

Contoh 4.5.1 Misal pdf bersama dari dua peubah acak X_1 dan X_2 adalah

$$f(x_1, x_2) = I(0 < x_1 < 1, 0 < x_2 < 1).$$

Maka

$$f_1(x_1) = \int_{-\infty}^{\infty} f(x_1, x_2) dx_2 = \int_0^1 dx_2 = I(0 < x_1 < 1),$$

dan

$$f_2(x_2) = \int_{-\infty}^{\infty} f(x_1, x_2) dx_1 = \int_0^1 dx_1 = I(0 < x_2 < 1).$$

Oleh karena itu kita peroleh

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

Jadi X_1 dan X_2 bebas.

Contoh 4.5.2 Misal pdf bersama dari dua peubah acak X_1 dan X_2 adalah

$$f(x_1, x_2) = (x_1 + x_2) I(0 < x_1 < 1, 0 < x_2 < 1).$$

Kita akan menunjukkan bahwa X_1 dan X_2 tidak bebas. Kita tentukan terlebih dahulu pdf marginal dari X_1 dan X_2 berturut-turut seperti berikut ini:

$$f_1(x_1) = \int_0^1 \left(x_1 + x_2\right) dx_2 = \left(x_1 + \frac{1}{2}\right) I(0 < x_1 < 1),$$

dan

$$f_2(x_2) = \int_0^1 \left(x_1 + x_2\right) dx_1 = \left(\frac{1}{2} + x_2\right) I(0 < x_2 < 1).$$

Oleh karena itu

$$f(x_1, x_2) \neq f_1(x_1) f_1(x_2).$$

Jadi X_1 dan X_2 tidak bebas.

Berdasarkan Definisi 4.5.1 untuk menentukan bebas tidaknya dua peubah acak, kita harus menentukan terlebih dahulu masing-masing pdf marginal dari X_1 dan X_2. Teorema berikut ini memungkinkan kita untuk menentukan bebas tidaknya X_1 dan X_2 dengan cara yang lebih praktis.

Teorema 4.5.1 Misal dua peubah acak X_1 dan X_2 mempunyai pdf bersama $f(x_1, x_2)$ dengan ruang peubah acak berturut-turut $\mathfrak{A}_1$ dan $\mathfrak{A}_2$. Maka X_1 dan X_2 dikatakan bebas jika dan hanya jika $f(x_1, x_2)$ dapat ditulis sebagai hasil kali fungsi tidak negatif yang hanya bergantung pada x_1 dan fungsi tidak negatif yang hanya bergantung pada x_2, yaitu,

$$f(x_1, x_2) = g(x_1)h(x_2),$$

di mana $g(x_1) > 0,\ x_1 \in \mathfrak{A}_1,\ h(x_2) > 0,\ x_2 \in \mathfrak{A}_2$.

Bukti

Jika X_1 dan X_2 bebas, maka berdasarkan Definisi 4.5.1

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

di mana $f_1(x_1)$ dan $f_2(x_2)$ berturut-turut pdf marginal dari X_1 dan X_2. Oleh karena itu jika kita ambil $g(x_1) = f_1(x_1)$ dan $h(x_2) = f_2(x_2)$.

Selanjunya kita buktikan konversnya, yaitu jika diketahui

$$f(x_1, x_2) = g(x_1)h(x_2).$$

Untuk peubah acak kontinu, kita punya

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

$$= g(x_1)\int_{-\infty}^{\infty} h(x_2)dx_2 = g(x_1)c_1,$$

di mana

$$c_1 = \int_{-\infty}^{\infty} h(x_2)dx_2.$$

Dengan cara serupa kita peroleh

$$f_2(x_2) = c_2 h(x_1),$$

di mana

$$c_2 = \int_{-\infty}^{\infty} g(x_1)dx_1.$$

Karena

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x_1,x_2)dx_1dx_2 = 1,$$

maka

$$1 = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} g(x_1)h(x_2)dx_1dx_2 = \left[\int_{-\infty}^{\infty} g(x_1)dx_1\right]\left[\int_{-\infty}^{\infty} h(x_2)dx_2\right] = c_2c_1.$$

Jadi

$$f(x_1,x_2) = g(x_1)h(x_2) = c_1g(x_1)c_2h(x_2) = f_1(x_1)f_2(x_2).$$

Untuk peubah acak diskret, tinggal mengganti integral menjadi jumlah, bukti selesai.

Contoh 4.5.3 Misal pdf bersama dari dua peubah acak kontinu X_1 dan X_2 adalah

$$f(x_1,x_2) = 12x_1x_2(1-x_2)I(0 < x_1 < 1, 0 < x_2 < 1).$$

Kita pilih

$$g(x_1) = 12x_1I(0 < x_1 < 1),$$

dan

$$h(x_2) = x_2(1-x_2)I(0 < x_2 < 1).$$

Maka

$$f(x_1,x_2) = g(x_1)h(x_2).$$

Jadi X_1 dan X_2 bebas.

Jika Contoh 4.5.2 kita selesaikan menggunakan Teorema 1, maka kita akan dengan mudah dapat melihat bahwa X_1 dan X_2 tidak bebas, karena

$$f(x_1,x_2) = \left(x_1 + x_2\right)I(0 < x_1 < 1, 0 < x_2 < 1),$$

tidak dapat dijadikan sebagai hasil kali $g(x_1)$ dan $h(x_2)$.

Contoh 4.5.4 Misal pdf bersama dari dua peubah acak X dan Y adalah

$$f(x,y) = 8xyI(0 < x < y < 1).$$

Jika kita lihat dari rumus $8xy$, tentu kita menduga bahwa X dan Y bebas. Meskipun demikian jika kita lihat dari daerah definisinya, kita melihat bahwa ruang dua peubah acak X dan Y, yaitu

$$\mathfrak{A} = \left\{(x,y): 0 < x < y < 1\right\}$$

tidak dapat dijadikan sebagai hasil kali dari ruang X, yaitu $\mathfrak{A}_1$, dan ruang Y, yaitu $\mathfrak{A}_2$. Jadi X dan Y tidak bebas.

Teorema berikut ini sering digunakan dalam menghitung peluang jika diketahui peubah-peubah acaknya bebas.

Teorema 4.5.2 Misal X_1 dan X_2 dua peubah acak yang bebas. Maka

$$P\left(a < X_1 < b, c < X_2 < d\right) = P(a < X_1 < b)P(c < X_2 < d),$$

untuk setiap $a < b$ dan $c < d$, di mana a, b, c, dan d suatu konstanta.

Bukti

Karena X_1 dan X_2 dua peubah acak bebas, maka

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

Oleh karena itu

$$\begin{aligned} P\left(a < X_1 < b, c < X_2 < d\right) &= \int_a^b \int_c^d f(x_1, x_2) dx_2 dx_1 \\ &= \int_a^b \int_c^d f_1(x_1) f_2(x_2) dx_2 dx_1 \\ &= \int_a^b f_1(x_1) dx_1 \int_c^d f_2(x_2) dx_2 \\ &= P(a < X_1 < b)P(c < X_2 < d), \end{aligned}$$

dalam hal X_1 dan X_2 kontinu. Jika X_1 dan X_2 diskret tinggal mengganti integral menjadi jumlah.

Kebebasan dua peubah acak berakibat pula pada perkalian ekspektasinya.

Teorema 4.5.3 Misal dua peubah acak X_1 dan X_2 bebas, dan misal $u(X_1)$ dan $v(X_2)$ berturut-turut adalah fungsi dari X_1 dan X_2. Maka

$$E\left[u(X_1)v(X_2)\right] = E\left[u(X_1)\right]E\left[v(X_2)\right].$$

Bukti

Kita buktikan untuk peubah acak kontinu, sedangkan untuk peubah acak diskret, tinggal mengganti integral dengan jumlah. Menggunakan definisi kebebasan dari peubah acak, yaitu $f(x_1, x_2) = f_1(x_1)f_1(x_2)$, dan definisi ekspektasi, kita peroleh

$$\begin{aligned} E\left[u(X_1)v(X_2)\right] &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} u(x_1)v(x_2)f(x_1, x_2)dx_1dx_2 \\ &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} u(x_1)v(x_2)f_1(x_1)f_2(x_2)dx_1dx_2 \\ &= \left[\int_{-\infty}^{\infty} u(x_1)f_1(x_1)dx_1\right]\left[\int_{-\infty}^{\infty} v(x_2)f_2(x_2)dx_2\right] \\ &= E\left[u(X_1)\right]E\left[v(X_2)\right], \end{aligned}$$

bukti selesai.

Contoh 4.5.5 Misal X dan Y peubah-peubah acak bebas. Maka

$$\begin{aligned} Cov(X,Y) &= E(XY) - E(X)E(Y) \\ &= E(X)E(Y) - E(X)E(Y) = 0. \end{aligned}$$

Lebih lanjut jika varians masing-masing peubah acak tersebut positif, maka $\rho_{XY} = 0.$

Contoh 4.5.5 tidak berlaku sebaliknya, hal ini dapat kita lihat pada Contoh 4.5.6 berikut ini.

Contoh 4.5.6 Misal X dan Y peubah-peubah acak dengan pdf bersama

$$f(x, y) = \frac{1}{3} I\left[(-1,1),(0,0),(1,1)\right].$$

Maka kita peroleh:

$$\mu_X = \sum_y \sum_x xf(x, y) = (-1).\frac{1}{3} + 0.\frac{1}{3} + 1.\frac{1}{3} = 0,$$

$$\mu_Y = \sum_y \sum_x y f(x,y) = (1).\frac{1}{3} + 0.\frac{1}{3} + 1.\frac{1}{3} = \frac{2}{3},$$

$$E(XY) = \sum_y \sum_x xy f(x,y) = (-1).1.\frac{1}{3} + 0.0.\frac{1}{3} + 1.1.\frac{1}{3} = 0,$$

$$\sigma_X^2 = E(X^2) - \mu_X^2 = \sum_y \sum_x x^2 f(x,y) - 0 = 1.\frac{1}{3} + 0.\frac{1}{3} + 1.\frac{1}{3} = \frac{2}{3},$$

$$\sigma_Y^2 = E(Y^2) - \mu_Y^2 = \sum_y \sum_x y^2 f(x,y) - \left(\frac{2}{3}\right)^2$$

$$= \left(1.\frac{1}{3} + 0.\frac{1}{3} + 1.\frac{1}{3}\right) - \left(\frac{2}{3}\right)^2 = \frac{2}{3} - \left(\frac{2}{3}\right)^2 = \frac{2}{9}.$$

Jadi

$$\rho = \frac{0 - 0.\frac{2}{3}}{\sqrt{\frac{2}{3}}.\sqrt{\frac{2}{9}}} = 0.$$

Ruang X dan Y, yaitu, $\mathfrak{A} = \{(x,y) : (x,y) = (-1,1), (0,0), (1,1)\}$ tidak dapat ditulis sebagai hasil kali dari $\mathfrak{A}_X$ dan $\mathfrak{A}_Y$, berturut-turut merupakan ruang dari X dan ruang dari Y. Oleh karena itu X dan Y tidak bebas. Jadi $\rho = 0$ tidak berakibat X dan Y bebas, atau

$$P(X = 0, Y = 0) \neq P(X = 0)P(Y = 0).$$

Teorema 4.5.4 Misal dua peubah acak X_1 dan X_2 bebas. Maka

$$M(t_1, t_2) = M(t_1, 0)M(0, t_2).$$

Bukti

$$M(t_1,t_2) = E\left[e^{t_1X_1 + t_2X_2}\right] = E\left[e^{t_1X_1}e^{t_2X_2}\right] = E\left[e^{t_1X_1}\right]E\left[e^{t_2X_2}\right] = M(t_1,0)M(0,t_2).$$

Kebebasan pada dua peubah acak dapat kita perumum untuk n peubah acak.

Definisi 4.5.2 Misal peubah-peubah acak $X_1, X_2, ..., X_n$ mempunyai pdf bersama $f(x_1, x_2, ..., x_n)$ dan pdf marginalnya berturut-turut adalah $f_1(x_1), f_2(x_2),...., f_n(x_n).$ Maka $X_1, X_2, ..., X_n$ dikatakan bebas jika

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

Sifat-sifat kebebasan pada dua peubah acak dapat pula diperumum untuk n peubah acak. Dalam hal peubah-peubah acak pada definisi 4.5.2 sebarannya identik, maka peubah-peubah acak $X_1, X_2, ..., X_n$ dikatakan sebagai ***sampel acak***.

Definisi 4.5.3 Misal $X_1, X_2, ..., X_n$ menyatakan n peubah acak bebas, masing-masing pdf-nya sama, $f(x),$ yaitu pdf dari $X_1, X_2, ..., X_n$ berturut-turut $f(x_1), f(x_2),...., f(x_n),$ sehingga pdf bersamanya adalah

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

Maka peubah-peubah acak $X_1, X_2, ..., X_n$ dikatakan sampel acak.

Contoh 4.5.7 Misal $X_1, X_2, ..., X_n$ sampel acak dari suatu populasi dengan pdf

$$f(x) = 2xI(0 < x < 1).$$

Maka pdf bersama dari $X_1, X_2, ..., X_n$ adalah

$$f(x_1, ..., x_n) = 2x_1 ... 2x_n = 2^n x_1 ... x_n I(0 < x_i < 1, i = 1, ..., n).$$

Contoh 4.5.8 Anggap waktu hidup bola lampu tertentu dari populasi yang bersebaran eksponensial dengan pdf

$$f(x) = e^{-x} I(0 < x < \infty),$$

diukur dalam tahun. Misalkan sampel acak berukuran 2 diambil dari populasi ini. Maka kita akan memperoleh pdf bersama dari $X_1, X_2,$

$$f(x_1, x_2) = f(x_1) f(x_2) = e^{-x_1} e^{-x_2} = e^{-(x_1 + x_2)} I(0 < x_i < \infty, i = 1, 2).$$

Misalkan kita ingin mengetahui peluang bahwa total waktu hidup kedua lampu kurang dari $0.5.$ Maka kita peroleh:

$$P\left[X_1+X_2\le 0.5\right]=\int_0^{0.5}\int_0^{0.5-x_2} e^{-(x_1+x_2)}dx_1dx_2=\int_0^{0.5}\left[-e^{-(x_1+x_2)}\right]_0^{0.5-x_2}dx_2$$
$$=\int_0^{0.5}\left[e^{-x_2}-e^{-0.5}\right]dx_2=\left[-e^{-x_2}-x_2e^{-0.5}\right]_0^{0.5}$$
$$=\left(-e^{-0,5}-0.5e^{-0.5}\right)-\left(-1-0\right)=1-0.5e^{-0.5}-e^{-0.5}\approx 0.09.$$

Latihan 4.5

4.5.1. Misal dua peubah acak X dan Y bebas sehingga $E(X)=2,$ $E(Y)=3,$ $Var(X)=4,$ dan $Var(Y)=16.$ Tentukan:
(a) $E(5X-Y).$
(b) $Var(5X-Y).$
(c) $Cov(3X+Y,Y).$
(d) $Cov(X,5X-Y).$

4.5.2. Misal dua peubah acak X dan Y mempunyai pdf bersama
$$f(x,y)=2e^{-x-y}I(0<x<y<\infty).$$
Selidiki apakah X dan Y bebas?

4.5.3. Misal dua peubah acak X dan Y mempunyai pdf bersama
$$f(x,y)=\frac{1}{16}I(x=1,2,3,4,y=1,2,3,4).$$
Selidiki apakah X dan Y bebas?

4.5.4. Misal dua peubah acak X dan Y mempunyai pdf bersama
$$f(x,y)=4x(1-y)I(0<x<1,0<y<1).$$
Hitung $P\left[0<X<\frac{1}{3},0<Y<\frac{1}{2}\right].$

4.5.5. Hitung peluang gabungan kejadian $a<X<b,-\infty<Y<\infty,$ dan $-\infty<X<\infty,c<Y<d,$ jika X dan Y bebas dengan $P(a<X<b)=\frac{2}{3},$ dan $P(c<Y<d)=\frac{5}{8}.$

4.5.6. Misal dua peubah acak X dan Y mempunyai pdf bersama

$$f(x, y) = e^{-x-y} I(0 < x < \infty, 0 < y < \infty).$$

Tunjukkan bahwa:

(a) X dan Y bebas.

(b) $E\left[e^{t(X+Y)}\right] = (1-t)^{-2}, t < 1.$

4.5.7. Misal $X_1, ..., X_n$ sampel acak dari suatu populasi dengan pdf

$$f(x) = 3x^2 I(0 < x < 1).$$

(a) Tentukan pdf bersama dari $X_1, X_2, ..., X_n$.

(b) Hitung $P\left[X_1 < 0.3\right]$.

(c) Hitung peluang bahwa semua pengamatan kurang dari 0.3.

(d) Paling sedikit dua pengamataan kurang dari 0.3.

4.6 Transformasi Peubah Acak

Misal X_1 dan X_2 adalah dua peubah acak sebarang. Anggap kita mengetahui sebaran gabungan dari X_1 dan X_2 dan kita ingin menentukan sebaran dari transformasi X_1 dan X_2, katakanlah, $Y = g(X_1, X_2)$. Kita mungkin dapat memperoleh cdf dari Y.

Misal $f_{X_1,X_2}(x_1, x_2)$ pdf bersama dari dua peubah acak diskret X_1 dan X_2 dengan $\mathfrak{A}$ himpunan titik-titik dalam bidang x_1x_2 sehingga $f_{X_1,X_2}(x_1, x_2) > 0$. Misal $y_1 = u_1(x_1, x_2)$ dan $y_2 = u_2(x_1, x_2)$ mendefinisikan transformasi satu-satu dan pada dari $\mathfrak{A}$ ke $\mathfrak{M}$, di mana $\mathfrak{M}$ menyatakan himpunan dalam bidang y_1y_2 sehingga $x_1 = w_1(y_1, y_2)$ dan $x_2 = w_2(y_1, y_2)$ transformasi satu-satu dan pada dari $\mathfrak{M}$ ke $\mathfrak{A}$. Maka pdf bersama dua peubah acak baru $Y_1 = u_1(X_1, X_2)$ dan $Y_2 = u_2(X_1, X_2)$ adalah

$$f_{Y_1,Y_2}(y_1, y_2) = f_{X_1,X_2}\left[w_1(y_1, y_2), w_2(y_1, y_2)\right] |J| I\left[(y_1, y_2) \in \mathfrak{M}\right].$$

Contoh 4.6.1 Misal X_1 dan X_2 mempunyai pdf bersama

$$f_{X_1,X_2}(x_1,x_2)=\frac{\mu_1^{x_1}\mu_2^{x_2}e^{-\mu_1}e^{-\mu_2}}{x_1!x_2!}I\left(x_1=0,1,2,...,\ x_2=0,1,2,....\right),$$

di mana μ_1 dan μ_2 bilangan-bilangan real positif tertentu. Maka ruang $\mathfrak{A}$ adalah himpunan titik (x_1,x_2). Misal kita ingin menentukan pdf dari peubah acak diskret $Y=X_1+X_2$. Jika kita menggunakan teknik transformasi peubah acak, maka kita memerlukan peubah acak Y_2. Karena Y_2 tidak tersedia pada permasalahan kita, maka kita pilih dia yang sederhana sehingga kita mempunyai transformasi satu-satu dan pada, yaitu kita pilih $Y_2=X_2$. Maka $y_1=x_1+x_2$ dan $y_2=x_2$ menyatakan trasformasi satu-satu dan pada dari $\mathfrak{A}$ ke $\mathfrak{M}$, di mana

$$\mathfrak{M}=\left\{(y_1,y_2):y_2=0,1,2,...,y_1,\ y_1=0,1,2,...\right\}.$$

Fungsi kebalikannya adalah $x_1=y_1-y_2$ dan $x_2=y_2$. Oleh karena itu pdf bersama dari dua peubah acak diskret Y_1 dan Y_2 adalah

$$f_{Y_1,Y_2}(y_1,y_2)=\frac{\mu_1^{y_1-y_2}\mu_2^{y_2}e^{-\mu_1}e^{-\mu_2}}{(y_1-y_2)!y_2!}I\left[(y_1,y_2)\in\mathfrak{M}\right].$$

Lebih lanjut kita dapat menentukan pdf marginalnya, misal pdf marginal dari Y_1 adalah

$$f_{Y_1}(y_1)=\sum_{y_2=0}^{y_1}f_{Y_1,Y_2}(y_1,y_2)=\frac{e^{-\mu_1}e^{-\mu_2}}{y_1!}\sum_{y_2=0}^{y_1}\frac{y_1!}{(y_1-y_2)!y_2!}\mu_1^{y_1-y_2}\mu_2^{y_2}$$

$$=\frac{(\mu_1+\mu_2)^{y_1}e^{-(\mu_1+\mu_2)}}{y_1!}I(y_1=0,1,2,...),$$

yang merupakan pdf dari peubah acak bersebaran Poisson dengan purata $\mu=\mu_1+\mu_2$, atau $Y_1 \sim POI(\mu_1+\mu_2)$. Oleh karena itu mgf dari adalah

$$M_{Y_1}(t)=e^{(u_1+\mu_2)(e^t-1)}.$$

Untuk peubah acak kontinu kita mulai dengan contoh berikut.

Contoh 4.6.2 Misal suatu percobaan memilih secara acak satu titik (X,Y) dari persegi satuan $\mathfrak{A}=\{(x,y):0<x<1,0<y<1\}$. Maka pdf bersama dari X dan Y adalah

$$f_{X,Y}(x,y)=I(0<x<1,0<y<1).$$

Misal kita ingin menentukan pdf dari $Z=X+Y$. Pertama-tama kita tentukan terlebih dahulu cdf dari Z seperti berikut

$$\begin{aligned}F_Z(z)&=P\left(X+Y\le z\right)\\&=\int_0^z\int_0^{z-x}dydx\,I(0\le z<1)+\left(1-\int_{z-1}^1\int_{z-x}^1 dydx\right)I(1\le z<2)+I(2\le z)\\&=\frac{z^2}{2}I(0\le z<1)+\left(1-\frac{(2-z)^2}{2}\right)I(1\le z<2)+I(2\le z).\end{aligned}$$

Karena $f_Z(z)=F'_Z(z)$ untuk semua z sehingga $f_Z(z)$ kontinu, maka

$$f_Z(z)=zI(0<z<1)+(2-z)I(1<z<2).$$

Sekarang kita bahas secara umum teknik transformasi untuk peubah acak kontinu. Misal $f_{X_1,X_2}(x_1,x_2)$ pdf bersama dari dua peubah acak kontinu X_1 dan X_2 dengan $\mathfrak{A}$ himpunan titik-titik dalam bidang x_1x_2 sehingga $f_{X_1,X_2}(x_1,x_2)>0$. Misal $y_1=u_1(x_1,x_2)$ dan $y_2=u_2(x_1,x_2)$ mendefinisikan transformasi satu-satu dan pada dari $\mathfrak{A}$ ke $\mathfrak{M}$, di mana $\mathfrak{M}$ menyatakan himpunan dalam bidang y_1y_2 sehingga $x_1=w_1(y_1,y_2)$ dan $x_2=w_2(y_1,y_2)$ transformasi satu-satu dari $\mathfrak{M}$ pada $\mathfrak{A}$. Determinan orde 2,

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

disebut ***Jacobi*** dari transformasi ini. Anggap turunan parsial pertama kontinu dan J tidak nol dalam $\mathfrak{M}$. Misal A himpunan bagian dari $\mathfrak{A}$, dan misal B adalah peta dari A di bawah transformasi satu-satu dan pada. Maka $\{(X_1,X_2)\in A\}$ ekivalen dengan $\{(Y_1,Y_2)\in B\}$. Oleh karena itu

$$P\left[\left(Y_1, Y_2\right) \in B\right] = P\left[\left(X_1, X_2\right) \in A\right] = \iint_A f_{X_1,X_2}(x_1, x_2) dx_1 dx_2.$$

Menggunakan sifat dalam analisis

$$\iint_A f_{X_1,X_2}(x_1, x_2) dx_1 dx_2 = \iint_B f_{X_1,X_2}\left[w_1(y_1, y_2), w_2(y_1, y_2)\right] | J | dy_1 dy_2$$

kita peroleh

$$P\left[\left(Y_1, Y_2\right) \in B\right] = \iint_B f_{X_1,X_2}\left[w_1(y_1, y_2), w_2(y_1, y_2)\right] | J | dy_1 dy_2,$$

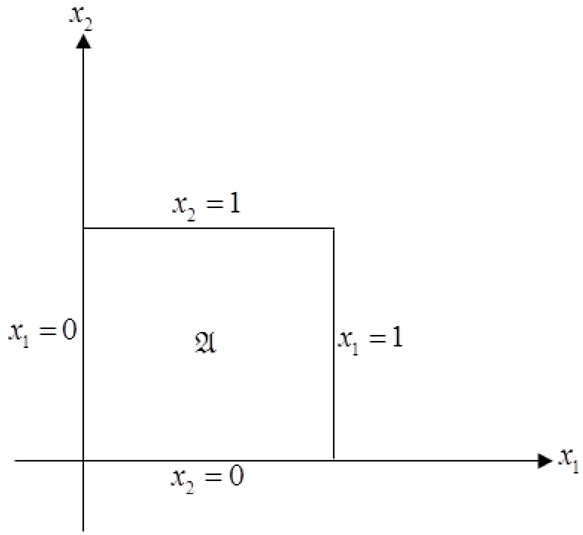
yang berakibat bahwa pdf bersama dari Y_1 dan Y_2 adalah

$$f_{Y_1,Y_2}(y_1, y_2) = f_{X_1,X_2}\left[w_1(y_1, y_2), w_2(y_1, y_2)\right] | J | I\left[(y_1, y_2) \in \mathfrak{M}\right].$$

Contoh 4.6.3 Misal dua peubah acak X_1 dan X_2 mempunyai pdf bersama

$$f_{X_1,X_2}(x_1, x_2) = I(0 < x_1 < 1, 0 < x_2 < 1).$$

Ruang dari dua peubah acak ini adalah
$\mathfrak{A} = \left\{\left(x_1, x_2\right) : 0 < x_1 < 1, 0 < x_2 < 1\right\}$, seperti terlihat pada Gambar 4.6.1.



Gambar 4.6.1

Misal $Y_1 = X_1 + X_2$ dan $Y_1 = X_1 - X_2$. Maka transformasi

$$y_1 = u_1(x_1, x_2) = x_1 + x_2,$$
$$y_2 = u_2(x_1, x_2) = x_1 - x_2,$$

adalah transformasi satu-satu. Kita tentukan $\mathfrak{M}$ dalam bidang $y_1 y_2$ yang merupakan peta dari $\mathfrak{A}$ di bawah transformasi ini, dan kebalikannya adalah

$$x_1 = w_1(y_1, y_2) = \tfrac{1}{2}\left(y_1 + y_2\right),$$
$$x_2 = w_2(y_1, y_2) = \tfrac{1}{2}\left(y_1 - y_2\right).$$

Untuk memperoleh $\mathfrak{M}$ yang merupakan peta dari transformasi satu-satu dan pada dari $\mathfrak{A}$ kita tentukan batas-batasnya seperti berikut ini:

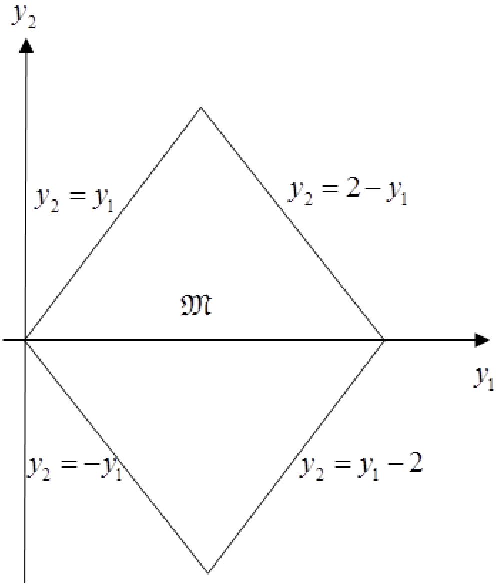
untuk $x_1 = 0,$ kita peroleh $y_2 = -y_1,$

untuk $x_1 = 1,$ kita peroleh $y_2 = 2 - y_1,$

untuk $x_2 = 0,$ kita peroleh $y_2 = y_1,$

untuk $x_2 = 1,$ kita peroleh $y_2 = y_1 - 2.$

Daerah dari $\mathfrak{M}$ ditunjukkan oleh Gambar 4.6.2.



Gambar 4.6.2

Jacobi dari transformasi ini adalah

$$J = \begin{vmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & -\frac{1}{2} \end{vmatrix} = -\frac{1}{2}.$$

Jadi pdf bersama dari Y_1 dan Y_2 adalah

$$f_{Y_1,Y_2}(y_1,y_2) = f_{X_1,X_2}\left[w_1(y_1,y_2), w_2(y_1,y_2)\right] |J| = \tfrac{1}{2} I\left[(y_1,y_2)\in\mathfrak{M}\right].$$

Lebih lanjut diperoleh pdf marginal dari Y_1,

$$f_{Y_1}(y_1) = \int_{-\infty}^{\infty} f_{Y_1,Y_2}(y_1,y_2)dy_2 = y_1 I(0 < y_1 < 1) + (2 - y_1)I(1 < y_1 < 2).$$

Hasil ini serupa dengan hasil dari Contoh 4.6.2. Ditinggalkan sebagai latihan untuk pdf marginal dari Y_2.

Transformasi yang baru saja kita bahas adalah transformasi untuk dua peubah acak kontinu. Perluasan untuk n peubah acak $X_1, X_2, ..., X_n$ adalah seperti berikut ini. Kita perhatikan integral bentuk

$$\int ... \int_A h(x_1, x_2, ..., x_n) dx_1 dx_2 ... dx_n,$$

Atas himpunan bagian A dari ruang berdimensi n, $\mathfrak{A}$. Misal

$$y_1 = u_1(x_1, x_2, ..., x_n), y_2 = u_2(x_1, x_2, ..., x_n), ..., y_n = u_n(x_1, x_2, ..., x_n),$$

dan misal kebalikan dari fungsi ini adalah

$$x_1 = w_1(y_1, y_2, ..., y_n), x_2 = w_2(y_1, y_2, ..., y_n), ..., x_n = w_n(y_1, y_2, ..., y_n),$$

yang mendefinisikan transformasi satu-satu dari $\mathfrak{A}$ pada $\mathfrak{M}$, di mana $\mathfrak{M}$ dalam ruang berdimensi n dari $y_1, y_2, ..., y_n$. Jacobi dari transformasi ini adalah

$$J = \begin{vmatrix} \frac{\partial x_1}{\partial y_1} & \frac{\partial x_1}{\partial y_2} & \cdots & \frac{\partial x_1}{\partial y_n} \\ \frac{\partial x_2}{\partial y_1} & \frac{\partial x_2}{\partial y_2} & \cdots & \frac{\partial x_2}{\partial y_n} \\ \vdots & \vdots & & \vdots \\ \frac{\partial x_n}{\partial y_1} & \frac{\partial x_n}{\partial y_2} & \cdots & \frac{\partial x_n}{\partial y_n} \end{vmatrix}$$

tidak nol dalam $\mathfrak{M}$. Jika B adalah himpunan bagian dari $\mathfrak{M}$ yang merupakan peta dari A, maka

$$\int ... \int_A h(x_1, x_2, ..., x_n) dx_1 dx_2 ... dx_n$$

$$= \int ... \int_B h\left[w_1(y_1, y_2, ..., y_n), w_2(y_1, y_2, ..., y_n), ..., w_n(y_1, y_2, ..., y_n)\right] | J | dy_1 dy_2 ... dy_n,$$

dan pdf bersama dari n peubah acak $X_1, X_2, ..., X_n$ adalah

$$g(y_1, y_2, ..., y_n)$$

$$= h\left[w_1(y_1, y_2, ..., y_n), w_2(y_1, y_2, ..., y_n), ..., w_n(y_1, y_2, ..., y_n)\right] | J | I\left[(y_1, y_2, ..., y_n) \in \mathfrak{M}\right].$$

Contoh 4.6.4 Misal X_1, X_2, X_3 mempunyai pdf

$$h(x_1, x_2, x_3) = 48 x_1 x_2 x_3 I(0 < x_1 < x_2 < x_3 < 1).$$

Jika $Y_1 = X_1 / X_2$, $Y_2 = X_2 / X_3$, dan $Y_3 = X_3$, maka kebalikan dari transformasi ini adalah

$$x_1 = y_1 y_2 y_3,\ x_2 = y_2 y_3, \text{ dan } x_3 = y_3,$$

dan Jacobi-nya adalah

$$J = \begin{vmatrix} y_2y_3 & y_1y_3 & y_1y_2 \\ 0 & y_3 & y_2 \\ 0 & 0 & 1 \end{vmatrix} = y_2y_3^2.$$

Jadi pdf bersama dari Y_1, Y_2, Y_3 adalah

$$g(y_1, y_2, y_3) = 48\left(y_1y_2y_3\right)\left(y_1y_2\right)y_3\left|y_2y_3{}^2\right| = 48y_1y_2^3y_3^5\, I(0 < y_i < 1, i = 1, 2, 3).$$

Latihan 4.6

4.6.1. Tentukan pdf marginal dari Y_2 dalam Contoh 4.6.3.

4.6.2. Misal X_1, X_2, X_3 merupakan sampel acak dari sebaran dengan pdf

$$f(x) = e^{-x} I(0 < x < \infty).$$

Misal peubah-peubah acak Y_1, Y_2, Y_3 didefinisikan sebagai

$$Y_1 = \frac{X_1}{X_1 + X_2 + X_3},\ Y_2 = \frac{X_2}{X_1 + X_2 + X_3},\ \text{dan } Y_3 = X_1 + X_2 + X_3.$$

Tentukan:

(a) pdf bersama dari Y_1, Y_2, Y_3.

(b) pdf marginal dari Y_1.

(c) pdf bersama dari Y_1 dan Y_2.

4.6.3. Ulangi Latihan 4.6.2 untuk peubah-peubah acak

$$Y_1 = \frac{X_1}{X_1 + X_2}, Y_2 = \frac{X_1 + X_2}{X_1 + X_2 + X_3},\ \text{dan } Y_3 = X_1 + X_2 + X_3.$$

4.6.4. Ulangi Latihan 4.6.2 untuk peubah-peubah acak

$$Y_1 = X_1, Y_2 = X_1 + X_2,\ \text{dan } Y_3 = X_1 + X_2 + X_3.$$

4.6.5. Jika $f(x) = \frac{1}{2} I(-1 < x < 1),$ tentukan pdf dari $Y = X^2$.

4.6.6. Jika $f(x) = \frac{1}{4} I(-1 < x < 3)$, tentukan pdf dari $Y = X^2$.

4.6.7. Misal X_1, X_2, X_3, X_4 mempunyai pdf bersama

$$(x_1, x_2, x_3, x_4) = 24I(0 < x_1 < x_2 < x_3 < x_4 < 1).$$

Tentukan pdf bersama dari $Y_1 = X_1 / X_2, Y_2 = X_2 / X_3, X_3 / X_4, Y_4 = X_4$, dan tunjukkan bahwa mereka bebas.

4.6.8. Misal X_1, X_2, X_3 sampel acak dari sebaran gamma, $X_i \sim GAM(1, \alpha_i)$. Misal peubah-peubah acak Y_1, Y_2, Y_3 didefinisikan sebagai

$$Y_1 = \frac{X_1}{X_1 + X_2 + X_3}, Y_2 = \frac{X_2}{X_1 + X_2 + X_3}, \text{ dan } Y_3 = X_1 + X_2 + X_3.$$

(a) Tentukan pdf bersama dari Y_1, Y_2, Y_3.

(b) Tunjukkan bahwa $Y_i \sim GAM(1, \alpha_1 + \alpha_2 + \alpha_3)$.

4.6.9. Misal X_1, X_2, X_3 sampel acak dengan mgf bersama $M(t) = \left(\frac{3}{4} + \frac{1}{4} e^t \right)^3$, untuk semua $t \in R$. Tentukan:

(a) $P\left(X_1 = k\right), k = 0,1,2.$

(b) mgf dari $Y = X_1 + X_2 + X_3$.

(c) $P\left(Y = k\right),\ k = 0, 1, 2, 3, 4, 5, 6.$

BAB 5 STATISTIK DAN SEBARAN SAMPEL

5.1 Pendahuluan

Pada Bab ini kita akan membahas fungsi dari peubah acak berkaitan dengan sebarannya, yang disebut ***sebaran turunan (derived distribution)*** atau ***sebaran sampel (sampling distribution)***.

Tujuan instruksional umum dari mempelajari bab ini adalah Anda diharapkan memahami konsep tentang sebaran sampel.

Tujuan instruksional khusus mempelajari bab ini adalah Anda diharapkan dapat menentukan: statistik, statistik urutan, sebaran multivariat normal, khi kuadrat, t, F, dan β.

5.2 Statistik

Definisi 5.2.1 Suatu fungsi dari peubah-peubah acak, $T = t(X_1, X_2, ..., X_n)$, yang tidak bergantung pada sebarang parameter disebut ***statistik.***

Berikut ini diberikan contoh fungsi peubah acak yang merupakan statistik dan contoh fungsi peubah acak yang bukan statistik.

Contoh 5.2.1 Misal $X_1, X_2, ..., X_n$ peubah-peubah acak dari sebarang sebaran. Maka fungsi

$$Y = \sum_{i=1}^{n} X_i,$$

dan

$$Y = \Pi_{i=1}^{n} X_i,$$

masing-masing adalah statistik.

Contoh 5.2.2 Misal $X_1, X_2, ..., X_n$ sampel acak berukuran n. Maka fungsi

$$Z = \sum_{i=1}^{n} \frac{(X_i - \mu)}{\sigma},$$

adalah bukan statistik, karena dia bergantung pada parameter μ dan σ.

Contoh 5.2.3 Misal $X_1, X_2, ..., X_n$ sampel acak berukuran *n*. Maka

$$\bar{X}_n = \sum_{i=1}^{n} \frac{X_i}{n},$$

dan

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

adalah statistik, berturut-turut disebut sebagai ***purata sampel*** dan ***varians sampel***. Pada referensi lain varians sampel dirumuskan sebagai

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

Perbedaan pada penyebut ini akan berakibat pada perbedaan ekspektasinya. Dan akan berakibat pula pada bias tidaknya S_n^2 sebagai penaksir dari σ^2.

Teorema-teorema berikut ini merupakan sifat penting dari purata sampel dan varians sampel.

Teorema 5.2.1 Misal $X_1, X_2, ..., X_n$ sampel acak berukuran *n*, dengan $E(X) = \mu,$ dan $Var(X) = \sigma^2$. Maka

$$E(\bar{X}_n) = \mu,$$

dan

$$Var(\bar{X}_n) = \frac{\sigma^2}{n}.$$

Bukti

Karena $X_1, X_2, ..., X_n$ sampel acak berukuran *n*, dengan $E(X) = \mu,$ dan $Var(X) = \sigma^2,$ maka

$$E(\bar{X}_n) = E\left[\sum_{i=1}^{n} \frac{X_i}{n}\right] = \frac{1}{n} E\left[\sum_{i=1}^{n} X_i\right] = \frac{1}{n}\left[\sum_{i=1}^{n} E(X_i)\right] = \frac{1}{n} n\mu = \mu,$$

dan

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

Teorema 5.2.2 Misal $X_1, X_2, ..., X_n$ sampel acak berukuran n, dengan $E(X) = \mu$, dan $Var(X) = \sigma^2$. Maka

(i) $E(S_n^2) = \sigma^2$,

(ii) $Var(S_n^2) = \left[E(X^4) - \frac{n-3}{n-1}\sigma^4\right]/n.$

Bukti

Mula-mula kita ubah rumus varians menjadi berikut ini

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

lihat Latihan 5.2.1. Oleh karena itu kita peroleh

$$E(S_n^2) = E\left[\frac{\sum_{i=1}^{n} X_i^2 - n\bar{X}_n^2}{n-1}\right] = \frac{1}{n-1} E\left[\sum_{i=1}^{n} X_i^2 - n\bar{X}_n^2\right]$$

$$= \frac{1}{n-1}\left[E\left(\sum_{i=1}^{n} X_i^2\right) - nE\left(\bar{X}_n^2\right)\right].$$

Kita tentukan terlebih dahulu $E\left(\sum_{i=1}^{n} X_i^2\right)$. Karena $X_1, X_2, ..., X_n$ sampel acak, maka

$$E(X_i^2) = E(X^2), i = 1, 2, ..., n.$$

Sehingga

$$E\left(\sum_{i=1}^{n} X_i^2\right) = nE(X^2).$$

Menggunakan persamaan varians,

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

atau

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

kita peroleh

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

Selanjutnya kita akan menentukan $E\left(\bar{X}^2\right)$. Karena

$$Var\left(\bar{X}_n\right) = E\left(\bar{X}_n^2\right) - \left[E\left(\bar{X}_n\right)\right]^2,$$

maka menggunakan Teorema 5.2.1, diperoleh

$$\frac{\sigma^2}{n} = E(\bar{X}_n^2) - \mu^2,$$

atau

$$E\left(\bar{X}_n^2\right) = \mu^2 + \frac{\sigma^2}{n}.$$

Jadi

$$E\left(S_n^2\right) = \frac{1}{n-1}\left[n\left(\mu^2+\sigma^2\right) - n\left(\mu^2 + \frac{\sigma^2}{n}\right)\right] = \frac{1}{n-1}\left[(n-1)\sigma^2\right] = \sigma^2,$$

bagian (i) terbukti. Bukti bagian (ii) diabaikan.

Untuk melengkapi pembahasan tentang statistik, akan kita bahas statistik urutan. Misal $X_1, X_2, ..., X_n$ sampel acak dari sebaran kontinu dengan pdf $f(x)$ positif pada $a < x < b.$ Misal

$$Y_1 = u_1(X_1, X_2, ..., X_n) = \min\left\{X_1, X_2, ..., X_n\right\}$$
$$Y_n = u_n(X_1, X_2, ..., X_n) = \max\left\{X_1, X_2, ..., X_n\right\}$$

Secara umum $Y_i = u_i(X_1, ..., X_n)$, yaitu nilai minimum ke i dari $X_1, X_2, ..., X_n$. Maka diperoleh hubungan

$$\begin{array}{ccccccc} Y_1 & \le & Y_2 & \le & ... & \le & Y_n \\ \updownarrow & & \updownarrow & & & & \updownarrow \\ X_{1,n} & \le & X_{2,n} & \le & ... & \le & X_{n,n} \end{array}$$

Peubah acak–peubah acak terurut ini, yaitu, $Y_1, Y_2, ..., Y_n$ atau $X_{1,n}, X_{2,n}, ..., X_{n,n}$, disebut ***statistik urutan,*** dan Y_i disebut statistik urutan ke i.

Teorema 5.2.3 Misal $Y_1, Y_2, ..., Y_n$ statistik urutan yang diperoleh dari sampel acak $X_1, X_2, ..., X_n$ dari sebaran kontinu dengan pdf $f(x)$ positif pada $(a < x < b)$. Maka pdf bersama dari $Y_1, Y_2, ..., Y_n$ adalah

$$g(y_1, y_2, ..., y_n) = n!\, f(y_1) f(y_2) ... f(y_n) I(a < y_1 < y_2 < ... < y_n < b).$$

Bukti

Kejadian-kejadian dari peubah-peubah acak $X_1, X_2, ..., X_n$ dapat dipartisi dalam $n!$ kejadian yang saling asing yang memetakan pada kejadian-kejadian dari peubah-peubah acak $Y_1, Y_2, ..., Y_n$, sebut

$$\{(y_1, y_2, ..., y_n) : a < y_1 < y_2 < ... < y_n < b\}.$$

Satu dari $n!$ kejadian-kejadian ini adalah $a < x_1 < x_2 < ... < x_n < b$, dan yang lainnya dapat ditentukan oleh permutasi n dari semua nilai x yang mungkin. Transformasi yang mengaitkan $x_1 = y_1, x_2 = y_2, ..., x_n = y_n$ mempunyai Jacobi sama dengan satu. Demikian juga untuk yang lainya dari masing-masing transformasi mempunyai Jacobi satu dari ± 1. Oleh karena itu

$$\begin{aligned} g(y_1, y_2, ..., y_n) &= \sum_{i=1}^{n!} f(y_1) f(y_2) ... f(y_n) \,|\, J_i \,| \\ &= n!\, f(y_1) f(y_2) ... f(y_n) I(a < y_1 < y_2 < ... < y_n < b). \end{aligned}$$

Contoh 5.2.4 Misal X menyatakan peubah acak kontinu dengan pdf $f(x)$ positif pada $a < x < b$. Maka cdf-nya adalah

$$F(x) = \int_a^x f(w)\, dw\, I(a \le x < b) + I(b \le x).$$

Oleh karena itu ada median m tunggal dari sebaran ini dengan $F(m) = \frac{1}{2}$. Misal X_1, X_2, X_3 sampel acak dari sebaran ini dan misal Y_1, Y_2, Y_3 statistik urutan dari sampel ini. Kita akan menghitung peluang dari $Y_2 \le m$. Berdasarkan Teorema 5.2.3 kita peroleh pdf bersama dari Y_1, Y_2, Y_3,

$$g(y_1, y_2, y_3) = 6 f(y_1, y_2, y_3) I(a < y_1 < y_2 < y_3 < b),$$

dan pdf dari Y_2,

$$h(y_2) = 6f(y_2)\int_{y_2}^{b}\int_{a}^{y_2} f(y_1)f(y_3)dy_1dy_3$$
$$= 6f(y_2)F(y_2)\left[1 - F(y_2)\right]I(a < y_2 < b).$$

Jadi

$$P(Y_2 \le m) = 6\int_a^m \left\{F(y_2)f(y_2) - \left[F(y_2)\right]^2 f(y_2)\right\}dy_2$$
$$= 6\left\{\frac{\left[F(y_2)\right]^2}{2} - \frac{\left[F(y_2)\right]^3}{3}\right\}_a^m = \frac{1}{2}.$$

Teori integral berikut ini akan memudahkan kita dalam menentukan pdf statistik urutan ke k, Y_k, yaitu

$$\int_a^x \left[F(w)\right]^{\alpha-1} f(w)dw = \frac{\left[F(x)\right]^{\alpha}}{\alpha}, \alpha > 0$$

dan

$$\int_x^b \left[1 - F(w)\right]^{\beta-1} f(w)dw = \frac{\left[1 - F(x)\right]^{\beta}}{\beta}, \beta > 0.$$

Sekarang kita tentukan pdf-nya seperti berikut

$$g_k(y_k) = \int_a^{y_k} \cdots \int_a^{y_2}\int_{y_k}^{b} \cdots \int_{y_{n-1}}^{b} n!\, f(y_1)f(y_2)\ldots f(y_n)dy_n\ldots dy_{k+1}dy_1\ldots dy_{k-1}$$
$$= \frac{n!}{(k-1)!(n-k)!}\left[F(y_k)\right]^{k-1}\left[1 - F(y_k)\right]^{n-k} f(y_k)I(a < y_k < b).$$

Khususnya untuk $k = 1$, Y_1 disebut sebagai ***statistik urutan terkecil,*** pdf-nya adalah

$$g_1(y_1) = \frac{n!}{(1-1)!(n-1)!}\left[F(y_1)\right]^{1-1}\left[1 - F(y_1)\right]^{n-1} f(y_1)$$
$$= n\left[1 - F(y_1)\right]^{n-1} f(y_1)I(a < y_1 < b), \qquad (5.2.1)$$

dan untuk $k = n$, Y_n disebut sebagai ***statistik urutan terbesar,*** pdf-nya adalah

$$g_n(y_n) = \frac{n!}{(n-1)!(n-n)!}\left[F(y_n)\right]^{n-1}\left[1 - F(y_n)\right]^{n-n} f(y_n)$$
$$= n\left[F(y_n)\right]^{n-1} f(y_n)I(a < y_n < b). \qquad (5.2.2)$$

Cara lain untuk mendapatkan pdf dari Y_1 dan Y_n, diperoleh dengan menentukan cdf-nya terlebih dahulu. Misal $G_1(y_1)$ cdf dari Y_1. Maka

$$\begin{aligned} G_1(y_1) &= P\left(Y_1 \le y_1\right) \\ &= 1 - P\left(Y_1 > y_1\right). \end{aligned}$$

Karena $Y_1 = \min\left\{X_1, X_2, ..., X_n\right\}$, maka $Y_1 > y_1$ ekivalen dengan

$$X_1 > y_1,\ \ X_2 > y_1,\ \ ...,\ \text{dan}\ \ X_n > y_1.$$

Oleh karena itu

$$P\left(Y_1 > y_1\right) = P\left(X_1 > y_1, X_2 > y_1, ..., \text{ dan } X_n > y_1\right).$$

Karena $X_1, X_2, ..., X_n$ sampel acak, maka

$$\begin{aligned} &P\left(X_1 > y_1, X_2 > y_1, ..., \text{ dan } X_n > y_1\right) \\ &= P\left(X_1 > y_1\right)P\left(X_2 > y_1\right)...P\left(X_n > y_1\right) \\ &= \left[1 - P\left(X_1 \le y_1\right)\right]\left[1 - P\left(X_2 \le y_1\right)\right]...\left[1 - P\left(X_n \le y_1\right)\right] \\ &= \left[1 - F(y_1)\right]\left[1 - F(y_1)\right]...\left[1 - F(y_1)\right] \\ &= \left[1 - F(y_1)\right]^n I(a < y_1 < b). \end{aligned}$$

Oleh karena itu

$$G_1(y_1) = \left[1 - \left[1 - F(y_1)\right]^n\right] I(a < y_1 < b).$$

Jadi pdf Y_1 adalah

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

Hasil ini sama dengan persamaan (5.2.1). Selanjutnya, jika $G_n(y_n)$ cdf dari Y_n. Maka

$$G_n(y_n) = P\left(Y_n \le y_n\right)$$

Karena $Y_n = \max\left\{X_1, X_2, ..., X_n\right\}$, maka $Y_n \le y_n$ ekivalen dengan

$$X_1 \le y_n,\ \ X_2 \le y_n,\ \ ...,\ \text{dan}\ \ X_n \le y_n.$$

Oleh karena itu

$$\begin{aligned}G_n(y_n) &= P(X_1 \le y_n, X_2 \le y_n, ..., \text{ dan } X_n \le y_n) \\ &= P(X_1 \le y_n)P(X_2 \le y_n)...P(X_n \le y_n) \\ &= F(y_n)F(y_n)...F(y_n) = [F(y_n)]^n I(a < y_n < b).\end{aligned}$$

Jadi pdf Y_n adalah

$$g_n(y_n) = n[F(y_n)]^{n-1} f(y_n)I(a < y_n < b).$$

Hasil ini sama dengan persamaan (5.2.2).

Contoh 5.2.5 Misal $Y_1, Y_2, ..., Y_n$ statistik urutan yang diperoleh dari sampel acak $X_1, ..., X_n$ dari populasi dengan pdf

$$f(x) = 2xI(0 < x < 1).$$

Maka

$$F(x) = \int_0^x 2tdt = x^2 I(0 < x < 1) + I(x \ge 1).$$

Sehingga cdf dan pdf dari Y_1 berturut-turut adalah

$$G_1(y_1) = 1 - [1 - F(y_1)]^n = \left[1 - \left(1 - y_1^{\,2}\right)^n\right] I(0 < y_1 < 1) + I(y_1 \ge 1).$$

dan

$$\begin{aligned}g_1(y_1) &= n[1 - F(y_1)]^{n-1} f(y_1) \\ &= n\left[1 - y_1^{\,2}\right]^{n-1} 2y_1 I(0 < y_1 < 1).\end{aligned}$$

Ditinggalkan sebagai latihan untuk cdf dan pdf dari Y_n.

Latihan 5.2

5.2.1 Tunjukkan bahwa varians sampel pada Contoh 5.2.3 dapat ditulis sebagai

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

5.2.2. Misal X_1, X_2 sampel acak berukuran 2. Tentukan c sehingga

$$S_n^2 = c\left(X_1 - X_2\right)^2.$$

5.2.3. Misal $x_i = i, i = 1,...,n.$ Hitung nilai dari $\overline{x}_n$ dan s_n^2.

5.2.4 Misal $x_i = 2^i, i = 1,...,n.$ Hitung nilai dari $\overline{x}_n$ dan s_n^2.

5.2.5. Misal $y_i = a + bx_i, i = 1,...,n,$ di mana a dan b suatu konstanta. Tentukan $\overline{y}_n = \sum_{i=1}^{n} \frac{y_i}{n}$ dan $s_y^2 = \frac{\sum_{i=1}^{n} y_i^2 - n\overline{y}_n^{\,2}}{n-1}$ dalam suku-suku $a, b, \overline{x}_n,$ dan s_x^2.

5.2.6. Misal $X_1,...,X_n$ sampel acak, dan anggap varians sampel dirumuskan sebagai $S_n^{\,2} = \sum_{i=1}^{n} \frac{\left(X_i - \overline{X}_n\right)^2}{n}.$

(a) Tunjukkan bahwa $S_n^{\,2} = \frac{\sum_{i=1}^{n} X_i^2 - n\overline{X}_n^{\,2}}{n}.$

(b) Tentukan $E\left(S_n^{\,2}\right).$

5.2.7. Misal $X_1,...,X_n$ peubah-peubah acak sehingga $E(X_i) = 0,$ $Var(X_i) = i^2,\ i = 1,...,30.$ Hitunglah $E\left(\sum_{i=1}^{100} X_i^{\,2}\right).$

5.2.8. Misal Y_1, Y_2, Y_3 adalah statistik urutan dari sebaran dengan pdf $f(x) = 2xI(0 < x < 1).$ Tentukan koefisien korelasi antara Y_2 dan Y_3.

5.2.9. Misal Y_1, Y_2, Y_3, Y_4, Y_5 adalah statistik urutan dari sebaran dengan pdf $f(x) = e^{-x} I(0 < x < \infty).$

(a) Tentukan pdf bersama dari Y_2 dan Y_4.

(b) Tunjukkan bahwa $Z_1 = Y_2$ dan $Z_2 = Y_4 - Y_2$ bebas.

5.2.10. Tentukan cdf dan pdf dari Y_n untuk Contoh 5.2.5.

5.2.11. Misal $X_1, X_2, X_3,$ sampel acak dari suatu populasi dengan pdf

$$f(x) = 3(1-x)^2 I(0 < x < 1).$$

Misal Y adalah minimum dari ketiga peubah acak tersebut. Tentukan:

(a) pdf dari Y,

(b) cdf dari Y.

5.3 Sebaran Normal Multivariat

Pembahasan pada subpokok bahasan ini diperlukan pemahaman tentang sifat-sifat mgf dan matriks. Sebelum membahas sebaran normal multivariat, untuk menyederhanakan permasalahan kita mulai dengan contoh berikut ini.

Contoh 5.3.1 Misal peubah-peubah acak X_1 dan X_2 bebas berturut-turut dari sebaran normal $N(\mu_1, \sigma_1^2)$ dan $N(\mu_2, \sigma_2^2),$ dan misal $Y = X_1 - X_2.$ Tentukan sebaran dari $Y.$

Penyelesaian

Misal $M_Y(t)$ mgf dari $Y.$ Maka

$$M_Y(t) = E\left(e^{t(X_1 - X_2)}\right) = E\left(e^{tX_1} e^{-tX_2}\right) = E\left(e^{tX_1}\right) E\left(e^{-tX_2}\right),$$

karena X_1 dan X_2 bebas. Diketahui X_1 dan X_2 berturut-turut dari sebaran normal $N(\mu_1, \sigma_1^2)$ dan $N(\mu_2, \sigma_2^2),$ maka

$$E\left(e^{tX_1}\right) = \exp\left(\mu_1 t + \frac{\sigma_1^2 t^2}{2}\right),$$

dan

$$E\left(e^{tX_2}\right) = \exp\left(\mu_2 t + \frac{\sigma_2^2 t^2}{2}\right),$$

untuk semua bilangan real $t.$ Untuk memperoleh $E\left(e^{-tX_2}\right)$ kita ganti t oleh $-t$ pada $E\left(e^{tX_2}\right)$. Sehingga

$$E\left(e^{-tX_2}\right) = \exp\left(-\mu_2 t + \frac{\sigma_2^2 t^2}{2}\right).$$

Oleh karena itu

$$M_Y(t) = \exp\left(\mu_1 t + \frac{\sigma_1^2 t^2}{2}\right)\exp\left(-\mu_2 t + \frac{\sigma_2^2 t^2}{2}\right)$$
$$= \exp\left((\mu_1 - \mu_2)t + \frac{(\sigma_1^2 + \sigma_2^2)t^2}{2}\right).$$
Jadi $Y \sim N(\mu_1 - \mu_2, \sigma_1^2 + \sigma_2^2)$.

Teorema berikut merupakan bentuk umum dari Contoh 5.3.1, yang merupakan salah satu sifat penting dalam sebaran normal.

Teorema 5.3.1 Misal $X_1, X_2, ..., X_n$ peubah-peubah acak bebas, dan $X_i \sim N(\mu_i, \sigma_i^2), i = 1,2, ..., n$, dan misal $Y = \sum_{i=1}^{n} k_i X_i$. Maka
$$Y \sim N\left(\sum_{i=1}^{n} k_i \mu_i, \sum_{i=1}^{n} k_i^2 \sigma_i^2\right).$$

Bukti

Misal $M_Y(t)$ mgf dari Y. Maka

Karena $X_1, ..., X_n$ bebas, maka
$$E\left[e^{tk_1X_1} e^{k_2X_2} ... e^{tk_nX_n}\right] = E\left[e^{tk_1X_1}\right] E\left[e^{tk_2X_2}\right] ... E\left[e^{tk_nX_n}\right].$$
Oleh karena itu kita peroleh
$$M_Y(t) = E\left[e^{tk_1X_1}\right] E\left[e^{tk_2X_2}\right] ... E\left[e^{tk_nX_n}\right].$$
Karena $X_i \sim N(\mu_i, \sigma_i^2)$, maka
$$E\left[e^{tX_i}\right] = \exp\left[\mu_i t + \frac{\sigma_i^2 t^2}{2}\right], i = 1, ..., n.$$
Sehingga
$$E\left[e^{tk_iX_i}\right] = \exp\left[\mu_i t k_i + \frac{\sigma_i^2 t^2 k_i^2}{2}\right] = \exp\left[k_i \mu_i t + \frac{k_i^2 \sigma_i^2 t^2}{2}\right].$$
Jadi

$$M_Y(t) = \exp\left[k_1\mu_1 t + \frac{k_1^2\sigma_1^2 t^2}{2}\right]\exp\left[k_2\mu_2 t + \frac{k_2^2\sigma_2^2 t^2}{2}\right]\ldots\exp\left[k_n\mu_n t + \frac{k_n^2\sigma_n^2 t^2}{2}\right]$$

$$= \exp\left[\left(\sum_{i=1}^{n} k_i\mu_i\right)t + \frac{\left(\sum_{i=1}^{n} k_i^2\sigma_i^2\right)t^2}{2}\right].$$

Hasil terakhir merupakan mgf dari peubah acak yang bersebaran $N\left(\sum_{i=1}^{n} k_i\mu_i, \sum_{i=1}^{n} k_i^2\sigma_i^2\right)$. Jadi $Y \sim N(\sum_{i=1}^{n} k_i\mu_i , \sum_{i=1}^{n} k_i^2\sigma_i^2)$.

Akibat 5.3.1 Misal $X_1, X_2, ..., X_n$ sampel acak dari sebaran $N\left(\mu, \sigma^2\right)$. Maka

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

Bukti dari Akibat 5.3.1 ini diperoleh dari Teorema 5.3.1, dengan mengambil $k_i = \frac{1}{n}, i = 1, ..., n.$

Akibat 5.3.2 Misal $X_1, ..., X_n$ sampel acak dari sebaran $N\left(\mu, \sigma^2\right)$. Maka

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

Bukti Akibat 5.3.2 ini diperoleh dari Akibat 5.3.1 dan Teorema 3.7.1.

Contoh 5.3.2 Misal peubah acak X menyatakan waktu hidup, dalam bulan, suatu baterai, dan anggap $X \sim N(0,1)$. Misal 25 baterai diambil dari populasi ini. Tentukan konstanta c sehingga peluang purata waktu hidup 25 baterai tersebut lebih dari c adalah 95%.

Penyelesaian

Misal $\bar{X}_n$ purata waktu hidup 25 baterai. Maka $E\left(\bar{X}_n\right)=60$, dan $Var(\bar{X}_n)=\frac{35}{25}$. Oleh karena itu menggunakan Akibat 1 kita peroleh bahwa $\bar{X}_n \sim N\left(60, \frac{35}{25}\right)$. Sehingga

$$P\left[\bar{X}_n > c\right] = P\left[\frac{\bar{X}_n - 60}{\sqrt{\frac{35}{25}}} > \frac{c-60}{\sqrt{\frac{35}{25}}}\right] = 1 - P\left[\frac{\bar{X}_n - 60}{\sqrt{\frac{35}{25}}} \le \frac{c-60}{\sqrt{\frac{35}{25}}}\right] = 1 - \Phi\left(\frac{c-60}{\sqrt{\frac{35}{25}}}\right) = 0.95.$$

Atau

$$\Phi\left(\frac{c-60}{\sqrt{\frac{35}{25}}}\right) = 0.05.$$

Dari Minitab diperoleh $\frac{c-60}{\sqrt{\frac{35}{25}}} = -1.6449.$ Jadi

$$c = 60 - 1.6449\sqrt{\frac{35}{25}} = 58.0537.$$

Sekarang kita membahas sebaran normal multivariat. Kita perhatikan vektor acak $\mathbf{Z} = (Z_1, Z_2, ..., Z_n)^t$ di mana $Z_1, Z_2, ..., Z_n$ sampel acak dari sebaran normal, $Z_i \sim N(0,1)$. Maka pdf bersama dari $\mathbf{Z}$ adalah

$$f_{\mathbf{Z}}(\mathbf{z}) = \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi}} \exp\left\{-\frac{1}{2} z_i^2\right\} = \left(\frac{1}{\sqrt{2\pi}}\right)^{n/2} \exp\left\{-\frac{1}{2}\sum_{i=1}^{n} z_i^2\right\}$$

$$= \left(\frac{1}{\sqrt{2\pi}}\right)^{n/2} \exp\left\{-\frac{1}{2}\mathbf{z}^t\mathbf{z}\right\}.$$

Karena $Z_1, Z_2, ..., Z_n$ sampel acak dari sebaran normal, $Z_i \sim N(0,1)$, maka

$$E\left[\mathbf{Z}\right] = \mathbf{0} \text{ dan } Cov\left[\mathbf{Z}\right] = \mathbf{I}_n, \qquad (5.3.1)$$

di mana $\mathbf{I}_n$ menyatakan matriks identitas urutan n. Fungsi pembangkit momen dari $\mathbf{Z}$ adalah

$$M_{\mathbf{Z}}(\mathbf{t}) = E\left[\exp\left\{\mathbf{t}^t\mathbf{Z}\right\}\right] = E\left[\prod_{i=1}^{n}\exp\left\{t_i Z_i\right\}\right] = \prod_{i=1}^{n} E\left[\exp\left\{t_i Z_i\right\}\right]$$
$$= \exp\left\{\frac{1}{2}\sum_{i=1}^{n} t_i^2\right\} = \exp\left\{\frac{1}{2}\mathbf{t}^t\mathbf{t}\right\}, \qquad (5.3.2)$$

untuk semua $\mathbf{t} \in R^n$. Kita katakan bahwa $\mathbf{Z}$ mempunyai ***sebaran normal multivariat*** dengan vektor purata $\mathbf{0}$ dan matriks kovarians $\mathbf{I}_n$, ditulis $\boldsymbol{Z} \sim N_n(\mathbf{0}, \boldsymbol{I}_n)$.

Secara umum, anggap $\mathbf{\Sigma}$ adalah matriks urutan $n \times n$, simetrik, dan ***semi definit positif***. Secara aljabar $\mathbf{\Sigma}$ dapat dinyatakan sebagai dekomposisi

$$\mathbf{\Sigma} = \mathbf{\Gamma}^t\mathbf{\Lambda}\mathbf{\Gamma},$$

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

$$\mathbf{\Lambda} = diag\left(\lambda_1, \lambda_2, ..., \lambda_n\right), \lambda_1 \geq \lambda_2 \geq ... \geq \lambda_n \geq 0$$

adalah nilai-nilai eigen dari $\mathbf{\Sigma}$, dan kolom-kolom dari $\mathbf{\Gamma}^t$, $\mathbf{v}_1, \mathbf{v}_2, ..., \mathbf{v}_n$, adalah vektor-vektor eigen. Matriks $\mathbf{\Gamma}$ adalah matriks ortogonal, yaitu $\mathbf{\Gamma}^{-1} = \mathbf{\Gamma}$. Oleh karena itu $\mathbf{\Gamma}\mathbf{\Gamma}^t = \mathbf{I}$. Kita dapat menulis dekomposisi ini dalam bentuk lain, seperti

$$\mathbf{\Sigma} = \mathbf{\Gamma}^t\mathbf{\Lambda}\mathbf{\Gamma} = \sum_{i=1}^{n}\lambda_i\mathbf{v}_i\mathbf{v}_i^t, \qquad (5.3.3)$$

ditinggalkan sebagai latihan. Kita definisikan akar kuadrat matriks semi definit positif $\mathbf{\Sigma}$ sebagai

$$\mathbf{\Sigma}^{1/2} = \mathbf{\Gamma}^t\mathbf{\Lambda}^{1/2}\mathbf{\Gamma},$$

di mana $\mathbf{\Lambda}^{1/2} = diag\left(\sqrt{\lambda_1}, \sqrt{\lambda_2}, ..., \sqrt{\lambda_n}\right)$. Oleh karena itu

$$\left(\mathbf{\Sigma}^{1/2}\right)^{-1} = \mathbf{\Gamma}^t\mathbf{\Lambda}^{-1/2}\mathbf{\Gamma}, \qquad (5.3.4)$$

ditinggalkan sebagai latihan.

Kita definisikan vektor acak $\mathbf{X}$ sebagai

$$\mathbf{X} = \mathbf{\Sigma}^{1/2}\mathbf{Z} + \boldsymbol{\mu}, \qquad (5.3.5)$$

di mana $\boldsymbol{\mu}$ vektor konstanta $n \times 1$. Menggunakan persamaaan (5.3.1) kita peroleh

$$E\left[\mathbf{X}\right] = E\left[\mathbf{\Sigma}^{1/2}\mathbf{Z} + \boldsymbol{\mu}\right] = \boldsymbol{\mu} \text{ dan } Cov\left[\mathbf{X}\right] = \mathbf{\Sigma}^{1/2}\mathbf{\Sigma}^{1/2} = \mathbf{\Sigma}.$$

Selanjutnya menggunakan persamaan (5.3.2) diperoleh mgf dari $\mathbf{X}$

$$\begin{aligned} M_{\mathbf{X}}(\mathbf{t}) = E\left[\exp\left\{\mathbf{t}^t\mathbf{X}\right\}\right] &= E\left[\exp\left\{\mathbf{t}^t\boldsymbol{\Sigma}^{1/2}\mathbf{Z}+\mathbf{t}^t\boldsymbol{\mu}\right\}\right] \\ &= \exp\left\{\mathbf{t}^t\boldsymbol{\mu}\right\} E\left[\exp\left\{\left(\boldsymbol{\Sigma}^{1/2}\mathbf{t}\right)^t \mathbf{Z}\right\}\right] \\ &= \exp\left\{\mathbf{t}^t\boldsymbol{\mu}\right\}\exp\left\{\frac{1}{2}\left(\boldsymbol{\Sigma}^{1/2}\mathbf{t}\right)^t\boldsymbol{\Sigma}^{1/2}\mathbf{t}\right\} \\ &= \exp\left\{\mathbf{t}^t\boldsymbol{\mu}\right\}\exp\left\{\frac{1}{2}\mathbf{t}^t\boldsymbol{\Sigma}\mathbf{t}\right\}. \end{aligned}$$

Hasil di atas didefinisikan seperti berikut.

Definisi 5.3.1 Misal $\mathbf{X}$ vektor acak berdimensi *n*. Maka $\mathbf{X}$ dikatakan bersebaran normal multivariat jika mgf-nya adalah

$$M_{\mathbf{X}}(\mathbf{t}) = \exp\left\{\mathbf{t}^t\boldsymbol{\mu}\right\}\exp\left\{\frac{1}{2}\mathbf{t}^t\boldsymbol{\Sigma}\mathbf{t}\right\},$$

untuk semua $\mathbf{t}\in R^n$ dan $\boldsymbol{\Sigma}$ adalah matriks simetrik, semi definit positif dan $\boldsymbol{\mu}\in R^n$, ditulis $\boldsymbol{X}\sim N_n(\mu,\boldsymbol{\Sigma})$.

Dari transformasi satu-satu pada (5.3.5) kita peroleh transformasi kebalikan

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

dengan Jacobi $\left|\boldsymbol{\Sigma}^{-1/2}\right| = \left|\boldsymbol{\Sigma}\right|^{-1/2}$. Oleh karena itu pdf dari $\mathbf{Z}$ adalah

$$f_{\mathbf{Z}}(\mathbf{z}) = \frac{1}{\left(2\pi\right)^{n/2}\left|\boldsymbol{\Sigma}\right|^{1/2}}\exp\left\{-\frac{1}{2}\left(\mathbf{x}-\boldsymbol{\mu}\right)^t\boldsymbol{\Sigma}^{-1}\left(\mathbf{x}-\boldsymbol{\mu}\right)\right\},$$

untuk semua $\mathbf{x}\in R^n$.

Teorema 5.3.2 Misal $\boldsymbol{X}\sim N_n(\mu,\boldsymbol{\Sigma})$ dan misal $\mathbf{Y} = \mathbf{AX}+\mathbf{b}$, di mana $\mathbf{A}$ adalah matriks $m\times n$ dan $\mathbf{b}\in R^m$. Maka

$$\boldsymbol{Y}\sim N_m(\boldsymbol{A}\mu+\boldsymbol{b},\boldsymbol{A\Sigma A}^t).$$

Bukti

Kita tentukan mgf dari $\mathbf{Y}$

$$
\begin{aligned}
M_{\mathbf{Y}}(\mathbf{t}) &= E\left[\exp\left\{\mathbf{t}^t\mathbf{Y}\right\}\right] \\
&= E\left[\exp\left\{\mathbf{t}^t\left(\mathbf{AX}+\mathbf{b}\right)\right\}\right] \\
&= \exp\left\{\mathbf{t}^t\mathbf{b}\right\} E\left[\exp\left\{\left(\mathbf{A}^t\mathbf{t}\right)^t \mathbf{X}\right\}\right] \\
&= \exp\left\{\mathbf{t}^t\mathbf{b}\right\} \exp\left\{\left(\mathbf{A}^t\mathbf{t}\right)^t \boldsymbol{\mu} + \frac{1}{2}\left(\mathbf{A}^t\mathbf{t}\right)^t \boldsymbol{\Sigma}\left(\mathbf{A}^t\mathbf{t}\right)\right\} \\
&= \exp\left\{\mathbf{t}^t\left(\mathbf{A}\boldsymbol{\mu}+\mathbf{b}\right) + \frac{1}{2}\mathbf{t}^t\mathbf{A}\boldsymbol{\Sigma}\mathbf{A}^t\mathbf{t}\right\}.
\end{aligned}
$$

Jadi $\boldsymbol{Y} \sim N_m(\boldsymbol{A}\mu + \boldsymbol{b}, \boldsymbol{A\Sigma A}^t)$.

Berdasarkan teorema ini kita dapat juga memperoleh sebaran marginal dari sebaran normal multivariat. Misal $\mathbf{X}_1$ adalah sebarang vektor bagian dari $\mathbf{X}$, katakan berdimensi $m < n$. Kita tulis $\mathbf{X}$, sebagai

$$
\mathbf{X} = \begin{bmatrix} \mathbf{X}_1 \\ \mathbf{X}_2 \end{bmatrix}, \tag{5.3.6}
$$

di mana $\mathbf{X}_2$ vektor berdimensi $p = n - m$. Secara sama kita partisi purata dan matriks kovarians dari $\mathbf{X}$ sebagai

$$
\boldsymbol{\mu} = \begin{bmatrix} \boldsymbol{\mu}_1 \\ \boldsymbol{\mu}_2 \end{bmatrix} \text{ dan } \boldsymbol{\Sigma} = \begin{bmatrix} \boldsymbol{\Sigma}_{11} & \boldsymbol{\Sigma}_{12} \\ \boldsymbol{\Sigma}_{21} & \boldsymbol{\Sigma}_{22} \end{bmatrix}, \tag{5.3.7}
$$

di mana $\boldsymbol{\Sigma}_{11}$ menyatakan matriks kovarians dari $\mathbf{X}_1$ dan $\boldsymbol{\Sigma}_{12}$ memuat semua kovarians antara komponen-komponen $\mathbf{X}_1$ dan $\mathbf{X}_2$. Kita definisikan $\mathbf{A}$ sebagai matriks,

$$
\mathbf{A} = \left[\mathbf{I}_m \,\vdots\, \mathbf{0}_{mp}\right],
$$

di mana $\mathbf{0}_{mp}$ adalah matriks $m \times p$ yang unsur-unsurnya adalah nol. Maka $\mathbf{X}_1 = \mathbf{AX}$. Oleh karena itu menggunakan Teorema 5.3.1. untuk sebaran ini, kita peroleh akibat berikut:

Akibat 5.3.3 Misal $\boldsymbol{X} \sim N_n(\mu, \boldsymbol{\Sigma})$ dipartisi seperti pada persamaan (5.3.6) dan (5.3.7). Maka $\boldsymbol{X} \sim N_m(\mu_i, \boldsymbol{\Sigma}_{11})$.

Hasil ini sangat penting, karena sebarang sebaran marginal dari $\boldsymbol{X} \sim N_n(\mu, \boldsymbol{\Sigma})$, juga bersebaran normal.

Contoh 5.3.3 Dalam contoh ini kita akan memeriksa kasus sebaran normal multivariat jika $n = 2$. Sebaran ini disebut sebagai ***sebaran normal bivariat***. Sekarang kita anggap

$$\boldsymbol{X} = (X_1, X_2)^t \sim N_2(\mu, \boldsymbol{\Sigma}),$$

di mana

$$\boldsymbol{\mu} = \begin{bmatrix} \mu_1 \\ \mu_2 \end{bmatrix} \text{ dan } \boldsymbol{\Sigma} = \begin{bmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{bmatrix}.$$

Oleh karena itu μ_1 dan σ_1^2 berturut-turut purata dan varians dari X_1; μ_2 dan σ_2^2 berturut-turut purata dan varians dari X_2; dan σ_{12} adalah kovarians antara X_1 dan X_2. Ingat bahwa $\sigma_{12} = \rho\sigma_1\sigma_2$, di mana ρ koefisien korelasi anara X_1 dan X_2. Kita peroleh kebalikan dari $\boldsymbol{\Sigma}$,

$$\boldsymbol{\Sigma}^{-1} = \frac{1}{\sigma_1^2\sigma_2^2\left(1-\rho^2\right)}\begin{bmatrix} \sigma_2^2 & -\rho\sigma_1\sigma_2 \\ -\rho\sigma_1\sigma_2 & \sigma_1^2 \end{bmatrix}.$$

Jadi pdf dari $\mathbf{X} = \left(X_1, X_2\right)^t$ adalah

$$f(x_1, x_2) = \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}}e^{-q/2},$$

di mana

$$q = \frac{1}{1-\rho^2}\left[\left(\frac{x_1-\mu_1}{\sigma_1}\right)^2 - 2\rho\left(\frac{x_1-\mu_1}{\sigma_1}\right)\left(\frac{x_2-\mu_2}{\sigma_2}\right) + \left(\frac{x_2-\mu_2}{\sigma_2}\right)^2\right].$$

Jika X_1 dan X_2 bebas, maka koefisiennya adalah nol dan pdf-nya

$$f(x_1, x_2) = \frac{1}{2\pi\sigma_1\sigma_2}e^{-\frac{1}{2}\left[\left(\frac{x_1-\mu_1}{\sigma_1}\right)^2 + \left(\frac{x_2-\mu_2}{\sigma_2}\right)^2\right]}.$$

Selanjutnya menggunakan Akibat 5.3.3 kita peroleh $X_1 \sim N(\mu_1, \sigma_1^2)$ dan $X_1 \sim N(\mu_2, \sigma_2^2)$.

Teorema 5.3.3 Misal $\boldsymbol{X} \sim N_n(\mu, \boldsymbol{\Sigma})$, dipartisi seperti pada persamaan (5.3.6) dan (5.3.7). Maka $\mathbf{X}_1$ dan $\mathbf{X}_2$ bebas jika dan hanya jika $\boldsymbol{\Sigma}_{12} = \mathbf{0}$.

Bukti

Mula-mula kita catat bahwa $\boldsymbol{\Sigma}_{21} = \boldsymbol{\Sigma}_{12}{}^t$. Suatu mgf dari $\mathbf{X}_1$ dan $\mathbf{X}_2$ adalah

$$M_{\mathbf{X}_1,\mathbf{X}_2}(\mathbf{t}_1,\mathbf{t}_2) = \exp\left\{\mathbf{t}_1^t\boldsymbol{\mu}_1 + \mathbf{t}_2^t\boldsymbol{\mu}_2 + \frac{1}{2}\left(\mathbf{t}_1^t\boldsymbol{\Sigma}_{11}\mathbf{t}_1 + \mathbf{t}_2^t\boldsymbol{\Sigma}_{22}\mathbf{t}_2 + \mathbf{t}_2^t\boldsymbol{\Sigma}_{21}\mathbf{t}_1 + \mathbf{t}_1^t\boldsymbol{\Sigma}_{12}\mathbf{t}_2\right)\right\},$$

di mana $\mathbf{t}^t = (\mathbf{t}_1^t, \mathbf{t}_2^t)$ adalah partisi seperti pada $\boldsymbol{\mu}$. Menggunakan Akibat 5.3.3 kita peroleh $\boldsymbol{X}_1 \sim N_m(\mu_1, \boldsymbol{\Sigma}_{11})$ dan $\boldsymbol{X}_2 \sim N_p(\mu_2, \boldsymbol{\Sigma}_{22})$. Oleh karena itu perkalian mgf marginalnya adalah

$$M_{\mathbf{X}_1}(\mathbf{t}_1)M_{\mathbf{X}_2}(\mathbf{t}_2) = \exp\left\{\mathbf{t}_1^t\boldsymbol{\mu}_1 + \mathbf{t}_2^t\boldsymbol{\mu}_2 + \frac{1}{2}\left(\mathbf{t}_1^t\boldsymbol{\Sigma}_{11}\mathbf{t}_1 + \mathbf{t}_2^t\boldsymbol{\Sigma}_{22}\mathbf{t}_2\right)\right\}.$$

Jadi $\mathbf{X}_1$ dan $\mathbf{X}_2$ bebas jika dan hanya jika $\boldsymbol{\Sigma}_{12} = \mathbf{0}$.

Kombinasi dari hasil Teorema 5.3.2 dan Teorema 5.3.3 kita peroleh teorema berikut ini.

Teorema 5.3.4 Misal $\boldsymbol{X} \sim N_n(\mu, \boldsymbol{\Sigma})$, dipartisi seperti pada persamaan (5.3.6) dan (5.3.7), dan misal $\boldsymbol{\Sigma}$ definit positif. Maka sebaran bersyarat dari $\mathbf{X}_1 \mid \mathbf{X}_2$ adalah

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

Bukti

Kita tentukan sebaran bersama dari vektor-vektor acak $\mathbf{W} = \mathbf{X}_1 - \boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1}\mathbf{X}_2$ dan $\mathbf{X}_2$. Sebaran ini diperoleh dari transformasi

$$\begin{bmatrix} \mathbf{W} \\ \mathbf{X}_2 \end{bmatrix} = \begin{bmatrix} \mathbf{I}_m & -\boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1} \\ \mathbf{0} & \mathbf{I}_p \end{bmatrix} \begin{bmatrix} \mathbf{X}_1 \\ \mathbf{X}_2 \end{bmatrix}.$$

Karena transformasi ini linear, maka menggunakan Teorema 5.3.2 diperoleh sebaran bersamanya adalah sebaran normal multivariat dengan

$$E[\mathbf{W}] = \boldsymbol{\mu}_1 - \boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1}\boldsymbol{\mu}_2,\ E(\mathbf{X}_2) = \boldsymbol{\mu}_2,$$

dan matriks kovariansnya adalah

$$\begin{bmatrix} \mathbf{I}_m & -\boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1} \\ \mathbf{0} & \mathbf{I}_p \end{bmatrix} \begin{bmatrix} \boldsymbol{\Sigma}_{11} & \boldsymbol{\Sigma}_{12} \\ \boldsymbol{\Sigma}_{21} & \boldsymbol{\Sigma}_{22} \end{bmatrix} \begin{bmatrix} \mathbf{I}_m & \mathbf{0} \\ -\boldsymbol{\Sigma}_{22}^{-1}\boldsymbol{\Sigma}_{21} & \mathbf{I}_p \end{bmatrix} = \begin{bmatrix} \boldsymbol{\Sigma}_{11} - \boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1}\boldsymbol{\Sigma}_{21} & \mathbf{0} \\ \mathbf{0} & \boldsymbol{\Sigma}_{22} \end{bmatrix}.$$

Menggunakan Teorema 5.3.3, diperoleh bahwa vektor-vektor acak $\mathbf{W}$ dan $\mathbf{X}_2$ bebas. Jadi sebaran bersyarat dari $\mathbf{W} \mid \mathbf{X}_2$ sama dengan sebaran marginal dari $\mathbf{W}$, yaitu

$$\boldsymbol{W} \sim N_m(\mu_1 - \boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1}\mu_2, \boldsymbol{\Sigma}_{11} - \boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1}\boldsymbol{\Sigma}_{21}).$$

Karena bebas, maka sebaran bersyarat dari $\mathbf{X}_1 = \mathbf{W} + \boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1}\mathbf{X}_2$ bila $\mathbf{X}_2$ adalah

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

ini seperti hasil yang kita inginkan.

Menggunakan Teorema 5.3.4, maka sebaran bersyarat dari $\mathbf{X}_1$ bila $\mathbf{X}_2 = x_2$ adalah

$$N\left[\mu_2 + \rho\frac{\sigma_2}{\sigma_1}(x_1 - \mu_1), \sigma_2^2(1-\rho^2)\right].$$

Secara sama diperoleh
sebaran bersyarat dari $\mathbf{X}_2$ bila $\mathbf{X}_1 = x_1$ adalah

$$N\left[\mu_1 + \rho\frac{\sigma_1}{\sigma_2}(x_2 - \mu_2), \sigma_1^2(1-\rho^2)\right].$$

Latihan 5.3

5.3.1. Misal peubah acak X menyatakan waktu hidup, dalam bulan, suatu baterai, dan anggap $X \sim N(60{,}25)$. Misal 25 baterai diambil dari populasi ini. Tentukan konstanta c sehingga peluang purata waktu hidup 25 baterai tersebut lebih dari c adalah 95%. Bandingkan hasil ini dengan hasil pada Contoh 5.3.2. Berikan penjelasan.

5.3.2. Misal peubah acak X berat dalam kilogram suatu dos yang berisi minuman, di mana $X \sim N(20{,}4)$. Berapakah peluang bahwa 50 dos yang diambil dari populasi ini kurang dari 1.2 ton?

5.3.3 Diberikan dua sampel acak yang bebas $X_1, \ldots, X_{n_1}$ dan $Y_1, \ldots, Y_{n_2}$, beturut-turut dari populasi bersebaran normal, $X_i \sim N(\mu_1, \sigma_1^2)$ dan $Y_i \sim N(\mu_2, \sigma_2^2)$. Tentukan sebaran dari $\bar{X}_{n_1} - \bar{Y}_{n_2}$.

5.3.4. Misal X_1, X_2 sampel acak dari sebaran normal, $X_i \sim N(\mu, \sigma^2)$, dan misal $Y_1 = X_1 + X_2$, dan $Y_2 = X_1 - X_2$. Tunjukkan bahwa Y_1 dan Y_2 bebas, masing-masing bersebaran normal.

5.3.5. Misal X_1 dan X_2 bebas berturut-turut dari sebaran $N(6,1)$, dan $N(7,1)$. Hitung:

(a) $P[X_1 > X_2]$. *Petunjuk*: Tulis $P[X_1 > X_2] = P[X_1 - X_2 > 0]$, kemudian tentukan sebaran dari $X_1 - X_2$.

(b) $P[X_1 + 2X_2 \leq 15]$.

5.3.6. Tunjukkan persamaan (5.3.3), yaitu $\Sigma = \Gamma^t \Lambda \Gamma = \sum_{i=1}^{n} \lambda_i \mathbf{v}_i \mathbf{v}_i^t$.

5.3.7. Tunjukkan persamaan (5.3.4), yaitu $\left(\Sigma^{1/2}\right)^{-1} = \Gamma^t \Lambda^{-1/2} \Gamma$.

5.3.8. Misal X_1 dan X_2 bersebaran normal bivariat, berturut-turut dengan parameter-parameter $\mu_1 = 2.8$, $\mu_2 = 110$, $\sigma_1^2 = 0.16$, $\sigma_2^2 = 100$, dan $\rho = 0.6$. Hitung:

(a) $P(106 < X_2 < 124)$.

(b) $P(106 < X_2 < 124 \mid X = 3.2)$.

5.3.9. Misal $\mathbf{X} = (X_1, X_2, X_3)^t$ mempunyai sebaran normal multivariat, $\boldsymbol{X} \sim N_3(\mathbf{0}, \boldsymbol{\Sigma})$, di mana

$$\boldsymbol{\Sigma} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 1 \\ 0 & 1 & 2 \end{bmatrix}.$$

Hitung $P(X_1 > X_2 + X_3 + 2)$. *Petunjuk:* Tentukan vektor $\mathbf{a}$ sehingga $\mathbf{a}^t \mathbf{X} = X_1 - X_2 - X_3$.

5.3.10. Misal X, Y, dan Z mempunyai pdf bersama

$$f(x,y,z)=\left(\frac{1}{2\pi}\right)^{3/2}\exp\left(-\frac{x^2+y^2+z^2}{2}\right)\left[1+xyz\exp\left(-\frac{x^2+y^2+z^2}{2}\right)\right].$$

Jelas bahwa $X, Y,$ dan Z tidak bebas. Tunjukkan bahwa masing-masing pasangan dari peubah-peubah acak tersebut mempunyai sebaran normal bivariat.

5.3.11. Misal $\mathbf{X}=(X_1,X_2,...,X_n)^t$ mempunyai sebaran normal multivariat, $\boldsymbol{X}\sim N_n(\mathbf{0},\mathbf{\Sigma})$, dan misal $\bar{X}=n^{-1}\sum_{i=1}^{n}X_i$.

(a) Tulis $\bar{X}$ sebagai $\mathbf{a}^t\mathbf{X}$ untuk vektor $\mathbf{a}$ yang sesuai dan gunakan Teorema 5.3.2 untuk menentukan sebaran dari $\bar{X}$.

(b) Tentukan sebaran dari $\bar{X}$, jika semua komponen dari X_i mempunyai purata μ.

5.4 Sebaran Khi Kuadrat

Sebaran khi kuadrat merupakan bentuk khusus dari sebaran gamma, seperti disajikan pada definisi berikut ini.

Definisi 5.4.1 Misal Y peubah acak bersebaran gamma dengan parameter $\theta=2$ dan $\kappa=r/2$. Maka peubah acak Y dikatakan ber***sebaran khi kuadrat*** dengan derajat bebas r, ditulis $Y\sim\chi^2(r)$.

Jika $Y\sim\chi^2(r)$, maka cdf dari Y ditulis $H(y;r)$. Nilai persentil ke $\gamma 100\%$ dari Y, dinyatakan sebagai $\chi^2_\gamma(r)$ dapat diperoleh dari tabel sebaran khi kuadrat atau perangkat lunak statistik, dalam buku ini menggunakan Minitab. Dengan kata lain kita dapat mencari $\chi^2_\gamma(r)$ jika r dan γ diketahui. Persentil ke $\gamma 100\%$ dari Y, dalam bentuk peluang ditulis sebagai

$$P\left[Y\le\chi^2_\gamma(r)\right]=\gamma.$$

Sifat-sifat pada sebaran gamma berlaku pula pada sebaran khi kuadrat, dengan mengganti parameter $\theta=2$ dan $\kappa=r/2$. Seperti pada Teorema 5.4.1 berikut ini.

Teorema 5.4.1 Misal $Y \sim \chi^2(r)$. Maka

(i) $f(y;r) = \dfrac{1}{2^{r/2}\Gamma(r/2)} x^{\frac{r}{2}-1} e^{-\frac{y}{2}} I(y>0).$

(ii) $E(Y) = r.$

(iii) $Var(Y) = 2r.$

(iv) $M_Y(t) = (1-2t)^{-r/2}, t < \frac{1}{2}.$

(v) $E(Y^k) = 2^k \dfrac{\Gamma\left(\frac{r}{2}+k\right)}{\Gamma\left(\frac{r}{2}\right)},\ k > -r/2.$

Contoh 5.4.1 Misal $Y \sim \chi^2(10)$. Tentukan:

(i) pdf dari Y

(ii) $E(Y)$

(iii) $Var(Y)$

(iv) $M_Y(t)$

(v) $E(Y^4)$

(vi) $P[Y \le 15]$

(vii) $P[8 \le Y \le 15]$

Penyelesaian

(i) $f(y;10) = \dfrac{1}{2^5\Gamma(5)} y^{5-1} e^{-y/2} I(y>0)$

(ii) $E(Y) = 10$

(iii) $Var(Y) = 20$

(iv) $M_Y(t) = (1-2t)^{-5}, t < \frac{1}{2}$

(v) $E(Y^4) = 2^4 \dfrac{\Gamma(9)}{\Gamma(5)}$

(vi) $P[Y \le 15] = H(15;10) = 0.8679,$ diperoleh dari Minitab.

(vii) $P[8 \le Y \le 15] = P[Y \le 15] - P[Y < 8] = H(15;10) - H(8;10)$

$= 0.8679 - 0.3712 = 0.4968,$

diperoleh dari Minitab.

Teorema 5.4.2 berikut ini memberi fasilitas kepada kita untuk mengubah peubah acak yang bersebaran gamma menjadi sebaran khi kuadrat.

Teorema 5.4.2 Misal $X \sim GAM(\theta, \kappa)$. Maka $Y = \frac{2X}{\theta} \sim \chi^2(2\kappa)$.

Bukti

Kita tentukan mgf dari Y seperti berikut

$$M_Y(t) = E\left[e^{tY}\right] = E\left[e^{t\frac{2X}{\theta}}\right] = E\left[e^{\frac{2t}{\theta}X}\right]$$

$$= M_X\left(\frac{2t}{\theta}\right) = \left(1 - \theta\frac{2t}{\theta}\right)^{-\kappa} = (1-2t)^{-\frac{2\kappa}{2}}.$$

Ini merupakan mgf dari peubah acak yang bersebaran khi kuadrat dengan derajat bebas 2κ.

Contoh 5.4.2 Misal waktu hidup (dalam tahun) komponen elektronik tertentu mengikuti sebaran gamma dengan parameter $\theta = 3$, dan $\kappa = 4$. Tentukan peluang bahwa komponen tersebut hidup kurang dari 5 tahun.

Penyelesaian

Misal X menyatakan waktu hidup komponen. Menggunakan Teorema 2 kita peroleh

$$P[X \le 5] = P\left[\frac{2X}{3} \le \frac{2 \cdot 5}{3}\right] = P\left[\frac{2X}{3} \le 3.33\right] = H(3.33; 8) = 0.0880,$$

dari Minitab.

Teorema berikut ini menyajikan hubungan antara sebaran normal baku dengan sebaran khi kuadrat.

Teorema 5.4.3 Misal $Z \sim N(0,1)$. Maka $Z^2 \sim \chi^2(1)$.

Bukti

Seperti pada pembuktian teorema sebelumnya, kita tentukan mgf dari Z^2, yaitu

$$M_{Z^2}(t) = E\left[e^{tZ^2}\right] = \int_{-\infty}^{\infty} e^{tz^2} f(z)dz = \int_{-\infty}^{\infty} e^{tz^2} \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}} dz$$

$$= \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{tz^2 - \frac{z^2}{2}} dz = \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{(1-2t)z^2}{2}} dz.$$

Misal $\sqrt{(1-2t)}z = y.$ Maka $\sqrt{(1-2t)}dz = dy.$ Atau $dz = \frac{1}{\sqrt{(1-2t)}} dy.$

Sehingga kita peroleh

$$M_{Z^2}(t) = \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{y^2}{2}} \frac{1}{\sqrt{(1-2t)}} dy = \frac{1}{\sqrt{(1-2t)}} \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{y^2}{2}} dy.$$

Kita ketahui bahwa $\frac{1}{\sqrt{2\pi}} e^{-\frac{y^2}{2}}$ merupakan pdf dari peubah acak yang bersebaran normal baku, sehingga

$$\int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{y^2}{2}} dy = 1.$$

Jadi

$$M_{Z^2}(t) = \frac{1}{\sqrt{(1-2t)}}.1 = \frac{1}{\sqrt{(1-2t)}} = (1-2t)^{-\frac{1}{2}}.$$

Ini merupakan mgf dari sebaran khi kuadrat dengan derajat bebas 1.

Akibat 5.4.1 Misal $X \sim N(\mu, \sigma^2)$. Maka $Y = \frac{(X-\mu)^2}{\sigma^2} \sim \chi^2(1)$.

Bukti dari Akibat 5.4.1 kita peroleh menggunakan Teorema 3.7.1 dan Teorema 5.4.3.

Teorema berikutnya menyatakan bahwa jumlah dari peubah acak yang bersebaran khi kuadrat juga bersebaran khi kuadrat.

Teorema 5.4.4 Misal peubah-peubah acak $Y_1, Y_2, ..., Y_n$ bebas, $Y_i \sim \chi^2(r_i)$. Maka

$$V = \sum_{i=1}^{n} Y_i \sim \chi^2\left(\sum_{i=1}^{n} Y_i\right).$$

Bukti

Kita tentukan mgf dari V, seperti berikut

$$M_V(t) = E\left[e^{tV}\right] = E\left[\exp\left(t\sum_{i=1}^{n} Y_i\right)\right]$$

$$= E\left[\exp\left(tY_1 + ... + tY_n\right)\right] = E\left[\exp\left(tY_1\right)...\exp\left(tY_n\right)\right].$$

Karena peubah-peubah acak $Y_1, Y_2, ..., Y_n$ bebas, maka

$$E\left[\exp\left(tY_1\right)...\exp\left(tY_n\right)\right] = E\left[\exp\left(tY_1\right)\right]...E\left[\exp\left(tY_n\right)\right]$$

Sehingga

$$M_V = M_{Y_1}(t)...M_{Y_n}(t) = (1-2t)^{-\frac{r_1}{2}}...(1-2t)^{-\frac{r_n}{2}} = (1-2t)^{-\frac{\sum_{i=1}^{n} r_i}{2}},$$

yang merupakan mgf dari peubah acak yang bersebaran $\chi^2\left(\sum_{i=1}^{n} r_i\right)$.

Berdasarkan Akibat 5.4.1 dan Teorema 5.4.4 kita memperoleh Akibat 5.4.2 berikut.

Akibat 5.4.2 Misal $X_1, X_2, ..., X_n$ sampel acak dari sebaran normal, $X_i \sim N(\mu, \sigma^2), i = 1,2, ..., n$. Maka

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

(ii) $\left(\frac{\bar{X}_n-\mu}{\sigma/\sqrt{n}}\right)^2 \sim \chi^2(1)$.

Pembahasan terakhir dari subpokok bahasan ini adalah menentukan sebaran dari $(n-1)S^2/\sigma^2$. Lebih rinci kita sajikan dalam teorema berikut.

Teorema 5.4.5 Misal $X_1, X_2, ..., X_n$ sampel acak berukuran n dari sebaran $N(\mu, \sigma^2)$. Maka

(i) $\bar{X}_n$ dan suku-suku $X_i - \bar{X}_n, i = 1, 2, ..., n$ adalah bebas

(ii) $\bar{X}_n$ dan $S_n^{\ 2}$ adalah bebas

(iii) $\frac{(n-1)S_n^2}{\sigma^2} \sim \chi^2(n-1)$.

Bukti

Kita buktikan menggunakan sifat-sifat pada sebaran normal multivariat. Mula-mula kita misalkan $\mathbf{X} = \left(X_1, X_2, ..., X_n\right)^t$. Karena $X_1, ..., X_n$ sampel acak dari sebaran normal, $X_i \sim N(\mu, \sigma^2), i = 1,2, ..., n$, maka $\boldsymbol{X} \sim N_n(\mu\mathbf{1}, \sigma^2\boldsymbol{I})$, di mana $\mathbf{1}$ menyatakan vektor yang semua komponennya adalah 1. Misal $\mathbf{v}^t = \left(1/n, 1/n, ..., 1/n\right)^t = \left(1/n\right)\mathbf{1}^t$, maka $\bar{X}_n = \mathbf{v}^t\mathbf{X}$. Kita definisikan vektor $\mathbf{Y}$ sebagai $\mathbf{Y} = \left(X_1 - \bar{X}_n, X_2 - \bar{X}_n, ..., X_n - \bar{X}_n\right)^t$. Kita definisikan transformasi:

$$\mathbf{W} = \begin{bmatrix} \bar{X}_n \\ \mathbf{Y} \end{bmatrix} = \begin{bmatrix} \mathbf{v}^t \\ \mathbf{I} - \mathbf{Iv}^t \end{bmatrix} \mathbf{X}.$$

Karena $\mathbf{W}$ adalah transformasi linear dari sebaran normal multivariat, maka menggunakan Teorema 5.3.2 dia mempunyai sebaran normal multivariat dengan purata

$$E\left[\mathbf{W}\right] = \begin{bmatrix} \mathbf{v}^t \\ \mathbf{I} - \mathbf{1v}^t \end{bmatrix} \mu\mathbf{1} = \begin{bmatrix} \mu \\ \mathbf{0}_n \end{bmatrix},$$

di mana $\mathbf{0}_n$ menyatakan vektor yang semua komponennya adalah 0, dan matriks kovarians

$$\mathbf{\Sigma} = \begin{bmatrix} \mathbf{v}^t \\ \mathbf{I} - \mathbf{1v}^t \end{bmatrix} \sigma^2\mathbf{I} \begin{bmatrix} \mathbf{v}^t \\ \mathbf{I} - \mathbf{1v}^t \end{bmatrix}^t = \sigma^2 \begin{bmatrix} \frac{1}{n} & \mathbf{0}_n^t \\ \mathbf{0}_n & \mathbf{I} - \mathbf{1v}^t \end{bmatrix}.$$

Karena $\underset{\cdot}{X}$ adalah komponen pertama dari $\mathbf{W}$, maka kita juga mendapatkan Akibat 5.3.1. Selanjutnya, karena kovarians-kovariansnya adalah 0, maka kita peroleh bagian (i). Dengan menyatakan S^2 sebagai $(n-1)^{-1}\mathbf{Y}^t\mathbf{Y}$, maka kita peroleh bagian (ii). Untuk membuktikan bagian (iii), kita perhatikan peubah acak

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

Menurut Akibat 5.4.2 bagian (i), $V \sim \chi^2(n)$. Menggunakan bagian (ii), maka kita peroleh dua suku pada ruas kanan persamaan terakhir adalah bebas. Menurut Akibat 5.4.2 bagian (ii), suku kedua ruas kanan persamaan terakhir bersebaran $\chi^2(1)$. Oleh karena itu mgf kedua sisi dapat kita nyatakan sebagai

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

Kita peroleh mgf dari $\dfrac{(n-1)S_n^{\,2}}{\sigma^2}$ sebagai

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

yang merupakan mgf dari peubah acak bersebaran $\chi^2(n-1)$. Hasil ini membuktikan bagian (iii).

Contoh 5.4.3 Misal X menyatakan waktu hidup baterai (dalam bulan) anggap bersebaran normal, $X \sim N(24,16)$. Misal 30 baterai diambil secara acak dari sebaran ini. Tentukan peluang bahwa variansnya lebih dari 20.

Penyelesaian

Kita gunakan Teorema 5.4.5 (iii) untuk menyelesaikannya, yaitu

$$\begin{aligned}
P\left[S_n^{\,2} \ge 20\right] = P\left[\frac{29S_n^{\,2}}{16} \ge \frac{29\times 20}{16}\right] &= 1-P\left[\frac{29S_n^{\,2}}{16} < 36.25\right] \\
&= 1-H(36.25;29) = 1-0.8336 = 0.1664.
\end{aligned}$$

Jadi peluang bahwa variansnya lebih dari 20 adalah 0.1664.

Latihan 5.4

5.4.1. Misal $Z \sim N(0,1)$. Hitung:
(a) $P[Z^2 < 3.84]$ menggunakan tabel sebaran normal,
(b) $P[Z^2 < 3.84]$ menggunakan tabel sebaran khi kuadrat.
(c) Beri komentar mengenai hasil (a) dan (b).

5.4.2. Misal $Z_1, ..., Z_n$ sampel acak dari sebaran normal baku, dan misal $\bar{Z}_n$ purata sampel tersebut. Hitung:
(a) $P\left[\sum_{i=1}^{16} Z_i^2 < 32\right]$.
(b) $P\left[\sum_{i=1}^{16} \left(Z_i - \bar{Z}_{16}\right)^2 < 32\right]$.

5.4.3. Anggap waktu hidup suatu komponen adalah bebas dan mengikuti sebaran eksponensial, $T_i \sim EXP(100)$.
(a) Apakah sebaran dari $\sum_{i=1}^{10} T_i$?
(b) Menggunakan transformasi khi kuadrat, hitung $P\left[\sum_{i=1}^{10} T_i \geq 1.2\right]$.

5.4.4. Ulangi Latihan 5.4.3 untuk $T_i \sim GAM(100, 1.2)$.

5.4.5. Anggap $X \sim \chi^2(m)$, $W = X + Y \sim \chi^2(m+n)$, di mana X dan Y bebas. Tunjukkan bahwa $W - X \sim \chi^2(n)$.

5.4.6. Misal $X_1, ..., X_{10}$ sampel acak dari sebaran normal, $X_i \sim N(12, 9)$. Tentukan:
(a) $P\left[10 < \bar{X}_n < 14\right]$
(b) $P\left[8 < S_n^2 < 10\right]$

5.4.7. Misal X_1, X_2 sampel acak bebas dari sebaran normal, $X_i \sim N(\mu, \sigma^2)$, $Y_1 = X_1 + X_2$, dan $Y_2 = X_1 - X_2$. Tunjukkan bahwa Y_1 dan Y_2 bebas.

5.4.8. Misal $Y \sim \chi^2(6)$. Hitung $P\left[\frac{Y}{1+Y} > \frac{11}{16}\right]$.

1.4.9. Misal vektor acak $\boldsymbol{X} \sim N_n(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, di mana $\boldsymbol{\Sigma}$ matriks definit positif. Buktikan bahwa vektor acak $W = (\boldsymbol{X} - \boldsymbol{\mu})^t \boldsymbol{\Sigma}^{-1} (\boldsymbol{X} - \boldsymbol{\mu}) \sim \chi^2(n)$.

5.5 Sebaran *t*, *F*, dan *Beta*

Tujuan dalam subpokok bahasan ini adalah mengkaji dua sebaran yang sangat penting dalam permasalahan inferens statistis, yaitu sebaran *t* (Student) dan sebaran *F*. Pembahasan dilengkapi dengan transformasi dari sebaran *F* yang disebut sebagai sebaran *beta*.

Misal Z peubah acak dari sebaran $N(0,1)$, dan V peubah acak dari sebaran $\chi^2(r)$, di mana Z dan V bebas. Maka pdf bersama dari Z dan V, katakanlah $\varphi(z,v)$, adalah perkalian dari pdf Z dan pdf V, yaitu

$$\varphi(z,v) = \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}} \frac{1}{\Gamma\left(\frac{r}{2}\right) 2^{\frac{r}{2}}} v^{\frac{r}{2}-1} e^{-\frac{v}{2}} I(-\infty < z < \infty, 0 < v < \infty).$$

Sekarang kita definisikan peubah acak T sebagai

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

Kita akan menggunakan sifat di atas untuk menentukan pdf dari T. Mula-mula kita bentuk persamaan

$$t = \frac{z}{\sqrt{v/r}} \text{ dan } u = v.$$

Persamaan ini mendefinisikan transformasi satu-satu dan pada dari

$$\mathfrak{A} = \{(z,v) : -\infty < z < \infty, 0 < v < \infty\}$$

ke

$$\mathfrak{M} = \{(t,u) : -\infty < t < \infty, 0 < u < \infty\}.$$

Kebalikan transformasi ini adalah

$$z = \frac{t\sqrt{u}}{\sqrt{r}} \text{ dan } v = u.$$

Oleh karena itu kita peroleh Jacobi

$$J = \begin{vmatrix} \frac{\partial z}{\partial t} & \frac{\partial z}{\partial u} \\ \frac{\partial v}{\partial t} & \frac{\partial v}{\partial u} \end{vmatrix} = \begin{vmatrix} \frac{\sqrt{u}}{\sqrt{r}} & \frac{t}{2\sqrt{ur}} \\ 0 & 1 \end{vmatrix} = \frac{\sqrt{u}}{\sqrt{r}}.$$

Misal $g(t,u)$ pdf bersama dari T dan U. Maka

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

Oleh karena itu pdf marginal dari T adalah

$$g_1(t) = \int_{-\infty}^{\infty} g(t,u)du = \int_0^{\infty} \frac{1}{\sqrt{2\pi r}\Gamma(r/2)2^{r/2}} u^{\frac{r+1}{2}-1} \exp\left[-\frac{u}{2}\left(1+\frac{t^2}{r}\right)\right] du.$$

Jika kita misalkan $y = \frac{u\left[1+\left(t^2/r\right)\right]}{2}$, maka kita peroleh

$$\begin{aligned} g_1(t) &= \int_0^{\infty} \frac{1}{\sqrt{2\pi r}\Gamma(r/2)2^{r/2}} \left(\frac{2y}{1+t^2/r}\right)^{\frac{r+1}{2}-1} e^{-y} \left(\frac{2}{1+t^2/r}\right) dy \\ &= \frac{\Gamma\left[(r+1)/2\right]}{\sqrt{\pi r}\Gamma(r/2)} \frac{1}{(1+t^2/r)^{\frac{r+1}{2}}} I(-\infty < t < \infty). \end{aligned}$$

Hasil di atas kita ringkas dalam teorema berikut ini.

Teorema 5.5.1 Misal $Z \sim N(0,1)$ dan $V \sim \chi^2(r)$, Z dan V bebas. Maka pdf dari

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

adalah

$$f(t;r) = \frac{\Gamma\left[(r+1)/2\right]}{\sqrt{\pi r}\,\Gamma(r/2)} \frac{1}{(1+t^2/r)^{\frac{r+1}{2}}} I(-\infty < t < \infty).$$

Peubah acak T pada Teorema 5.5.1 dikatakan ber***sebaran t*** dengan derajat bebas r, ditulis $T \sim t(r)$. Seperti halnya pada sebaran khi kuadrat, nilai persentil ke $\gamma 100\%$ dari T, dinyatakan sebagai $t_\gamma(r)$ dapat diperoleh dari tabel sebaran t atau perangkat lunak statistik, dalam buku ini menggunakan Minitab.

Contoh 5.5.1 Misal T peubah acak yang bersebaran t dengan derajat bebas 12. Hitung $P(|T| > 2.30)$.

Penyelesaian

$$\begin{aligned} P(|T| > 2.30) &= P(T < -2.30 \text{ atau } T > 2.30) = P(T < -2.30) + P(T > 2.30) \\ &= P(T < -2.30) + [1 - P(T \le 2.30)] \\ &= 0.0201 + (1 - 0.9799) = 0.0402. \end{aligned}$$

diperoleh dari Minitab.

Salah satu terapan dari sebaran t adalah untuk menduga parameter μ berdasarkan sampel dari sebaran normal jika σ^2 tidak diketahui. Hal ini diberikan oleh teorema berikut ini.

Teorema 5.5.2 Misal $X_1, ..., X_n$ sampel acak dari sebaran $N(\mu, \sigma^2)$. Maka

$$\frac{\bar{X}_n - \mu}{S_n/\sqrt{n}} \sim t(n-1).$$

Bukti

Peubah acak di atas kita tulis kembali dalam bentuk berikut

$$\frac{\bar{X}_n - \mu}{S_n / \sqrt{n}} = \frac{\sqrt{n}\left(\bar{X}_n - \mu\right)/\sigma}{\sqrt{\dfrac{(n-1)S_n^{\ 2}/\sigma^2}{(n-1)}}}.$$

Dari Akibat 5.3.2 diperoleh pembilang pada ruas kanan persamaan bersebaran $N(0,1)$. Dari Torema 5.4.5 diperoleh akar penyebut pada ruas kanan persamaan bersebaran $\chi^2(n-1)$ dibagi dengan derajat bebasnya. Oleh karena itu menggunakan Teorema 5.5.1 kita peroleh

$$\frac{\bar{X}_n - \mu}{S_n/\sqrt{n}} \sim t(n-1).$$

Bukti selesai.

Selanjutnya kita tentukan purata dan varians dari T seperti berikut ini. Kita tulis

$$T = Z\left(V/r\right)^{-1/2},$$

di mana $Z \sim N(0,1)$, $V \sim \chi^2(r)$, Z dan V bebas. Oleh karena itu momen ke k dari T adalah

$$E\left(T^k\right) = E\left[Z^k\left(\frac{V}{r}\right)^{-k/2}\right] = E\left(Z^k\right)E\left[\left(\frac{V}{r}\right)^{-k/2}\right]$$

$$= E\left(Z^k\right)\frac{2^{-k/2}\Gamma\left(\frac{r}{2}-\frac{k}{2}\right)}{\Gamma\left(\frac{r}{2}\right)r^{-k/2}},\ k < r.$$

Jika $k = 1$, maka $E(Z) = 0$. Jadi purata dari T adalah nol, atau $E(T) = 0$. Untuk menentukan varians kita pilih $k = 2$. Dalam hal ini tentunya $r > 2$. Karena $E(Z^2) = 1$, maka

$$Var(T) = E\left(T^2\right) = \frac{r}{r-2}.$$

Jadi jika $T \sim t(r)$, di mana $r > 2$, maka purata dan varians dari T berturut-turut adalah nol dan $\dfrac{r}{r-2}$.

Contoh 5.5.2 Anggap sebaran dari nilai ujian Statitistika Matematis adalah normal dengan purata 70 dan simpangan baku tidak diketahui. Misal sampel acak berukuran 40 diambil dari populasi ini dan diketahui variansnya 16. Hitung peluang bahwa purata sampel acaknya kurang dari 85.

Penyelesaian Misal $\bar{X}_n$ menyatakan purata sampel acak dari ukuran 40. Maka

$$P\left[\bar{X}_n < 85\right] = P\left[\frac{\sqrt{40}(\bar{X}_n - 70)}{4} < \frac{\sqrt{40}(85-70)}{4}\right]$$
$$= P\left[T < 23.72\right] = 1,$$

dari Minitab.

Sebaran berikutnya subpokok bahasan ini adalah sebaran F. Kita perhatikan dua peubah acak bebas U dan V berturut-turut dari sebaran khi kuadrat dengan derajat bebas r_1 dan r_2. Misal $\varphi(u,v)$ pdf bersama dari U dan V. Maka

$$\varphi(u,v) = \frac{1}{\Gamma(r_1/2)\Gamma(r_2/2)2^{(r_1+r_2)/2}} u^{\frac{r_1}{2}-1} v^{\frac{r_2}{2}-1} e^{-\frac{u+v}{2}} I(0 < u < \infty, 0 < v < \infty).$$

Kita definiskan peubah acak

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

Kita akan menentukan $g_1(w)$ yang merupakan pdf marginal W dari pdf bersama w dan z, di mana

$$w = \frac{u/r_1}{v/r_2} \text{ dan } h = v.$$

Persamaan ini mendefinisikan transformasi satu-satu dan pada dari

$$\mathfrak{A} = \left\{(u,v) : 0 < u < \infty, 0 < v < \infty\right\}$$

ke

$$\mathfrak{M} = \left\{(w,h) : 0 < w < \infty, 0 < h < \infty\right\}.$$

Kebalikan transformasi ini adalah

$$u = \frac{r_1 h}{r_2} w \text{ dan } v = h.$$

Oleh karena itu kita peroleh Jacobi

$$J = \begin{vmatrix} \frac{\partial u}{\partial w} & \frac{\partial u}{\partial h} \\ \frac{\partial v}{\partial w} & \frac{\partial v}{\partial h} \end{vmatrix} = \begin{vmatrix} \frac{r_1 h}{r_2} & \frac{r_1}{r_2} w \\ 0 & 1 \end{vmatrix} = \frac{r_1 h}{r_2}.$$

Misal $g(w,h)$ pdf bersama dari W dan H. Maka

$$g(w,h) = \frac{1}{\Gamma(r_1/2)\Gamma(r_2/2)2^{(r_1+r_2)/2}} \left(\frac{r_1 h}{r_2} w \right)^{\frac{r_1}{2}-1} \times$$

$$h^{\frac{r_2}{2}-1} e^{-\frac{h}{2}\left(\frac{r_1}{r_2}w+1\right)} \frac{r_1}{r_2} h \; I\left[(w,h) \in \mathrm{B}\right].$$

Oleh karena itu pdf marginal $g_1(w)$ dari W adalah

$$g_1(w) = \int_{-\infty}^{\infty} g(w,h)\,dh$$

$$= \int_0^{\infty} \frac{1}{\Gamma(r_1/2)\Gamma(r_2/2)2^{(r_1+r_2)/2}} \left(\frac{r_1 h}{r_2} w \right)^{\frac{r_1}{2}-1} h^{\frac{r_2}{2}-1} e^{-\frac{h}{2}\left(\frac{r_1}{r_2}w+1\right)} \frac{r_1}{r_2} h \; dh$$

$$= \int_0^{\infty} \frac{\left(r_1/r_2\right)^{\frac{r_1}{2}} w^{\frac{r_1}{2}-1}}{\Gamma(r_1/2)\Gamma(r_2/2)2^{(r_1+r_2)/2}} h^{\frac{r_1+r_2}{2}-1} e^{-\frac{h}{2}\left(\frac{r_1}{r_2}w+1\right)} \; dh.$$

Misal

$$y = \frac{h}{2}\left(\frac{r_1}{r_2}w+1\right).$$

Oleh karena itu

$$g_1(w) = \int_0^{\infty} \frac{\left(r_1/r_2\right)^{\frac{r_1}{2}} w^{\frac{r_1}{2}-1}}{\Gamma(r_1/2)\Gamma(r_2/2)2^{(r_1+r_2)/2}} \left(\frac{2y}{\frac{r_1}{r_2}w+1} \right)^{\frac{r_1+r_2}{2}-1} e^{-y} \left(\frac{2}{\frac{r_1}{r_2}w+1} \right) dy$$

$$= \frac{\Gamma\left(\frac{r_1+r_2}{2}\right)\left(r_1/r_2\right)^{\frac{r_1}{2}}}{\Gamma(r_1/2)\Gamma(r_2/2)} \frac{w^{\frac{r_1}{2}-1}}{\left(\frac{r_1}{r_2}w+1\right)^{\frac{r_1+r_2}{2}}} \; I(0 < w < \infty).$$

Hasil di atas kita ringkas dalam Teorema 5.5.3 berikut ini

Teorema 5.5.3 Misal $U \sim \chi^2(r_1)$ dan $V \sim \chi^2(r_2)$ U dan V bebas. Maka pdf dari

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

adalah

$$\phi(w; r_1, r_2) = \frac{\Gamma\left(\frac{r_1 + r_2}{2}\right)(r_1 / r_2)^{\frac{r_1}{2}}}{\Gamma(r_1 / 2)\Gamma(r_2 / 2)} \frac{w^{\frac{r_1}{2}-1}}{\left(\frac{r_1}{r_2} w + 1\right)^{\frac{r_1+r_2}{2}}} I(0 < w < \infty).$$

Peubah acak W pada Teorema 5.5.3 dikatakan ber***sebaran F*** dengan derajat bebas r_1 dan r_2, ditulis $W \sim F(r_1, r_2)$. Seperti halnya pada sebaran-sebaran sebelumnya, nilai persentil ke $\gamma 100\%$ dari W, dinyatakan sebagai $F_\gamma(r_1, r_2)$ dapat diperoleh dari tabel sebaran F atau perangkat lunak statistik, dalam buku ini menggunakan Minitab.

Karena peubah acak W merupakan perbandingan dari dua peubah acak yang bersebaran khi kuadrat, maka peubah acak $1/W$ juga merupakan perbandingan dari dua peubah acak yang bersebaran khi kuadrat. Oleh karena itu dia juga bersebaran F. Hal ini diberikan oleh teorema berikut, yang buktinya ditinggalkan sebagai latihan.

Teorema 5.5.4 Jika $W \sim F(r_1, r_2)$, maka $\frac{1}{W} \sim F(r_2, r_1)$.

Teorema di atas berakibat pada kenyataan berikut ini.

Akibat 5.5.1 Misal $F_\gamma(r_1, r_2)$ persentil ke $\gamma 100\%$ dari $W \sim F(r_1, r_2)$ yaitu

$$P\left[W \le F_\gamma\left(r_1, r_2\right)\right] = \gamma.$$

Maka persentil ke $(1-\gamma)100\%$ dari $W \sim F(r_1, r_2)$ adalah

$$F_{1-\gamma}(r_1, r_2) = \frac{1}{F_\gamma(r_2, r_1)}.$$

Bukti
Mula–mula kita nyatakan $F_{1-\gamma}(r_1, r_2)$ sebagai

$$1-\gamma = P\left[W \le F_{1-\gamma}(r_1, r_2)\right] = P\left[\frac{1}{W} \ge \frac{1}{F_{1-\gamma}(r_1, r_2)}\right] = 1 - P\left[\frac{1}{W} < \frac{1}{F_{1-\gamma}(r_1, r_2)}\right].$$

Oleh karena itu

$$P\left[\frac{1}{W} < \frac{1}{F_{1-\gamma}(r_1, r_2)}\right] = \gamma.$$

Dari Teorema 5.5.4 diperoleh bahwa $\frac{1}{W} \sim F(r_2, r_1)$, sehingga

$$F_{\gamma}(r_2, r_1) = \frac{1}{F_{1-\gamma}\left(r_1, r_2\right)},$$

atau

$$F_{1-\gamma}(r_1, r_2) = \frac{1}{F_{\gamma}\left(r_2, r_1\right)}.$$

Bukti selesai.

Contoh 5.5.3 Misal W mempunyai sebaran F dengan derajat bebas $r_1 = 10$ dan $r_2 = 5$. Tentukan a dan b sehingga:

(i) $P(W \le b) = 0.9.$

(ii) $P(W \le a) = 0.1.$

(iii) $P(a \le W \le b) = 0.8.$

Penyelesaian
Nilai a dan b berikut ini ditentukan menggunakan Minitab.

(i) $b = F_{0.9}(10,5) = 3.2974$

(ii) $a = F_{0.1}(10,5) = 0.3966$. Nilai ini dapat pula diperoleh menggunakan Akibat 5.5.1, yaitu

$$a = F_{0.1}(10,5) = \frac{1}{F_{0.9}(5,10)} = \frac{1}{2.5216} = 0.3966.$$

(iii) $P(a \le W \le b) = 0.8$ dapat ditulis sebagai $P(W \le b) - P(W < a) = 0.8$. Salah satu kemungkinan nilai-nilai ini adalah $P(W \le b) = 0.9$ dan $P(W < a) = 0.1$. Jadi $b = F_{0.9}(10,5) = 3.2974$ dan $a = F_{0.1}(10,5) = 0.3966$.

Selanjutnya kita tentukan purata dari W, dengan menulisnya sebagai

$$W = \left(\frac{r_2}{r_1}\right)\left(\frac{U}{V}\right),$$

di mana U dan V bebas, berturut-turut dari sebaran $\chi^2(r_1)$ dan $\chi^2(r_2)$. Oleh karena itu momen ke k dari W adalah

$$E\left(W^k\right) = \left(\tfrac{r_2}{r_1}\right)^k E\left(U^k\right)E\left(V^{-k}\right),\ k > -(r_1/2) \text{ dan } r_2 > 2k.$$

Untuk $k = 1$, kita peroleh

$$E(W) = \frac{r_2}{r_1} r_1 \frac{\Gamma\left(\frac{r_2}{2} - 1\right)}{2\Gamma\left(\frac{r_2}{2}\right)} = \frac{r_2}{r_2 - 2}.$$

Jadi untuk r_2 cukup besar, purata dari W mendekati 1. Ditinggalkan sebagai latihan untuk menunjukkan

$$Var(W) = \frac{2r_2^2\left(r_1 + r_2 - 2\right)}{r_1\left(r_2 - 2\right)^2\left(r_2 - 4\right)}.$$

Suatu peubah acak yang bersebaran F dapat ditransformasi sehingga bersebaran *beta*. Misal $W \sim F(r_1, r_2)$ dan

$$Y = \frac{\left(r_1/r_2\right)W}{1 + \left(r_1/r_2\right)W}.$$

Ditinggalkan sebagai latihan untuk menunjukkan bahwa pdf dari Y adalah

$$f(y; r_1, r_2) = \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} y^{a-1}(1-y)^{b-1} I(0 < y < 1),$$

di mana $a = r_1/2$ dan $b = r_2/2$. Peubah acak Y yang mempunyai pdf ini disebut ber***sebaran beta*** dengan parameter-parameter $a > 0$ dan $b > 0$, dinyatakan sebagai $Y \sim VETA(a, b)$.

Purata dan varians dari Y beruturut-turut adalah

$$E(Y) = \frac{a}{a+b} \text{ dan } Var(Y) = \frac{ab}{(a+b+1)(a+b)^2},$$

ditinggalkan sebagai Latihan.

Latihan 5.5

5.5.1. Misal T mempunyai sebaran t dengan derajat bebas 10. Tentukan:

(a) $P(T \le 2.3)$.

(b) a sehingga $P(T \le a) = 0.95$.

(c) b sehingga $P(T \le b) = 0.05$.

(d) c sehingga $P(|T| \le c) = 0.90$.

5.5.2. Anggap sebaran dari nilai ujian Statitistika Matematis adalah normal dengan purata 70 dan simpangan baku tidak diketahui. Misal sampel acak berukuran 30 diambil dari populasi ini sehingga diperoleh nilai-nilai seperti berikut:

68	78	78	90	69	78	56	35	59	42	70	76	50	65	51
82	56	62	59	41	40	83	70	88	65	65	54	76	61	86

Hitung peluang bahwa purata sampel acaknya antara 60 dan 80.

5.5.3. Misal X_1 dan X_2 bebas dan $X_i \sim N(\mu_i, \sigma_i^2), 1,2$. Tentukan peubah acak T yang merupakan fungsi dari X_1 dan X_2 sehingga T bersebaran t, tentukan pula derajat bebasnya.

5.5.4. Buktikan Teorema 5.5.4.

5.5.5. Jika $T \sim t(v)$, maka tentukan sebaran dari T^2.

5.5.6. Misal W mempunyai sebaran F dengan derajat bebas 10 dan 12. Tentukan:

(a) $P(W \le 2,3)$.

(b) Nilaii a sehingga $P(W \le a) = 0.95$.

(c) Nilai b sehingga $P(W \le b) = 0.05$.

(d) Nilai c sehingga $P(c < W < d) = 0.90$.

(e) $P\left(\frac{1}{W} \le 0.25\right)$.

5.5.7. Misal X_1 dan X_2 bebas dan $X_i \sim N(\mu_i, \sigma_i^2), i = 1,2$. Tentukan peubah acak W yang merupakan fungsi dari X_1 dan X_2 sehingga W bersebaran F, tentukan pula derajat bebasnya.

5.5.8. Misal X_1 dan X_2 sampel acak dari sebaran yang mempunyai pdf $f(x) = e^{-x} I(0 < x < \infty)$. Tunjukkan bahwa $W = \dfrac{X_1}{X_2}$ mempunyai sebaran F.

5.5.9. Misal $X_1, ..., X_{n_1}$ dan $Y_1, ..., Y_{n_2}$ dua sampel acak bebas dari berturut-turut dari dua sebaran normal, $X_i \sim N(\mu_1, \sigma_1^2)$dan $Y_i \sim N(\mu_2, \sigma_2^2)$.

(a) Tunjukkan bahwa $\frac{S_1^2 \sigma_2^2}{S_2^2 \sigma_1^2} \sim F(n_1 - 1, n_2 - 1)$.

(b) Tentukan f_1 sehingga $P\left[\dfrac{S_1^2 \sigma_2^2}{S_2^2 \sigma_1^2} \le f_1\right] = 0.95$, jika $n_1 = 16$, dan $n_2 = 21$.

5.5.10. Misal $W \sim F(r_1, r_2)$. Buktikan varians dari W adalah $\dfrac{2r_2^2(r_1 + r_2 - 2)}{r_1(r_2 - 2)^2(r_2 - 4)}$.

5.5.11. Misal $W \sim F(r_1, r_2)$, dan

$$Y = \frac{(r_1 / r_2)W}{1 + (r_1 / r_2)W}.$$

Tunjukkan bahwa pdf dari Y adalah

$$f(y; r_1, r_2) = \frac{\Gamma(a + b)}{\Gamma(a)\Gamma(b)} y^{a-1}(1 - y)^{b-1} I(0 < y < 1),$$

di mana $a = r_1 / 2$ dan $b = r_2 / 2$, purata dan variansnya beturut-turut adalah

$$E(Y) = \frac{a}{a + b} \text{ dan } Var(Y) = \frac{ab}{(a + b + 1)(a + b)^2}.$$

BAB 6
SEBARAN BATAS

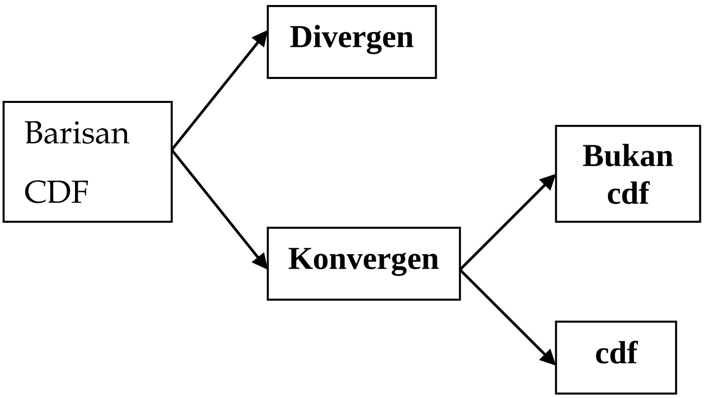
6.1 Pendahuluan

Pada bab sebelumnya kita telah melihat bahwa statistik urutan, khususnya statistik urutan terkecil, $Y_1 = X_{1,n}$, dan statistik urutan terbesar, $Y_n = X_{n,n}$ masing-masing cdf-nya bergantung pada n. Permasalahan yang muncul adalah apakah untuk n mendekati tak hingga cdf ini konvergen ke suatu cdf.

Secara umum jika kita mempunyai n peubah acak, $X_1,\ X_2,\ ...,\ X_n$, maka kita dapat melihat fenomena berikut ini.

$$\begin{array}{cccc} X_1, & X_2, & ..., & X_n \\ \updownarrow & \updownarrow & & \updownarrow \\ f_1(x), & f_2(x), & ..., & f_n(x) \\ \updownarrow & \updownarrow & & \updownarrow \\ F_1(x), & F_2(x), & ..., & F_n(x) \end{array}$$

Baris terakhir merupakan barisan cdf. Hal-hal yang mungkin dapat terjadi dari barisan cdf ini untuk n yang besar sekali atau mendekati tak hingga dapat dilihat pada bagan berikut ini.



Kenyataan di atas merupakan dasar dari gagasan tentang konsep sebaran batas dan kekonvergenan dalam sebaran.

Tujuan instruksional umum dari mempelajari bab ini adalah Anda diharapkan memahami konsep tentang sebaran batas.

Tujuan instruksional khusus dari mempelajari bab ini adalah Anda diharapkan dapat: menentukan sebaran batas barisan peubah acak, menyelidiki kekonvergenan barisan peubah acak, menentukan sebaran batas barisan peubah acak menggunakan mgf, menerapkan teorema limit pusat, dan menentukan pendekatan sebaran diskret dengan sebaran normal.

6.2 Barisan Peubah Acak

Kita perhatikan barisan peubah acak $Y_1, Y_2, Y_3, \ldots$ yang berkorespondensi dengan barisan cdf $G_1(y), G_2(y), G_3(y), \ldots$. Untuk masing-masing $n = 1, 2, 3, \ldots$ menurut definisi cdf

$$G_n(y) = P(Y_n \leq y).$$

Kekonvergenan barisan peubah acak ini jika dilihat dari barisan cdf-nya disajikan pada definisi berikut ini.

Definisi 6.2.1 Misalkan $G_n(y)$ cdf dari Y_n untuk masing-masing $n = 1, 2, 3, \ldots$ dan misalkan $G(y)$ cdf dari Y sehingga

$$\lim_{n \to \infty} G_n(y) = G(y)$$

untuk semua nilai y *sehingga* $G(y)$ *kontinu*. Maka barisan peubah acak $Y_1, Y_2, Y_3, \ldots$ dikatakan ***konvergen dalam sebaran*** ke Y, dinyatakan $Y_n \xrightarrow{d} Y$. Dan $G(y)$ disebut ***sebaran batas*** dari Y_n.

Agar lebih mudah dalam memahami definisi di atas marilah kita ikuti satu-persatu contoh-contoh berikut ini.

Contoh 6.2.1 Misal $X_1, X_2, \ldots, X_n$ sampel acak dari sebaran seragam, $X_i \sim UNIF(0,1)$, dan misal $Y_n = X_{n,n}$, yaitu statistik urutan terbesar. Tentukan sebaran batas dari $Y_n = X_{n,n}$.

Penyelesaian

Karena $X_1, X_2, ..., X_n$ sampel acak dan $X_i \sim UNIF(0,1)$, maka pdf dari X_i

$$f(x) = I(0 < x < 1).$$

Oleh karena itu cdf-nya

$$F(x) = xI(0 \le x < 1) + I(1 \le x).$$

Misal $G_n(y)$ fungsi sebaran dari Y_n. Maka menggunakan sifat cdf statistik urutan terbesar pada Bab sebelumnya, yaitu

$$G_n(y) = [F(y)]^n,$$

kita memperoleh

$$G_n(y) = y^n I(0 \le y < 1) + I(1 \le y).$$

Untuk menentukan sebaran batasnya kita limitkan $G_n(y)$. Menggunakan sifat barisan bilangan real,

$$\lim_{n\to\infty} y^n = 0, 0 < y < 1$$

kita memperoleh

$$\lim_{n\to\infty} G_n(y) = I(1 \le y).$$

Menggunakan sifat-sifat cdf kita dapat mengetahui bahwa

$$G(y) = I(1 \le y)$$

adalah cdf. Jadi Y_n mempunyai sebaran batas, dan fungsi sebaran batasnya adalah

$$G(y) = I(1 \le y).$$

Fungsi sebaran batas di atas terpusat pada satu nilai, yaitu $y = 1$. Peubah acak dengan fungsi sebaran yang demikian disebut ber***sebaran merosot*** pada $y = 1$. Lebih jelas tentang fungsi sebaran merosot, marilah kita lihat definisi berikut ini.

Definisi 6.2.2 Fungsi $G(y)$ adalah cdf dari ***sebaran merosot*** pada $y = c$, jika

$$G(y) = I(c \le y).$$

Dengan kata lain, $G(y)$ adalah cdf dari sebaran diskret yang peluangnya satu pada nilai $y = c$ dan nol untuk yang lainnya.

Contoh 6.2.2 Misal $\bar{X}_n$ mempunyai cdf

$$F_n(\bar{x}) = \int_{-\infty}^{\bar{x}} \frac{1}{\sqrt{2\pi/n}} e^{-nw^2/2} dw$$

Tentukanlah sebaran batas dari $\bar{X}_n$.

Penyelesaian

Jika kita ubah variabel dari w menjadi $v = \sqrt{n}w,$ maka diperoleh

$$dv = \sqrt{n}dw,$$

dan batas-batasnya menjadi:

(i) batas bawah: w mendekati $-\infty \rightarrow v$ mendekati $-\infty$

(ii) batas atas: $w = \bar{x} \rightarrow v = \sqrt{n}\bar{x}$

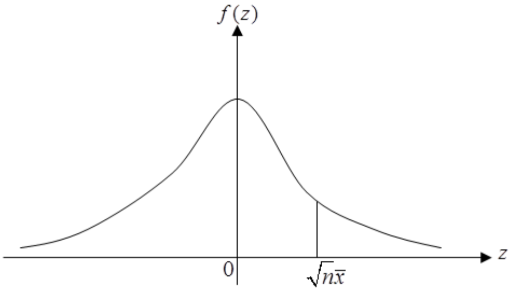
Sehingga kita peroleh

$$F_n(\bar{x}) = \int_{-\infty}^{\sqrt{n}\bar{x}} \frac{1}{\sqrt{2\pi}} e^{-v^2/2} dv.$$

Kita ingat bahwa pdf peubah acak yang bersebaran normal baku $N(0,1)$ adalah

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

yang kurvanya simetri terhadap $z = 0$. Oleh karena itu $F_n(\bar{x})$ merupakan luas daerah yang dibatasi oleh kurva $f(z)$, dan sumbu z dari $z = -\infty$ sampai $z = \sqrt{n}\bar{x}$ seperti gambar di bawah ini.



Gambar 6.2.1

Dengan mengingat bahwa luas daerah keseluruhan sama dengan 1, maka kita peroleh

$$\lim_{n\to\infty} F_n(\bar{x}) = \frac{1}{2} I(\bar{x}=0) + I(0<\bar{x}).$$

Sekarang jika kita ambil fungsi sebaran

$$F(\bar{x}) = I(0 \le \bar{x})$$

maka

$$\lim_{n\to\infty} F_n(\bar{x}) = F(\bar{x}),$$

untuk setiap $\bar{x}$ pada saat $F(\bar{x})$ kontinu. Jadi $\bar{X}_n$ mempunyai sebaran batas dengan fungsi sebaran $F(\bar{x})$ yang merosot pada $\bar{x}=0.$ Kita perhatikan bahwa

$$\lim_{n\to\infty} F_n(0) \ne F(0),$$

hal ini bukan merupakan gangguan untuk menerapkan Definisi 6.2.1, karena pada $\bar{x}=0$, $F(\bar{x})$ tidak kontinu.

Tidak semua sebaran batas adalah sebaran merosot, seperti terlihat pada contoh berikut ini. Meskipun demikian sebelumnya kita tuliskan teori tentang limit yang sering digunakan pada berbagai permasalahan sebaran batas:

(i) $\lim_{n\to\infty}\left(1+\dfrac{c}{n}\right)^{nb} = e^{cb},$

(ii) $\lim_{n\to\infty}\left(1+\dfrac{c}{n}+\dfrac{d(n)}{n}\right)^{nb} = e^{cb},$ jika $\lim_{n\to\infty} d(n) = 0.$

Contoh 6.2.3 Misal $X_1, X_2, ..., X_n$ sampel acak dari sebaran Pareto, $X_i \sim PAR(1{,}1)$ dan misal $Y_n = nX_{1,n}$, di mana $X_{1,n}$ adalah statistik urutan terkecil dari $X_1, X_2, ..., X_n$. Tentukan sebaran batas dari Y_n.

Penyelesaian

Karena $X_1, X_2, ..., X_n$ sampel acak dari sebaran Pareto, $X_i \sim PAR(1{,}1)$, maka cdf dari X_i adalah

$$F(x) = \left[1-(1+x)^{-1}\right] I(x>0).$$

Misal $G_n(y)$ adalah cdf dari Y_n. Maka

$$G_n(y) = P\left[Y_n \le y\right] = P\left[nX_{1,n} \le y\right] = P\left[X_{1,n} \le \frac{y}{n}\right] = F_{1,n}\left(\frac{y}{n}\right)$$

di mana $F_{1,n}$ adalah cdf dari statistik urutan terkecil $X_{1,n}$. Sehingga menggunakan sifat cdf urutan terkecil pada Bab sebelumnya, yaitu

$$F_{1,n}(x) = 1 - \left[1 - F(x)\right]^n$$

kita memperoleh

$$G_n(y) = 1 - \left[1 - \left(1 - (1+\frac{y}{n})^{-1}\right)\right]^n = 1 - \left((1+\frac{y}{n})^{-1}\right)^n = \left[1 - \left(1 + \frac{y}{n}\right)^{-n}\right] I(y > 0).$$

Menggunakan sifat limit pada (i) diperoleh

$$\lim_{n\to\infty} G_n(y) = \left(1 - e^{-y}\right) I(y > 0).$$

Jadi sebaran batas dari Y_n adalah

$$G(y) = \left(1 - e^{-y}\right) I(y > 0),$$

yang merupakan cdf dari peubah acak yang bersebaran eksponensial, $EXP(1)$. Atau dapat pula dinyatakan sebagai

$$Y_n \xrightarrow{d} Y, Y \sim EXP(1).$$

Tidak selalu barisan peubah acak mempunyai sebaran batas. Contoh berikut ini merupakan contoh dari barisan peubah acak yang tidak mempunyai sebaran batas.

Contoh 6.2.4 Misal $X_1, X_2, ..., X_n$ sampel acak dari sebaran Pareto, $X_i \sim PAR(1,1)$, dan misal $Y_n = X_{n,n}$, yang merupakan statistik urutan terbesar dari $X_1, X_2, ..., X_n$. Tunjukkan bahwa Y_n tidak mempunyai sebaran batas.

Penyelesaian

Seperti halnya pada Contoh 6.2.3, karena $X_1, X_2, ..., X_n$ sampel acak dari sebaran Pareto, $X_i \sim PAR(1,1)$, maka cdf dari X_i adalah

$$F(x) = \left[1 - (1+x)^{-1}\right] I(x > 0).$$

Misal $G_n(y)$ adalah cdf dari statistik urutan terbesar Y_n. Maka

$$G_n(y) = [F(y)]^n = \left[1-(1+y)^{-1}\right]^n = \left(\frac{y}{1+y}\right)^n I(y>0).$$

Karena untuk $y > 0$,

$$0 < \left(\frac{y}{1+y}\right)^n < 1,$$

maka $\lim_{n\to\infty} G_n(y) = 0$, untuk semua y, dan ini bukanlah cdf untuk semua y. Jadi Y_n tidak mempunyai sebaran batas.

Latihan 6.2

6.2.1. Misal $X_1, X_2, ..., X_n$ sampel acak dari sebaran seragam, $X_i \sim UNIF(0,\theta)$, dan misal $Y_n = X_{n,n}$, yaitu statistik urutan terbesar. Tentukan sebaran batas dari $Y_n = X_{n,n}$.

6.2.2. Misal $X_1, X_2, ..., X_n$ sampel acak dari sebaran seragam, $X_i \sim UNIF(0,\theta)$, dan misal $Y_n = X_{n,n}$, yaitu statistik urutan terbesar. Tentukan sebaran batas dari $Z_n = n(\theta - Y_n)$.

6.2.3. Misal $X_1, X_2, ..., X_n$ sampel acak dari sebaran eksponensial, $X_i \sim EXP(\theta)$, dan misal $Y_n = X_{1,n}$ merupakan statistik urutan terkecil. Tentukan sebaran batas dari $Y_n = X_{1,n}$.

6.2.4. Misal X_n peubah acak yang pdf-nya bergantung pada n, yaitu

$$f_n(x) = I\left(x = 2 + \tfrac{1}{n}\right).$$

Tentukan sebaran batas dari X_n.

6.2.5. Misal $\bar{X}_n$ menyatakan purata dari sampel acak berukuran n dari sebaran normal, $N(\mu, \sigma^2)$. Tentukan sebaran batas dari $\bar{X}_n$.

6.2.6. Misal pdf dari Y_n adalah

$$f_n(y) = I(y = n),$$

Tunjukkan bahwa Y_n tidak mempunyai sebaran batas.

6.2.7. Misal $X_1, X_2, ..., X_n$ sampel acak dari sebaran normal, $X_i \sim N(\mu, \sigma^2)$, di mana $\mu > 0$. Tunjukkan bahwa $Z_n = \sum_{i=1}^{n} X_i$ tidak mempunyai sebaran batas.

6.2.8. Misal $X_1, X_2, ..., X_n$ sampel acak dari sebaran dengan cdf

$$F(x) = \left(1 - x^{-1}\right) I(x > 1).$$

Selidiki apakah barisan peubah acak berikut ini mempunyai sebaran batas, jika ya tentukan sebaran batasnya.

(a) Statistik urutan terkecil $X_{1,n}$.

(b) $X_{1,n}^{n}$.

6.2.9. Misal $X_1, X_2, ..., X_n$ sampel acak dari sebaran dengan cdf

$$F(x) = \left(1 - x^{-2}\right) I(x > 1).$$

Selidiki apakah barisan peubah acak berikut ini mempunyai sebaran batas, jika ya tentukan sebaran batasnya.

(a) Statistik urutan terkecil $X_{1,n}$.

(b) Statistik urutan terbesar $X_{n,n}$.

(c) $n^{-1/2} X_{n,n}$.

6.2.10. Misal $X_1, X_2, ..., X_n$ sampel acak dari sebaran dengan cdf

$$F(x) = 1 - e^{-x},$$

untuk semua bilangan reak x. Selidiki apakah barisan peubah acak berikut ini mempunyai sebaran batas, jika ya tentukan sebaran batasnya.

(a) Statistik urutan terbesar $X_{n,n}$.

(b) $X_{n,n} - \ln n.$

6.3 Kekonvergenan dalam Peluang

Barisan peubah acak pada Contoh 6.2.1 dan Contoh 6.2.2 mempunyai sebaran batas yang merosot berturut-turut pada $y = 1$, dan $\bar{x} = 0.$ Hal ini mengantar kita pada konsep tentang konvergen dalam peluang, seperti definisi di bawah ini.

Definisi 6.3.1 Suatu Barisan peubah acak $Y_1, Y_2, Y_3, \ldots$ dikatakan ***konvergen dalam peluang*** ke konstanta c jika mempunyai sebaran batas yang merosot pada $y = c$.

Oleh karena itu barisan peubah acak pada Contoh 6.2.1 dan Contoh 6.2.2 konvergen dalam peluang berturut-turut ke $y = 1$, dan $\bar{x} = 0.$

Teorema berikut ini menyajikan kriteria alternatif untuk menunjukkan kekonvergenan dalam peluang suatu barisan peubah acak. Penerapannya akan dibantu oleh ketaksamaan Chebyshev.

Teorema 6.3.1 Barisan peubah acak $Y_1, Y_2, Y_3, \ldots$ konvergen dalam peluang ke c jika dan hanya jika untuk setiap $\varepsilon > 0,$

$$\lim_{n\to\infty} P\left[|Y_n - c| < \varepsilon \right] = 1,$$

atau, ekivalen dengan

$$\lim_{n\to\infty} P\left[|Y_n - c| \geq \varepsilon \right] = 0.$$

Bukti

Pertama, anggap bahwa untuk setiap $\varepsilon > 0,$

$$\lim_{n\to\infty} P\left[|Y_n - c| < \varepsilon \right] = 1.$$

Kita akan membuktikan bahwa barisan peubah $Y_1, Y_2, Y_3, \ldots$ konvergen dalam peluang ke c**.** Dengan kata lain, jika $G_n(y)$ cdf dari Y_n, maka kita akan menunjukkan bahwa

$$\lim_{n\to\infty} G_n(y) = I(y > c).$$

Jika hal ini dapat dipenuhi, maka kita dapat mengatakan bahwa sebaran batas dari Y_n merosot pada $y = c$, yaitu

$$G(y) = I(y \geq c).$$

Oleh karena itu dengan Definisi 1, kita memperoleh barisan peubah acak $Y_1, Y_2, Y_3, ...$ konvergen dalam peluang ke c. Sekarang perhatikan kenyataan

$$\begin{aligned} P\left[|Y_n - c| < \varepsilon\right] &= P\left[c - \varepsilon < Y_n < c + \varepsilon\right] \\ &= P\left[Y_n < c + \varepsilon\right] - P\left[Y_n \leq c - \varepsilon\right] \\ &= G_n\left[(c+\varepsilon)_-\right] - G_n(c - \varepsilon), \end{aligned}$$

di mana $G_n\left[(c+\varepsilon)_-\right]$ adalah limit kiri dari $G_n(y)$ pada $y = c + \varepsilon$. Oleh karena itu kita mempunyai

$$1 = \lim_{n\to\infty} P\left[|Y - c| < \varepsilon\right] = \lim_{n\to\infty} G_n\left[(c+\varepsilon)_-\right] - \lim_{n\to\infty} G_n(c - \varepsilon).$$

Karena $0 \leq G_n(y) \leq 1$ untuk semua y dan untuk semua bilangan bulat positif n, maka haruslah

$$\lim_{n\to\infty} G_n(c - \varepsilon) = 0, \text{ dan } \lim_{n\to\infty} G_n\left[(c+\varepsilon)_-\right] = 1.$$

Karena kesamaan ini berlaku untuk setiap $\varepsilon > 0$, maka

$$\lim_{n\to\infty} G_n(y) = I(y > c).$$

Jadi barisan peubah $Y_1, Y_2, Y_3, ...$ konvergen dalam peluang ke c.

Kedua, kita anggap bahwa barisan peubah $Y_1, Y_2, Y_3, ...$ konvergen dalam peluang ke c. Hal ini sama saja dengan menganggap bahwa

$$\lim_{n\to\infty} G_n(y) = I(y > c),$$

atau

$$\lim_{n\to\infty} G_n(c - \varepsilon) = 0, \lim_{n\to\infty} G_n\left[(c+\varepsilon)_-\right] = 1.$$

Kita akan membuktikan bahwa untuk setiap $\varepsilon > 0$

$$\lim_{n\to\infty} P\left[|Y_n - c| < \varepsilon\right] = 1.$$

Sekarang ambil sebarang $\varepsilon > 0$, maka

$$P\left[|Y_n - c| < \varepsilon\right] = G_n\left[(c+\varepsilon)_-\right] - G_n(c - \varepsilon).$$

Oleh karena itu

$$\lim_{n\to\infty} P\left[|Y_n - c| < \varepsilon\right] = \lim_{n\to\infty} G_n\left[(c+\varepsilon)_-\right] - \lim_{n\to\infty} G_n(c-\varepsilon) = 1 - 0 = 1.$$

Jadi setiap $\varepsilon > 0$

$$\lim_{n\to\infty} P\left[|Y_n - c| < \varepsilon\right] = 1.$$

Hasil terakhir ini telah melengkapi bukti dari Teorema 6.3.1.

Barisan yang konvergen dalam peluang ke konstanta c, ditulis sebagai $Y_n \xrightarrow{P} c.$

Contoh 6.3.1 Misal $X_1, X_2, ..., X_n$ sampel acak dari sebaran Bernoulli, $X_i \sim BIN(1, p)$, dan misal $\hat{p}_n = \dfrac{\sum_{i=1}^{n} X_i}{n}$. Tunjukkan bahwa $\hat{p}_n \xrightarrow{P} p.$

Penyelesaian

Karena $X_1, X_2, ..., X_n$ sampel acak bersebaran Bernoulli**,** $X_i \sim BIN(1, p)$, maka $E(X_i) = p$ dan $Var(X_i) = pq$, untuk setiap $i = 1, 2, ..., n$. Oleh karena itu menggunakan Teorema 5.4.1 kita peroleh

$$\mu = E(\hat{p}_n) = p, \text{ dan } \sigma^2 = Var(\hat{p}_n) = \frac{pq}{n}.$$

Sehingga menggunakan ketaksamaan Chebyshev kita peroleh

$$P\left[|\hat{p}_n - p| < \varepsilon\right] \geq 1 - \frac{pq}{\varepsilon^2 n},$$

dalam hal ini $k\sigma = \varepsilon$, atau $k = \dfrac{\varepsilon}{\sigma} = \dfrac{\varepsilon}{\sqrt{pq/n}}$, untuk sebarang $\varepsilon > 0.$

Sehingga

$$\lim_{n\to\infty} P\left[|\hat{p}_n - p| < \varepsilon\right] \geq \lim_{n\to\infty}\left[1 - \frac{pq}{\varepsilon^2 n}\right] = 1.$$

Karena nilai peluang paling besar 1, maka kita peroleh

$$\lim_{n\to\infty} P\left[|\hat{p}_n - p| < \varepsilon\right] = 1.$$

Jadi $\hat{p}_n \xrightarrow{P} p.$ Hal yang perlu kita perhatikan dalam hasil ini adalah bahwa $\hat{p}_n$, yang merupakan proporsi sampel acak, konvergen dalam peluang ke p, yang merupakan proporsi populasi. Artinya untuk n yang sangat besar proporsi populasi dapat didekati oleh proporsi sampel. Hasil ini disebut sebahai ***Hukum Bilangan Besar***.

Contoh 6.3.2 Misal $X_1, X_2, ..., X_n$ sampel acak dari sebaran normal, $X_i \sim N(\mu, \sigma^2)$, dan misal $\bar{X}_n = \frac{\sum_{i=1}^{n} X_i}{n}$. Tunjukkan bahwa $\bar{X}_n \xrightarrow{P} u.$

Penyelesaian

Dengan cara serupa pada Contoh 6.3.1 kita peroleh $E(\bar{X}_n) = \mu$, dan $Var(\bar{X}_n) = \frac{\sigma^2}{n}.$ Sehingga untuk sebarang $\varepsilon > 0$,

$$P\left[|\bar{X}_n - \mu| \geq \varepsilon\right] \leq \frac{\sigma^2}{n\varepsilon^2},$$

di mana $\varepsilon = \frac{k\sigma}{\sqrt{n}}$, atau $k = \frac{\varepsilon\sqrt{n}}{\sigma}.$ Oleh karena itu

$$\lim_{n\to\infty} P\left[|\bar{X}_n - \mu| \geq \varepsilon\right] \leq \lim_{n\to\infty} \frac{\sigma^2}{n\varepsilon^2} = 0.$$

Karena nilai peluang paling kecil 0, maka

$$\lim_{n\to\infty} P\left[|\bar{X}_n - \mu| \geq \varepsilon\right] = 0.$$

Jadi $\bar{X}_n \xrightarrow{P} u.$

Teorema berikut ini merupakan bentuk umum dari contoh-contoh di atas.

Teorema 6.3.2 Hukum Lemah Bilangan Besar. Misal $X_1, X_2, ..., X_n$ adalah sampel acak dari sebaran dengan purata μ dan varians $\sigma^2 < \infty.$ Misal $\bar{X}_n$ adalah purata sampel. Maka $\bar{X}_n \xrightarrow{P} \mu.$

Bukti

Dari Teorema 5.2.1 kita peroleh purata dan varians dari $\bar{X}_n$ berturut-turut adalah μ dan $\sigma^2/n.$ Oleh karena itu menggunakan ketaksamaan Chebyshev untuk sebarang $\varepsilon > 0$ kita mempunyai,

$$P\left[\left|\bar{X}_n - \mu\right| \geq \varepsilon\right] \leq \frac{\sigma^2}{n\varepsilon^2} \to 0,$$

bukti selesai.

Beberapa teorema kekonvergenan dalam peluang berikut sangat berguna pada pembahasan selanjutnya.

Teorema 6.3.3 Anggap $X_n \xrightarrow{P} X$ dan $Y_n \xrightarrow{P} Y.$ Maka

$$X_n + Y_n \xrightarrow{P} X + Y.$$

Bukti

Ambil $\varepsilon > 0$ sebarang. Menggunakan ketaksamaan segitiga kita dapat menulis

$$\begin{aligned}\left|X_n - X\right| + \left|Y_n - Y\right| &\geq \left|X_n - X + Y_n - Y\right| \\ &= \left|\left(X_n + Y_n\right) - \left(X + Y\right)\right| \geq \varepsilon.\end{aligned}$$

Karena P adalah relatif monoton untuk perluasan himpunan, maka

$$\begin{aligned}P\left[\left|\left(X_n + Y_n\right) - \left(X + Y\right)\right| \geq \varepsilon\right] &\leq P\left[\left|X_n - X\right| + \left|Y_n - Y\right| \geq \varepsilon\right] \\ &\leq P\left[\left|X_n - X\right| \geq \varepsilon/2\right] + P\left[\left|Y_n - Y\right| \geq \varepsilon/2\right].\end{aligned}$$

Menggunakan hipotesis dari teorema, dan dua suku terakhir konvergen ke nol, maka bukti selesai.

Teorema 6.3.4 Anggap $X_n \xrightarrow{P} X$ dan a suatu konstanta. Maka

$$aX_n \xrightarrow{P} aX.$$

Bukti

Jika $a = 0,$ maka hasil ini langsung diperoleh. Anggap $a \neq 0.$ Ambil $\varepsilon > 0$ sebarang. Maka

$$P\left[\left|aX_n - aX\right| \geq \varepsilon\right] = P\left[|a|\,|X_n - X| \geq \varepsilon\right] = P\left[\left|X_n - X\right| \geq \varepsilon/\left|a\right|\right],$$

menggunakan hipotesis suku terakhir konvergen ke nol, bukti selesai.

Teorema 6.3.5 Anggap $X_n \xrightarrow{P} a$ dan fungsi real g kontinu pada a. Maka

$$g\left(X_n\right) \xrightarrow{P} g(a).$$

Bukti

Ambil $\varepsilon > 0$ sebarang. Karena g kontinu pada a, maka ada $\delta > 0$ sehingga

$$|x-a| < \delta \rightarrow |g(x)-g(a)| < \varepsilon.$$

Oleh karena itu

$$|g(x)-g(a)| \geq \varepsilon \rightarrow |x-a| \geq \delta.$$

Substitusi X_n untuk x pada implikasi terakhir, kita peroleh

$$P\left[|g(X_n)-g(a)| \geq \varepsilon\right] \leq P\left[|X_n - a| \geq \delta\right].$$

Bukti selesai.

Teorema ini memberi kita banyak penggunaan. Sebagai contoh, jika $X_n \xrightarrow{P} a$, maka

$$X_n^2 \xrightarrow{P} a^2,$$

$$1/X_n \xrightarrow{P} 1/a, \text{ asalkan } a \neq 0,$$

$$\sqrt{X_n} \xrightarrow{P} \sqrt{a}, \text{ asalkan } a \geq 0.$$

Teorema 6.3.6 Anggap $X_n \xrightarrow{P} X$ dan $Y_n \xrightarrow{P} Y$. Maka

$$X_n Y_n \xrightarrow{P} XY.$$

Bukti

Menggunakan hasil sebelumnya, kita peroleh

$$\begin{aligned} X_n Y_n &= \frac{1}{2} X_n^2 + \frac{1}{2} Y_n^2 - \frac{1}{2}\left(X_n - Y_n\right)^2 \\ &\xrightarrow{P} \frac{1}{2} X^2 + \frac{1}{2} Y^2 - \frac{1}{2}(X - Y)^2 \\ &= XY. \end{aligned}$$

Definisi 6.3.2 Kita katakan barisan peubah-peubah acak $\{X_n\}$ ***terbatas dalam peluang***, jika untuk semua $\varepsilon > 0$ ada konstanta B_ε dan bilangan bulat positif N_ε sehingga

$$n > N_\varepsilon \Rightarrow P\left[|X_n| \le B_\varepsilon\right] \ge 1 - \varepsilon.$$

Teorema 6.3.7 Misal $\{X_n\}$ adalah barisan peubah-peubah acak terbatas dalam peluang dan misal $\{Y_n\}$ adalah barisan peubah-peubah acak yang konvergen dalam peluang ke nol. Maka

$$X_n Y_n \xrightarrow{P} 0.$$

Bukti

Misal $\varepsilon > 0$ sebarang. Pilih $B_\varepsilon > 0$ dan bilangan N_ε sehingga

$$n > N_\varepsilon \Rightarrow P\left[|X_n \le B_\varepsilon\right] \ge 1 - \varepsilon.$$

Oleh karena itu

$$\begin{aligned}\lim_{n\to\infty} P\left[|X_nY_n| \ge \varepsilon\right] &\le \lim_{n\to\infty} P\left[|X_nY_n| \ge \varepsilon, |X_n| \le B_\varepsilon\right] + \lim_{n\to\infty} P\left[|X_nY_n| \ge \varepsilon, |X_n| > B_\varepsilon\right] \\ &\le \lim_{n\to\infty} P\left[|Y_n| \ge \varepsilon / B_\varepsilon\right] + \varepsilon = \varepsilon.\end{aligned}$$

Bukti selesai.

Latihan 6.3

6.3.1 Misal peubah acak Y_n bersebaran $BIN(n, p)$. Tunjukkan bahwa:

(a) $\dfrac{Y_n}{n} \xrightarrow{P} p,$

(b) $\left(1 - \dfrac{Y_n}{n}\right) \xrightarrow{P} (1 - p).$

6.3.2. Misal $X_1, X_2, ..., X_n$ peubah-peubah acak bebas, $X_i \sim BIN(1, p_i)$, dan misal $Y_n = \dfrac{\sum_{i=1}^{n}(X_i - p_i)}{n}$. Tunjukkan bahwa $Y_n \xrightarrow{P} 0.$

6.3.3. Misal W_n menyatakan peubah acak dengan purata μ dan varians $\frac{b}{n^a}$, di mana $a > 0, \mu,$ dan b adalah konstanta. Tunjukkan bahwa $W_n \xrightarrow{P} \mu.$

6.3.4. Misal Y_n adalah statistik urutan terbesar dari sampel acak berukuran n dari sebaran seragam pada selang $(0,\theta)$, dan $Z_n = \sqrt{Y_n}$. Tunjukkan bahwa $Z_n \xrightarrow{P} \sqrt{\theta}.$

6.3.5. Misal $X_1, X_2, ..., X_n$ adalah sampel acak dari sebaran dengan pdf

$$f(x) = e^{-(x-\theta)} I(x > \theta, -\infty < \theta < \infty).$$

Suatu pdf yang demikian disebut ***eksponensial tergeser.*** Misal $Y_n = X_{1,n}$ statistik urutan terkecil dari $X_1, X_2, ..., X_n$.

(a) Tentukan cdf dan pdf dari Y_n.

(b) Buktikan bahwa $Y_n \xrightarrow{P} \theta,$

6.4 Teorema Limit Pusat

Materi tentang Teorema Limit Pusat sering sekali digunakan dalam bidang Statistika, baik Statistika Matematis maupaun Statistika Terapan. Teorema Limit Pusat dibuktikan dengan memanfaatkan barisan mgf, atau fungsi pembangkit momen. Oleh karena itu pembahasan akan diawali oleh kekonvergenan dari mgf.

Untuk menentukan sebaran batas dari peubah acak Y_n selama ini kita menggunakan definisi dari sebaran batas, yaitu sebaran batas diperoleh langsung dari $G_n(y)$, untuk n mendekati tak hingga. Berikut ini kita ingin melihatnya dari sisi lain, yaitu dari mgf-nya, atau fungsi pembangkit momennya, tentunya jika mgf-nya ada. Teorema berikut ini akan menyajikan hal itu, diberikan tanpa bukti.

Teorema 6.3.1 Misal $Y_1, Y_2, Y_3, ...$ barisan peubah acak berturut-turut dengan cdf $G_1(y), G_2(y), G_3(y), ...$ dan mgf $M_1(t), M_2(t), M_3(t), ...$. Misal $M(t)$ adalah mgf dari suatu cdf $G(y)$, dan

$$\lim_{n\to\infty} M_n(t) = M(t),$$

untuk semua t pada selang yang memuat nol, $-h < t < h$. Maka

$$\lim_{n\to\infty} G_n(y) = G(y),$$

untuk semua y, sehingga $G(y)$ kontinu.

Teorema di atas tidak berlaku sebaliknya. Contoh 6.2.1, mgf-nya tidak konvergen, lihat Latihan 6.4.1.

Contoh 6.4.1 Misal $X_1, X_2, ..., X_n$ sampel acak dari sebaran Bernoulli, $X_i \sim BIN(1, p)$, dan $Y_n = \sum_{i=1}^{n} X_i$. Tunjukkan bahwa

$$Y_n \xrightarrow{d} Y, Y \sim POI(\mu),$$

Penyelesaian

Kita gunakan Teorema 6.4.1 untuk menyelesaikan permasalahan ini. Karena $X_1, X_2, ..., X_n$ sampel acak dari sebaran Bernoulli**,** $X_i \sim BIN(1, p)$, dan $Y_n = \sum_{i=1}^{n} X_i$, maka $Y_n \sim BIN(n, p)$. Oleh karena itu $\mu = np$, atau $p = \frac{\mu}{n}$. Sehingga

$$M_n(t) = \left(pe^t + q\right)^n = \left(\frac{\mu}{n}e^t + (1 - \frac{\mu}{n})\right)^n = \left(1 + \frac{\mu(e^t - 1)}{n}\right)^n.$$

Menggunakan sifat limit yang diberikan pada Subpokok Bahasan 6.2 kita memperoleh

$$\lim_{n\to\infty} M_n(t) = e^{\mu(e^t - 1)}$$

yang merupakan mgf dari peubah acak bersebaran Poisson dengan purata μ. Jadi $Y_n \xrightarrow{d} Y, Y \sim POI(\mu)$. Dapat pula dikatakan bahwa jika $Y_n \sim BIN(n, p)$, dan jika $n \to \infty$, $p \to 0$, dan $np = \mu$ maka $Y_n \sim POI(\mu)$

Contoh 6.4.2 Misal $X_1, X_2, ..., X_n$ sampel acak dari sebaran Bernoulli, $X_i \sim BIN(1, p)$ dan $\hat{p}_n = \frac{Y_n}{n} = \frac{\sum_{i=1}^{n} X_i}{n}$. Tunjukkan bahwa

$$\hat{p}_n \xrightarrow{p} p.$$

Penyelesaian

Kita lihat mgf dari $\hat{p}_n$ dengan memanfaatkan mgf dari Y_n seperti berikut ini.

$$M_{\hat{p}_n}(t) = E\left[e^{t\frac{Y_n}{n}}\right] = E\left[e^{\frac{t}{n}Y_n}\right] = M_{Y_n}\left(\frac{t}{n}\right) = \left(pe^{\frac{t}{n}} + q\right)^n$$

Menggunakan ekspansi deret $e^x = 1 + x + \frac{x^2}{2} + \ldots,$ dan dengan memandang $x = \frac{t}{n},$ kita peroleh

$$\begin{aligned} M_{\hat{p}_n}(t) &= \left[p\left(1 + \frac{t}{n} + \frac{t^2}{2n^2} + \ldots\right) + q\right]^n \\ &= \left[p + q + \frac{pt}{n} + \frac{pt^2}{2n^2} + \ldots\right]^n \\ &= \left[1 + \frac{pt}{n} + \frac{d(n)}{n}\right]^n, \end{aligned}$$

di mana

$$d(n) = \frac{pt^2}{2n} + \frac{pt^3}{6n^2} + \ldots$$

Oleh karena itu menggunakan sifat limit kita peroleh

$$\lim_{n\to\infty} M_{\hat{p}_n}(t) = e^{pt}.$$

Ini merupakan mgf dari peubah acak yang sebarannya merosot pada p. Hal ini dapat dilihat seperti berikut:

Misal peubah acak Y mempunyai pdf

$$g(y) = I(y = p).$$

Maka cdf-nya adalah

$$G(y) = I(y \geq p),$$

dan ini merupakan cdf merosot pada $y = p$. Lebih lanjut diperoleh mgf

$$M_Y(t) = E(e^{tY}) = \sum_y e^{ty} f(y) = e^{tp} 1 = e^{pt}.$$

Jadi $\hat{p}_n \xrightarrow{p} p$. Hasil ini persis sama dengan hasil pada Contoh 6.3.1, yang disebut sebagai ***Hukum Bilangan Besar***.

Contoh 6.4.3 Misal $X_1, X_2, ..., X_n$ sampel acak dari sebaran Bernoulli**,** $X_i \sim BIN(1, p)$ $Y_n = \sum_{i=1}^{n} X_i$, dan $Z_n = \frac{Y_n - np}{\sqrt{npq}}$. Tunjukkan bahwa

$$Z_n \xrightarrow{d} Z, Z \sim N(0,1).$$

Penyelesaian

Kita ingat kembali sifat bahwa jika $Y_n \sim BIN(n, p)$, maka $\mu = np$ dan $\sigma = \sqrt{npq}$. Selanjutnya kita selesaikan dengan prosedur yang serupa dengan Contoh 6.4.2. Mula-mula kita ubah bentuk dari Z_n, yaitu

$$Z_n = \frac{Y_n}{\sigma} - \frac{np}{\sigma}.$$

Oleh karena itu

$$M_{Z_n}(t) = E\left[e^{tZ_n}\right] = E\left[e^{t\frac{Y_n}{\sigma} - t\frac{np}{\sigma}}\right] = E\left[e^{t\frac{Y_n}{\sigma}} e^{-\frac{tnp}{\sigma}}\right] = e^{-\frac{ptn}{\sigma}} E\left[e^{\frac{t}{\sigma}Y_n}\right]$$

$$= e^{-\frac{ptn}{\sigma}} M_{Y_n}\left(\frac{t}{\sigma}\right) = e^{-\frac{ptn}{\sigma}} \left(pe^{\frac{t}{\sigma}} + q\right)^n = \left[e^{-\frac{pt}{\sigma}}\left(pe^{\frac{t}{\sigma}} + q\right)\right]^n.$$

Seperti halnya pada Contoh 6.4.2, jika kita gunakan ekspansi deret maka kita peroleh

$$M_{Z_n}(t) = \left[\left(1 - \frac{pt}{\sigma} + \frac{p^2t^2}{2\sigma^2} - ...\right)\left(p + \frac{pt}{\sigma} + \frac{pt^2}{2\sigma^2} + ... + q\right)\right]^n = \left[1 + \frac{t^2}{2n} + \frac{d(n)}{n}\right]^n,$$

di mana $\frac{t^2}{2n}$ diperoleh dari bentuk perkalian

$$-\frac{pt}{\sigma}.\frac{pt}{\sigma} + \frac{p^2t^2}{2\sigma^2}.1 + 1.\frac{pt^2}{2\sigma^2} = \frac{-2p^2t^2 + p^2t^2 + pt^2}{2\sigma^2} = \frac{pt^2 - p^2t^2}{2\sigma^2}$$

$$= \frac{p(1-p)t^2}{2npq} = \frac{pqt^2}{2npq} = \frac{t^2}{2n},$$

sedangkan $\frac{d(n)}{n}$ diperoleh dari sisa perkalian yang lainnya. Kita peroleh hasil akhir

$$\lim_{n\to\infty} M_{Z_n}(t) = e^{\frac{t^2}{2}},$$

ini merupakan mgf dari peubah acak yang bersebaran normal baku. Jadi $Z_n \xrightarrow{d} Z,\ Z \sim N(0,1)$.

Hasil dari Contoh 6.4.3 dapat kita perumum untuk sebarang sampel acak dari sebaran dengan purata μ dan varians σ^2, yang disebut sebagai ***Teorema Limit Pusat/Central Limit Theorem*** seperti disajikan pada teorema berikut.

Teorema 6.4.2 Teorema Limit Pusat **(CLT).** Misal $X_1, X_2, ..., X_n$ sampel acak dari sebaran dengan purata μ dan varians $\sigma^2 < \infty$, dan misal

$$Z_n = \frac{\sum_{i=1}^{n} X_i - n\mu}{\sqrt{n}\sigma}.$$

Maka sebaran batas dari Z_n adalah normal baku, ditulis $Z_n \xrightarrow{d} Z$, $Z \sim N(0,1)$.

Bukti

Pembuktian di bawah ini menganggap bahwa mgf dari sampel acak $X_1, ..., X_n$ ada. Namun demikian, jika mgf-nya tidak ada bukti dapat dimodifikasi menggunakan fungsi karakteristik $\varphi(t) = E(e^{itX})$ (Bhat, 1981:195).

Misal $m(t)$ menyatakan mgf dari $X - \mu$, atau $m(t) = M_{X-\mu}(t)$. Maka

$$m(0) = E(e^{0(X-\mu)}) = E(e^0) = E(1) = 1,$$
$$m'(0) = E(X - \mu) = 0,$$

dan

$$m''(0) = E(X - \mu)^2 = \sigma^2.$$

Menggunakan ekspansi deret Taylor sekitar 0, dan untuk ξ antara 0 dan t, kita peroleh

$$m(t) = m(0) + m'(0)t + \frac{m''(\xi)t^2}{2} = 1 + 0 + \frac{m''(\xi)t^2}{2}$$

$$= 1 + \frac{\sigma^2 t^2}{2} + \frac{(m''(\xi) - \sigma^2)t^2}{2},$$

hasil terakhir ini diperoleh dari menambah dan mengurangi ruas kanan dengan $\frac{\sigma^2 t^2}{2}$.

Selanjutnya jika kita tulis Z_n dalam bentuk

$$Z_n = \frac{\sum_{i=1}^{n}(X_i - \mu)}{\sqrt{n}\sigma},$$

maka akan kita peroleh

$$M_{Z_n}(t) = E\left[e^{tZ_n}\right] = E\left[\exp\left\{t\frac{\sum_{i=1}^{n}(X_i - \mu)}{\sqrt{n}\sigma}\right\}\right]$$

$$= E\left[\exp\left\{\frac{t}{\sqrt{n}\sigma}\sum_{i=1}^{n}(X_i - \mu)\right\}\right]$$

$$= E\left[\exp\left\{\frac{t}{\sqrt{n}\sigma}(X_1 - \mu)\right\}\exp\left\{\frac{t}{\sqrt{n}\sigma}(X_2 - \mu)\right\}\ldots\exp\left\{\frac{t}{\sqrt{n}\sigma}(X_n - \mu)\right\}\right]$$

Karena $X_1, X_2, ..., X_n$ sampel acak, maka $X_1, X_2, ..., X_n$ bebas sehingga kita peroleh

$$E\left[\exp\left\{\frac{t}{\sqrt{n}\sigma}(X_1 - \mu)\right\}\exp\left\{\frac{t}{\sqrt{n}\sigma}(X_2 - \mu)\right\}\ldots\exp\left\{\frac{t}{\sqrt{n}\sigma}(X_n - \mu)\right\}\right]$$

$$= E\left[\exp\left\{\frac{t}{\sqrt{n}\sigma}(X_1 - \mu)\right\}\right]E\left[\exp\left\{\frac{t}{\sqrt{n}\sigma}(X_2 - \mu)\right\}\right]\ldots E\left[\exp\left\{\frac{t}{\sqrt{n}\sigma}(X_n - \mu)\right\}\right].$$

Kita perhatikan bahwa

$$E\left[\exp\left\{\frac{t}{\sqrt{n}\sigma}(X_i - \mu)\right\}\right] = M_{X_i - \mu}\left(\frac{t}{\sqrt{n}\sigma}\right).$$

Sekali lagi karena $X_1, X_2, ..., X_n$ sampel acak, sebaran dari $X_1, X_2, ..., X_n$ identik sehingga mgf dari X_i sama untuk setiap $i = 1, 2, ..., n$. Hal ini akan berakibat pula bahwa

$$M_{X_1-\mu}\left(\frac{t}{\sqrt{n}\sigma}\right) = ... = M_{X_n-\mu}\left(\frac{t}{\sqrt{n}\sigma}\right) = M_{X-\mu}\left(\frac{t}{\sqrt{n}\sigma}\right)$$

Sehingga kita peroleh

$$M_{Z_n}(t) = \left[M_{X-\mu}\left(\frac{t}{\sqrt{n}\sigma}\right)\right]^n = \left[m\left(\frac{t}{\sqrt{n}\sigma}\right)\right]^n$$

$$= \left[1 + \frac{\sigma^2 t^2}{2n\sigma^2} + \frac{\left(m''(\xi) - \sigma^2\right)t^2}{2n\sigma^2}\right]^n, 0 < |\xi| < \frac{|t|}{\sqrt{n}\sigma}.$$

Karena jika $n \to \infty$, maka $\frac{t}{\sqrt{n}\sigma} \to 0$, akibatnya $\xi \to 0$. Oleh karena itu $m''(\xi) \to m''(0) = \sigma^2$, atau $\left(m''(\xi) - \sigma^2\right) \to 0$. Sehingga kita dapat mengambil fungsi

$$d(n) = \frac{\left(m''(\xi) - \sigma^2\right)t^2}{2\sigma^2}.$$

Oleh karena itu kita peroleh

$$M_{Z_n}(t) = \left[1 + \frac{t^2}{2n} + \frac{d(n)}{n}\right]^n.$$

Menggunakan teori limit kita mendapatkan hasil

$$\lim_{n\to\infty} M_{Z_n}(t) = e^{\frac{t^2}{2}}.$$

Jadi $Z_n \xrightarrow{d} Z,\ Z \sim N(0,1)$.

Perlu kita catat bahwa Z_n pada Teorema 6.4.2 di atas dapat pula ditulis sebagai

$$Z_n = \frac{\bar{X}_n - \mu}{\sigma / \sqrt{n}}.$$

Persamaan terakhir ini sering digunakan untuk memperoleh nilai pendekatan purata populasi oleh purata sampel, jika purata populasi tersebut tidak diketahui.

Contoh 6.4.4 Misal $X_1, X_2, ..., X_n$ sampel acak dari sebaran seragam, $X_i \sim UNIF(0,1)$, dan misal $Y_n = \sum_{i=1}^{n} X_i$. Tunjukkan bahwa $Y_{12} - 6$ mendekati sebaran normal baku.

Penyelesaian

Karena diketahui $X_1, X_2, ..., X_n$ sampel acak dari sebaran seragam, $X_i \sim UNIF(0,1)$, maka $\mu = \frac{0+1}{2} = \frac{1}{2}$, dan $\sigma^2 = \frac{(1-0)^2}{12} = \frac{1}{12}$. Sekarang kita perhatikan persamaan berikut:

$$Y_{12} - 6 = \sum_{i=1}^{12} X_i - 6 = 12 \frac{\sum_{i=1}^{12} X_i}{12} - 6$$

$$= 12\left(\bar{X}_{12} - \frac{1}{2} \right) = \sqrt{12}\sqrt{12}\left(\bar{X}_{12} - \frac{1}{2} \right)$$

$$= \frac{\bar{X}_{12} - \frac{1}{2}}{\sqrt{\frac{1}{12}} / \sqrt{12}}.$$

Anggap $n = 12$ suatu nilai yang cukup besar, maka menggunakan CLT kita diperoleh $Y_{12} - 6 \xrightarrow{d} Z,\ Z \sim N(0,1)$.

Dengan kata lain $Y_{12} - 6$ mendekati sebaran normal baku.

Contoh 6.4.5 Ulangi Contoh 6.4.4. Jika $\bar{X}_n = \sum_{i=1}^{n} X_i$, untuk tentukan nilai pendekatan dari $P\left(0.40 < \bar{X}_{12} < 0.60\right)$.

Penyelesaian

Seperti halnya pada Contoh 6.4.4, anggap $n = 12$, suatu nilai yang cukup besar, maka menggunakan prosedur serupa akan kita peroleh nilai pendekatan

$$P(0.40 < \bar{X}_{12} < 0.60) = P\left[12\left(0.40 - \tfrac{1}{2}\right) < 12\left(\bar{X}_{12} - \tfrac{1}{2}\right) < 12\left(0.60 - \tfrac{1}{2}\right)\right]$$

$$= P\left[-1.2 < Z < 1.2\right] = P\left[Z < 1.2\right] - P\left[Z < -1.2\right]$$

$$= 2P\left[Z < 1.2\right] - 1 = 2 \times 0.885 = 0.77,$$

dari tabel sebaran normal baku.

Latihan 6.4

6.4.1. Tentukan mgf peubah acak pada Contoh 6.2.1. Tunjukkan bahwa mgf tersebut tidak konvergen.

6.4.2. Misal $Z_i \sim N(0,1)$ dan misal $Z_1, Z_2, Z_3, \ldots$ bebas. Gunakan Teorema 6.4.1 untuk menentukan sebaran batas dari $\sum_{i=1}^{n} \dfrac{Z_i - 1/n}{\sqrt{n}}$.

6.4.3. Misal $X_n \sim \chi^2(n)$, dan misal $Z_n = \dfrac{X_n - n}{\sqrt{2n}}$. Gunakan Teorema 6.4.1 untuk menunjukkan bahwa $Z_n \xrightarrow{d} Z,\ Z \sim N(0,1)$.

6.4.4. Misal peubah acak $Z_n \sim POI(n)$, dan misal $Y_n = \dfrac{Z_n - n}{\sqrt{n}}$. Gunakan Teorema 6.4.1 untuk menunjukkan bahwa $Y_n \xrightarrow{d} Z,\ Z \sim N(0,1)$.

6.4.5. Misal $\bar{X}_n$ menyatakan purata dari sampel berukuran n dari sebaran dengan pdf $f(x) = e^{-x} I(0 < x < \infty)$.

(a) Misal $Y_n = \sqrt{n}(\bar{X}_n - 1)$. Tunjukkan bahwa

$$M_{Y_n}(t) = \left[e^{t/\sqrt{n}} - (t/\sqrt{n})e^{t/\sqrt{n}}\right]^{-n},\ t < \sqrt{n}.$$

(b) Tentukan sebaran batas dari Y_n.

6.4.6. Misal $\bar{X}$ menyatakan purata sampel acak berukuran 100 dari sebaran $\chi^2(50)$. Hitunglah nilai pendekatan dari $P(49 < \bar{X} < 51)$.

6.4.7. Misal $\bar{X}$ menyatakan purata sampel acak berukuran 128 dari sebaran gamma dengan $\theta = 2$, dan $\kappa = 4$. Hitunglah nilai pendekatan dari $P(7 < \bar{X} < 9)$.

6.4.8. Misal $X_1, X_2, ..., X_{100}$ sampel acak dari sebaran eksponensial, $X_i \sim EXP(1)$, dan misal $Y = \sum_{i=1}^{100} X_i$. Tentukan nilai pendekatan normal dari:

(a) $P(Y > 110)$.

(b) $P(1.1 < Y < 1.2)$.

6.5 Pendekatan Sebaran Diskret dengan Sebaran Normal

Dalam Contoh 6.4.2 kita telah melihat bahwa barisan proporsi peubah acak dari sebaran binomial dapat kita dekati dengan sebaran normal baku, yang disebut sebagai pembakuan. Padahal kita ketahui bahwa sebaran binomial dari sebaran diskret, sedangkan sebaran normal dari dari sebaran kontinu.

Hal tersebut tentu akan menimbulkan masalah dalam perhitungan peluangnya, karena dalam peluang kedua sebaran ini mempunyai perbedaan yang cukup besar. Sebagai contoh, jika X peubah acak dari sebaran seragam diskret yang terdefinisi pada nilai-nilai $x = 1, 2, 3, 4, 5$, maka nilai dari

$$P[2 \le X \le 4] = P[1 < X < 5] = P[1.1 < X < 4.5],$$

dan sebagainya. Tetapi

$$P[2 \le X \le 4] \ne P[2 < X < 4] \ne P[2 \le X < 4] \ne P[2 < X \le 4].$$

Di pihak lain, jika Y peubah acak dari sebaran seragam kontinu yang terdefinisi pada selang $1 < x < 5$, maka

$$P[2 \le X \le 4] \ne P[1 < X < 5] \ne P[1.1 < X < 4.5],$$

dan sebagainya. Tetapi

$$P[2 \le X \le 4] = P[2 < X < 4] = P[2 \le X < 4] = P[2 < X \le 4].$$

Hal ini dapat terjadi karena untuk sebarang peubah acak kontinu peluang peubah acak tersebut sama dengan konstanta tertentu adalah nol.

Berangkat dari kenyataan di atas, maka diperlukan suatu prosedur sehingga kebenaran pendekatan sebaran diskret oleh sebaran kontinu dapat dipertanggungjawabkan. Untuk itu marilah kita ikuti Contoh 6.5.1 berikut.

Contoh 6.5.1 Taufik Hidayat seorang pemain bulu tangkis nasional dalam pertandingan tertentu peluang smesnya mematikan lawan anggap 0.6. Jika dalam pertandingan tersebut dia melakukan smes sebanyak 20 kali, berapakah peluang bahwa paling sedikit 12 smesnya mematikan lawan.

Penyelesaian

Permasalahan ini dapat kita selesaian dengan dua cara. Pertama dengan menghitung peluang eksaknya, yaitu

$$P\left[Y_{20} \geq 12\right] = 1 - P\left[Y_{20} \leq 11\right] = 1 - B(11; 20, 0.6) = 1 - 0.4044 = 0.5956,$$

menggunakan tabel sebaran binomial. Kedua menggunakan pendekatan normal, yaitu

$$\begin{aligned} P\left[Y_{20} \geq 12\right] = 1 - P\left[Y_{20} \leq 11\right] &= 1 - P\left[\frac{Y_{20} - 12}{\sqrt{4.8}} \leq \frac{11 - 12}{\sqrt{4.8}}\right] \\ &= 1 - \Phi\left(\frac{11-12}{\sqrt{4.8}}\right) = 1 - \Phi(-0.46) = 1 - (1 - \Phi(0.46)) \\ &= \Phi(0.46) = 0.6772, \end{aligned}$$

dari tabel sebaran normal baku. Hasil ini tidaklah memuaskan, karena perbedaannya cukup besar. Jika kita mencoba menghitungnya menggunakan koreksi kekontinuan, kita peroleh

$$\begin{aligned} P\left[Y_{20} \geq 12\right] = 1 - P\left[Y_{20} \leq 11\right] &= 1 - P\left[Y_{20} \leq 11,5\right] \\ &= 1 - P\left[\frac{Y_{20} - 12}{\sqrt{4.8}} \leq \frac{11.5 - 12}{\sqrt{4.8}}\right] = 1 - \Phi\left(\frac{11.5-12}{\sqrt{4.8}}\right) \\ &= 1 - \Phi(-0.23) = 1 - (1 - \Phi(0.23)) = \Phi(0.23) = 0.5910, \end{aligned}$$

dari tabel sebaran normal baku. Hasil terakhir sesuai dengan harapan kita.

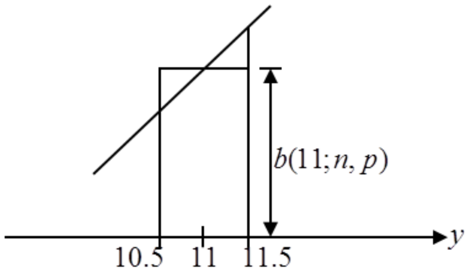
Dari hasil yang kita peroleh pada Contoh 6.5.1, tentu timbul pertanyaan, mengapa koreksinya harus 0.5? Jawaban dari pertanyaan ini dapat kita lihat pada uraian berikut.

Misal $Y \sim BIN(n, p)$. Untuk menyederhanakan permasalahan kita ambil kasus $n = 20, p = 0.6.$ Maka

$P(Y = 11) = b(11; n, p) = 0.1594,$

diperoleh dari Minitab. Nilai pdf, $b(11; n, p) = 0.1594,$ sama dengan luas daerah persegi panjang dengan tinggi $b(11; n, p) = 0.1594,$ dan panjang alas 1. Panjang alas ini dapat dipandang sebagai panjang selang

$[11-0.5, 11+0.5]$. Luas daerah persegi panjang ini dapat didekati oleh luas daerah di bawah pdf $Y \sim N(\mu, \sigma^2)$, di mana $\mu = np$, dan $\sigma^2 = npq$. Hal ini dapat kita lihat pada Gambar 6.5.1 berikut ini.



Gambar 6.5.1

Sehingga pendekatan normal dari $P(Y=11)$ adalah

$$P(Y=11)=P(10.5 \le Y \le 11.5)=P\left[\frac{10.5-12}{\sqrt{4.8}} \le \frac{Y-12}{\sqrt{4.8}} \le \frac{11.5-12}{\sqrt{4.8}}\right]$$

$$=\Phi[-0.23]-\Phi[-0.68]=0.4090-0.2483=0.1607,$$

dari Minitab. Akurasi pendekatan ini akan memberikan hasil terbaik untuk $p=0.5$, karena sebaran binomial ini akan simetris untuk nilai tersebut.

Koreksi kontinuitas ini dapat juga digunakan untuk sebaran diskret yang lainnya.

Contoh 6.5.2 Misal $X_1, X_2, ..., X_n$ sampel acak dari sebaran Poisson, $X_i \sim POI(1)$ dan misal $Y_n=\sum_{i=1}^{n} X_i$. Tunjukkan bahwa:

(a) $Y_n \sim POI(n)$.

(b) Hitung nilai eksak dari $P[Y_{20} \le 20]$, dan $P[15 \le Y_{20} \le 30]$

(c) Hitung nilai pendekatan normal dari $P[Y_{20} \le 20]$, dan $P[15 \le Y_{20} \le 30]$

Penyelesaian

(a) Karena $X_1, X_2, ..., X_n$ sampel acak dari sebaran Poisson, $X_i \sim POI(1)$, maka

$M_{X_i}(t) = e^{1(e^t-1)} = e^{(e^t-1)}.$

Oleh karena itu

$M_{Y_n}(t) = E[e^{tY_n}] = M_{X_1}(t)...M_{X_n}(t) = e^{(e^t-1)}...e^{(e^t-1)} = e^{n(e^t-1)}.$

Hasil terakhir ini merupakan mgf dari peubah acak yang bersebaran Poisson dengan parameter n. Jadi $Y_n \sim POI(n)$

(b) Karena $Y_n \sim POI(n)$ maka $Y_{20} \sim POI(20)$ Oleh karena itu kita peroleh nilai eksak

$P[Y_{20} \le 20] = 0.5591,$

dan

$P[15 \le Y_{20} \le 30] = 0.9865 - 0.1049 = 0.8816,$

dari Minitab.

(c) Karena $Y_{20} \sim POI(20)$, maka $\mu = 20,$ dan $\sigma^2 = 20.$ Oleh karena itu kita peroleh nilai pendekatan normal

$$P[Y_{20} \le 20] = P[Y_{20} \le 20.5] = P\left[\frac{Y_{20}-20}{\sqrt{20}} \le \frac{20.5-20}{\sqrt{20}}\right]$$

$$= \Phi(0.11) = 0.5438,$$

dan

$$P[15 \le Y_{20} \le 30] = P[14.5 \le Y_{20} \le 30.5] = P\left[\frac{14.5-20}{\sqrt{20}} \le \frac{Y_{20}-20}{\sqrt{20}} \le \frac{30.5-20}{\sqrt{20}}\right]$$

$$= \Phi(2.35) - \Phi(-1.23) = 0.9906 - 0.1093 = 0.8813,$$

dari Minitab.

Latihan 6.5

6.5.1. Misal $X_1, X_2, ..., X_n$ sampel acak dari sebaran binomial, $Y_n \sim BIN(1, p)$, di mana $p = 0.5$, dan misal $Y_n = \sum_{i=1}^{n} X_i$. Hitung

(a) nilai eksak dari $P[40 \le Y_{100} \le 55]$.

(b) nilai pendekatan normal dari $P[40 \le Y_{100} \le 55]$.

6.5.2. Misal $X_1, X_2, ..., X_n$ sampel acak dari sebaran binomial, $X_i \sim BIN(1, p)$, di mana $p = \frac{1}{5}$, dan misal $Y_n = \frac{\sum_{i=1}^{n} X_i}{n}$.

(a) Hitung nilai pendekatan normal dari $P[0.25 < Y_{400}]$.

(b) Tentukan n sehingga diperoleh nilai pendekatan normal $P[Z_n > 0.5] \geq 0.95$.

6.5.3. Misal $X_1, X_2, ..., X_n$ sampel acak dari sebaran Poisson, $X_i \sim POI(np)$, di mana $p = 0.5$, dan misal $Y_n = \sum_{i=1}^{n} X_i$. Hitung nilai pendekatan normal dari $P[40 \leq Y_{100} \leq 55]$.

2.5.4. Misal Y menyatakan jumlah nilai-nilai dari sampel acak berukuran 12 dengan pdf

$$f(x) = \frac{1}{6} I(x = 1, 2, 3, 4, 5, 6).$$

Hitung nilai pendekatan normal dari $P[36 \leq Y \leq 48]$.

BAB 7
TEORI PENDUGAAN

7.1 Pendahuluan

Pada bab-bab sebelumnya kita telah membahas pengembangan dari konsep peluang dan peubah acak untuk membangun model matematika dari fenomena alam. Karakteristik numerik tertentu dari fenomena alam menjadi perhatian kita, tetapi nilainya tidak dapat dihitung langsung. Malahan kita mungkin mengamati satu atau lebih peubah acak, yang sebarannya bergantung pada karakteristik tertentu. Tujuan kita dalam bab selanjutnya adalah mengembangkan metode untuk menganalisis nilai pengamatan dari peubah acak untuk mendapatkan informasi tentang karakteristik yang tidak diketahui.

Dalam beberapa kasus adalah mungkin untuk mendapatkan suatu model khusus berdasarkan pada anggapan atau pengetahuan tentang populasi. Lebih sering, kita tidak dapat mengetahui spesifikasi pdf secara lengkap, tetapi mungkin hanya menganggap bahwa $f(x,\theta)$ adalah suatu anggota dari beberapa keluarga sebaran yang diketahui (seperti normal, gamma, Poisson) dan bahwa θ adalah parameter yang tidak diketahui seperti purata atau varians dari suatu sebaran.

Tujuan instruksional umum dari mempelajari bab ini adalah Anda diharapkan memahami teori pendugaan dari parameter θ berdasarkan data pengamatan suatu populasi.

Tujuan instruksional khusus dari mempelajari bab ini adalah Anda diharapkan dapat menentukan: penduga menggunakan metode momen, penduga menggunakan metode kemungkinan maksimum, penduga takbias, batas bawah Rao-Cramer, penduga efisien, keefisienan penduga, MVUE, selang kepercayaan dua sisi suatu parameter, selang kepercayaan satu sisi suatu parameter, selang kepercayaan parameter untuk permasalahan dua sampel dan penduga Bayes.

Kita anggap bahwa himpunan n peubah acak bebas, $X_1, X_2, ..., X_n$, masing-masing dengan pdf $f(x,\theta)$, akan diamati untuk menghasilkan himpunan data $x_1, x_2, ..., x_n$. Tentu, ini memungkinkan kita untuk menyatakan pdf bersama dari sampel acak seperti

$$f(x_1, x_2, ..., x_n;\theta) = f(x_1;\theta)\, f(x_2;\theta) ... f(x_n;\theta).$$

pdf bersama ini menyediakan hubungan antara data pengamatan dan model matematika untuk populasi.

Anggap bahwa sebaran dari populasi dapat dinyatakan dengan anggota dari beberapa keluarga pdf, $f(x,\theta)$. Kita misalkan Ω sebagai ***ruang parameter***, yaitu semua nilai yang mungkin dari parameter θ. Dan misal $X_1, X_2, ..., X_n$ sampel acak dari sebaran dengan pdf $f(x,\theta)$, dan bahwa $\tau(\theta)$ adalah fungsi dari θ.

Definisi 7.1.1 Suatu statistik $T = \ell(X_1, X_2, ..., X_n)$ yang digunakan untuk menduga nilai dari $\tau(\theta)$ disebut ***penduga*** dari $\tau(\theta)$, dan nilai pengamatan dari statistik ini, yaitu $t = \ell(x_1, x_2, ..., x_n)$ disebut ***dugaan*** dari $\tau(\theta)$.

Tentu, ini termasuk pendugaan dari nilai parameter itu sendiri, yaitu $\tau(\theta) = \theta$.

Kita akan menggunakan tiga jenis huruf yang berbeda dalam notasi kita. Huruf kapital T menyatakan statistik yang merupakan penduga, huruf kecil t menyatakan nilai pengamatan dari statistik atau disebut dugaan, dan huruf skrip ℓ menyatakan fungsi yang dipakai untuk peubah acak. Di pihak lain, untuk membedakan antara parameter tidak diketahui θ dengan penduganya dipakai tanda topi, yaitu ditulis sebagai $\hat{\theta}$.

Dua metode yang sering digunakan dalam permasalahan pendugaan diberikan pada subpokok bahasan berikut ini.

7.2 Pendugaan Metode Momen

Pada bab sebelumnya kita telah menggunakan purata sampel $\bar{X}_n$ sebagai penduga takbias dari purata populasi μ. Untuk yang pertama disebut sebagai momen sampel pertama, dan untuk yang kedua disebut momen populasi pertama. Secara umum, momen populasi diduga oleh momen sampel. Metode pendugaan ini disebut sebagai ***penduga metode momen/method of moments estimators (MMEs)***, dan hasil dari pendugaan dengan metode ini ditulis sebagai MME.

Kita perhatikan pdf dari populasi, $f(x;\theta_1,\theta_2,...,\theta_k)$, yang bergantung pada satu atau lebih parameter $\theta_1,\theta_2,...,\theta_k$. Kita telah mengetahui bahwa momen ke j di sekitar nol, dinyatakan sebagai $\mu_j^{'}$, didefinisikan sebagai

$$\mu_j^{'}(\theta_1,\theta_2,...,\theta_k) = E(X^j),\ j = 1,2,...$$

Secara khusus $\mu_1^{'}(\theta_1,\theta_2,...,\theta_k) = \mu.$

Untuk sampai pada pendugaan metode momen, mula-mula kita definisikan momen sampel ke j berikut ini.

Definisi 7.2.1 Misal $X_1, X_2,..., X_n$ sampel acak dari sebaran dengan pdf $f(x;\theta_1,\theta_2,...,\theta_k)$. Maka ***momen sampel*** ke j diberikan oleh

$$M_j^{'} = \frac{\sum_{i=1}^{n} X_i^j}{n},\ j = 1,2,3,...$$

Prinsip dari MMEs adalah memilih penduga dari parameter $\theta_1,\theta_2,...,\theta_k$ oleh penduga $\hat{\theta}_1,\hat{\theta}_2,...,\hat{\theta}_k$ sehingga diperoleh dugaan yang khas dari $\mu_j^{'}(\theta_1,\theta_2,...,\theta_k) = E(X^j),\quad j = 1,2,3,...$ adalah $M_j^{'}$ di mana

$$M_j^{'} = \mu_j^{'}(\hat{\theta}_1,\hat{\theta}_2,...,\hat{\theta}_k),\ j = 1,2,3,....$$

Kembali pada pembahasan awal bahwa penduga dari μ adalah $\bar{X}$. Pandang μ sebagai $\mu_1^{'}(\theta)$, atau $\mu = \mu_1^{'}(\theta)$. Dengan MMEs diperoleh bahwa $\mu_1^{'}(\hat{\theta}) = M_1^{'} = \bar{X}_n$, atau $\hat{\mu} = \bar{X}_n$.

Contoh 7.2.1 Misal $X_1, X_2,..., X_n$ suatu sampel acak dari sebaran dengan dua parameter tidak diketahui, purata μ dan varians σ^2. Dari hasil sebelumnya kita peroleh MME dari $\mu = \mu_1^{'}$ adalah $\hat{\mu} = \bar{X}_n$. Sekarang perhatikan bahwa

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

Sehingga

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

Karena $E(X^2)=\mu_2'$, maka MME-nya adalah

$$\hat{\mu}_2' = M'_2 = \frac{\sum_{i=1}^{n} X_i^2}{n}.$$

Oleh karena itu

$$\frac{\sum_{i=1}^{n} X_i^2}{n} = \hat{\sigma}^2 + \bar{X}_n^{\,2}.$$

Jadi

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

Contoh 7.2.2 Misal $X_1, X_2, ..., X_n$ suatu sampel acak dari sebaran eksponensial, $X_i \sim EXP(\theta)$, dan misal kita ingin menentukan MME dari peluang

$p = p(\theta) = P(X \geq 1) = e^{-\frac{1}{\theta}}$. Karena $X_i \sim EXP(\theta)$, maka $\mu_1'(\theta) = \mu = \theta$, sehingga MME-nya adalah $\hat{\theta} = \bar{X}_n$. Karena $p = e^{-\frac{1}{\theta}}$, maka $\theta = -\dfrac{1}{\ln p}$. Oleh karena itu menggunakan MME $\hat{\theta} = \bar{X}_n$, kita peroleh $\bar{X} = -\dfrac{1}{\ln \hat{p}}$.

Jadi $\hat{p} = e^{-\frac{1}{\bar{X}}} = e^{-\frac{1}{\hat{\theta}}} = p(\hat{\theta})$.

Dari hasil ini kita peroleh bahwa MME dari $p = p(\theta)$ adalah $\hat{p} = \hat{p}(\theta) = p(\hat{\theta})$. Secara umum

$$\hat{\tau}(\theta_1, \theta_2, ..., \theta_n) = \tau(\hat{\theta}_1, \hat{\theta}_2, ..., \hat{\theta}_n).$$

Sifat semacam ini disebut sebagai ***sifat invarians.***

Contoh 7.2.3 Misal $X_1, X_2, ..., X_n$ suatu sampel acak dari sebaran gamma, $X_i \sim GAM(\theta, \kappa)$. Karena $\mu = \kappa\theta$, maka dengan MMEs dan sifat invarians kita peroleh bahwa $\bar{X}_n = \hat{\kappa}\hat{\theta}$, atau

$$\hat{\kappa} = \frac{\bar{X}_n}{\hat{\theta}} \tag{7.2.1}$$

Karena

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

maka

$$E(X^2) = \kappa\theta^2 + \kappa^2\theta^2 = \kappa(1+\kappa)\theta^2.$$

Sekali lagi dengan MMEs dan sifat invarians kita peroleh

$$\frac{\sum_{i=1}^{n} X_i^2}{n} = \hat{\kappa}(1+\hat{\kappa})\hat{\theta}^2 \tag{7.2.2}$$

Substitusi persamaan (7.2.1) ke persamaan (7.2.2) kita peroleh

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

Persamaan tersebut ekivalen dengan

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

Sehingga diperoleh penyelesaian

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

dan

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

Latihan 7.2

7.2.1 Tentukan MME dari θ berdasarkan pada sampel acak $X_1, X_2, ..., X_n$ untuk masing-masing pdf berikut ini:

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

(b) $f(x;\theta) = (\theta+1)x^{-\theta-2} I(1 < x, 0 < \theta).$

(c) $f(x;\theta)=\theta^2 x e^{-\theta x} I(0<x, 0<\theta)$.

(d) $f(x;\theta)=\theta^x (1-\theta)^{1-x} I(x=0,1; 0\le\theta\le 1)$.

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

7.2.2 Tentukan MME dari parameternya berdasarkan pada sampel acak $X_1, X_2, ..., X_n$ untuk masing-masing sebaran berikut ini:

(a) $X_i \sim BIN(1, p)$.

(b) $X_i \sim BIN(3, p)$.

(c) $X_i \sim GAM(2, \kappa)$.

(d) $X_i \sim WEI(\theta, \frac{1}{2})$.

(e) $X_i \sim PAR(\theta, \kappa)$, θ dan κ kedua-duanya tidak diketahui.

7.2.3. Misal $X_1, X_2, ..., X_n$ sampel acak dari ***sebaran eksponensial dua parameter***, $X_i \sim EXP(1, \eta)$. Tentukan MME dari η. *Catatan*: Jika $X \sim EXP(1, \eta)$ maka $\mu = 1+\eta$.

7.2.4. Misal $X_1, X_2, ..., X_n$ sampel acak dari sebaran eksponensial dua parameter, $X_i \sim EXP(\theta, \eta)$. Tentukan MME dari η dan θ. *Catatan*: Jika $X \sim EXP(\theta, \eta)$ maka $\mu = \theta + \eta$ dan $\sigma^2 = \theta^2$.

7.2.5. Misal $X_1, X_2, ..., X_n$ sampel acak dari sebaran geometrik, $X_i \sim GEO(p)$. Tentukan MME dari:

(a) $E(X) = \frac{1}{p}$.

(b) $Var(X) = \frac{1-p}{p^2}$.

(c) $P[X > k] = (1-p)^k, k = 1, 2, ...$

Petunjuk: Gunakan sifat invarians.

7.3 Pendugaan Metode Kemungkinan Maksimum

Kita sekarang sampai pada sutau metode yang sering digunakan dalam menduga parameter. Ide dari metode ini adalah menenentukan nilai peluang maksimum berdasarkan data pengamatan untuk menduga parameter yang tidak diketahui. Untuk lebih jelasnya kita ikuti contoh berikut ini.

Contoh 7.3.1 Misal sekeping uang logam yang tidak seimbang, dan misal purata proporsi dari muncul muka adalah satu dari ketiga nilai $p = 0.2,\ 0.3,$ atau $0.8.$ Suatu percobaan dilakukan untuk melempar uang logam tersebut sebanyak dua kali dan diamati banyak muncul muka. Dari percobaan tersebut dapat dibuat model matematika seperti berikut ini.

Misal X_1, X_2 sampel acak dari sebaran Bernoulli, $X_i \sim BIN(1, p),\ \ i = 1, 2,$ dengan ruang parameter $\Omega = \{0.2, 0.3, 0.8\}.$ Karena $X_i \sim BIN(1, p),\ \ i = 1, 2,$ maka $\mu = \mu_1' = E(X) = p.$ Sehingga MME dari p adalah $\bar{X}$. Dalam contoh ini terlihat bahwa MMEs tidak memberikan hasil yang diharapkan, karena dari percobaan tersebut kita peroleh $\bar{x}$:

0 , jika $(x_1, x_2) = (0, 0)$,

$0.5,$ jika $(x_1, x_2) = (0,1), (1, 0),$

$1,$ jika $(x_1, x_2) = (1,1).$

Semua nilai tersebut bukan anggota dari ruang parameter Ω. Oleh karena itu diperlukan metode lain yang diharapkan dapat memberikan hasil yang baik, yaitu ***metode kemungkinan maksimum***.

Sekarang kita perhatikan pdf bersama sampel acak X_1, X_2,

$$\begin{aligned} f(x_1, x_2; p) &= f(x_1; p) f(x_2; p) \\ &= p^{x_1}(1-p)^{1-x_1}.p^{x_2}(1-p)^{1-x_2} I(x_i = 0,1; i = 1, 2) \\ &= p^{x_1+x_2}(1-p)^{2-x_1-x_2} I(x_i = 0,1; i = 1, 2). \end{aligned}$$

Nilai lengkap dari $f(x_1, x_2; p)$ disajikan pada Tabel 7.3.1 berikut ini, untuk berbagai nilai dari pasangan (x_1, x_2) dan p.

TABEL 7.3.1

p	(x_1, x_2)			
	(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

Sekarang anggap bahwa percobaan menghasilkan pasangan $(x_1, x_2) =$ $(0,0)$. Dari Tabel 7.3.1 dapat kita lihat bahwa $p = 0.20$ memberikan nilai pdf yang lebih besar dibanding dua nilai p yang lainnya. Oleh karena itu $\hat{p} = 0.20$, jika $(x_1, x_2) = (0,0)$. Dengan cara yang sama kita mendapatkan $\hat{p} = 0.30$, jika $(x_1, x_2) = (0,1)$ atau $(1,0)$. Dan terakhir $\hat{p} = 0.80$ jika $(x_1, x_2) = (1,1)$. Jadi dugaan kemungkinan maksimum dari p berdasarkan pengamatan (x_1, x_2) adalah

$$\hat{p} = \begin{cases} 0.20 & \text{jika } (x_1, x_2) = (0,0) \\ 0.30 & \text{jika } (x_1, x_2) = (0,1), (1,0) \\ 0.80 & \text{jika } (x_1, x_2) = (1,1) \end{cases}$$

Definisi 7.3.1 Suatu pdf bersama dari n peubah acak $X_1, X_2, ..., X_n$ yang bergantung pada parameter θ disebut ***fungsi kemungkinan,*** ditulis

$$L(\theta) = f(x_1, x_2, ..., x_n; \theta).$$

Dalam hal $X_1, X_2, ..., X_n$ sampel acak, maka kita mempunyai

$$L(\theta) = f(x_1; \theta) f(x_2; \theta) ... f(x_n; \theta).$$

Definisi 7.3.2 Misal $x_1, x_2, ..., x_n$ nilai pengamatan berturut-turut dari peubah acak $X_1, X_2, ..., X_n$. Nilai $\hat{\theta}$ dalam Ω yang merupakan fungsi dari $x_1, x_2, ..., x_n$ disebut ***dugaan kemungkinan maksimum*** dari θ jika $L(\hat{\theta})$ nilai maksimum dari $L(\theta)$ untuk setiap θ dalam Ω.

Kita catat di sini bahwa bahasa asli dari dugaan kemungkinan maksimum adalah ***maximum likelihood estimate.*** Untuk selanjutnya dugaan kemungkinan maksimum ditulis sebagai MLE.

Selanjutnya jika Ω merupakan selang terbuka, dan jika $L(\theta)$ dapat diturunkan dan memenuhi asumsi maksimum pada Ω, maka MLE merupakan solusi dari persamaan

$$\frac{d}{d\theta}L(\theta)=0.$$

Menentukan solusi persamaan di atas sama saja dengan menentukan solusi persamaan

$$\frac{d}{d\theta}\ln L(\theta)=0.$$

Hal ini dapat kita mengerti, karena $\ln$ merupakan fungsi naik sejati. Sehingga nilai $\hat{\theta}_1$ yang memberikan $L(\theta)$ maksimum sama dengan nilai $\hat{\theta}_2$ yang memberikan $\ln L(\theta)$ maksimum. Persamaan terakhir disebut sebagai persamaan kemungkinan, yang akan sering kita gunakan untuk mempermudah proses perhitungan.

Contoh 7.3.2 Misal $X_1, X_2, ..., X_n$ sampel acak dari sebaran Poisson dengan parameter θ, $X_i \sim POI(\theta)$. Kita akan menentukan MLE dari θ. Pertama-tama kita tentukan $L(\theta)$. Karena $X_1, X_2, ..., X_n$ sampel acak dari sebaran Poisson dengan parameter θ, $X_i \sim POI(\theta)$, maka fungsi kemungkinannya adalah

$$L(\theta)=f(x_1;\theta)f(x_2;\theta)...f(x_n;\theta)=\frac{e^{-\theta}\theta^{x_1}}{x_1!}\frac{e^{-\theta}\theta^{x_2}}{x_2!}...\frac{e^{-\theta}\theta^{x_n}}{x_n!}=\frac{e^{-n\theta}\theta^{\sum_{i=1}^{n}x_i}}{\Pi_{i=1}^{n}x_i!}.$$

Lambang $\Pi_{i=1}^{n}x_i!$ berarti perkalian dari $x_1!, ..., x_n!$. Logaritma dari fungsi kemungkinannya adalah

$$\ln L(\theta)=-n\theta+\sum_{i=1}^{n}x_i\ln\theta-\ln\left(\Pi_{i=1}^{n}x_i!\right).$$

Lebih lanjut kita peroleh persamaan kemungkinan

$$\frac{d}{d\theta}\ln L(\theta)=-n+\frac{\sum_{i=1}^{n}x_i}{\theta}=0.$$

Oleh karena itu kita mendapatkan solusi

$$\hat{\theta} = \frac{\sum_{i=1}^{n} x_i}{n} = \overline{x}_n.$$

Jadi MLE dari θ adalah $\hat{\theta} = \overline{x}_n$.

Untuk menyelidiki bahwa solusi di atas memang memberikan hasil yang maksimum untuk $\ln L(\theta)$, maka secara kalkulus kita harus dapat menunjukkan bahwa turunan kedua dari $\frac{d^2}{d\theta^2}\ln L(\theta)$ pada $\theta = \overline{x}_n$ negatif, seperti dapat kita lihat berikut ini.

$$\frac{d^2}{d\theta^2}\ln L(\theta) = -\frac{\sum_{i=1}^{n} x_i}{\theta^2} = -\frac{\sum_{i=1}^{n} x_i}{\overline{x}_n^{\,2}} = -\frac{\frac{n}{n}\sum_{i=1}^{n} x_i}{\overline{x}_n^{\,2}} = -\frac{n\overline{x}_n}{\overline{x}_n^{\,2}} = -\frac{n}{\overline{x}_n}.$$

Karena $X_1, X_2, ..., X_n$ sampel acak dari sebaran Poisson, maka $\overline{x} > 0$. Oleh karena itu

$$\frac{d^2 \ln L(\theta)}{d\theta^2} < 0.$$

Contoh 7.3.3 Misal $X_1, X_2, ..., X_n$ sampel acak dari sebaran dengan pdf

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

Maka kita peroleh fungsi kemungkinan

$$L(\theta) = \theta^{\sum_{i=1}^{n} x_i} (1-\theta)^{n-\sum_{i=1}^{n} x_i} I(x_i = 0,1,\ i = 1,2,3,...,\ 0 \le \theta \le 1).$$

Logaritma dari fungsi kemungkinan ini adalah

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

Sehingga kita peroleh persamaan kemungkinan

$$\frac{d}{d\theta}\ln L(\theta) = \frac{\sum_{i=1}^{n} x_i}{\theta} + \frac{n-\sum_{i=1}^{n} x_i}{1-\theta}(-1) = \frac{\sum_{i=1}^{n} x_i}{\theta} - \frac{n-\sum_{i=1}^{n} x_i}{1-\theta} = 0.$$

Persamaan ini ekivalen dengan persamaan

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

yang solusinya merupakan MLE dari $\theta,$ yaitu

$$\hat{\theta} = \frac{\sum_{i=1}^{n} x_i}{n} = \overline{x}_n.$$

Contoh 7.3.4 Misal $X_1, X_2, ..., X_n$ sampel acak dari sebaran normal, $X \sim N(\mu, \theta)$. Kita akan menentukan MLE dari parameter μ dan $\theta.$ Kita mempunyai pdf

$$f(x;\mu,\theta) = \frac{1}{\sqrt{2\pi\theta}} e^{-\frac{(x-\mu)^2}{2\theta}}.$$

Oleh karena itu fungsi kemungkinannya

$$L(\mu,\theta) = \frac{1}{\sqrt{2\pi\theta}} e^{-\frac{(x_1-\mu)^2}{2\theta}} \cdots \frac{1}{\sqrt{2\pi\theta}} e^{-\frac{(x_n-\mu)^2}{2\theta}} = (2\pi\theta)^{-\frac{n}{2}} \exp\left[-\frac{\sum_{i=1}^{n}(x_i-\mu)^2}{2\theta}\right].$$

Logaritma fungsi kemungkinan ini adalah

$$\ln L(\theta) = -\frac{n}{2}\ln(2\pi) - \frac{n}{2}\ln\theta - \frac{\sum_{i=1}^{n}(x_i-\mu)^2}{2\theta}.$$

Karena terdapat dua parameter yang akan kita duga, maka kita perlukan prosedur turunan parsial. Mula-mula kita turunkan terhadap $\mu,$ kemudian terhadap $\theta,$ dan masing-masing turunan tersebut kita buat nol. Solusi dari kedua persamaan tersebut merupakan MLE dari μ dan $\theta.$ Lebih jelasnya adalah berikut ini.

$$\frac{\partial \ln L(\mu,\theta)}{\partial\mu} = \frac{2\sum_{i=1}^{n}(x_i-\mu)}{2\theta} = 0.$$

Dari persamaan ini kita peroleh MLE dari μ

$$\hat{\mu} = \overline{x}_n.$$

Berikutnya adalah

$$\frac{\partial \ln L(\mu,\theta)}{\partial \theta} = -\frac{n}{2\theta} + \frac{\sum_{i=1}^{n}(x_i-\mu)^2}{2\theta^2} = 0.$$

Persamaan tersebut ekivalen dengan

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

Dari persamaan ini dan dari hasil $\hat{\mu} = \bar{x}$, kita peroleh MLE dari θ adalah

$$\hat{\theta} = \frac{\sum_{i=1}^{n}(x_i-\bar{x}_n)^2}{n}.$$

Dari ketiga contoh di atas kita mendapatkan kesan bahwa untuk mendapatkan MLE, maka prosedur yang dipakai adalah dengan cara penurunan. Meskipun demikian tidak selalu prosedur tersebut yang digunakan, terutama untuk parameter yang bergantung pada data pengamatan. Berikut ini adalah contoh bahwa untuk mendapatkan MLE tidak menggunakan proses penurunan.

Contoh 7.3.5 Misal $X_1, X_2, ..., X_n$ sampel acak dari sebaran dengan pdf

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

Maka fungsi kemungkinannya adalah

$$L(\theta) = \frac{1}{\theta^n} I(0 < x_i \le \theta, i = 1, 2, ..., n).$$

Dari persyaratan bahwa fungsi tersebut terdefinisi pada $x_i \le \theta,$ untuk masing-masing $i = 1, 2, ..., n,$ maka $\hat{\theta}$ haruslah maksimum dari x_i, $i = 1, 2, ..., n.$ Dengan kata lain penduga dari θ adalah statistik urutan terbesar dari $X_1, X_2, ..., X_n$.

Seperti halnya pada MME, pada MLE berlaku pula sifat invarians, yaitu

$$\hat{\tau}(\theta_1, \theta_2, ..., \theta_n) = \tau(\hat{\theta}_1, \hat{\theta}_2 ..., \hat{\theta}_n).$$

Contoh 7.3.6 Misal $X_1, X_2, ..., X_n$ sampel acak dari sebaran Poisson dengan parameter θ, $X_i \sim POI(\theta)$. Kita akan menentukan MLE dari

$$\tau = \tau(\theta) = P[X = 0] = e^{-\theta}.$$

Dari Contoh 7.3.2 sudah kita peroleh bahwa $\hat{\theta} = \bar{x}_n$. Oleh karena itu menggunakan sifat invarians kita dapatkan

$$\hat{\tau} = \tau(\hat{\theta}) = e^{-\hat{\theta}} = e^{-\bar{x}_n}.$$

Misal θ_0 menyatakan *nilai benar* dari θ. Teorema 7.3.1 memberi alasan pada kemaksimuman fungsi kemungkinan. Untuk membuktikannya diperlukan suatu asumsi yang disebut sebagai ***kondisi keregularan (regularity conditions).***

Asumsi 7.3.1 Kondisi Kereguleran (Regularity Conditions).

(R0): pdf-pdf adalah berbeda, yaitu, $\theta \neq \theta' \Rightarrow f(x_i;\theta) \neq f(x_i;\theta')$.

(R1): pdf-pdf tedefinisi pada daerah yang sama untuk semua θ.

(R2): Titik θ_0 adalah titik dalam dari Ω.

Teorema 7.3.1 Misal θ_0 adalah parameter benar. Di bawah asumsi (R0) dan (R1),

$$\lim_{n\to\infty} P_{\theta_0}\left[L(\theta_0, \mathbf{X}) > L(\theta, \mathbf{X})\right] = 1, \text{ untuk semua } \theta \neq \theta_0.$$

Bukti

Logaritma dari ketaksamaan $L(\theta_0, \mathbf{X}) > L(\theta, \mathbf{X})$ adalah ekivalen dengan

$$\frac{1}{n}\sum_{i=1}^{n} \ln\left[\frac{f(X_i;\theta)}{f(X_i;\theta_0)}\right] < 0.$$

Karena jumlah-jumlah tersebut identik dengan ekspektasi terhingga dan $\phi(x) = -\ln(x)$ adalah cembung sempurna, maka menggunakan Hukum Bilangan Besar dan Ketaksamaan Jensen (Susiswo, Teori Peluang: Teorema 6.5.8), bahwa, jika θ_0 parameter benar, maka

$$\frac{1}{n}\sum_{i=1}^{n} \ln\left[\frac{f(X_i;\theta)}{f(X_i;\theta_0)}\right] \xrightarrow{P} E_{\theta_0}\left[\ln\frac{f(X_i;\theta)}{f(X_i;\theta_0)}\right] < \ln E_{\theta_0}\left[\frac{f(X_i;\theta)}{f(X_i;\theta_0)}\right].$$

Tetapi

$$E_{\theta_0}\left[\frac{f(X_i;\theta)}{f(X_i;\theta_0)}\right]=\int\frac{f(X_i;\theta)}{f(X_i;\theta_0)}f(X_i;\theta_0)dx=1.$$

Karena $\ln 1=0,$ maka bukti selesai.

Teorema 7.3.1 mengatakan bahwa keasimtotan fungsi kemungkinan maksimum pada nilai benar θ_0.

Latihan 7.3

7.3.1. Tentukan MLE dari θ berdasarkan sampel acak $X_1, X_2,..., X_n$ untuk masing-masing pdf berikut ini:

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

(b) $f(x;\theta)=(\theta+1)x^{-\theta-2}I(1<x, 0<\theta).$

(c) $f(x;\theta)=\theta^2 xe^{-\theta x}I(0<x, 0<\theta).$

(d) $f(x;\theta)=\theta^x e^{-\theta}/x!I(x=0,1,2,..., 0\le\theta<\infty),$ di mana $f(0;0)=1.$

7.3.2. Tentukan MLE dari θ berdasarkan sampel acak $X_1, X_2,..., X_n$ untuk masing-masing sebaran berikut ini:

(a) $X_i \sim BIN(1,p)$.

(b) $X_i \sim BIN(3,p)$.

(c) $X_i \sim GAM(2,\kappa)$.

(d) $X_i \sim WEI(\theta,\frac{1}{2})$.

(e) $X_i \sim PAR(1,\kappa)$.

7.3.3. Misal $X_1, X_2,..., X_n$ sampel acak dari sebaran eksponensial dua parameter, $X_i \sim EXP(1,\eta)$. Tentukan MLE dari η. *Catatan*; Jika $X \sim EXP(\theta,\eta)$, maka $f(x)=\dfrac{1}{\theta}e^{-\frac{(x-\eta)}{\theta}}I(0<\theta,\eta<x).$

7.3.4. Misal $X_1, X_2,..., X_n$ sampel acak dari sebaran eksponensial dua parameter, $X_i \sim EXP(\theta,\eta)$. Tentukan MLE dari η dan θ.

7.3.5. Misal $X_1, X_2, ..., X_n$ sampel acak dari sebaran geometrik, $X_i \sim GEO(p)$. Tentukan MLE dari:

(a) $E(X) = \dfrac{1}{p}$.

(b) $Var(X) = \dfrac{1-p}{p^2}$.

(c) $P[X > k] = (1-p)^k,\ k = 1, 2, 3, ...$

Petunjuk: Gunakan sifat invarians.

7.3.6. Misal $X_1, X_2, ..., X_n$ sampel acak dari sebaran dengan pdf

$f(x;\theta) = \dfrac{1}{\theta} e^{-\frac{x}{\theta}} I(0 < x < \infty).$ Tentukan MLE dari $P(X \leq 2).$

7.3.7. Diberikan tabel berikut

x	0	1	2	3	4	5
Frekuensi	6	10	14	13	6	1

yang merupakan ringkasan dari sampel berkuran 50 dari sebaran binomial dengan $n = 5.$ Tentukan MLE dari $P(X \geq 3).$

7.3.8. Diberikan tabel berikut

x	0	1	2	3	4	5
Frekuensi	7	14	12	13	6	3

yang merupakan ringkasan dari sampel berkuran 50 dari sebaran Poisson. Tentukan MLE dari $P(X = 2).$

7.3.9. Misal $X_1, X_2, ..., X_5$ sampel acak dari ***sebaran Cauchy*** dengan median $\theta,$ yaitu, dengan pdf

$$f(x;\theta)=\frac{1}{\pi}\frac{1}{1+(x-\theta)^2},\ -\infty<x<\infty,\ -\infty<\theta<\infty.$$

Jika $x_1=-1.94$, $x_2=0.59$, $x_3=-5.98$, $x_4=-0.08$, dan $x_5=-0.77$, dengan metode numerik tentukan MLE dari θ.

7.3.10. Misal $X_1, X_2, ..., X_n$ sampel acak dari sebaran normal, $X_i \sim N(\theta, \sigma^2)$ di mana σ^2 tetap, dan $0 \le \theta < \infty$. Tunjukkan bahwa MLE dari θ adalah $\hat{\theta}=\max\{0, \bar{X}_n\}$.

7.4 Penduga Takbias dan Konsisten

Beberapa sifat dari penduga mensyaratkan bahwa penduga tersebut harus merupakan penduga takbias. Konsep tentang penduga takbias diberikan oleh Definisi 7.4.1 berikut ini.

Definisi 7.4.1 Suatu penduga T dikatakan ***penduga takbias*** dari $\tau(\theta)$ jika

$$E(T)=\tau(\theta),$$

untuk semua $\theta \in \Omega$. Dan jika lain dikatakan bahwa T penduga bias dari $\tau(\theta)$.

Pada Teorema 5.2.1 dan Teorema 5.2.2 berturut-turut kita telah memperoleh $E(\bar{X}_n)=\mu$ dan $E(S_n^2)=\sigma^2$. Oleh karena itu $\bar{X}_n$ dan S_n^2 berturut-turut merupakan penduga takbias dari μ dan σ^2.

Contoh 7.4.1 Misal $X_1, X_2, ..., X_n$ sampel acak dari sebaran eksponensial, $X_i \sim EXP(\theta)$. Dari Contoh 7.2.2 kita telah memperoleh bahwa MME dari θ adalah $\bar{X}_n$. Karena

$$E(\bar{X}_n)=\mu=\theta,$$

maka $\bar{X}_n$ merupakan penduga takbias dari θ.

Contoh 7.4.2 Misal $X_1, X_2, ..., X_n$ sampel acak dari sebaran seragam, $X_i \sim UNIF(0,\theta)$. Kita akan menyelidiki penduga dari θ. Secara intuitif dugaan dari θ adalah maksimum dari sampel ini. Misal $Y_n = X_{n,n}$, yaitu statistik urutan terbesar dari $X_1, X_2, ..., X_n$. Maka cdf dan pdf dari Y_n berturut-turut adalah

$$F_{Y_n}(y) = \left(\frac{y}{\theta}\right)^n I(0 \le y < \theta) + I(\theta \le y),$$

dan

$$f_{Y_n}(y) = \frac{n}{\theta^n} y^{n-1} I(0 < y < \theta).$$

Berdasarkan pdf ini, mudah untuk menunjukkan bahwa $E(Y_n) = \frac{n}{n+1}\theta.$ Jadi Y_n penduga bias dari θ. Meskipun demikian $\frac{n+1}{n}Y_n$ adalah penduga takbias dari θ.

Contoh 7.4.3 Misal $X_1, X_2, ..., X_n$ sampel acak dari sebaran dengan pdf, $f(x)$. Anggap $E(X) = \mu$ terhingga dan anggap $f(x)$ simetrik terhadap μ. Dari Teorema 5.2.1 telah kita peroleh $\bar{X}_n$ adalah penduga takbias dari μ. Bagaimana tentang median? Misal $T = T(X_1, X_2, ..., X_n) = \text{median}\{X_1, X_2, ..., X_n\}$. Median sampel memenuhi dua sifat: (1) jika kita naikkan (turunkan) butir-butir sampel dengan b maka median sampel naik (turun) dengan b, dan (2) jika kita kalikan masing-masing butir sampel dengan -1 maka median sampel juga kelipatan -1. Kita dapat meringkas kedua sifat tersebut seperti berikut:

$$T = T(X_1 + b, X_2 + b, ..., X_n + b) = T\left(X_1, X_2, ..., X_n\right) + b,$$

$$T = T(-X_1, -X_2, ..., -X_n) = -T\left(X_1, X_2, ..., X_n\right).$$

Vektor acak $(X_1 - \mu, X_2 - \mu, ..., X_n - \mu)$ bersebaran sama dengan $\left(-(X_1 - \mu), -(X_2 - \mu), ..., -(X_n - \mu)\right)$. Secara khusus ekspektasi kedua vektor acak ini sama. Oleh karena itu kita mempunyai:

$$
\begin{aligned}
E[T]-\mu &= E[T(X_1, X_2, \ldots, X_n)]-\mu \\
&= E[T(X_1-\mu, X_2-\mu, \ldots, X_n-\mu)] \\
&= E\left[T\left(-(X_1-\mu), -(X_2-\mu), \ldots, -(X_n-\mu)\right)\right] \\
&= -E[T(X_1-\mu, X_2-\mu, \ldots, X_n-\mu)] \\
&= -E[T(X_1, X_2, \ldots, X_n)]+\mu \\
&= -E[T]+\mu.
\end{aligned}
$$

Oleh Karena itu kita peroleh $2E[T]=2\mu,$ sehingga $E[T]=\mu.$ Jadi sampel median adalah juga penduga takbias dari $\theta.$ Manakah di antara kedua penduga tersebut yang lebih baik? Jawaban dari pertanyaan ini dibahas di belakang.

Pembahasan **s**elanjutnya tentang penduga konsisten, yang diberikan oleh Definisi 7.4.2 berikut ini.

Definisi 7.4.2 Misal X adalah peubah acak dengan cdf $F(x;\theta),\ \theta\in\Omega.$ Misal $X_1, X_2, \ldots, X_n$ sampel acak dari sebaran ini, dan misal T_n suatu statistik. Kita katakan T_n ***penduga konsisten*** dari θ jika $T_n \xrightarrow{P} \theta.$

Contoh 7.4.4 Misal $X_1, X_2, \ldots, X_n$ adalah sampel acak dari sebaran dengan purata μ dan varians $\sigma^2<\infty.$ Maka menggunakan Hukum Lemah Bilangan Besar purata sampel, $\bar{X}_n,$ merupakan penduga konsisten dari $\mu.$ Apakah varians sampel juga penduga konsisten dari σ^2? Kita lihat seperti berikut ini.

$$
\begin{aligned}
S_n^2 &= \frac{1}{n-1}\sum_{i=1}^{n}\left(X_i-\bar{X}_n\right)^2 = \frac{n}{n-1}\left(\frac{1}{n}\sum_{i=1}^{n}X_i^2-\bar{X}_n^2\right) \\
&\xrightarrow{P} 1.\left[E\left(X_1^2\right)-\mu^2\right]=\sigma^2.
\end{aligned}
$$

Jadi S_n^2 penduga konsisten dari $\theta.$

Contoh 7.4.5 Ulangi Contoh 7.4.2, di mana $X_1, X_2, ..., X_n$ sampel acak dari sebaran seragam, $X_i \sim UNIF(0, \theta)$. Kita akan menyelidiki penduga dari θ. Berdasarkan cdf dari $Y_n = X_{n,n}$, mudah dilihat bahwa $Y_n \xrightarrow{P} \theta$. Jadi Y_n penduga konsisten dari θ. Demikian juga penduga takbias dari θ, yaitu $\frac{n+1}{n} Y_n$, dia juga penduga konsisten dari θ.

Latihan 7.4

7.4.1. Misal $X_1, X_2, ..., X_n$ sampel acak dari sebaran geometrik dengan parameter θ, $X_i \sim GEO(\theta)$. Tunjukkan bahwa $\bar{X}_n$ penduga takbias dari $\tau(\theta) = \frac{1}{\theta}$.

7.4.2. Perhatikan sampel acak berukuran n dari sebaran normal, $X_i \sim N(\mu, 9)$.
(a) Apakah MLE dari μ merupakan penduga takbias?
(b) Apakah MLE dari persentil ke 95 merupakan penduga takbias?

7.4.3. Perhatikan sampel acak berukuran n dari sebaran normal, $X_i \sim N(0, \theta)$. Apakah MLE dari θ, merupakan penduga takbias?

7.4.4. Ulangi Latihan 6.7.5 untuk memperoleh purata dari Y_n. Apakah Y_n penduga takbias dari θ? Tentukan penduga takbias dari θ berdasarkan pada Y_n.

7.4.5. Lihat kembali Latihan 5.3.1 s/d 5.3.4. Statistik manakah yang merupakan penduga konsisten dari parameternya?

7.5 Batas Bawah Rao-Cramer dan Keefisienan

Dalam subpokok bahasan ini kita akan membuktikan suatu ketaksamaan yang biasa yang disebut sebagai batas bawah Rao-Cramer yang memberikan batas bawah pada varians untuk sebarang penduga takbias.

Kita akan menunjukkan bahwa di bawah kondisi kereguleran, varians dari penduga kemungkinan maksimum menuju batas bawah ini ***asimtotik***.

Misal X peubah acak dengan pdf $f(x;\theta),\ \theta \in \Omega,$ di mana ruang parameter Ω adalah selang buka. Kita memerlukan kondisi kereguleran tambahan berikut ini.

Asumsi 7.5.1 Kondisi Kereguleran Tambahan (Additional Regularity Conditions).

(R3): pdf $f(x;\theta)$ dapat diturunkan dua kali atas θ.

(R4): $\int f(x;\theta)dx$ dapat diturunkan dua kali atas θ di bawah tanda integral.

Kondisi (R1)-(R4) mempunyai arti bahwa parameter θ tidak muncul dalam ujung selang di mana $f(x;\theta) > 0$ dan kita dapat mengubah pengintegralan dan penurunan berkenaan dengan θ. Pembahasan kita untuk kasus kontinu, untuk kasus diskret dapat diperoleh serupa. Kita mulai dengan identitas berikut

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

Jika identitas tersebut kita turunkan terhadap $\theta,$ maka kita peroleh

$$0 = \int_{-\infty}^{\infty} \frac{\partial f(x;\theta)}{\partial \theta} dx.$$

Kita ubah identitas tersebut menjadi

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

ekivalen dengan

$$0 = \int_{-\infty}^{\infty} \frac{\partial \ln f(x;\theta)}{\partial \theta} f(x;\theta)dx. \qquad (7.5.1)$$

Kita tulis persamaan terakhir menjadi bentuk ekspektasi,

$$E\left[\frac{\partial \ln f(X;\theta)}{\partial \theta}\right] = 0, \qquad (7.5.2)$$

yang berarti bahwa purata dari peubah acak $\dfrac{\partial \ln f(X;\theta)}{\partial \theta}$ adalah 0. Jika persamaan (7.5.1) kita turunkan terhadap $\theta,$ maka kita peroleh

$$0=\int_{-\infty}^{\infty}\frac{\partial^2 \ln f(x;\theta)}{\partial\theta^2} f(x;\theta)dx+\int_{-\infty}^{\infty}\frac{\partial \ln f(x;\theta)}{\partial\theta}\frac{\partial \ln f(x;\theta)}{\partial\theta} f(x;\theta)dx.$$

(7.5.3)

Suku kedua pada ruas kanan persamaan (7.5.3) dapat kita tulis sebagai ekspektasi, yang disebut ***Informasi Fisher*** yang dinyatakan sebagai $I(\theta)$, yaitu,

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

(7.5.4.)

Menggunakan persamaan (7.5.3) informasi Fisher dapat dihitung dari

$$I(\theta)=-\int_{-\infty}^{\infty}\frac{\partial^2 \ln f(x;\theta)}{\partial\theta^2} f(x;\theta)dx. \qquad (7.5.5)$$

Menggunakan persamaan (7.5.2), maka informasi Fisher merupakan varians dari $\frac{\partial \ln f(X;\theta)}{\partial\theta}$, yaitu

$$I(\theta)=\mathrm{var}\left(\frac{\partial \ln f(X;\theta)}{\partial\theta}\right). \qquad (7.5.6)$$

Fungsi

$$\frac{\partial \ln f(x;\theta)}{\partial\theta}$$

disebut sebagai ***fungsi skor.*** Ingat bahwa dalam menentukan MLE dari θ, yaitu $\hat{\theta}$, adalah solusi dari

$$\sum_{i=1}^{n}\frac{\partial \ln f(x_i;\theta)}{\partial\theta}=0.$$

Contoh 7.5.1 Misal suatu peubah acak dari sebaran Bernoulli, yaitu $X\sim BIN(1,\theta)$ $X \sim BIN(1,\theta)$. Maka

$$\ln f(x;\theta)=x\ln\theta+(1-x)\ln(1-\theta).$$

Turunan pertama dan kedua persamaan tersebut terhadap θ beruturut-turut adalah

$$\frac{\partial \ln f(x;\theta)}{\partial\theta}=\frac{x}{\theta}-\frac{(1-x)}{(1-\theta)},$$

dan

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

Jadi

$$I(\theta) = -E\left[-\frac{X}{\theta^2} - \frac{(1-X)}{(1-\theta)^2}\right]$$

$$= \frac{\theta}{\theta^2} + \frac{(1-\theta)}{(1-\theta)^2} = \frac{1}{\theta} + \frac{1}{(1-\theta)} = \frac{1}{\theta(1-\theta)},$$

yang nilainya akan menjadi besar untuk θ yang mendekati 0 atau 1.

Contoh 7.5.2 Misal $X_1, X_2, ..., X_n$ sampel acak sehingga

$$X_i = \theta + e_i,\ i = 1, 2, ..., n,$$

di mana $e_1, e_2, ..., e_n$ mempunyai sampel acak dari sebaran dengan pdf $f(x)$ yang terdefinisi pada $(-\infty, \infty)$. Maka pdf dari X_i adalah $f_{X_i}(x_i;\theta) = f(x_i - \theta),$ yang disebut sebagai ***model lokasi***. Anggap $f(x)$ memenuhi kondisi kereguleran. Maka

$$I(\theta) = \int_{-\infty}^{\infty}\left(\frac{f'(x-\theta)}{f(x-\theta)}\right)^2 f(x-\theta)dx$$

$$= \int_{-\infty}^{\infty}\left(\frac{f'(z)}{f(z)}\right)^2 f(z)dz.$$

Ini berarti model lokasi informasi Fisher-nya tidak bergantung pada parameter $\theta.$ Secara khusus jika $X_1, X_2, ..., X_n$ sampel acak dari ***sebaran Laplace,*** yaitu suatu sebaran yang pdf-nya

$$f(x;\theta) = \frac{1}{2}e^{-|x-\theta|}, -\infty < x < \infty, -\infty < \theta < \infty.$$

Maka X_i dapat ditulis sebagai

$$X_i = \theta + e_i,\ i = 1, 2, ..., n,$$

di mana $e_1, e_2, ..., e_n$ mempunyai sampel acak dari sebaran dengan pdf

$$f(z) = \frac{1}{2}e^{-|z|}, -\infty < z < \infty, -\infty < \theta < \infty.$$

Dalam analisis kita punya $\frac{d}{dz}|z| = \text{sgn}(z),$ di mana $\text{sgn}(z) = 1,\ 0,$ atau -1 bergantung pada apakah $z > 0,\ z = 0,$ atau $z < 0.$ Maka

$$f\,'(z) = -\frac{1}{2}\text{sgn}(z)e^{-|z|}, -\infty < z < \infty, -\infty < \theta < \infty.$$

Oleh karena itu

$$I(\theta) = \int_{-\infty}^{\infty}\left(\frac{f\,'(z)}{f(z)}\right)^2 f(z)dz = \int_{-\infty}^{\infty} f(z)dz = 1.$$

Dari persamaan 7.5.6, untuk sampel berukuran satu, katakanlah $X_1,$ informasi Fisher merupakan varians peubah acak $\frac{\partial \ln f(X_1;\theta)}{\partial\theta}.$ Bagaimana untuk sampel berukuran n? Misal $X_1, X_2, ..., X_n$ sampel acak dari sebaran dengan pdf $f(x;\theta).$ Jika fungsi kemungkinan $L(\theta)$ adalah pdf dari sampel ini dan peubah acak yang variansnya adalah informasi, dalam sampel ini diberikan oleh

$$\frac{\partial \ln L(\theta;\mathbf{X})}{\partial\theta} = \sum_{I=1}^{n}\frac{\partial \ln f(X_i;\theta)}{\partial\theta}.$$

Jumlah ini adalah dari ***iid (independent and identically distributed)*** dengan varians $I(\theta).$ Oleh karena itu informasi sampel ini adalah

$$Var\left(\frac{\partial \ln L(\theta,\mathbf{X})}{\partial\theta}\right) = nI(\theta). \qquad (7.5.7)$$

Informasi dari sampel acak berukuran n ini merupakan n kali dari informasi sampel berukuran satu. Sehingga dalam Contoh 7.5.1, informasi Fisher untuk sampel acak berukuran n dari sebaran Bernoulli $BIN(1,\theta)$ adalah $n/\left[\theta(1-\theta)\right].$

Kita sekarang siap untuk memperoleh batas bawah Rao-Cramer yang diberikan oleh teorema berikut.

Teorema 7.5.1 Batas Bawah Rao-Cramer. Misal $X_1, X_2, ..., X_n$ adalah sampel acak dari sebaran dengan pdf $f(x;\theta)$ untuk $\theta \in \Omega.$ Anggap bahwa kondisi kereguleran (R0)-(R4) terpenuhi. Misal

$Y = u(X_1, X_2, ..., X_n)$ adalah statistik dengan purata $E(Y) = E[u(X_1, X_2, ..., X_n)] = k(\theta).$ Maka

$$Var(Y) \geq \frac{[k'(\theta)]^2}{nI(\theta)}. \tag{7.5.8}$$

Bukti

Kita buktikan untuk kasus kontinu, dan untuk kasus diskret dapat diperoleh dengan cara serupa. Kita tulis ekspektasi dari Y sebagai

$$k(\theta) = \int_{-\infty}^{\infty} ... \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} u(x_1, x_2, ..., x_n) f(x_1;\theta) f(x_2;\theta) ... f(x_n;\theta) dx_i dx_2 ... dx_n.$$

Jika persamaan di atas kita turunkan terhadap $\theta,$ maka kita peroleh

$$\begin{aligned} k'(\theta) &= \int_{-\infty}^{\infty} ... \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} u(x_1, x_2, ..., x_n) \left[\sum_{i=1}^{n} \frac{1}{f(x_i;\theta)} \frac{\partial f(x_i;\theta)}{\partial \theta} \right] \\ &\qquad \times f(x_1;\theta) f(x_2;\theta) ... f(x_n;\theta) dx_i dx_2 ... dx_n. \\ &= \int_{-\infty}^{\infty} ... \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} u(x_1, x_2, ..., x_n) \left[\sum_{i=1}^{n} \frac{\partial \ln f(x_i;\theta)}{\partial \theta} \right] \\ &\qquad \times f(x_1;\theta) f(x_2;\theta) ... f(x_n;\theta) dx_i dx_2 ... dx_n. \end{aligned} \tag{7.5.9}$$

Kita definisikan peubah acak

$$Z = \sum_{i=1}^{n} \left[\frac{\partial \ln f(X_i;\theta)}{\partial \theta} \right].$$

Berdasarkan persamaan (7.5.2) dan (7.5.7) berturut-turut kita peroleh $E(Z) = 0$ dan $Var(Z) = nI(\theta).$ Sehingga persamaan (7.5.9) dapat dinyatakan sebagai $k'(\theta) = E(YZ).$ Oleh karena itu

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

di mana ρ merupakan koefisien korelasi antara Y dan $Z.$ Menggunakan $E(Z) = 0,$ kita mempunyai

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

Karena $\rho^2 \leq 1,$ maka

$$\left[\frac{k'(\theta)}{\sigma_Y \sqrt{nI(\theta)}} \right]^2 \leq 1,$$

yang dapat kita susun seperti pada persamaan (7.5.8), bukti selesai.

Akibat 7.5.1 Di bawah asumsi pada Teorema 7.5.1, jika

$$Y = u(X_1, X_2, ..., X_n)$$

penduga dari θ, sehingga $k(\theta) = \theta$, maka ketaksamaan Rao-Cramer menjadi

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

Perhatikan kembali model Bernoulli dengan peluang sukses θ yang telah dicoba pada Contoh 7.5.1. Telah ditunjukkan bahwa $1/nI(\theta) = \theta(1-\theta)/n.$ Dari Latihan 7.3.2 (a) MLE dari θ adalah $\bar{X}$. Purata dan varians dari sebaran ini berturut-turut adalah θ dan $\theta(1-\theta)$. Oleh karena itu purata dan varians dari $\bar{X}$ berturut-turut adalah θ dan $\theta(1-\theta)/n.$ Dalam kasus ini varians dari MLE mencapai batas bawah Rao-Cramer.

Definisi 7.5.1 Misal Y penduga takbias dari parameter θ. Statistik Y disebut ***penduga efisien*** dari θ jika varians dari Y mencapai batas bawah Rao-Cramer.

Definisi 7.5.2 Dalam hal kita dapat menurunkan terhadap suatu parameter di bawah tanda integral atau jumlah, maka nisbah dari batas bawah Rao-Cramer dan varians dari sebarang penduga takbias dari suatu parameter disebut ***keefisienan*** *(efficiency)* penduga.

Contoh 7.5.3 Misal $X_1, X_2, ..., X_n$ sampel acak dari sebaran Poisson yang mempunyai purata $\theta > 0.$ Dari Contoh 7.3.2 kita telah memperoleh $\bar{X}$ adalah MLE dari θ. Kita akan menunjukkan bahwa penduga ini adalah penduga efisien dari θ. Kita mempunyai

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

Maka

$$E\left[\left(\frac{\partial \ln f(X;\theta)}{\partial\theta}\right)^2\right]=E\left[\left(\frac{X-\theta}{\theta}\right)^2\right]=\frac{\sigma^2}{\theta^2}=\frac{\theta}{\theta^2}=\frac{1}{\theta}.$$

Oleh karena itu kita peroleh batas bawah Rao-Cramer $1/\left[n(1/\theta)\right]=\dfrac{\theta}{n}$, yang juga merupakan varians dari $\bar{X}$. Jadi $\bar{X}$ adalah penduga efisien dari θ.

Contoh 7.5.4 Misal $X_1, X_2, ..., X_n$ sampel acak berukuran $n>2$ dari sebaran $BETA(\theta,1)$, yaitu sebaran dengan pdf

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

Kita mempunyai

$$\frac{\partial \ln f(x;\theta)}{\partial\theta}=\ln x+\frac{1}{\theta}.$$

Dan turunan keduanya adalah

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

Oleh karena itu informasi Fisher $I(\theta)=\theta^{-2}$. Jadi kita peroleh batas bawah Rao-Cramer $1/\left[nI(\theta)\right]=\dfrac{\theta^2}{n}$. Dari Latihan 7.1.3 (a) kita dapat memperoleh bahwa MLE dari θ adalah $\hat{\theta}=-n/\sum_{i=1}^{n}\ln X_i$. Untuk memperoleh sebaran dari $\hat{\theta}$, kita misalkan $Y_i=-\ln X_i$. Transformasi ini menghasilkan $Y_i \sim GAM(1/\theta, 1)$. Dari Latihan 5.4.4 kita peroleh $W=\sum_{i=1}^{n} Y_i \sim GAM(\frac{1}{\theta}, n)$. Oleh karena itu menggunakan Teorema 5.4.1 kita peroleh

$$E\left[W^k\right]=\frac{(n+k-1)!}{\theta^k(n-1)!},$$

untuk $k>n$. Secara khusus untuk $k=-1$, kita peroleh

$$E\left[\hat{\theta}\right]=nE\left[W^{-1}\right]=\theta\frac{n}{n-1}.$$

Oleh karena itu $\hat{\theta}$ penduga bias, tetapi bias ini hilang jika $n\to\infty$. Untuk $k=-2$, kita memperoleh

$$E\left[\hat{\theta}^2\right] = n^2 E\left[W^{-2}\right] = \theta^2 \frac{n^2}{(n-1)(n-2)}.$$

Jadi

$$Var(\hat{\theta}) = \theta^2 \frac{n^2}{(n-1)^2(n-2)}.$$

Varians ini lebih besar dari pada batas bawah Rao-Cramer. Jadi $\hat{\theta}$ bukan penduga efisien dari θ.

Asumsi 7.5.2 Kondisi Kereguleran Tambahan (Additional Regularity Conditions).

(R5): pdf $f(x;\theta)$ dapat diturunkan tiga kali sebagai fungsi dari θ. Selanjutnya, untuk semua $\theta \in \Omega$, ada konstanta c dan fungsi $M(x)$ sehingga

$$\left| \frac{\partial^3}{\partial\theta^3} \ln f(x;\theta) \right| \le M(x),$$

dengan $E_{\theta_0}\left[M(x)\right] < \infty$, untuk semua $\theta_0 - c < \theta < \theta_0 + c$ dan untuk semua x dalam ruang peubah acak X.

Teorema 7.5.2 Misal $X_1, X_2, ..., X_n$ sampel acak dari suatu sebaran dengan pdf $f(x;\theta_0)$ untuk $\theta_0 \in \Omega$ sehingga kondisi kereguleran (R1)-(R5) terpenuhi. Anggap informasi Fisher memenuhi $0 < I(\theta) < \infty$. Maka untuk sebarang barisan dari solusi MLE memenuhi

$$\sqrt{n}\left(\hat{\theta} - \theta_0\right) \xrightarrow{d} N\left(0, \frac{1}{I(\theta_0)}\right). \qquad (7.5.10)$$

Bukti

Kita ekspansi fungsi turunan dari kemungkinan $L(\theta)$ ke dalam deret Taylor sekitar θ_0 pada $\hat{\theta}_n$

$$L'(\hat{\theta}_n) = L'(\theta_0) + \left(\hat{\theta}_n - \theta_0\right) L''(\theta_0) + \frac{1}{2}\left(\hat{\theta}_n - \theta_0\right)^2 L'''(\theta_n^*), \qquad (7.5.11)$$

di mana θ_n^* antara θ_0 dan $\hat{\theta}_n$. Karena $L'(\hat{\theta}_n) = 0$, maka kita peroleh

$$\sqrt{n}(\hat{\theta}-\theta_0)=\frac{n^{-1/2}L'(\theta_0)}{-n^{-1}L''(\theta_0)-(2n)^{-1}(\hat{\theta}-\theta_0)L'''(\theta_n^*)}. \qquad (7.5.12)$$

Menggunakan CLT kita peroleh

$$\frac{1}{\sqrt{n}}L'(\theta_0)=\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\frac{\partial \ln f(X_i;\theta_0)}{\partial\theta}\xrightarrow{D}N(0,I(\theta_0)), \qquad (7.5.13)$$

karena jumlah ini iid dengan $Var\left(\frac{\partial \ln f(X_i;\theta_0)}{\partial\theta}\right)=I(\theta_0)<\infty.$ Juga menggunakan Hukum Bilangan Besar,

$$-\frac{1}{n}L'(\theta_0)=-\frac{1}{n}\sum_{i=1}^{n}\frac{\partial^2 \ln f(X_i;\theta_0)}{\partial\theta^2}\xrightarrow{P}I(\theta_0). \qquad (7.5.14)$$

Untuk melengkapi bukti ini kita perlu hanya menunjukkan bahwa suku kedua dalam penyebut pada persamaan (7.5.12) konvergen dalam peluang ke nol. Karena $\hat{\theta}_n-\theta_0\xrightarrow{P}0,$ oleh Teorema 6.3.7, dapat diperoleh jika kita dapat menunjukkan bahwa $n^{-1}L'''(\theta_0^*)$ adalah terbatas dalam peluang. Misal c_0 adalah konstanta dalam kondisi (R5). Catatan bahwa $\left|\hat{\theta}_n-\theta_0\right|<c_0$ berakibat $\left|\theta_n^*-\theta_0\right|<c_0,$ yang menggunakan kondisi (R5) berakibat

$$\left|-\frac{1}{n}L'''(\theta_n^*)\right|\le\frac{1}{n}\sum_{i=1}^{n}\left|\frac{\partial^3 f(X_i;\theta)}{\partial\theta^3}\right|\le\frac{1}{n}\sum_{i=1}^{n}M(X_i). \qquad (7.5.15)$$

Oleh kondisi (R5), $E_{\theta_0}\left[M(X)\right]<\infty.$ Oleh karena itu

$$\frac{1}{n}\sum_{i=1}^{n}M(X_i)\xrightarrow{P}E_{\theta_0}\left[M(X)\right].$$

Untuk batas, kita pilih $1+E_{\theta_0}\left[M(X)\right].$ Misal $\varepsilon>0$ sebarang. Pilih N_1 dan N_2 sehingga

$$n\ge N_1\Rightarrow P\left[\left|\hat{\theta}_n-\theta_0\right|<c_0\right]\ge 1-\frac{\varepsilon}{2}. \qquad (7.5.16)$$

$$n\ge N_2\Rightarrow P\left[\left|\frac{1}{n}\sum_{i=1}^{n}M(X_i)-E_{\theta_0}\left[M(X)\right]\right|<1\right]\ge 1-\frac{\varepsilon}{2}. \qquad (7.5.17)$$

Selanjutnya dari persamaan (7.5.15), (7.5.16) dan (7.5.17) bahwa

$$n \ge \max\{N_1, N_2\} \Rightarrow P\left[\left|-\frac{1}{n}L'''(\theta_n^*)\right| \le 1 + E_{\theta_0}\left[M(X)\right]\right] \ge 1 - \frac{\varepsilon}{2},$$

maka $n^{-1}L'''(\theta_n^*)$ adalah terbatas dalam peluang.

Kita merampat (*generalize*) Definisi 7.5.1 dan Definisi 7.5.2 mengenai keefisienan asimtotik.

Definisi 7.5.3 Misal $X_1, X_2, ..., X_n$ sampel acak dari sebaran dengan pdf $f(x;\theta)$. Anggap $\hat{\theta}_{1n} = \hat{\theta}_{1n}\left(X_1, X_2, ..., X_n\right)$ adalah penduga dari θ_0 sehingga

$$\sqrt{n}\left(\hat{\theta}_{1n} - \theta_0\right) \xrightarrow{D} N\left(0, \sigma^2_{\hat{\theta}_{1n}}\right).$$

Maka

(a) ***Keefisienan asimtotik*** dari $\hat{\theta}_{1n}$ didefinisikan sebagai

$$e\left(\hat{\theta}_n\right) = \frac{1/I(\theta_0)}{\sigma^2_{\hat{\theta}_{1n}}}.$$

(b) Penduga $\hat{\theta}_{1n}$ dikatakan ***efisien secara asimtotik*** jika nisbah dalam (a) adalah satu.

(c) Misal $\hat{\theta}_{2n}$ penduga yang lain sehingga

$$\sqrt{n}\left(\hat{\theta}_{2n} - \theta_0\right) \xrightarrow{D} N\left(0, \sigma^2_{\hat{\theta}_{2n}}\right).$$

Maka ***keefisienan relatif asimtotik*** dari $\hat{\theta}_{1n}$ ke $\hat{\theta}_{1n}$ adalah

$$e\left(\hat{\theta}_{1n}, \hat{\theta}_{2n}\right) = \frac{\sigma^2_{\hat{\theta}_{2n}}}{\sigma^2_{\hat{\theta}_{1n}}}.$$

Oleh karena itu dengan Teorema 7.5.2, di bawah kondisi kereguleran, penduga kemungkinan maksimum adalah penduga efisien secara asimtotik. Ini adalah hasil yang baik. Juga, jika dua penduga adalah normal secara asimtotik dengan purata asimtotik sama, maka secara intuitif penduga dengan varians asimtotik yang lebih kecil yang akan kita pilih dari yang lainnya sebagai penduga yang lebih baik. Dalam hal ini, keefisienan relatif asimtotik dari penduga terpilih ke yang tidak terpilih adalah lebih besar dari satu.

Latihan 7.5

7.5.1. Buktikan bahwa $\bar{X}_n$, yang merupakan purata sampel acak dari sebaran normal, $X_i \sim N(\theta, \sigma^2)$, $-\infty < \theta < \infty$, untuk setiap $\sigma^2 > 0$ diketahui, adalah penduga efisien dari θ.

7.5.2. Misal $X_1, X_2, ..., X_n$ sampel acak dari sebaran dengan pdf $f(x;\theta) = \frac{1}{\theta} I(0 < x < \theta, 0 < \theta)$. Hitung

$$n\left\{\left[\frac{\partial \ln f(X;\theta)}{\partial \theta}\right]^2\right\}.$$

7.5.3. Misal $X_1, X_2, ..., X_n$ sampel acak dari ***sebaran Cauchy***, $X_i \sim CAU(1, \theta)$, yaitu sebaran dengan pdf

$$f(x;\theta) = \frac{1}{\pi\left[1 + (x - \theta)^2\right]}, -\infty < x < \infty, -\infty < x < \infty.$$

Tunjukkan bahwa batas bawah Rao-Cramer sebaran ini adalah $2/n$. Apakah MLE dari θ merupakan penduga efisien dari θ? Jelaskan!

7.5.4. Misal $X_1, X_2, ..., X_n$ sampel acak dari sebaran gamma, $X_i \sim GAM(\theta, \kappa)$, di mana $\kappa = 4$, dan $\theta > 0$.
(a) Tentukan informasi Fisher $I(\theta)$.
(b) Tunjukkan bahwa MLE dari θ adalah penduga efisien dari θ.
(c) Apakah sebaran asimtotik dari $\sqrt{n}\left(\hat{\theta} - \theta\right)$

7.5.5. Misal $X_1, X_2, ..., X_n$ sampel acak dari sebaran dengan pdf

$$f(x;\theta) = \frac{3\theta^3}{(x + \theta)^4} I(0 < x < \infty, 0 < \theta < \infty).$$

Tunjukkan bahwa $Y = 2\bar{X}$ adalah penduga takbias dari θ, dan tentukan keefisienannya.

7.5.6. Misal $X_1, X_2, ..., X_n$ sampel acak dari sebaran normal, $X_i \sim N(\theta, \sigma^2)$, di mana $-\infty < \theta < \infty$, dan $\sigma^2 > 0$. Anggap σ^2 diketahui. Tunjukkan bahwa $\bar{X} - \frac{\sigma^2}{n}$ adalah penduga takbias dari θ^2 dan tentukan keefisienannya.

7.5.7. Misal $X_1, X_2, ..., X_n$ sampel acak dari sebaran normal, $X_i \sim N(0, \theta)$.

(a) Tentukan penduga dari simpangan baku $\sqrt{\theta}$.

(b) Tentukan konstanta c sehingga $Y = c\sum_{i=1}^{n} |X_i|$ adalah penduga takbias dari $\sqrt{\theta}$ dan tentukan keefisienannya.

(c) Tunjukkan bahwa $\sum_{i=1}^{n} X_i^2 / n$ adalah penduga takbias dari θ dan mempunyai varians $2\theta^2 / n$.

7.5.8. Ulangi Contoh 7.5.3 untuk menunjukkan bahwa $2\theta W \sim \chi^2(2n)$.

7.6 Ukuran dari Kualitas Penduga

Anggap bahwa $f(x;\theta)$, untuk $\theta \in \Omega$, adalah pdf dari peubah acak X. Perhatikan penduga titik $Y_n = u(X_1, X_2, ..., X_n)$ berdasarkan pada sampel acak $X_1, X_2, ..., X_n$. Dalam subpokok bahasan sebelumnya telah didiskusikan beberapa sifat dari penduga. Ingat bahwa Y_n penduga konsisten jika dia konvergen dalam peluang ke θ. Sifat lainnya adalah tentang penduga takbias dari θ, yaitu jika $E(Y_n) = \theta$.

Jika dua penduga dari θ sama-sama takbias, maka kita akan memilih penduga yang mempunyai varians lebih kecil, seperti diberikan oleh definisi berikut ini.

Definisi 7.6.1 Untuk bilangan bulat positif n, $Y = u(X_1, X_2, ..., X_n)$ disebut ***penduga takbias varians minimum***

(minimum variance unbiased estimator/MVUE), jika Y takbias dan varians dari Y kurang dari atau sama dengan varians dari setiap penduga takbias dari θ.

Contoh 7.6.1 Misal $X_1, X_2, ..., X_9$ sampel acak dari sebaran normal $N(\theta, \sigma^2)$, di mana $-\infty < \theta < \infty$. Karena statistik $\bar{X}_n = \frac{\sum_{i=1}^{n} X_i}{9} \sim N(\theta, \frac{\sigma^2}{9})$, maka $\bar{X}_n$ penduga takbias dari θ. Statistik $X_i \sim N(\theta, \sigma^2)$, juga penduga takbias dari θ. Meskipun demikian varians dari $\bar{X}$, $\frac{\sigma^2}{9}$, lebih kecil daripada varians dari X_1, σ^2. Kita tidak dapat mengatakan bahwa $\bar{X}_n$ adalah MVUE dari θ, karena definisi di atas berlaku untuk setiap penduga takbias dari θ.

Sekarang kita akan membahas permasalahan pendugaan titik dari sudut pandang yang berbeda. Misal $X_1, X_2, ..., X_n$ sampel acak dari sebaran yang mempunyai pdf $f(x;\theta)$, $\theta \in \Omega$. Misal $Y = u(X_1, X_2, ..., X_n)$ akan dijadikan sebagai dasar untuk dugaan titik dari θ. Misal $\delta(y)$ adalah fungsi dari nilai pengamatan statistik Y yang merupakan dugaan titik dari θ. Maka fungsi δ *memutuskan* suatu nilai dugaan dari θ dan δ disebut ***fungsi keputusan*** atau ***kaidah keputusan.*** Nilai dari keputusan, katakanlah $\delta(y)$, disebut sebagai ***keputusan***. Suatu keputusan mungkin benar mungkin juga salah. Kita akan menggunakan ukuran keseriusan perbedaan antara θ dan $\delta(y)$. Oleh karena itu dari masing-masing pasangan terurut $[\theta, \delta(y)]$, $\theta \in \Omega$, kita akan mengaitkan bilangan tidak negatif $\Lambda[\theta, \delta(y)]$ yang mencerminkan keseriusan ini. Kita sebut fungsi Λ sebagai ***fungsi kerugian.*** Nilai ekspektasi (purata) dari fungsi kerugian ini disebut sebagai ***risiko.*** Oleh karena itu, jika $f_Y(y;\theta)$, $\theta \in \Omega$, pdf dari Y, maka ***fungsi risiko*** $R(\theta, \delta)$ diberikan oleh

$$R(\theta, \delta) = E\{\Lambda[\theta, \delta(Y)]\} = \int_{-\infty}^{\infty} \Lambda[\theta, \delta(Y)] f_Y(y;\theta)\, dy,$$

jika Y peubah acak kontinu. Kita ingin memilih fungsi keputusan sehingga meminimumkan risiko $R(\theta, \delta)$ untuk semua nilai $\theta \in \Omega$.

Contoh 7.6.2 Misal $X_1, X_2, ..., X_{25}$ sampel acak dari sebaran normal, $X_i \sim N(\theta, 1)$, untuk $-\infty < \theta < \infty$. Misal $Y = \bar{X}_n$, adalah purata sampel acak ini, dan misal $\Lambda\left[\theta, \delta(y)\right] = \left[\theta - \delta(y)\right]^2$. Kita akan akan membandingkan dua fungsi keputusan yang diberikan oleh $\delta_1(y) = y$ dan $\delta_2(y) = 0$, untuk $-\infty < \theta < \infty$. Fungsi-fungsi keputusan tersebut berkorespondensi dengan fungsi-fungsi risiko berturut-turut

$$R(\theta, \delta_1) = E\left[\left(\theta - Y\right)^2\right] = \frac{1}{25},$$

dan

$$R(\theta, \delta_2) = E\left[\left(\theta - 0\right)^2\right] = \theta^2.$$

Jelas bahwa jika $\theta = 0$, maka $\delta_2(y) = 0$ adalah keputusan yang sangat baik dan $R(0, \delta_2) = 0$. Meskipun demikian, jika $\theta \neq 0$, maka $\delta_2 = 0$ adalah keputusan yang lemah. Sebagai contoh, jika $\theta = 2$, maka

$$R(2, \delta_2) = 4 > R(2, \delta_1) = \frac{1}{25}.$$

Secara umum, kita melihat bahwa $R(2, \delta_2) < R(2, \delta_1)$, asalkan $-\frac{1}{5} < \theta < \frac{1}{5}$ dan yang lain $R(2, \delta_2) \geq R(2, \delta_1)$. Jika kita membatasi fungsi keputusan δ sehingga $E\left[\delta(Y)\right] = \theta$ untuk semua $\theta \in \Omega$, maka fungsi keputusan $\delta_2(y) = 0$ tidak dapat digunakan. Di bawah batas ini dan dengan diketahui $\Lambda\left[\theta, \delta(y)\right]$, fungsi risiko adalah varians dari penduga takbias $\delta(Y)$.

Misal kita tidak membatasi fungsi keputusan δ, sehingga $E\left[\delta(Y)\right] = \theta$ untuk semua $\theta \in \Omega$. Kita mengatakan bahwa fungsi keputusan yang meminimumkan maksimum dari fungsi risiko adalah fungsi keputusan terbaik. Karena, dalam contoh ini, $R\left(\theta, \delta_2\right) = \theta^2$ tidak terbatas, maka $\delta_2(y) = 0$ tidak dapat digunakan sebagai fungsi keputusan yang baik. Di pihak lain, jika $-\infty < \theta < \infty$, kita mempunyai

$$\max_{\theta} R\left(\theta, \delta_1\right) = \max_{\theta}\left(\tfrac{1}{25}\right) = \tfrac{1}{25}.$$

Oleh karena itu $\delta_1(y) = y = \bar{x}$ terlihat sebagai keputusan yang baik.

Berdasarkan contoh ini kita mendapatkan:

(1) Tanpa membatasi fungsi keputusan, sulit untuk menentukan fungsi keputusan yang mempunyai fungsi risiko yang kurang dari fungsi risiko dari keputusan yang lain secara seragam.

(2) Satu prinsip untuk memilih suatu fungsi keputusan disebut ***prinsip minimax (minimax principle).*** Prinsip ini menyatakan bahwa: Jika fungsi keputusan diberikan oleh $\delta_0(y)$ sehingga, untuk semua $\theta \in \Omega$,

$$\max_{\theta} R\left(\theta, \delta_0(y)\right) \leq \max_{\theta} R\left(\theta, \delta(y)\right),$$

untuk setiap fungsi keputusan yang lain $\delta(y)$, maka $\delta_0(y)$ disebut ***fungsi keputusan minimax (minimax decision function).***

Dengan pembatasan $E\left[\delta(Y)\right] = \theta$ dan fungsi kerugian

$$\Lambda\left[\theta, \delta(y)\right] = \left[\theta - \delta(y)\right]^2,$$

fungsi keputusan yang meminimumkan fungsi risiko sebagai penduga takbias dengan varians minimum. Jika pembatasan $E\left[\delta(Y)\right] = \theta$ diganti dengan kondisi yang lain, fungsi keputusan $\delta(Y)$, jika ada, dengan meminimumkan

$$E\left\{\Lambda\left[\theta, \delta(Y)\right]\right\} = E\left\{\left[\theta - \delta(Y)\right]^2\right\},$$

secara seragam dalam θ disebut sebagai ***penduga galat purata kuadrat minimum (minimum mean-squared-error).***

Latihan 7.6

7.6.1. Misal $Y_1 < Y_2 < Y_3$ adalah statistik urutan dari sampel acak berukuran 3 dari sebaran dengan pdf $f(x;\theta) = \dfrac{1}{\theta} I(0 < x < \theta < \infty)$.

(a) Tunjukkan bahwa $4Y_1$, $2Y_2$, dan $\frac{4}{3}Y_3$ adalah penduga-penduga takbias dari θ.

(b) Tentukan varians masing-masing penduga-penduga takbias tersebut.

7.6.2. Misal Y_1 dan Y_2 dua penduga takbias dari θ dan bebas. Misal varians dari Y_1 dua kali varians dari Y_2. Tentukan konstanta k_1 dan k_2 sehingga $k_1Y_1 + k_2Y_2$ penduga takbias dengan varians terkecil untuk sebarang kombinasi linear.

7.6.3. Dalam Contoh 7.6.2, ambil $\Lambda\left[\theta, \delta(y)\right] = |\theta - \delta(y)|$.

(a) Tunjukkan bahwa $R\left(\theta, \delta_1\right) = \frac{1}{5}\sqrt{2/\pi}$ dan $R\left(\theta, \delta_2\right) = |\theta|$.

(b) Di antara dua fungsi keputusan δ_1 dan δ_2, yang risiko maksimumnya terkecil.

7.6.4. Misal $X_1, X_2, ..., X_n$ sampel acak dari sebaran Poisson, $X_i \sim POI(\theta)$, $0 < \theta < \infty$. Misal $Y = \sum_{i=1}^{n} X_i$ dan misal $\Lambda\left[\theta, \delta(y)\right] = \left[\theta - \delta(y)\right]^2$. Jika kita batasi fungsi kuasa $\delta(y) = b + y/n$, di mana b tidak bergantung pada y, tunjukkan bahwa $R\left(\theta, \delta\right) = b^2 + \theta / n$. Apakah fungsi keputusan dari bentuk ini adalah suatu risiko terkecil seragam daripada setiap fungsi keputusan yang lain dari bentuk ini? Dengan penyelesaian ini, katakanlah δ, dan $0 < \theta < \infty$, tentukanlah $\max_{\theta} R\left(\theta, \delta\right)$ jika dia ada.

3.6.5. Misal $X_1, X_2, ..., X_n$ sampel acak dari sebaran normal, $X_i \sim N(\mu, \theta)$, $0 < \theta < \infty$, di mana μ tidak diketahui. Misal $Y = \sum_{i=1}^{n} \frac{(X_i - \bar{X}_n)^2}{n}$ dan misal $\Lambda\left[\theta, \delta(y)\right] = \left[\theta - \delta(y)\right]^2$. Jika fungsi kuasa $\delta(y) = by$, di mana b tidak bergantung pada y, tunjukkan bahwa

$$R\left(\theta, \delta\right) = \left(\theta^2 / n^2\right)\left[\left(n^2 - 1\right)b^2 - 2n(n-1)b + n^2\right].$$

Tunjukkan bahwa $b = n/(n+1)$ adalah adalah fungsi keputusan risiko minimum dari bentuk ini. Catatan bahwa $nY/(n+1)$ adalah bukan penduga takbias dari θ. Dengan $\delta(y) = ny/(n+1)$ dan $0 < \theta < \infty$, tentukan $\max_{\theta} R\left(\theta, \delta\right)$ jika dia ada.

7.7 Statistik Cukup suatu Parameter

Misal $X_1, X_2, ..., X_n$ sampel acak dari sebaran dengan pdf $f(x;\theta),\ \theta \in \Omega.$ Dalam subpokok bahasan ini kita membahas statistik yang dinyatakan sebagai $Y_1 = u_1(X_1, X_2, ..., X_n),$ yang disebut sebagai ***statistik cukup (sufficient statistic).*** Statistik cukup ini merupakan partisi dari ruang sampel yang diberikan oleh

$$\left(X_1, X_2, ..., X_n\right) \in \left\{\left(x_1, x_2, ..., x_n\right) : u_1\left(x_1, x_2, ..., x_n\right) = y_1\right\},$$

dan peluang bersyarat dari $X_1, X_2, ..., X_n,$ bila $Y_1 = y_1$ tidak bergantung pada $\theta.$ Secara intuitif, ini berarti himpunan yang ditentukan oleh $Y_1 = y_1$ adalah tetap, sebaran dari statistik lain, katakanlah $Y_2 = u_2(X_1, X_2, ..., X_n),$ tidak bergantung pada parameter θ karena sebaran bersyarat dari $X_1, X_2, ..., X_n$ tidak bergantung pada $\theta.$ Oleh karena itu tidak mungkin menggunakan $Y_2,$ bila $Y_1 = y_1,$ untuk membuat inferens statistis tentang $\theta.$ Ini alasan bahwa $Y_1 = u_1(X_1, X_2, ..., X_n)$ disebut statistik cukup.

Untuk lebih jelasnya tentang statistik cukup ini pembahasan kita mulai dengan contoh berikut ini.

Contoh 7.7.1 Misal $X_1, X_2, ..., X_n$ sampel acak dari sebaran yang mempunyai pdf

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

Statistik $Y_1 = X_1 + X_2 + ... + X_n$ mempunyai pdf

$$f_{Y_1}(y_1;\theta) = \binom{n}{y_1} \theta^{y_1} (1-\theta)^{n-y_1} I(y_1 = 0,1,...,n;\ 0 < \theta < 1).$$

Apakah peluang bersyarat

$$P\left(X_1 = x_1, X_2 = x_2, ..., X_n = x_n \mid Y_1 = y_1\right) = P(A \mid B),$$

katakanlah, di mana $y_1 = 0,1,...,n$? Kecuali jika jumlah dari bilangan-bilangan bulat $x_1, x_2, ..., x_n$ (yang masing-masing sama dengan nol atau satu) adalah sama untuk $y_1,$ peluang bersyaratnya jelas sama dengan nol karena $A \cap B = \varnothing.$ Tetapi dalam kasus $y_1 = \sum_{i=1}^{n} x_i,$ maka $A \subset B,$ sehingga $A \cap B = A$ dan

$$P(A \mid B) = P(A)/P(B).$$

Oleh karena itu peluang bersyaratnya sama dengan

$$\frac{\theta^{x_1}(1-\theta)^{1-x_1}\theta^{x_2}(1-\theta)^{1-x_2}\ldots\theta^{x_n}(1-\theta)^{1-x_n}}{\binom{n}{y_1}\theta^{y_1}(1-\theta)^{n-y_1}} = \frac{\theta^{\sum_{i=1}^{n} x_i}(1-\theta)^{n-\sum_{i=1}^{n} x_i}}{\binom{n}{\sum_{i=1}^{n} x_i}\theta^{\sum_{i=1}^{n} x_{i_1}}(1-\theta)^{n-\sum_{i=1}^{n} x_i}}$$

$$= \frac{1}{\binom{n}{\sum_{i=1}^{n} x_i}}.$$

Secara umum, misal $f_{Y_1}(y_1;\theta)$ adalah pdf dari $Y_1 = u_1(X_1, X_2, ..., X_n),$ di mana $X_1, X_2, ..., X_n$ sampel acak dari sebaran diskret dengan pdf $f(x;\theta),\ \theta \in \Omega.$ Peluang bersyarat dari $X_1 = x_1, X_2 = x_2, ..., X_n = x_n$ bila diketahui $Y_1 = y_1$ sama dengan

$$\frac{f(x_1;\theta)\,f(x_2;\theta)\ldots f(x_n;\theta)}{f_{Y_1}\left[u_1(x_1, x_2, ..., x_n);\theta\right]},$$

asalkan $x_1, x_2, ..., x_n$ sehingga $y_1 = u_1(x_1, x_2, ..., x_n)$ tetap dan sama dengan nol untuk yang lainnya. Kita katakana $Y_1 = u_1(X_1, X_2, ..., X_n)$ adalah statistik cukup untuk θ jika nisbah ini tidak bergantung pada $\theta.$

Definisi 7.7.1 Misal $X_1, X_2, ..., X_n$ sampel acak dari sebaran diskret dengan pdf $f(x;\theta),\ \theta \in \Omega.$ Misal $Y_1 = u_1(X_1, X_2, ..., X_n)$ adalah statistik dengan pdf $f_{Y_1}(y_1;\theta).$ Maka Y_1 adalah statistik cukup untuk θ jika

$$\frac{f(x_1;\theta)\,f(x_2;\theta)\ldots f(x_n;\theta)}{f_{Y_1}\left[u_1(x_1, x_2, ..., x_n);\theta\right]} = H(x_1, x_2, ..., x_n),$$

di mana $H(x_1, x_2, ..., x_n)$ tidak bergantung pada $\theta \in \Omega.$

Contoh 7.7.2 Misal $X_1, X_2, ..., X_n$ sampel acak dari sebaran gamma, $X_i \sim GAM(\theta, 2)$, $\theta > 0$. Karena mgf dari sebaran ini adalah

$$M(t) = (1-\theta t)^{-2}, t < 1/\theta,$$

maka mgf dari $Y_1 = \sum_{i=1}^{n} X_i$ adalah

$$M_{Y_1}(t) = E\left[e^{t(X_1+X_2+...+X_n)}\right] = E\left(e^{X_1}\right)E\left(e^{X_2}\right)...E\left(e^{X_n}\right)$$

$$= \Pi_{i=1}^{n}\left[(1-\theta t)^{-2}\right] = (1-\theta t)^{-2n}.$$

Jadi $Y_i \sim GAM(\theta, 2n)$. Oleh karena itu pdf-nya adalah

$$f_{Y_1}(y_1;\theta) = \frac{1}{\Gamma(2n)\theta^{2n}} y_1^{2n-1} e^{-y_1/\theta} I(0 < y_1 < \infty)$$

Kita peroleh nisbah

$$\frac{f(x_1;\theta)f(x_2;\theta)...f(x_n;\theta)}{f_{Y_1}\left[u_1(x_1, x_2, ..., x_n);\theta\right]} = \frac{\left[\dfrac{x_1^{2-1}e^{-x_1/\theta}}{\Gamma(2)\theta^2}\right]\left[\dfrac{x_2^{2-1}e^{-x_2/\theta}}{\Gamma(2)\theta^2}\right]...\left[\dfrac{x_n^{2-1}e^{-x_n/\theta}}{\Gamma(2)\theta^2}\right]}{\left[\dfrac{(x_1+x_2+...+x_n)^{2n-1}e^{-(x_1+x_2+...+x_n)/\theta}}{\Gamma(2n)\theta^{2n}}\right]}$$

$$= \frac{\Gamma(2n)(x_1x_2...x_n)}{\left[\Gamma(2)\right]^n(x_1+x_2+...+x_n)^{2n-1}},$$

di mana $0 < x_i < \infty$, $i = 1, 2, ..., n.$ Karena nisbah ini tidak bergantung pada θ, maka $Y_1 = \sum_{i=1}^{n} X_i$ adalah statistik cukup untuk θ.

Contoh 7.7.3 Misal $Y_1 < Y_2 < ... < Y_n$ menyatakan statistik urutan dari suatu sebaran yang mempunyai pdf

$$f(x;\theta) = e^{-(x-\theta)} I(\theta < x < \infty).$$

Oleh karena itu pdf dari Y_1 adalah

$$f_{Y_1}(y_1;\theta) = ne^{-n(y_1-\theta)} I(\theta < y_1 < \infty).$$

Jadi

$$\frac{\prod_{i=1}^{n} e^{-(x_i-\theta)} I(\theta < x_i < \infty)}{n e^{-n(\min x_i - \theta)} I(\theta < \min x_i < \infty)} = \frac{e^{-\sum_{i=1}^{n} x_i}}{n e^{-n \min x_i}}.$$

Nisbah ini tidak bergantung pada θ. Jadi statistik urutan terkecil Y_1 adalah statistik cukup untuk θ.

Teorema 7.7.1 Faktorisasi Neyman. Misal $X_1, X_2, ..., X_n$ sampel acak dari sebaran yang mempunyai pdf $f(x;\theta),\ \theta \in \Omega$. Statistik $Y_1 = u_1(X_1, X_2, ..., X_n)$ adalah statistik cukup untuk θ jika dan hanya jika kita dapat menentukan dua fungsi tidak negatif, k_1 dan k_2, sehingga

$$f(x_1;\theta) f(x_2;\theta) ... f(x_n;\theta) = k_1 \left[u_1(x_1, x_2, ..., x_n);\theta\right] k_2(x_1, x_2, ..., x_n),$$

di mana $k_2(x_1, x_2, ..., x_n)$ tidak bergantung pada θ.

Bukti

Kita akan membuktikannya untuk kasus kontinu. Anggap bahwa faktorisasi pada teorema benar. Kita buat transformasi satu-satu

$$y_1 = u_1(x_1, x_2, ..., x_n),\ \ y_2 = u_2(x_1, x_2, ..., x_n),\ \ldots,\ y_n = u_n(x_1, x_2, ..., x_n),$$

mempunyai balikan

$$x_1 = w_1(y_1, y_2, ..., y_n),\ \ x_2 = w_2(y_1, y_2, ..., y_n),\ \ldots,\ x_n = w_n(y_1, y_2, ..., y_n),$$

dan Jacobi J. Oleh karena itu pdf dari statistik $Y_1, Y_2, ..., Y_n$ adalah

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

di mana $w_i = w_i(y_1, y_2, ..., y_n),\ i = 1, 2, ..., n.$ Misal $f_{Y_1}(y_1;\theta)$ adalah pdf dari Y_1. Maka

$$f_{Y_1}(y_1;\theta) = \int_{-\infty}^{\infty} ... \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} g(y_1, y_2, ..., y_n;\theta) dy_2 dy_3 ... dy_n$$

$$= k_1(y_1;\theta) \int_{-\infty}^{\infty} ... \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} k_2(w_1, w_2, ..., w_n) \,|\, J \,|\, dy_2 dy_3 ... dy_n.$$

Fungsi k_2 maupun J tidak bergantung pada θ. Oleh karena itu $(n-1)$ kali integral dalam ruas kanan adalah fungsi dari y_1 sendiri, misal $m(y_1)$. Jadi

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

Jika $m(y_1) = 0$, maka $f_{Y_1}(y_1;\theta) = 0$. Jika $m(y_1) > 0$, kita dapat menulis

$$k_1\left[u_1(x_1, x_2, ..., x_n);\theta\right] = \frac{f_{Y_1}\left[u_1(x_1, x_2, ..., x_n);\theta\right]}{m\left[u_1(x_1, x_2, ..., x_n)\right]},$$

dan asumsi faktorisasi menjadi

$$f(x_1;\theta)f(x_2;\theta)...f(x_n;\theta) = f_{Y_1}\left[u_1(x_1, x_2, ..., x_n);\theta\right]\frac{k_2(x_1, x_2, ..., x_n)}{m\left[u_1(x_1, x_2, ..., x_n)\right]}.$$

Karena k_2 dan m keduanya tidak bergantung pada θ, maka Y_1 adalah statistik cukup untuk θ.

Kebalikannya, jika Y_1 statistik cukup untuk θ, maka faktorisasi dapat dinyatakan sebagai fungsi k_1 untuk pdf dari Y_1, sebut fungsi f_{Y_1}, dan k_2 sebagai H. Bukti selesai.

Selanjutnya kita akan membahas statistik cukup digunakan untuk menentukan MVUE. Pertama kita catat bahwa penduga cukup tidak tunggal dalam sebarang pengertian. Sebagai contoh, jika $Y_1 = u_1(X_1, X_2, ..., X_n)$ statistik cukup dan $Y_2 = g(Y_1)$ di mana $g(x)$ adalah fungsi satu-satu, maka

$$\begin{aligned} f(x_1;\theta)f(x_2;\theta)...f(x_n;\theta) &= k_1\left[u_1(y_1);\theta\right]k_2(x_1, x_2, ..., x_n) \\ &= k_1\left[u_1(g^{-1}(y_2));\theta\right]k_2(x_1, x_2, ..., x_n). \end{aligned}$$

Oleh karena itu, Y_2 juga statistik cukup. Meskipun demikian, seperti ditunjukkan teorema berikutnya, kecukupan dapat dimodifikasi sehingga menjadi penduga yang sangat baik.

Misal statistik cukup Y_1 sebagai X_1 dan Y_2, dan statistik takbias dari θ, sebagai X_2, dan misal $E(Y_2 \mid y_1) = \varphi(y_1)$. Maka kita peroleh

$$\theta = E(Y_2) = E\left[\varphi(Y_1)\right]$$

dan

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

Oleh karena itu fungsi $\varphi(Y_1)$ dari statistik cukup Y_1 adalah penduga takbias dari θ dan mempunyai varians lebih kecil daripada penduga takbias Y_2. Kita ringkas hasil di atas secara formal dalam teorema, yang disebut sebagai Rao-Blackwell.

Teorema 7.7.2 Rao-Blackwell. Misal $X_1, X_2, ..., X_n$ adalah sampel acak dari suatu sebaran yang mempunyai pdf $f(x;\theta),\ \theta \in \Omega.$ Misal $Y_1 = u_1(X_1, X_2, ..., X_n)$ adalah statistik cukup untuk $\theta,$ dan misal $Y_2 = u_2(X_1, X_2, ..., X_n),$ bukan fungsi dari Y_1 sendiri, adalah penduga takbias dari $\theta.$ Maka $E(Y_2 \mid y_1) = \varphi(y_1)$ mendefinisikan suatu statistik $\varphi(Y_1).$ Statistik $\varphi(Y_1)$ ini adalah fungsi dari statistik cukup untuk $\theta;$ dia adalah penduga takbias dari $\theta;$ dan variansnya kurang dari varians $Y_2.$

Teorema ini mengatakan kepada kita bahwa dalam mencari MVUE dari suatu parameter, adalah memungkinkan jika statistik cukup untuk parameter ada. Kita mulai dengan penduga takbias Y_2 sendiri, kemudian kita dapat memperbaiki dengan menghitung $E(Y_2 \mid y_1) = \varphi(y_1)$ sehingga $\varphi(Y_1)$ adalah penduga takbias dengan varians yang lebih kecil daripada varians $Y_2.$

Hubungan antara statistik cukup dan penduga kemungkinan maksimum (MLE) diberikan oleh teorema berikut ini.

Teorema 7.7.3 Misal $X_1, X_2, ..., X_n$ adalah sampel acak dari suatu sebaran yang mempunyai pdf $f(x;\theta),\ \theta \in \Omega.$ Misal statistik cukup $Y_1 = u_1(X_1, X_2, ..., X_n)$ untuk θ ada dan misal MLE $\hat{\theta}$ dari θ juga ada dan tunggal. Maka $\hat{\theta}$ adalah fungsi dari $Y_1 = u_1(X_1, X_2, ..., X_n).$

Bukti

Misal $f_{Y_1}(y_1;\theta)$ adalah pdf dari $Y_1.$ Maka oleh definisi kecukupan, kita peroleh fungsi kemungkinan

$$\begin{aligned} L(\theta; x_1, x_2, ..., x_n) &= f(x_1;\theta) f(x_2;\theta) ... f(x_n;\theta) \\ &= f_{Y_1}\left[u_1(x_1, x_2, ..., x_n);\theta\right] H(x_1, x_2, ..., x_n), \end{aligned}$$

di mana $H(x_1, x_2, ..., x_n)$ tidak bergantung pada $\theta.$ Oleh karena itu L dan $f_{Y_1},$ sebagai fungsi-fungsi dari $\theta,$ adalah serentak maksimum. Karena ada satu dan hanya satu nilai dari θ sehingga L maksimum dan oleh karena itu $f_{Y_1}\left[u_1(x_1, x_2, ..., x_n);\theta\right],$ maka nilai dari θ haruslah fungsi dari

$u_1(x_1, x_2, ..., x_n)$. Jadi MLE $\hat{\theta}$ adalah suatu fungsi dari statistik cukup $Y_1 = u_1(X_1, X_2, ..., X_n)$.

Dari subpokok bahasan 7.5 secara umum MLE adalah penduga takbias asimtotik dari θ. Oleh karena itu, ini dapat kita proses untuk menentukan statistik cukup dan kemudian MLE.

Contoh 7.7.4 Misal $X_1, X_2, ..., X_n$ adalah sampel acak dari suatu sebaran yang mempunyai pdf

$$f(x;\theta) = \theta e^{-\theta x} I(0 < x < \infty).$$

Kita ingin menentukan MVUE dari θ. Fungsi kemungkinan sampel ini adalah

$$L(\theta; x_1, x_2, ..., x_n) = \theta^n e^{-\theta \sum_{i=1}^{n} x_i}$$

untuk $x_i > 0,\ i = 1, 2, ..., n$. Oleh karena itu dengan teorema Faktorisasi Neyman, statistik $Y_1 = \sum_{i=1}^{n} X_i$ adalah statistik cukup untuk θ. Logarima dari fungsi kemungkinan adalah

$$\ln L(\theta) = n \ln \theta - \theta \sum_{i=1}^{n} x_i.$$

Dengan menurunkannya terhadap θ, dan menyamakannya dengan nol kita peroleh MLE dari θ yang memberikan

$$Y_2 = \frac{1}{\bar{X}}.$$

Statistik Y_2 asimtotik takbias. Karena $X_1, X_2, ..., X_n$ sampel acak dari sebaran gamma, $X_i \sim GAM(\frac{1}{\theta}, 1)$, maka $Y_i = \sum_{i=1}^{n} X_i \sim GAM(\frac{1}{\theta}, n)$. Oleh karena itu

$$E(Y_2) = E\left[\frac{1}{\bar{X}}\right] = nE\left[\frac{1}{\sum_{i=1}^{n} X_i}\right] = n\int_0^\infty \frac{\theta^n}{\Gamma(n)} x^{-1} x^{n-1} e^{-\theta x} dx,$$

dengan mengubah peubah $z = \theta x$, kita peroleh

$$E(Y_2) = E\left[\frac{1}{\bar{X}}\right] = \theta\frac{n}{(n-1)!}\Gamma(n-1) = \theta\frac{n}{n-1}.$$

Jadi statistik $(n-1)/\left(\sum_{i=1}^{n} X_i\right)$ adalah MVUE dari $\theta.$

Penduga takbias $\varphi(Y_1) = E(Y_2 \mid y_1),$ mempunyai varians lebih kecil daripada penduga takbias Y_2 dari $\theta.$ Misal fungsi $\Upsilon(y_3) = E\left[\varphi(Y_1) \mid Y_3 = y_3\right],$ di mana Y_3 statistik yang lain, yang tidak cukup untuk $\theta.$ Dengan teorema Rao-Blackwell, kita mempunyai $E\left[\Upsilon(y_3)\right] = \theta$ dan $\Upsilon(Y_3)$ mempunyai varians lebih kecil daripada $\varphi(Y_1).$ Maka $\Upsilon(Y_3)$ haruslah lebih baik daripada $\varphi(Y_1)$ sebagai penduga takbias. Tetapi ini tidak benar, karena Y_3 bukan statistik cukup.

Contoh 7.7.5 Misal X_1, X_2, X_3 sampel acak dari sebaran eksponensial dengan purata $\theta > 0.$ Maka pdf bersamanya adalah

$$\left(\frac{1}{\theta}\right)^3 e^{-(x_1+x_2+x_3)/\theta}\, I(0 < x_i < \infty,\ i = 1,2,3).$$

Dari teorema faktorisasi, kita melihat bahwa $Y_1 = X_1 + X_2 + X_3$ adalah statistik cukup untuk $\theta.$ Karena

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

maka $Y_1/3 = \bar{X}$ adalah fungsi dari statistik cukup yang takbias dari $\theta.$

Lebih lanjut, misal $Y_2 = X_2 + X_3$ dan $Y_3 = X_3.$ Transformasi satu-satu didefinisikan oleh

$$x_1 = y_1 - y_{2,}\ \ x_2 = y_2 - y_3,\ x_3 = y_3$$

mempunyai Jacobi sama dengan satu dan pdf bersama dari Y_1, Y_2, Y_3 adalah

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

Maka diperoleh pdf marginal dari Y_1 dan Y_3

$$g_{13}(y_1, y_3;\theta) = \left(\frac{1}{\theta}\right)^3 (y_1 - y_2)e^{-y_1/\theta}\, I(0 < y_3 < y_1 < \infty),$$

dan pdf marginal dari Y_3

$$g_3(y_3;\theta) = \frac{1}{\theta}e^{-y_3/\theta}\, I(0 < y_3 < \infty).$$

Oleh karena itu pdf bersyarat dari Y_1, bila $Y_3 = y_3$, adalah

$$g_{1|3}(y_1 \mid y_3) = \frac{g_{13}(y_1, y_2;\theta)}{g_3(y_3;\theta)} = \left(\frac{1}{\theta}\right)^2 (y_1 - y_3)e^{-(y_1-y_3)/\theta}\, I(0 < y_3 < y_1 < \infty).$$

Jadi

$$E\left(\frac{Y_1}{3} \mid y_3\right) = E\left(\frac{Y_1 - Y_3}{3} \mid y_3\right) + E\left(\frac{Y_3}{3} \mid y_3\right)$$

$$= \frac{1}{3}\int_{y_3}^{\infty}\left(\frac{1}{\theta}\right)^2 (y_1 - y_3)^2 e^{-(y_1-y_3)/\theta} dy_1 + \frac{y_3}{3}$$

$$= \frac{1}{3}\frac{\Gamma(3)\theta^3}{\theta^2} + \frac{y_3}{3} = \frac{2\theta + y_3}{3} = \Upsilon(y_3).$$

Jelas bahwa $E\left[\Upsilon(Y_3)\right] = \theta$ dan $Var\left[\Upsilon(Y_3)\right] \le Var(Y_1/3)$, tetapi $\Upsilon(Y_3)$ bukan statistik dan tidak dapat digunakan sebagai penduga dari θ.

Latihan 7.7

7.7.1. Misal $X_1, X_2, ..., X_n$ sampel acak dari sebaran normal, $X_i \sim N(0, \theta)$.

(a) Tunjukkan bahwa $\sum_{i=1}^{n} X_i^2$ adalah statistik cukup untuk θ.

(b) Tunjukkan bahwa MLE dari θ adalah fungsi dari statistik cukup untuk θ.

7.7.2. Misal $X_1, X_2, ..., X_n$ sampel acak dari sebaran Poisson, $X_i \sim POI(\theta)$.

(a) Tunjukkan bahwa $\sum_{i=1}^{n} X_i$ adalah statistik cukup untuk θ.

(b) Tunjukkan bahwa MLE dari θ adalah fungsi dari statistik cukup untuk θ.

7.7.3. Misal $X_1, X_2, ..., X_n$ sampel acak dari sebaran seragam kontinu $X_i \sim UNIF(0, \theta),\ 0 < \theta < \infty$.

(a) Tunjukkan bahwa statistik urutan terbesar dari sebaran ini adalah statistik cukup untuk θ.

(b) Tunjukkan bahwa MLE dari θ adalah fungsi dari statistik cukup untuk θ.

7.7.4. Misal $X_1, X_2, ..., X_n$ sampel acak dari sebaran geometrik, $X_i \sim GEO(\theta),\quad 0 < \theta < 1$.

(a) Tunjukkan bahwa $\sum_{i=1}^{n} X_i$ adalah statistik cukup untuk θ.

(b) Tunjukkan bahwa MLE dari θ adalah fungsi dari statistik cukup untuk θ.

7.7.5. Misal $Y_1 < Y_2 < Y_3 < Y_4 < Y_5$ adalah statistik urutan dari sampel acak berukuran 5 dari sebaran seragam kontinu yang mempunyai pdf

$$f(x;\theta) = 1/\theta\, I(0 < x < \theta < \infty).$$

(a) Tunjukkan bahwa $2Y_3$ adalah statistik cukup untuk θ.

(b) Tentukan pdf bersama dari Y_3 dan statistik cukup Y_5 untuk θ.

(c) Tentukan ekspektasi bersyarat $E\left(2Y_3 \mid y_5\right) = \varphi(y_5)$.

(d) Bandingkan varians dari $2Y_3$ dengan $\varphi(Y_5)$.

7.7.6. Misal peubah-peubah acak X dan Y mempunyai pdf bersama

$$f(x, y) = \frac{2}{\theta^2} e^{-(x+y)/\theta}\, I(0 < x < y < \infty).$$

(a) Tunjukkan bahwa purata dan varians dari Y berturut-turut adalah $3\theta/2$ dan $5\theta^2/4$.

(b) Tunjukkan bahwa $E(Y \mid x) = x + \theta$.

(c) Tunjukkan bahwa varians dari $X + \theta$ adalah $\theta^2/4$.

7.8 Selang Kepercayaan

Sejauh ini kita telah membahas penduga suatu parameter beserta sifat-sifatnya. Meskipun demikian penduga yang telah kita bahas adalah penduga berupa titik atau disebut penduga titik. Pada subpokok bahasan selanjutnya kita akan membahas penduga yang berupa selang, atau disebut penduga selang, yang menghasilkan suatu selang kepercayaan.

Misal $X_1, X_2, ..., X_n$ sampel acak dengan pdf bersama $f(x_1, x_2, ..., x_n; \theta)$. Misal W dan U statistik sehingga $W = w(X_1, X_2, ..., X_n)$ dan $U = u(X_1, X_2, ..., X_n)$. Jika suatu percobaan menghasilkan data $x_1, x_2, ..., x_n$, maka kita mempunyai nilai pengamatan untuk kedua statistik ini masing-masing $w(x_1, x_2, ..., x_n)$ dan $u(x_1, x_2, ..., x_n)$.

Definsi 7.8.1 Suatu selang

$$\left(w(x_1, x_2, ..., x_n), u(x_1, x_2, ..., x_n)\right)$$

disebut ***selang kepercayaan*** $100\gamma\%$ untuk θ jika

$$P\left[w(X_1, X_2, ..., X_n) < \theta < u(X_1, X_2, ..., X_n)\right] = \gamma$$

di mana $0 < \gamma < 1$. Nilai-nilai $w(x_1, x_2, ..., x_n)$ dan $u(x_1, x_2, ..., x_n)$ berturut-turut disebut ***batas bawah kepercayaan*** dan ***batas atas kepercayaan*** untuk θ, dan $100\gamma\%$ disebut ***taraf kepercayaan*** untuk θ.

Kita akan menggunakan dua notasi yang berbeda, yaitu huruf besar (W, U) untuk ***selang penduga***, atau dalam beberapa literatur disebut ***selang acak***, sedangkan huruf kecil $(w(x), u(x))$ untuk selang kepercayaan, yang merupakan nilai pengamatan dari selang penduga.

Contoh 7.8.1 Misal $X_1, X_2, ..., X_n$ sampel acak berukuran n dari sebaran eksponensial, $X_i \sim EXP(\theta)$. Kita akan menentukan selang kepercayaan $100\gamma\%$ dari θ. Sebaran eksponensial merupakan bentuk khusus dari sebaran gamma, dengan $\kappa = 1$. Oleh karena itu

$$\frac{2X}{\theta} \sim \chi^2(2),$$

dan

$$\frac{2n\bar{X}_n}{\theta} = \sum_{i=1}^{n} \frac{2X_i}{\theta} \sim \chi^2(2n).$$

Kita perhatikan hubungan antara γ dengan α, yang pada pembahasan selanjutnya disebut sebagai ***taraf signifikan***, yaitu $\gamma = 1-\alpha$. Oleh karena itu

$$\gamma = 1-\alpha = P\left[\chi^2_{\alpha/2}(2n) \le \frac{2n\bar{X}}{\theta} \le \chi^2_{1-\alpha/2}(2n)\right]$$

$$= P\left[\frac{2n\bar{X}}{\chi^2_{1-\alpha/2}(2n)} \le \theta \le \frac{2n\bar{X}}{\chi^2_{\alpha/2}(2n)}\right].$$

Jadi selang kepercayaan $100\gamma\%$ untuk θ adalah

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

Kita ingat sekali lagi bahwa selang kepercayaan di atas merupakan nilai pengamatan dari selang penduga atau selang acak

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

yang berarti bahwa peluang parameter θ terletak pada selang tersebut sebesar $100\gamma\%$. Lebih lanjut jika kita menggunakan taraf signifikan $\alpha = 0.10$ dan $n = 40$, maka kita peroleh $\gamma = 0.90$. Sehingga selang kepercayaan 90% untuk θ adalah

$$\left(\frac{2n\bar{x}}{\chi^2_{1-\alpha/2}(2n)}, \frac{2n\bar{x}}{\chi^2_{\alpha/2}(2n)}\right) = \left(\frac{80\bar{x}}{\chi^2_{0.95}(80)}, \frac{80\bar{x}}{\chi^2_{0.05}(80)}\right)$$

$$= \left(\frac{80\bar{x}}{101.8795}, \frac{80\bar{x}}{60.3915}\right),$$

nilai $\chi^2_{0.95}(80) = 101.8795$, dan $\chi^2_{0.05}(80) = 60.3915$ masing-masing diperoleh dari Minitab 12. Sehingga jika purata pengamatan sampel kita ketahui, maka nilai dari selang kepercayaan tersebut juga dapat kita peroleh.

Kita telah dapat menentukan selang kepercayaan suatu parameter. Dengan kata lain kita telah dapat menentukan batas bawah dan batas atas kepercayaan secara bersama-sama. Batas kepercayaan yang demikian disebut ***batas kepercayaan dua sisi***.

Dalam beberapa hal sering pula kita menginginkan salah satu dari batas kepercayaan, yaitu batas bawah kepercayaan atau batas atas kepercayaan saja, tetapi tidak keduanya. Hal yang demikian ini disebut ***batas kepercayaan satu sisi***. Kita perhatikan definisi berikut ini.

Definisi 7.8.2

1. Selang kepercayaan $\left(w(x_1,x_2,...,x_n),\infty\right)$ disebut selang kepercayaan satu sisi $100\gamma\%$ untuk θ jika
$$P\left[w(X_1,X_2,...,X_n)<\theta\right]=\gamma$$
di mana $0<\gamma<1$. Nilai-nilai $w(x_1,x_2,...,x_n)$ disebut batas bawah kepercayaan satu sisi untuk θ.
2. Selang kepercayaan $\left(-\infty,u(x_1,x_2,...,x_n)\right)$ disebut selang kepercayaan satu sisi $100\gamma\%$ untuk θ jika
$$P\left[\theta<u(X_1,X_2,...,X_n)\right]=\gamma$$
di mana $0<\gamma<1$. Nilai-nilai $u(x_1,x_2,...,x_n)$ disebut batas atas kepercayaan satu sisi untuk θ.

Contoh 7.8.2 Misal $X_1,X_2,...,X_n$ sampel acak berukuran n dari sebaran eksponensial, $X_i \sim EXP(\theta)$. Tentukan:

1. Batas bawah kepercayaan satu sisi $100\gamma\%$ untuk θ.
2. Batas atas kepercayaan satu sisi $100\gamma\%$ untuk θ.

Penyelesaian

1. Seperti halnya pada Contoh 7.8.1, kita dapat memperoleh
$$\frac{2n\bar{X}_n}{\theta}=\sum_{i=1}^{n}\frac{2X_i}{\theta}\sim\chi^2(2n).$$

   Dan hubungan antara γ dengan α, adalah $\gamma=1-\alpha$. Oleh karena itu
$$\gamma=1-\alpha=P\left[\frac{2n\bar{X}}{\theta}\le\chi^2_{1-\alpha}(2n)\right]=P\left[\frac{2n\bar{X}}{\chi^2_{1-\alpha}(2n)}\le\theta\right].$$
   Jadi batas bawah kepercayaan satu sisi $100\gamma\%$ untuk θ adalah $w(x)=\dfrac{2n\bar{x}}{\chi^2_{1-\alpha}(2n)}$. Lebih lanjut jika kita menggunakan taraf signifikan $\alpha=0.10$ dan $n=40$, maka kita peroleh $\gamma=0.90$. Sehingga batas bawah kepercayaan satu sisi 90% untuk θ adalah

$$w(x)=\frac{2n\overline{x}}{\chi^2_{1-\alpha}(2n)}=\frac{80\overline{x}}{\chi^2_{0.90}(80)}=\frac{80\overline{x}}{96.5782}.$$

Dan kita peroleh pula selang kepercayaan satu sisi $100\gamma\%$ untuk θ adalah $\left(\frac{80\overline{x}}{96.5782},\infty\right)$.

2. Dengan cara serupa kita peroleh batas atas kepercayaan satu sisi $100\gamma\%$ untuk θ adalah

$$u(x)=\frac{2n\overline{x}}{\chi^2_{\alpha}(2n)}=\frac{80\overline{x}}{\chi^2_{0.1}(80)}=\frac{80\overline{x}}{64.2778}.$$

Dan selang kepercayaannya $\left(-\infty,\frac{80\overline{x}}{64.2778}\right)$.

Selanjutnya kita akan menentukan selang kepercayaan untuk μ dari sampel acak yang bersebaran normal jika variansnya diketahui.

Contoh 7.8.3 Misal $X_1,X_2,...,X_n$ sampel acak dari sebaran normal, $X_i \sim N(\mu,\sigma^2)$, dan σ^2 diketahui. Kita akan menentukan selang kepercayaan $100(1-\alpha)\%$ untuk μ. Kita punya

$$Z=\frac{\sqrt{n}(\bar{X}_n-\mu)}{\sigma}\sim N(0,1).$$

Oleh karena itu kita peroleh

$$1-\alpha=P\left[z_{\alpha/2}<Z<z_{1-\alpha/2}\right]=P\left[z_{\alpha/2}<\frac{\sqrt{n}(\bar{X}-\mu)}{\sigma}<z_{1-\alpha/2}\right]$$

$$=P\left[\bar{X}+z_{\alpha/2}\frac{\sigma}{\sqrt{n}}<\mu<\bar{X}+z_{1-\alpha/2}\frac{\sigma}{\sqrt{n}}\right],$$

menggunakan sifat kesimetrisan pada sebaran normal, $z_{\alpha/2}=-z_{1-\alpha/2}$. Selain bentuk di atas kita peroleh pula bentuk

$$1-\alpha=P\left[\bar{X}-z_{1-\alpha/2}\frac{\sigma}{\sqrt{n}}<\mu<\bar{X}+z_{1-\alpha/2}\frac{\sigma}{\sqrt{n}}\right].$$

Hasil mana yang kita gunakan bergantung pada tabel yang tersedia, dalam hal tabel menyediakan seluruh nilai maka persamaan pertama bisa kita pakai. Di pihak lain dalam hal tabel hanya menyediakan nilai-nilai tidak negatif persamaan kedua yang kita pakai. Lebih lanjut kita peroleh selang kepercayaan $100(1-\alpha)\%$ untuk μ adalah

$$\left(\bar{x} - z_{1-\alpha/2}\frac{\sigma}{\sqrt{n}}, \bar{x} + z_{1-\alpha/2}\frac{\sigma}{\sqrt{n}}\right).$$

Sebagai contoh, selang kepercayaan 95% untuk μ adalah

$$\left(\bar{x} - 1.96\frac{\sigma}{\sqrt{n}}, \bar{x} + 1.96\frac{\sigma}{\sqrt{n}}\right).$$

Berikut ini kita akan menentukan selang kepercayaan untuk μ dari sampel acak yang bersebaran normal jika variansnya ***tidak*** diketahui.

Contoh 7.8.4 Misal $X_1, X_2, ..., X_n$ sampel acak dari sebaran normal, $X_i \sim N(\mu, \sigma^2)$, dan σ^2 tidak diketahui. Kita akan menentukan selang kepercayaan $100(1-\alpha)\%$ untuk μ. Berdasarkan Teorema 5.5.2 kita mempunyai

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

Oleh karena itu kita peroleh

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

Kita dapat pula menggunakan sifat kesimetrikan, dalam hal ini pada sebaran t, yaitu $t(n-1)_{\alpha/2} = -t(n-1)_{1-\alpha/2}$. Sehingga kita peroleh pula bentuk lain

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

Hasil mana yang kita gunakan bergantung pada tabel yang tersedia, dalam hal tabel menyediakan seluruh nilai maka persamaan pertama yang kita pakai. Di pihak lain dalam hal tabel hanya menyediakan nilai-

nilai tidak negatif persamaan kedua yang kita pakai. Lebih lanjut kita peroleh selang kepercayaan $100(1-\alpha)\%$ untuk μ adalah

$$\left(\bar{x}-t_{1-\alpha/2}(n-1)\frac{s}{\sqrt{n}}, \bar{x}+t_{1-\alpha/2}(n-1)\frac{s}{\sqrt{n}}\right).$$

Contoh berikut ini masih dari sebaran normal, tetapi kita akan menentukan selang kepercayaan untuk σ^2.

Contoh 7.8.5 Misal $X_1, X_2, ..., X_n$ sampel acak dari sebaran normal, $X_i \sim N(\mu, \sigma^2)$. Kita mempunyai

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

Oleh karena itu kita peroleh

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

Jadi selang kepercayaan $100(1-\alpha)\%$ untuk σ^2 adalah

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

Kita telah menentukan selang kepercayaan beberapa parameter. Meskipun demikian kita belum menentukan selang kepercayaan untuk parameter dari peubah acak diskret. Contoh berikut ini merupakan selang kepercayaan untuk parameter peubah acak diskret. Kita gunakan pendekatan normal.

Contoh 7.8.6 Misal $X_1, X_2, ..., X_n$ sampel acak dari sebaran Bernoulli, $X_i \sim BIN(1, p)$. Tentukan selang kepercayaan $100(1-\alpha)\%$ untuk p.

Penyelesaian

Karena $X_1, X_2, ..., X_n$ sampel acak dari sebaran Bernoulli, $X_i \sim BIN(1, p)$, maka $\mu = 1 \times p = p.$ Oleh karena itu MMEs dari p adalah $\hat{p} = \frac{\sum_{i=1}^{n} X_i}{n}.$

Menggunakan CLT kita peroleh

$$\frac{\hat{p} - p}{\sqrt{p(1-p)/n}} \xrightarrow{d} Z \sim N(0,1).$$

Oleh karena itu untuk $n \to \infty,$ berlaku

$$P\left[z_{\alpha/2} < \frac{\hat{p} - p}{\sqrt{p(1-p)/n}} < z_{1-\alpha/2} \right] = 1 - \alpha.$$

Selanjutnya kita gunakan $\hat{p}$ lagi sebagai penduga dari p untuk menentukan selang kepercayaan dari $p,$ seperti berikut ini.

$$\frac{\hat{p} - p}{\sqrt{\hat{p}(1-\hat{p})/n}} \xrightarrow{d} Z \sim N(0,1).$$

Sehingga untuk untuk $n \to \infty,$ berlaku

$$P\left[z_{\alpha/2} < \frac{\hat{p} - p}{\sqrt{\hat{p}(1-\hat{p})/n}} < z_{1-\alpha/2} \right] = 1 - \alpha.$$

Jadi selang kepercayaan $100(1-\alpha)\%$ untuk p adalah

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

Atau

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

Latihan 7.8

7.8.1. Misal nilai pengamatan $\overline{X}$ dari purata sampel acak berukuran 20 dari sebaran $N(\mu, 80)$ adalah $81.2.$ Tentukan:

(a) Batas bawah kepercayaan satu sisi 95% untuk $\mu.$

(b) Batas atas kepercayaan satu sisi 95% untuk $\mu.$

(c) Selang kepercayaan 95% untuk $\mu.$

7.8.2. Misal $\bar{X}$ purata sampel acak berukuran n dari sebaran $N(\mu, 9)$. Tentukan n sehingga $P\left[\bar{X}-1<\mu<\bar{X}+1\right]=0.90.$

7.8.3. Misal $X_1, X_2, ..., X_n$ sampel acak dari sebaran $EXP(\theta)$.

(a) Jika $\bar{x}=17.9,$ dan maka tentukan batas bawah kepercayaan satu sisi 95% untuk θ.

(b) Tentukan batas bawah kepercayaan satu sisi 95% untuk $P(X>t)=e^{-t/\theta}$.

7.8.4. Misal $X_1, X_2, ..., X_9$ sampel acak dari sebaran $N(\mu, \sigma^2)$, dan σ^2 tidak diketahui. Tentukan selang kepercayaan 95% untuk μ.

7.8.5. Misal sampel acak berukuran 17 dari sebaran $N(\mu, \sigma^2)$ mempunyai nilai-nilai pengamatan $\bar{x}=4.7$ dan $s^2=5.76$. Tentukan selang kepercayaan 90% untuk μ.

7.8.6. Misal $Y \sim BIN(n, p)$. Tentukan selang kepercayaan 95% untuk p.

7.8.7. Misal $Y \sim BIN(300, p)$. Jika nilai pengamatan Y adalah $y=75$, tentukan selang kepercayaan 90% untuk p.

7.8.8. Jika 8.6 , 7.9 , 8.3 , 6.4 , 8.4 , 9.8 , 7.2 , 7.8 , 7.5 adalah nilai-nilai pengamatan dari sampel acak berukuran 9 dari sebaran $N(\mu, \sigma^2)$, σ^2 tidak diketahui, maka tentukanlah selang kepercayaan 90% untuk μ.

7.8.9. Jika 8.6 , 7.9 , 8.3 , 6.4 , 8.4 , 9.8 , 7.2 , 7.8 , 7.5 adalah nilai-nilai pengamatan dari sampel acak berukuran 9 dari sebaran $N(8, \sigma^2)$, maka tentukanlah selang kepercayaan 90% untuk σ^2.

7.8.10. Misal $X_1, X_2, ..., X_n$ sampel acak dari sebaran Weibull, $X_i \sim WEI(\theta, 20)$.

(a) Tunjukkan bahwa $Q = 2\sum_{i=1}^{n} \frac{X_i^2}{\theta^2} \sim \chi^2(2n)$.

(b) Gunakan Q untuk menentukan selang kepercayaan $100(1-\alpha)\%$ untuk θ.

7.9 Permasalahan Dua Sampel

Sering kali pengambilan sampel acak dimaksudkan untuk membandingkan dua atau lebih populasi. Salah satunya mungkin untuk membandingkan purata dari dua proses, atau di pihak lain untuk membandingkan varians dari dua proses. Selang kepercayaan akan memberikan informasi dalam membuat nisbah dua proses ini.

Kita perhatikan dua sampel acak berukuran n_1, dan n_2 dari dua populasi berturut-turut bersebaran normal $X_i \sim N(\mu_1, \sigma_1^2)$, dan $Y_i \sim N(\mu_2, \sigma_2^2)$. Kita akan menentukan selang kepercayaan untuk nisbah $\dfrac{\sigma_1^2}{\sigma_2^2}$. Karena

$$\frac{s_1^2\sigma_2^2}{s_2^2\sigma_1^2} \sim F(n_1, n_2),$$

maka

$$P\left[f_{\alpha/2}(n_1-1, n_2-1) < \frac{S_1^2\sigma_2^2}{S_2^2\sigma_1^2} < f_{1-\alpha/2}(n_1-1, n_2-1) \right] = 1-\alpha,$$

atau

$$P\left[\frac{S_2^2}{S_1^2} f_{\alpha/2}(n_1-1, n_2-1) < \frac{\sigma_2^2}{\sigma_1^2} < \frac{S_2^2}{S_1^2} f_{1-\alpha/2}(n_1-1, n_2-1) \right] = 1-\alpha.$$

Jadi selang kepercayaan $100(1-\alpha)\%$ untuk $\dfrac{\sigma_2^2}{\sigma_1^2}$ adalah

$$\left(\frac{s_2^2}{s_1^2} f_{\alpha/2}(n_1-1, n_2-1), \frac{s_2^2}{s_1^2} f_{1-\alpha/2}(n_1-1, n_2-1) \right).$$

Contoh 7.9.1 Anggap dua sampel acak masing-masing berukuran $n_1 = 16$ dan $n_2 = 21$ dengan varians berturut-turut $s_1^2 = 0.60$ dan $s_2^2 = 0.20$. Jika $\alpha = 0.1$, maka

$$f_{1-\alpha/2}(n_1-1, n_2-1) = f_{0.95}(15, 20) = 2.20$$

dan

$$f_{\alpha/2}(n_1-1, n_2-1) = f_{0.05}(15,20) = \frac{1}{f_{0.95}(20,15)} = \frac{1}{2.33} = 0.429.$$

Oleh karena itu selang kepercayaan 90% untuk $\dfrac{\sigma_2^2}{\sigma_1^2}$ adalah $(0.143, 0.733)$.

Catatan: karena selang ini tidak memuat 1, maka kita dapat menyimpulkan bahwa $\sigma_2^2 \neq \sigma_1^2$. Permasalahan ini nanti akan dibahas lebih dalam pada pengujian hipotesis.

Pembahasan berikutnya masih dari dua populasi yang bersebaran normal, yaitu menentukan selang kepercayaan untuk selisih dua purata. Kita bahas terlebih dahulu dalam hal kedua varians diketahui, yaitu σ_1^2 dan σ_2^2 masing-masing diketahui. Kita gunakan kenyataan bahwa

$$\bar{Y}_{n_2} - \bar{X}_{n_1} \sim N\left(\mu_2 - \mu_1, \frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}\right).$$

Oleh karena itu

$$Z = \frac{(\bar{Y}_{n_2} - \bar{X}_{n_1}) - (\mu_2 - \mu_1)}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}} \sim N(0,1).$$

Selanjutnya untuk menyederhanakan penulisan $\bar{X}_{n_1}$ ditulis sebagai $\bar{X}$ demikian juga $\bar{Y}_{n_2}$ ditulis sebagai $\bar{Y}$. Seperti halnya pada pembahasan sebelumnya, untuk mendapatkan selang kepercayaan $100(1-\alpha)\%$, kita tentukan seperti berikut ini.

$$P\left[z_{\alpha/2} < Z < z_{1-\alpha/2}\right] = 1-\alpha$$

$$P\left[z_{\alpha/2} < \frac{(\bar{Y} - \bar{X}) - (\mu_2 - \mu_1)}{\sqrt{\sigma_1^2/n_1 + \sigma_2^2/n_2}} < z_{1-\alpha/2}\right] = 1-\alpha.$$

Kita peroleh selang kepercayaan $100(1-\alpha)\%$ untuk $(\mu_2 - \mu_1)$ adalah

$$\left(\bar{y} - \bar{x} + z_{\alpha/2}\sqrt{\sigma_1^2/n_1 + \sigma_2^2/n_2}, \bar{y} - \bar{x} + z_{1-\alpha/2}\sqrt{\sigma_1^2/n_1 + \sigma_2^2/n_2}\right),$$

atau

$$\left(\bar{y} - \bar{x} - z_{1-\alpha/2}\sqrt{\sigma_1^2/n_1 + \sigma_2^2/n_2}, \bar{y} - \bar{x} + z_{1-\alpha/2}\sqrt{\sigma_1^2/n_1 + \sigma_2^2/n_2}\right).$$

Selanjutnya kita akan menentukan selang kepercayaan selisih purata jika jika kedua variansnya, yaitu σ_1^2 dan σ_2^2 tidak diketahui, tetapi $\sigma_1^2 = \sigma_2^2 = \sigma^2$. Seperti halnya pada selang kepercayaan untuk satu parameter μ, di sini kita juga akan menggunakan statistik T. Mula-mula kita gunakan kenyataan bahwa sekelompok penduga dari varians bersama

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

Jika

$$V = \frac{(n_1+n_2-2)S_p^2}{\sigma^2},$$

maka

$$V = \frac{(n_1-1)S_1^2}{\sigma^2} + \frac{(n_2-1)S_2^2}{\sigma^2} \sim \chi^2(n_1+n_2-2).$$

Oleh karena itu

$$T = \frac{(\bar{Y}-\bar{X})-(\mu_2-\mu_1)}{S_p\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}} = \frac{Z}{\sqrt{\frac{V}{n_1+n_2-2}}} \sim t(n_1+n_2-2),$$

bukti dijadikan sebagai latihan. Seperti halnya pada pembahasan sebelumnya kita akan dengan mudah mendapakan hasil bahwa selang kepercayaan $100(1-\alpha)\%$ untuk $\mu_2-\mu_1$ adalah

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

atau

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

Contoh 7.9.2 Anggap dua sampel acak masing-masing berukuran $n_1 = 16$ dan $n_2 = 21$ dengan purata masing-masing $\bar{x} = 4.31,$ $\bar{y} = 5.22,$ dan varians masing-masing $s_1^2 = 0.12,$ $s_2^2 = 0.10.$ Maka

$$s_p^2 = \left[(15)(0.12)+(20)(0.10)\right]/35 = 0.109,$$

sehingga $s_p = 0.330.$ Jika $\alpha = 0.05,$ maka

$$t_{1-\alpha/2}(35) = t_{0.975}(35) = 2.032.$$

Oleh karena itu selang kepercayaan 95% untuk $\mu_2 - \mu_1$ adalah $(0.68, 1.133)$.

Seperti halnya pada catatan sebelumnya, karena selang kepercayaan ini tidak memuat nilai 0, maka kita tidak dapat mengatakan bahwa $\mu_2 \neq \mu_1$. Permasalahan ini akan dibahas lebih dalam pada pengujian hipotesis.

Selanjutnya kita akan menentukan selang kepercayaan selisih purata jika jika kedua variansnya, yaitu σ_1^2 dan σ_2^2 tidak diketahui, tetapi $\sigma_1^2 \neq \sigma_2^2$. Untuk n_1 dan n_2 mendekati takhingga,

$$\frac{(\bar{Y}-\bar{X})-(\mu_2-\mu_1)}{\sqrt{\frac{s_1^2}{n_1}+\frac{s_2^2}{n_2}}} \xrightarrow{d} Z \sim N(0,1) \qquad (7.9.1)$$

Jadi untuk ukuran sampel besar, pendekatan batas kepercayaat untuk $\mu_2 - \mu_1$ dapat diperoleh dari persamaan (7.9.1). Pendekatan ini berlaku juga untuk sampel besar tidak dari sebaran normal.

Untuk ukuran sampel kecil dari sebaran normal, sebaran dari peubah acak pada (7.9.1) bergantung pada σ_1^2 dan σ_2^2, dapat digunakan pendekatan berdasarkan pada sebaran t,

$$T = \frac{(\bar{Y}-\bar{X})-(\mu_2-\mu_1)}{\sqrt{\frac{s_1^2}{n_1}+\frac{s_2^2}{n_2}}} \sim t(v),$$

derajat bebasnya adalah diduga oleh

$$\nu = \frac{\left(s_1^2/n_1 + s_2^2/n_2\right)^2}{\left[\left(s_1^2/n_1\right)^2/\left(n_1-1\right)\right] + \left[\left(s_2^2/n_2\right)^2/\left(n_2-1\right)\right]}.$$

Secara umum derajat bebas ini tidak menghasilkan bilangan bulat positif, tetapi dapat digunakan interpolasi untuk menentukan selang kepercayaannya. Permasalahan umum untuk membuat inferens tentang $\mu_2 - \mu_1$ dengan varians tidak sama dikemukakan dalam permasalahan Behrens-Fisher.

Semua hasil pada pembahasan kita tentang selang kepercayaan di atas dengan asumsi bahwa sampel-sampel acak tersebut bebas. Di bawah ini kita akan membahas dua sampel yang tidak bebas, yaitu dua sampel yang dibangkitkan dari satu sampel, atau sampel yang berpasangan. Sebagai contoh jika kita ingin menyelidiki pengaruh obat tertentu terhadap berat badan, maka kita dapat memandangnya sebagai dua sampel yang berpasangan.

Misal kita mempunyai sampel acak dari n pasang, yaitu (X_i, Y_i), dan misal $D_i = Y_i - X_i$, $i = 1, 2, \ldots, n$, bersebaran normal dengan purata $\mu_D = \mu_2 - \mu_1$ dan varians $\sigma_D^2 = \sigma_1^2 + \sigma_2^2 - 2\rho\sigma_1\sigma_2$, atau

$$D_i \sim N(\mu_D, \sigma_D^2).$$

Misal

$$\bar{D} = \sum_{i=1}^{n} \frac{D_i}{n} = \bar{Y} - \bar{X},$$

dan

$$S_D^2 = \frac{n\sum_{i=1}^{n} D_i^2 - \left(\sum_{i=1}^{n} D_i\right)^2}{n(n-1)}.$$

Oleh karena itu

$$T = \frac{\bar{D} - \mu_D}{S_D\sqrt{n}} \sim t(n-1),$$

sehingga selang kepercayaan $(1-\alpha)100\%$ untuk μ_D adalah

$$\bar{d} \pm t_{1-\alpha/2}(n-1)s_D / \sqrt{n}.$$

Latihan7.9

7.9.1. Misal dua sampel acak bebas masing-masing bersebaran normal $N(\mu_1, 8)$ dan $N(\mu_2, 8)$ dengan $\bar{x} = 4.8$, $\bar{y} = 5.6$, $n_1 = 20$, dan $n_2 = 18$. Tentukan selang kepercayaan 95% untuk $\mu_2 - \mu_1$.

7.9.2. Misal dua sampel acak bebas masing-masing bersebaran normal $N(\mu_1, \sigma^2)$ dan $N(\mu_2, \sigma^2)$ dengan $\bar{x} = 4.8$, $s_1^2 = 8.64$, $\bar{y} = 5.6$,

$s_2^2 = 7.88,\quad n_1 = 15,$ dan $n_2 = 20.$ Tentukan selang kepercayaan 95% untuk $\mu_2 - \mu_1.$

7.9.3. Misal dua sampel acak bebas masing-masing bersebaran normal $N\left(\mu_1, \sigma_1^2\right)$ dan $N\left(\mu_2, \sigma_2^2\right)$ dengan $\sigma_1^2 \neq \sigma_2^2,\quad \bar{x} = 4.8,\quad s_1^2 = 8.64,$ $\bar{y} = 5.6,\quad s_2^2 = 7.88,\quad n_1 = 15,$ dan $n_2 = 20.$ Tentukan pendekatan selang kepercayaan 95% untuk $\mu_2 - \mu_1.$

7.9.4. Misal dua sampel acak bebas masing-masing bersebaran normal $N\left(\mu_1, \sigma_1^2\right)$ dan $N\left(\mu_2, \sigma_2^2\right)$ dengan $n_1 = 15,\quad n_2 = 16,\quad s_1^2 = 8.64,$ dan $s_2^2 = 7.88.$ Tentukan:

a. Selang kepercayaan 95% untuk $\dfrac{\sigma_2^2}{\sigma_1^2}.$

b. Batas bawah 95% untuk $\dfrac{\sigma_2^2}{\sigma_1^2}.$

c. Batas atas 95% untuk $\dfrac{\sigma_2^2}{\sigma_1^2}.$

7.9.5. Misal dua peubah acak saling bebas Y_1 dan Y_2 dengan sebaran binomial dengan $n_1 = n_2 = 100,$ dengan nilai pengamatan masing-masing $y_1 = 50$ dan $y_2 = 40.$ Tentukan selang kepercayaan 95% untuk $p_1 - p_2.$

7.9.6. Misal dua sampel acak bebas masing-masing bersebaran eksponensial, $X_i \sim EXP(\theta_1)$, dan $Y_i \sim EXP(\theta_2)$, $i = 1, \ldots, n_1;\ \ j = 1, \ldots, n_2.$

(a) Tunjukkan bahwa $\frac{\theta_2 \bar{X}}{\theta_1 \bar{Y}} \sim F(2n_1, 2n_2).$

(b) Turunkan rumus selang kepercayaan $100(1-\alpha)\%$ untuk $\dfrac{\theta_2}{\theta_1}.$

7.9.7. Misal dua populasi masing-masing bersebaran normal dengan $\sigma_1^2 = \sigma_2^2 = \sigma^2$, dan varians sampel bersamanya, yaitu S_1^2 dan S_2^2 mempunyai purata

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

Dan misal

$$V = \frac{(n_1 + n_2 - 2)S_p^2}{\sigma^2}.$$

Buktikan bahwa

(c) $V = \frac{(n_1-1)S_1^2}{\sigma^2} + \frac{(n_2-1)S_2^2}{\sigma^2} \sim \chi^2(n_1 + n_2 - 2)$.

(d) $T = \frac{(\bar{Y}-\bar{X})-(\mu_2-\mu_1)}{S_p\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}} = \frac{Z}{\sqrt{\frac{V}{n_1+n_2-2}}} \sim t(n_1 + n_2 - 2)$.

7.9.8. Misal diketahui sampel acak dari n pasang, yaitu (X_i, Y_i), di mana X_i adalah berat badan orang ke i sebelum mengkonsumsi obat tertentu, dan Y_i adalah berat badan orang ke i setelah mengkonsumsi obat tertentu. Misal dari 30 orang yang diuji coba diperoleh:

Data sebelum mengkonsumsi obat tertentu:

48.054	109.032	114.768	90.417	73.561	58.895	79.541
95.201	90.681	56.128	49.073	32.904	55.195	61.777
23.960	69.621	55.193	68.434	118.749	74.387	84.930
26.862	38.594	104.271	58.314	44.227	107.826	39.181
18.349	116.810					

Data setelah mengkonsumsi obat tertentu:

57.8940 68.0018 86.7309 58.6977 80.4364 70.9561
64.1523
73.5232 79.2312 51.1944 86.1404 79.8889 54.3916
30.3850
62.0062 48.5749 48.6728 45.6783 64.6570 65.3762
52.4895
73.8399 75.6171 71.1410 48.5505 69.9978 59.1688
64.0067
66.7467 63.9632

Dengan menganggap $D_i \sim N(\mu_D, \sigma_D^2)$ tentukan selang kepercayaan 95% untuk $u_D = \mu_1 - \mu_2$. Dan beri interpretasi dari hasil ini.

7.10 Penduga Bayes

Dalam subpokok bahasan sebelumnya kita telah mendapatkan dua statistik, katakanlah W dan U, di mana $W < U$, sehingga besar peluang bahwa selang acak (W, U) memuat parameter adalah γ. Kita sekarang akan membahas pendekatan lain dari penduga selang. Pendekatan ini mempertimbangkan pengetahuan awal dari suatu percobaan yang disebut ***penduga Bayes.***

Misal peubah acak X mempunyai sebaran peluang yang bergantung pada θ, $\theta \in \Omega$. Sebagai contoh, jika θ menyatakan purata dari sebaran normal, maka Ω merupakan garis real. Kita perkenalkan peubah acak Θ yang mempunyai sebaran peluang atas himpunan Ω, dan seperti pembahasan sebelumnya, bahwa x adalah nilai dari peubah acak X, kita anggap θ nilai yang mungkin dari peubah acak Θ. Oleh karena itu X bergantung pada θ, suatu peubah acak yang menentukan peubah acak Θ. Kita nyatakan pdf dari Θ sebagai $h(\theta)$ dan kita ambil $h(\theta) = 0$ jika θ bukan anggota Ω. Pdf ini disebut sebagai ***pdf awal (prior)*** dari Θ. Kita nyatakan pdf dari X sebagai $f(x \mid \theta)$ karena kita pandang sebagai pdf bersyarat dari X, bila $\Theta = \theta$. Selanjutnya kita gunakan model

$$X|\theta \sim f(x;\theta)$$
$$\Theta \sim h(\theta).$$

Sekarang anggap bahwa $X_1, X_2, ..., X_n$ adalah sampel acak dari sebaran dari X, bila $\Theta = \theta$ dengan pdf $f(x \mid \theta)$. Misal $\mathbf{X}^t = \left(X_1, X_2, ..., X_n\right)$ dan $\mathbf{x}^t = \left(x_1, x_2, ..., x_n\right)$. Maka kita dapat menulis pdf bersyarat dari $\mathbf{X}$, bila $\Theta = \theta$, sebagai

$$L(\mathbf{x} \mid \theta) = f(x_1 \mid \theta) f(x_2 \mid \theta) ... f(x_n \mid \theta).$$

Jadi pdf bersama dari $\mathbf{X}$ dan Θ adalah

$$g(\mathbf{x}, \theta) = L(\mathbf{x} \mid \theta) h(\theta).$$

Jika Θ adalah peubah acak kontinu, maka pdf marginal dari $\mathbf{X}$ adalah

$$g_1(\mathbf{x}) = \int_{-\infty}^{\infty} g(\mathbf{x}, \theta) d\theta.$$

Jika Θ adalah peubah diskret, integral diganti dengan jumlah. Dalam kasus lain pdf bersyarat dari $\Theta,$ bila $\mathbf{X} = \mathbf{x},$ adalah adalah

$$k(\theta \,|\, \mathbf{x}) = \frac{g(\mathbf{x}, \theta)}{g_1(\mathbf{x})} = \frac{L(\mathbf{x} \,|\, \theta) h(\theta)}{g_1(\mathbf{x})}, \;\; g_1(\mathbf{x}) > 0.$$

Hubungan ini merupakan bentuk dari kaidah Bayes pada Teorema 5.4.4. Sebaran yang didefinisikan oleh pdf bersyarat $k(\theta \,|\, \mathbf{x})$ disebut ***pdf posterior*** dari $\Theta.$ Disebut demikian karena $h(\theta)$ adalah pdf dari Θ awal untuk pengamatan $\mathbf{X},$ sedangkan $k(\theta \,|\, \mathbf{x})$ adalah pdf dari Θ setelah pengamatan Y telah dilakukan. Dalam beberapa kejadian, $h(\theta)$ tidak diketahui, dipilih $h(\theta)$ yang berpengaruh pada pdf $k(\theta \,|\, \mathbf{x}).$

Contoh 7.10.1 Perhatikan model

$$X_i | \theta \sim \text{iid Poisson}, POI(\theta)$$
$$\Theta \sim GAM(\beta, \alpha),\ \alpha \text{ dan } \beta \text{ diketahui.}$$

Oleh karena sampel acak ini dari sebaran Poisson dengan purata θ dan sebaran awal adalah sebaran $GAM\left(\beta, \alpha\right).$ Misal $\mathbf{X}^t = \left(X_1, X_2, ..., X_n\right).$ Maka pdf bersyarat dari $\mathbf{X},$ bila $\Theta = \theta,$ adalah

$$L(\mathbf{x} \,|\, \theta) = \frac{\theta^{x_1} e^{-\theta}}{x_1!} \frac{\theta^{x_2} e^{-\theta}}{x_2!} ... \frac{\theta^{x_n} e^{-\theta}}{x_n!} I(x_i = 0, 1, 2, ...\, , i = 1, 2, ..., n),$$

dan pdf awal adalah

$$h(\theta) = \frac{\theta^{\alpha - 1} e^{-\theta/\beta}}{\beta^{\alpha} \Gamma(\alpha)} I(0 < \theta < \infty).$$

Sehingga pdf bersama dari $\mathbf{X}$ dan $\Theta,$ adalah

$$\begin{aligned} g(\mathbf{x}, \theta) &= L(\mathbf{x} \,|\, \theta) h(\theta) \\ &= \left[\frac{\theta^{x_1} e^{-\theta}}{x_1!} \frac{\theta^{x_2} e^{-\theta}}{x_2!} ... \frac{\theta^{x_n} e^{-\theta}}{x_n!} \right] \left[\frac{\theta^{\alpha - 1} e^{-\theta/\beta}}{\beta^{\alpha} \Gamma(\alpha)} \right] I(x_i = 0, 1, 2, ...\, , \\ & \quad i = 1, 2, ..., n,\ 0 < \theta < \infty). \end{aligned}$$

Maka pdf marginal dari $\mathbf{X},$ adalah

$$g_1(\mathbf{x}) = \int_0^\infty \frac{\theta^{\sum_{i=1}^{n} x_i + \alpha - 1} e^{-(n+1/\beta)\theta}}{x_1!x_2!\ldots x_n!\beta^\alpha \Gamma(\alpha)} d\theta = \frac{\Gamma\left(\sum_{i=1}^{n} x_i + \alpha\right)}{x_1!x_2!\ldots x_n!\beta^\alpha \Gamma(\alpha)(n+1/\beta)^{\sum_{i=1}^{n} X_i + \alpha}}.$$

Dan pdf posterior dari $\Theta,$ bila $\mathbf{X} = \mathbf{x},$ adalah

$$k(\theta \mid \mathbf{x}) = \frac{L(\mathbf{x} \mid \theta)h(\theta)}{g_1(\mathbf{x})} = \frac{\theta^{\sum_{i=1}^{n} X_i + \alpha - 1} e^{-\theta/[\beta/(n\beta+1)]}}{\Gamma\left(\sum_{i=1}^{n} X_i + \alpha\right)[\beta/(n\beta+1)]^{\sum_{i=1}^{n} x_i + \alpha}} I(0 < \theta < \infty).$$

Dalam Contoh 7.10.1 kita tidak perlu menentukan $g_1(\mathbf{x})$ untuk menentukan $k(\theta \mid \mathbf{x}).$ Jika kita bagi $L(\mathbf{x} \mid \theta)h(\theta)$ dengan $g_1(\mathbf{x})$ kita mendapatkan perkalian faktor, yang bergantung pada $\mathbf{x}$ tetapi tidak bergantung pada $\theta,$ katakanlah $c(\mathbf{x}),$ dan

$$\theta^{\sum_{i=1}^{n} x_i + \alpha - 1} e^{-\theta/[\beta/(n\beta+1)]}.$$

Sehingga

$$k(\theta \mid \mathbf{x}) = c(\mathbf{x})\theta^{\sum_{i=1}^{n} x_i + \alpha - 1} e^{-\theta/[\beta/(n\beta+1)]},$$

asalkan $0 < \theta < \infty,$ dan $x_i = 0, 1, 2, \ldots, n,$ $i = 1, 2, \ldots, n.$ Meskipun demikian, $c(\mathbf{x})$ haruslah merupakan konstanta yang diperlukan untuk membuat $k(\theta \mid \mathbf{x})$ suatu pdf, sebut

$$c(\mathbf{x}) = \frac{1}{\Gamma\left(\sum_{i=1}^{n} x_i + \alpha\right)[\beta/(n\beta+1)]^{\sum_{i=1}^{n} x_i + \alpha}}.$$

Oleh karena itu seringkali $k(\theta \mid \mathbf{x})$ ditulis sebagai nisbah untuk $L(\mathbf{x} \mid \theta)h(\theta),$ yaitu,

$$k(\theta \mid y) \infty L(\mathbf{x} \mid \theta)h(\theta).$$

Selanjutnya anggap kita ingin menentukan penduga dari $\theta.$ Dari sudut pandang Bayes, adalah memilih fungsi keputusan δ sehingga $\delta(\mathbf{x})$ diprediksi nilai dari θ (sebagai nilai percobaan dari peubah acak Θ) jika kedua nilai $\mathbf{x}$ dan $k(\theta \mid \mathbf{x})$ diketahui. Secara umum, bagaimana kita membuat prediksi nilai percobaan dari sebarang peubah acak, katakanlah W, jika kita ingin prediksi kita cukup layak untuk nilai pengamatan? Banyak ahli statistika ingin memprediksi purata, $E(W)$, dari sebaran W; lainnya ingin memprediksi dari median (berharap tunggal) dari sebaran W; beberapa ingin prediksi dari modus (berharap tunggal) dari distirbusi W; dan beberapa ingin prediksi yang lainnya. Meskipun demikian, semuanya menginginkan memilih fungsi keputusan yang bergantung pada fungsi kerugian $\Lambda\left[\theta,\delta(\mathbf{x})\right]$. ***Dugaan Bayes*** adalah fungsi keputusan δ yang meminimumkan

$$E\left\{\Lambda\left[\Theta,\delta(\mathbf{x})\right] \mid \mathbf{X}=\mathbf{x}\right\}=\int_{-\infty}^{\infty}\Lambda\left[\Theta,\delta(\mathbf{x})\right]k\left(\theta \mid \mathbf{x}\right)d\theta,$$

jika Θ adalah peubah acak kontinu. Mengganti integral dengan jumlah jika Θ adalah peubah acak diskret. Peubah acak yang berkaitan dengan dugaan Bayes $\delta(\mathbf{X})$ disebut ***penduga Bayes.*** Sebagai contoh, fungsi kerugian diberikan oleh

$$\Lambda\left[\theta,\delta(\mathbf{x})\right]=\left[\theta-\delta(\mathbf{x})\right]^2,$$

maka dugaan Bayes adalah $\delta(\mathbf{x})=E\left(\Theta \mid \mathbf{x}\right)$, yang merupakan purata dari sebaran bersyarat Θ, bila $\mathbf{X}=\mathbf{x}$. Dari sifat ekspektasi bahwa $E\left[\left(W-b\right)^2\right]$, jika ada, minimum jika $b=E(X)$. Jika fungsi kerugian diberikan oleh $\Lambda\left[\theta,\delta(\mathbf{x})\right]=\left|\theta-\delta(\mathbf{x})\right|$, maka median sebaran bersyarat dari Θ, bila $\mathbf{X}=\mathbf{x}$, adalah dugaan Bayes. Sekali lagi dari sifat ekspektasi bahwa $E\left(|W-b|\right)$, jika ada, adalah minimum jika b adalah median dari sebaran W.

Ekspektasi bersyarat dari kerugian, bila $\mathbf{X}=\mathbf{x}$, mendefinisikan peubah acak yang merupakan fungsi dari $\mathbf{X}$. Nilai ekspektasi dari fungsi $\mathbf{X}$, diberikan oleh

$$\int_{-\infty}^{\infty}\left\{\int_{-\infty}^{\infty}\Lambda\left[\theta,\delta(\mathbf{x})\right]k(\theta \mid y)d\theta\right\}g_1(\mathbf{x})d\mathbf{x}$$

$$=\int_{-\infty}^{\infty}\left\{\int_{-\infty}^{\infty}\Lambda\left[\theta,\delta(\mathbf{x})\right]L(\mathbf{x} \mid \theta)d\mathbf{x}\right\}h(\theta)d\theta,$$

dalam kasus kontinu. Integral dalam kurung persamaan, untuk setiap $\theta \in \Omega,$ suatu fungsi risiko $R(\theta, \delta).$ Oleh karena itu persamaan tersebut merupakan purata dari risiko. Karena dugaan Bayes meminimumkan

$$\int_{-\infty}^{\infty} \Lambda\left[\theta, \delta(\mathbf{x})\right] k\left(\theta \mid \mathbf{x}\right) d\theta,$$

untuk setiap $\mathbf{x}$ di mana $g_1(\mathbf{x}) > 0,$ jelas bahwa dugaan Bayes $\delta(\mathbf{x})$ meminimumkan purata risiko ini.

Contoh 7.10.2 Perhatikan model

$$X_i|\theta \sim \text{iid binomial}, BIN(1, \theta)$$

$$\Theta \sim BETA(\alpha, \beta),\ \alpha \text{ dan } \beta \text{ diketahui.}$$

Maka pdf awalnya adalah

$$h(\theta) = \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} \theta^{\alpha-1}(1-\theta)^{\beta-1} I(0 < \theta < 1),$$

di mana α dan β konstanta positif.

Kita akan mencari fungsi keputusan δ yang merupakan dugaan Bayes. Statistik cukup $Y = \sum_{i=1}^{n} X_i$ mempunyai sebaran binomial, $Y \sim BIN(n, \theta).$

Oleh karena itu pdf bersyarat dari $Y,$ bila $\Theta = \theta,$ adalah

$$g(y \mid \theta) = \binom{n}{y} \theta^y (1-\theta)^{n-y} I(y = 0, 1, \ldots, n),$$

dan pdf bersyarat dari $\Theta,$ bila $Y = y,$ adalah

$$k(\theta \mid y) \infty \theta^y (1-\theta)^{n-y} \theta^{\alpha-1}(1-\theta)^{\beta-1} I(0 < \theta < 1).$$

Jadi

$$k(\theta \mid y) = \frac{\Gamma(n+\alpha+\beta)}{\Gamma(\alpha+y)\Gamma(n+\beta-y)} \theta^{\alpha+y-1}(1-\theta)^{\beta+n-y-1} I(0 < \theta < 1, y = 0, 1, \ldots, n).$$

Selanjutnya kita ambil fungsi kerugian $\Lambda\left[\theta, \delta(y)\right] = \left[\theta - \delta(y)\right]^2.$ Karena Y peubah acak diskret, sedangkan Θ peubah acak kontinu, maka risiko yang diharapkan,

$$\int_0^1 \left\{ \sum_{i=1}^{n} \left[\theta - \delta(y)\right]^2 \binom{n}{y} \theta^y (1-\theta)^{n-y} \right\} h(\theta) d\theta = \sum_{i=1}^{n} \left\{ \int_0^1 \left[\theta - \delta(y)\right]^2 k(\theta \mid y) d\theta \right\} k_1(y).$$

Dugaan Bayes $\delta(y)$ adalah purata dari sebaran bersyarat dari $\Theta,$ bila $Y = y,$

$$\delta(y) = \int_0^1 \theta k(\theta | y) d\theta$$
$$= \frac{\Gamma(n+\alpha+\beta)}{\Gamma(\alpha+y)\Gamma(n+\beta-y)} \int_0^1 \theta^{\alpha+y}(1-\theta)^{\beta+n-y-1} d\theta$$
$$= \frac{\alpha+y}{\alpha+\beta+n}.$$

Dugaan Bayes ini dapat ditulis sebagai

$$\delta(y) = \left(\frac{n}{\alpha+\beta+n}\right)\frac{y}{n} + \left(\frac{\alpha+y}{\alpha+\beta+n}\right)\frac{\alpha}{\alpha+\beta},$$

yang merupakan purata bobot dari dugaan kemungkinan maksimum y/n dari θ dan purata $\alpha/(\alpha+\beta)$ pdf awal dari parameternya.

Contoh 7.10.3 Kita perhatikan model normal

$$X_i \sim \text{iid N}(\theta, \sigma^2),\ \sigma^2 \text{ diketahui}$$
$$\Theta \sim N(\theta_0, \sigma_0^2),\ \theta_0 \text{ dan } \sigma_0^2 \text{ diketahui.}$$

Maka $Y = \bar{X}$ adalah statistik cukup. Oleh karena itu model di atas ekivalen dengan model

$$Y|\theta \sim N\left(\theta, \frac{\sigma^2}{n}\right), \text{ di mana } \sigma^2 \text{ diketahui}$$
$$\Theta \sim N(\theta_0, \sigma_0^2),\ \theta_0 \text{ dan } \sigma_0^2 \text{ diketahui.}$$

Oleh karena itu kita mempunyai pdf posterior

$$k(\theta | y) \propto \frac{1}{\sqrt{2\pi\sigma}/\sqrt{n}} \frac{1}{\sqrt{2\pi\sigma_0}} \exp\left[-\frac{(y-\theta)^2}{2(\sigma^2/n)} - \frac{(\theta-\theta_0)^2}{2\sigma_0^2}\right].$$

Jika kita eliminasi semua kostanta (berisi faktor yang hanya memuat y), kita peroleh

$$k(\theta | y) \propto \exp\left[-\frac{(\sigma_0^2 + \sigma^2/n)\theta^2 - 2(y\sigma_0^2 + \theta_0\sigma^2/n)\theta}{2(\sigma^2/n)\sigma_0^2}\right]$$
$$= \exp\left[\frac{\left(\theta - \dfrac{y\sigma_0^2 + \theta_0\sigma^2/n}{\sigma_0^2 + \sigma^2/n}\right)^2}{\dfrac{2(\sigma^2/n)\sigma_0^2}{\sigma_0^2 + \sigma^2/n}}\right].$$

Yaitu, pdf posterior dari parameter adalah normal dengan purata

$$\frac{y\sigma_0^2 + \theta_0\sigma^2/n}{\sigma_0^2 + \sigma^2/n} = \left(\frac{\sigma_0^2}{\sigma_0^2 + \sigma^2/n}\right)y + \left(\frac{\sigma^2/n}{\sigma_0^2 + \sigma^2/n}\right)\theta_0$$

dan varians $(\sigma^2/n)\sigma_0^2/(\sigma_0^2 + \sigma^2/n)$. Jika fungsi kerugian galat kuadrat digunakan, purata posterior ini adalah penduga Bayes. Kita catat lagi bahwa dia merupakan purata bobot dari dugaan kemungkinan maksimum $y = \bar{x}$ dan purata awal θ_0.

Dugaan selang dari θ ditentukan dari dua fungsi $w(\mathbf{x})$ dan $u(\mathbf{x})$ sehingga peluang bersyarat

$$P\left[w(\mathbf{x}) < \Theta < u(\mathbf{x}) \mid \mathbf{X} = \mathbf{x}\right] = \int_{w(\mathbf{x})}^{u(\mathbf{x})} k(\theta \mid \mathbf{x})d\theta$$

adalah besar, sebagai contoh 0.95. Maka selang $w(\mathbf{x})$ ke $u(\mathbf{x})$ adalah dugaan selang dari θ yang berarti bahwa peluang bersyarat dari Θ anggota dari selang tersebut adalah 0.95. Selang semacam ini disebut ***selang terpercaya (credible)*** atau ***selang peluang***, disebut demikian supaya tidak rancu dengan selang kepercayaan.

Contoh 7.10.4 Dari Contoh 7.10.3, di mana $X_1, X_2, ..., X_n$ adalah sampel acak dari sebaran normal $N(\theta, \sigma^2)$, di mana σ^2 diketahui, dan sebaran awal adalah sebaran normal $N(\theta_0, \sigma_0^2)$. Statistik $Y = \bar{X}$ adalah statistik cukup. Selang terpercaya ditentukan berdasarkan purata dari sebaran posterior dengan menambah dan mengurangi dengan 1.96 dari simpangan bakunya, yaitu,

$$\frac{y\sigma_0^2 + \theta_0\sigma^2/n}{\sigma_0^2 + \sigma^2/n} \pm 1.96\sqrt{\frac{(\sigma^2/n)\sigma_0^2}{\sigma_0^2 + \sigma^2/n}}$$

yang merupakan selang terpercaya dari peluang 0.95 untuk θ.

Latihan 7.10

7.10.1. Misal $X_1, X_2, ..., X_n$ sampel acak dari sebaran binomial, $BIN(1, \theta)$. Misal sebaran awal dari Θ adalah *beta* dengan parameter α dan β. Tunjukkan bahwa pdf posterior $k(\theta \mid x_1, x_2, ..., x_n)$ adalah sama seperti $k(\theta \mid y)$ pada Contoh 7.10.2.

7.10.2. Misal $X_1, X_2, ..., X_n$ sampel acak dari sebaran normal, $N(\theta, \sigma^2)$, di mana $-\infty < \theta < \theta$ dan $\sigma^2 > 0$. Misal $Y = \bar{X}$ purata sampel acak. Ambil fungsi kerugian $\Lambda[\theta, \delta(y)] = |\theta - \delta(y)|$. Jika θ merupakan nilai pengamatan dari peubah acak Θ, yaitu, $N(\mu, \tau^2)$, di mana $\tau^2 > 0$ dan μ diketahui, tentukan raksiran Bayes $\delta(y)$ untuk dugaan titik θ.

7.10.3. Misal $X_1, X_2, ..., X_n$ sampel acak dari sebaran Poisson, $POI(\theta)$, di mana $0 < \theta < \infty$. Misal $Y = \sum_{i=1}^{n} X_i$. Gunakan fungsi kerugian $\Lambda[\theta, \delta(y)] = [\theta - \delta(y)]^2$. Jika θ merupakan nilai pengamatan dari peubah acak Θ, yang mempunyai pdf

$$h(\theta) = \theta^{\alpha-1} e^{-\theta/\beta} / \beta^{\alpha} \Gamma(\alpha) I(0 < \theta < \infty),$$

di mana $\beta > 0$ dan $\alpha > 0$ diketahui, tentukan dugaan Bayes $\delta(y)$ untuk dugaan titik θ.

7.10.4. Misal Y_n statistik urutan ke n dari sampel acak berukuran n dari sebaran dengan pdf

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

Gunakan fungsi kerugian $\Lambda[\theta, \delta(y_n)] = [\theta - \delta(y_n)]^2$. Jika θ merupakan nilai pengamatan dari peubah acak Θ, yang mempunyai pdf

$$h(\theta) = \frac{\beta \alpha^{\beta}}{\theta^{\beta+1}} I(\alpha < \theta < \infty),$$

di mana $\beta > 0$ dan $\alpha > 0$ diketahui, tentukan dugaan Bayes $\delta(y_n)$ untuk dugaan titik θ.

7.10.5. Dalam Contoh 7.10.2 misal $n = 10$, $\alpha = 10$, dan $\beta = 5$ sehingga $\delta(y) = (10 - y)/45$ adalah dugaan Bayes dari θ. Jika $Y \sim BIN(30, \theta)$, hitung risiko $E\left\{[\theta - \delta(Y)]^2\right\}$.

BAB 8
UJI HIPOTESIS

8.1 Pendahuluan

Dalam banyak kegiatan seringkali kita berhadapan dengan pertanyaan-pertanyaan di antaranya seperti berikut: Apakah metode pembelajaran yang baru lebih baik dibandingkan dengan metode yang sudah ada? Apakah obat tertentu lebih efektif dibandingkan dengan obat yang lain? Apakah komponen listrik tertentu lebih tahan lama dibandingkan dengan komponen yang lain? Diperlukan suatu percobaan untuk menjawab pertanyaan-pertanyaan tersebut. Permasalahan ini dalam Statistika disebut sebagai ***uji hipotesis.***

Tujuan instruksional umum dari mempelajari bab ini adalah Anda diharapkan memahami pengujian suatu hipotesis.

Tujuan instruksional khusus dari mempelajari bab ini adalah Anda diharapkan dapat menentukan: definisi suatu hipotesis, daerah kritis suatu uji, daerah kritis yang lebih baik, nilai peluang galat jenis I, nilai peluang galat jenis II, definisi fungsi kuasa, hubungan nilai peluang galat jenis I, nilai peluang galat jenis II, dan fungsi kuasa, uji hipotesis untuk purata jika simpangan baku diketahui, uji hipotesis untuk purata jika simpangan baku tidak diketahui, uji hipotesis untuk varians, uji hipotesis untuk nisbah dua varians, uji hipotesis untuk selisih dua purata, uji terkuasa, uji paling kuasa seragam, dan uji nisbah kemungkinan.

8.2 Daerah Kritis

Misal berdasarkan pengalaman diperoleh purata nilai Statistika Matematis mahasiswa Jurusan Matematika FMIPA Universtas Negeri Malang adalah 60. Misalkan kita ingin membuat suatu inovasi pembelajaran. Dengan inovasi tersebut tentu kita berkeinginan bahwa purata nilainya lebih dari 60, misal 62. Pernyataan pertama di atas disebut sebagai ***hipotesis nol***, ditulis sebagai H_0, merupakan hipotesis yang hendak ditolak. Pernyataan kedua disebut ***hipotesis alternatif*** atau ***hipotesis satu***, ditulis sebagai H_1.

Definisi 8.2.1 ***Hipotesis statistika*** adalah pernyataan tentang sebaran dari satu atau lebih peubah acak. Jika hipotesis statistika lengkap, maka disebut ***hipotesis sederhana***, dan jika yang lain disebut ***hipotesis komposit***.

Pernyataan pada permasalahan nilai Statistika Matematis tersebut secara matematis ditulis sebagai $H_0 : \mu = \mu_0 = 60,$ dan $H_1 : \mu = \mu_1 = 62,$ di mana μ merupakan parameter yang tidak diketahui. Himpunan semua nilai yang mungkin dari suatu parameter disebut sebagai ruang parameter, ditulis sebagai Ω. Untuk hipotesis tersebut $\Omega = \{\mu : \mu = 60, 62\}$.

Suatu percobaan telah dilakukan, secara khusus dalam contoh sebelumnya misal kita telah melakukan inovasi pembelajaran, sehingga diperoleh data atau nilai sampel $x_1, x_2, ..., x_n$. Maka berdasarkan data tersebut kita dapat memutuskan apakah hipotesis nol ditolak atau tidak ditolak. Daerah C sehingga jika $(x_1, x_2, ..., x_n) \in C$ hipotesis nol ditolak disebut ***daerah kritis*** atau ***daerah penolakan.***

Definisi 8.2.2 Daerah kritis untuk suatu pengujian hipotesis adalah himpunan bagian dari ruang sampel yang berkorespondensi dengan penolakan hipotesis nol.

Dalam contoh kita, $\overline{X}$ adalah statistik cukup untuk μ, kita gunakan $\overline{X}$ ini sebagai ***statistik uji***. Karena $H_1 : \mu_1 > \mu_0$, maka bentuk dari daerah kritis dalam permasalahan ini adalah $C = \{(x_1, x_2, ..., x_n) : \overline{x} \geq c\}$, untuk suatu konstanta c. Oleh karena itu kita akan menolak H_0 jika $\overline{x} \geq c$, dan tidak menolak H_0 jika $\overline{x} < c$. Terdapat dua galat yang mungkin kita buat dalam prosedur ini. Kita mungkin menolak H_0 padahal H_0 benar, atau kita mungkin tidak menolak H_0 padahal H_0 salah. Dua galat tersebut dinyatakan seperti berikut ini:

1. ***Galat jenis*** I: menolak H_0 padahal H_0 benar.
2. ***Galat jenis*** II: tidak menolak H_0 padahal H_0 salah.

Statistik uji atau daerah kritis yang baik adalah daerah kritis yang memberikan peluang membuat dua galat sekecil-kecilnya. Kita ambil notasi-notasi berikut untuk peluang masing-masing dua jenis galat:

1. $\alpha = P[\text{Galat jenis I}]$
2. $\beta = P[\text{Galat jenis II}]$

Definisi 8.2.3 Untuk hipotesis nol sederhana, $\alpha = P[\text{Galat jenis I}]$ dikatakan sebagai ***taraf signifikan*** suatu uji. Untuk hipotesis nol komposit, taraf signifikan suatu uji, α, adalah nilai maksimum peluang galat jenis I.

Taraf signifikan disebut juga ukuran daerah kritis, dan taraf signifikan yang sering digunakan dalam suatu uji adalah $\alpha = 0.05$ atau $\alpha = 0.01$. Di antara semua daerah kritis ukuran α, kita pilih daerah kritis yang memberikan nilai β terkecil.

Kembali ke permasalahan nilai Statistika Matematis dan kita misalkan sebagai X. Anggap nilai tersebut bersebaran normal, $X \sim N(\mu, 100)$. Akan kita tentukan α dan β untuk pengujian hipotesis $H_0 : \mu = 60$ lawan $H_1 : \mu = 62$ berdasarkan daerah kritis $C = \{(x_1, ..., x_{25}) \mid \bar{x} \geq 61\}$, seperti berikut ini:

$$\alpha = P\left[\bar{X} \geq 61 \mid H_0 : \mu = 60\right] = P\left[Z \geq \frac{5(61-60)}{10}\right]$$

$$= 1 - P[Z < 0.5] = 1\text{- } 0.6915 = 0.3085,$$

dan

$$\beta = P\left[\bar{X} < 61 \mid H_1 : \mu = 62\right] = P\left[Z < \frac{5(61-62)}{10}\right]$$

$$= P[Z < -0.5] = 0.3085.$$

Kita peroleh nilai α dan β yang cukup kecil. Jadi daerah kritis C cukup baik.

Selanjutnya jika kita tetapkan $\alpha = 0.05$ dan $\beta = 0.10,$ maka

$$0.05 = \alpha = P[X \geq c \mid H_0 : \mu = 60] = P\left[\frac{\sqrt{n}\left(\bar{X}-60\right)}{10} \geq \frac{\sqrt{n}\left(c-60\right)}{10}\right],$$

yang menghasilkan persamaan

$$1.645 = \frac{\sqrt{n}\left(c-60\right)}{10}. \qquad (8.2.1)$$

Dan

$$0.10 = \beta = P[X < c \mid H_1 : \mu = 62] = P\left[\frac{\sqrt{n}\left(\bar{X}-62\right)}{10} < \frac{\sqrt{n}\left(c-62\right)}{10}\right],$$

yang menghasilkan persamaan

$$-1.282 = \frac{\sqrt{n}\left(c-62\right)}{10}. \qquad (8.2.2)$$

Penyelesaian simultan dari persamaan (8.2.1) dan (8.2.2) adalah $n \approx 214,$ dan $c = 61.23.$ Jadi daerah kritisnya

$$C = \left\{\left(x_1, ..., x_{214}\right) \mid \bar{x} \geq 61.23\right\}.$$

Secara umum, jika kita menguji $H_0 : \mu = \mu_0$ lawan $H_1 : \mu = \mu_1$ (di mana $\mu_0 < \mu_1$) pada taraf signifikan $\alpha,$ maka berdasarkan statistik $Z_0,$ di mana

$$Z_0 = \frac{\sqrt{n}\left(\bar{X}-\mu_0\right)}{\sigma}$$

kita peroleh

$$\begin{aligned}\alpha &= P\left[\bar{X} \geq c \mid H_0 : \mu = \mu_0\right] \\ &= P\left[\frac{\sqrt{n}\left(\bar{X}-\mu_0\right)}{\sigma} \geq \frac{\sqrt{n}\left(c-\mu_0\right)}{\sigma}\right] \\ &= P\left[Z_0 \geq z_{1-\alpha}\right],\end{aligned}$$

di mana

$$z_{1-\alpha} = \frac{\sqrt{n}\left(c-\mu_0\right)}{\sigma}.$$

Oleh karena itu kita akan menolak H_0 jika

$$z_0 \geq z_{1-\alpha},$$

di mana z_0 merupakan nilai dari statistik Z_0. Dikatakan bahwa daerah kritis ukuran α ini terletak pada ekor kanan dari kurva sebaran normal. Uji hipotesis semacam ini disebut sebagai uji hipotesis satu arah. Selanjutnya

$$\begin{aligned}
\beta &= P\left[\bar{X} < c \mid H_1 : \mu = \mu_1\right] \\
&= P\left[\frac{\sqrt{n}\left(\bar{X}-\mu_1\right)}{\sigma} < \frac{\sqrt{n}\left(c-\mu_1\right)}{\sigma}\right] \\
&= P\left[\frac{\sqrt{n}\left(\bar{X}-\mu_1\right)}{\sigma} < \frac{\sqrt{n}\left(c-\mu_0+\mu_0-\mu_1\right)}{\sigma}\right] \\
&= P\left[\frac{\sqrt{n}\left(\bar{X}-\mu_1\right)}{\sigma} < \frac{\sqrt{n}\left(c-\mu_0\right)}{\sigma} + \frac{\sqrt{n}\left(\mu_0-\mu_1\right)}{\sigma}\right] \\
&= P\left[\frac{\sqrt{n}\left(\bar{X}-\mu_1\right)}{\sigma} < z_{1-\alpha} + \frac{\sqrt{n}\left(\mu_0-\mu_1\right)}{\sigma}\right].
\end{aligned}$$

Sehingga diperoleh hubungan

$$z_\beta = -z_{1-\beta} = z_{1-\alpha} + \frac{\sqrt{n}\left(\mu_0-\mu_1\right)}{\sigma}.$$

Dari persamaan terakhir kita peroleh

$$n = \frac{\left(z_{1-\alpha}+z_{1-\beta}\right)^2 \sigma^2}{\left(\mu_0-\mu_1\right)^2}.$$

Adakalanya taraf signifikan diperoleh dari fungsi kuasa suatu uji. Berikut ini akan diberikan definisi fungsi kuasa suatu uji.

Definisi 8.2.4 Untuk suatu nilai parameter tertentu θ, ***fungsi kuasa***, $\pi(\theta)$, dari suatu uji H_0 adalah peluang menolak H_0 jika parameter θ benar.

Untuk hipotesis sederhana $H_0 : \theta = \theta_0$ lawan $H_1 : \theta = \theta_1$ taraf signifikannya adalah $\alpha = \pi(\theta_0)$ dan $\pi(\theta_1) = 1-\beta$. Sedangkan untuk hipotesis komposit $H_0 : \theta \in \Omega_0$ lawan $H_1 : \theta \in \Omega - \Omega_0$ taraf signifikannya ialah $\alpha = \max_{\theta \in \Omega_0} \pi(\theta)$ dan jika $\theta \in \Omega - \Omega_0$, maka $\pi(\theta) = 1 - \beta$.

Kembali ke permasalahan peubah acak X yang bersebaran normal, $X \sim N(\mu, \sigma^2)$, di mana σ^2 diketahui. Kita akan menguji hipotesis nol sederhana lawan hipotesis alternatif komposit, yaitu $H_0 : \mu = \mu_0$ lawan $H_1 : \mu > \mu_0$. Dalam pembahasan sebelumnya kita telah memperoleh taraf signifikan α, yang menolak H_0, jika

$$z_0 = \frac{\sqrt{n}\left(\bar{x} - \mu_0\right)}{\sigma} \geq z_{1-\alpha},$$

atau

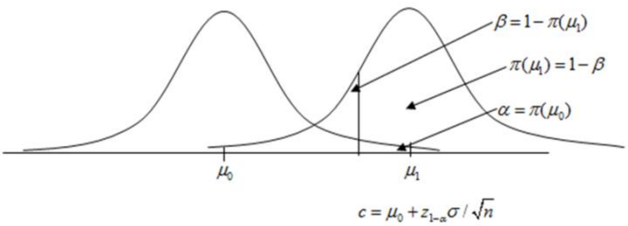
$$C = \left\{ \bar{x} \geq \mu_0 + \frac{\sigma}{\sqrt{n}} z_{1-\alpha} \right\}.$$

Fungsi kuasa untuk sebarang μ adalah

$$\pi(\mu) = P\left[\bar{X} \geq \mu_0 + \frac{\sigma}{\sqrt{n}} z_{1-\alpha} \mid \mu \right] = P\left[\frac{\left(\bar{X} - \mu\right)}{\sigma / \sqrt{n}} \geq \frac{\mu_0 + \frac{\sigma}{\sqrt{n}} z_{1-\alpha} - \mu}{\sigma / \sqrt{n}} \right]$$

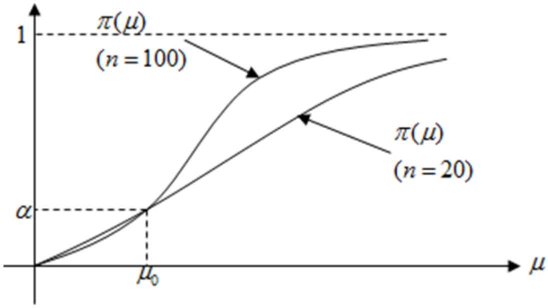
$$= P\left[Z \geq z_{1-\alpha} + \frac{\sqrt{n}\left(\mu_0 - \mu\right)}{\sigma} \right] = 1 - \Phi\left(z_{1-\alpha} + \frac{\sqrt{n}\left(\mu_0 - \mu\right)}{\sigma} \right)$$

Sehingga untuk $\mu = \mu_0$ kita peroleh $\pi(\mu_0) = \alpha$ dan untuk sebarang $\mu > \mu_0$ kita peroleh $\pi(\mu) = 1 - \beta$. Gambar 8.2.1 memberikan ilustrasi tentang hubungan α, β dan $\pi(\mu_1)$.



Gambar 8.2.1

Selanjutnya kita lihat pengujian hipotesis nol komposit lawan hipotesis alternatif komposit, yaitu $H_0 : \mu \leq \mu_0$ lawan $H_1 : \mu > \mu_0$. Dalam hal ini taraf signifikan $\alpha = \max_{\mu \leq \mu_0} \pi(\mu)$. Karena peluang menolak H_0 untuk sebarang $\mu \leq \mu_0$ adalah $\pi(\mu)$, maka dari Gambar 8.2.1 dengan memandang μ_1 sebagai μ_0 dan μ_0 sebagai μ, diperoleh $\pi(\mu) \leq \pi(\mu_0)$. Oleh karena itu kita peroleh $\alpha = \pi(\mu_0)$. Secara umum gambar fungsi kuasa untuk $n = 20$ dan $n = 100$ diberikan oleh gambar berikut ini.



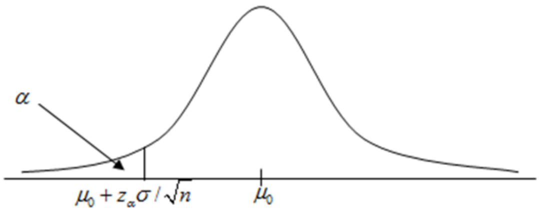
Gambar 8.2.2

Untuk pengujian $H_0 : \mu \geq \mu_0$ lawan $H_1 : \mu < \mu_0$, diperoleh hasil yang serupa, yaitu menolak H_0 jika

$$z_0 \leq z_\alpha .$$

Catatan: tidak semua tabel sebaran normal menyediakan nilai z_α. Jika demikian yang terjadi maka untuk menentukannya kita gunakan sifat $z_\alpha = -z_{1-\alpha}$.

Dikatakan bahwa daerah kritis ukuran α ini terletak pada ekor kiri dari kurva sebaran normal. Uji hipotesis semacam ini disebut juga sebagai uji hipotesis satu arah. Lebih jelasnya seperti terlihat pada gambar berikut ini.

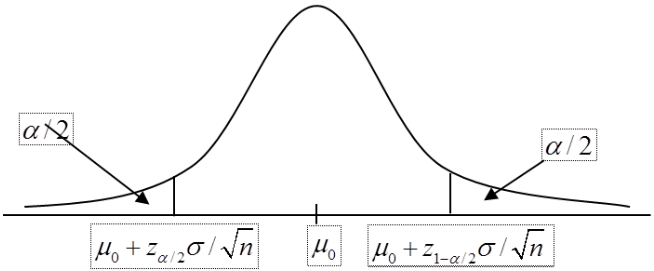


Gambar 8.2.3

Di pihak lain untuk pengujian $H_0 : \mu = \mu_0$ lawan $H_1 : \mu \neq \mu_0$, diperoleh hasil menolak H_0 jika

$$z_0 \leq z_{\alpha/2} \text{ atau } z_0 \geq z_{1-\alpha/2}.$$

Dikatakan bahwa daerah kritis ukuran α ini terletak pada dua ekor kiri dan kanan pada kurva sebaran normal. Uji hipotesis semacam ini disebut juga sebagai uji hipotesis dua arah. Lebih jelasnya seperti terlihat pada gambar berikut ini.



Gambar 8.2.4

Oleh karena itu kita tidak menolak H_0 jika nilai dari μ_0 memenuhi

$$\bar{x} + z_{\alpha/2} \sigma / \sqrt{n} < \mu_0 < \bar{x} + z_{1-\alpha/2} \sigma / \sqrt{n},$$

sedangkan untuk yang lain H_0 ditolak.

Fungsi kuasa dari uji hipotesis dua arah ini diperoleh dari

$$\begin{aligned}\pi(\mu) &= 1 - P\left[z_{\alpha/2} < Z_0 < z_{1-\alpha/2} \mid \mu\right] \\ &= 1 - \left[P\left[Z_0 < z_{1-\alpha/2} \mid \mu\right] - P\left[Z_0 < z_{\alpha/2} \mid \mu\right]\right] \\ &= 1 - P\left[Z_0 < z_{1-\alpha/2} \mid \mu\right] + P\left[< Z_0 < z_{\alpha/2} \mid \mu\right] \\ &= 1 - P\left[\frac{\sqrt{n}\left(\bar{X}-\mu_0\right)}{\sigma} < z_{1-\alpha/2} \mid \mu\right] + P\left[\frac{\sqrt{n}\left(\bar{X}-\mu_0\right)}{\sigma} < z_{\alpha/2} \mid \mu\right] \\ &= 1 - P\left[\frac{\sqrt{n}\left(\bar{X}-\mu\right)}{\sigma} < z_{1-\alpha/2} + \frac{\sqrt{n}\left(\mu_0-\mu\right)}{\sigma}\right] + P\left[\frac{\sqrt{n}\left(\bar{X}-\mu\right)}{\sigma} < z_{\alpha/2} + \frac{\sqrt{n}\left(\mu_0-\mu\right)}{\sigma}\right] \\ &= 1 - \Phi\left[z_{1-\alpha/2} + \frac{\sqrt{n}\left(\mu_0-\mu\right)}{\sigma}\right] + \Phi\left[z_{\alpha/2} + \frac{\sqrt{n}\left(\mu_0-\mu\right)}{\sigma}\right].\end{aligned}$$

Hasil-hasil di atas dapat kita rangkum pada teorema berikut ini.

Teorema 8.2.1 Misal $x_1, x_2, ..., x_n$ adalah nilai pengamatan dari sampel acak $X_1, X_2, ..., X_n$ yang bersebaran normal, $X_i \sim N(\mu, \sigma^2)$, di mana σ^2 diketahui, dan misal

$$z_0 = \frac{\sqrt{n}(\bar{x}-\mu_0)}{\sigma}.$$

1. Untuk pengujian $H_0 : \mu \leq \mu_0$ lawan $H_1 : \mu > \mu_0$ dengan taraf signifikan α menolak H_0 jika $z_0 \geq z_{1-\alpha}$. Fungsi kuasa dari uji ini adalah
$$\pi(\mu) = 1 - \Phi\left(z_{1-\alpha} + \frac{\sqrt{n}\left(\mu_0-\mu\right)}{\sigma}\right).$$
2. Untuk pengujian $H_0 : \mu \geq \mu_0$ lawan $H_1 : \mu < \mu_0$ dengan taraf signifikan α menolak H_0 jika $z_0 \leq z_{\alpha}$. Fungsi kuasa dari uji ini adalah
$$\pi(\mu) = \Phi\left(z_{\alpha} + \frac{\sqrt{n}\left(\mu_0-\mu\right)}{\sigma}\right).$$
3. Untuk pengujian $H_0 : \mu = \mu_0$ lawan $H_1 : \mu \neq \mu_0$ dengan taraf signifikan α menolak H_0 jika $z_0 \leq z_{\alpha/2}$ atau $z_0 \geq z_{1-\alpha/2}$. Fungsi kuasa dari uji ini adalah

$$\pi(\mu)=1-\Phi\left[z_{1-\alpha/2}+\frac{\sqrt{n}\left(\mu_0-\mu\right)}{\sigma}\right]+\Phi\left[z_{\alpha/2}+\frac{\sqrt{n}\left(\mu_0-\mu\right)}{\sigma}\right].$$

4. Ukuran sampel yang diperlukan untuk menerima taraf signifikan α yang ditentukan, dengan kuasa $1-\beta$ adalah

$$n=\frac{\left(z_{1-\alpha}+z_{1-\beta}\right)^2\sigma^2}{\left(\mu_0-\mu_1\right)^2}$$

untuk uji satu arah, dan

$$n=\frac{\left(z_{1-\alpha/2}+z_{1-\beta}\right)^2\sigma^2}{\left(\mu_0-\mu_1\right)^2}$$

untuk uji dua arah.

Pada prakteknya untuk menentukan ditolak atau tidaknya suatu hipotesis, menggunakan perangkat lunak statistik kita cukup hanya melihat nilai p. Kemudian membandingkan nilai p tersebut dengan taraf signifikan α yang telah kita tentukan. Jika $p\geq\alpha$, maka hipotesis tidak ditolak, jika sebaliknya hipotesis ditolak atau disebut signifikan. ***Nilai*** p ini diperoleh berdasarkan data pengamatan, yang didefinisikan sebagai nilai terkecil ukuran α sehingga H_0 ditolak.

Kembali ke permasalahan tentang nilai Statitika Matematis sebelumnya, yang kita misalkan sebagai X, dan anggap bahwa $X\sim N(60{,}100)$. Berikut ini kita lihat contoh hasil pengolahan data menggunakan Minitab nilai Statistika Matematis semester tahun akademik 2005-2006 kelas G mahasiswa Jurusan Matematika FMIPA Universitas Negeri Malang:

Data

33	63	90	56	58	80	43	55	49	50
81	61	64	47	63	78	61	62	62	63
65	68	70	70	58					

One-Sample Z: C1

Test of mu = 60 vs mu > 60

The assumed sigma = 10

Variable	N	Mean	StDev	SE Mean
C1	25	62.00	12.52	2.00

Variable	95.0% Lower Bound	Z	p
C1	58.71	1.00	0.159

Jika kita tetapkan $\alpha = 0.05,$ maka berdasarkan hasil analisis data di atas kita peroleh kesimpulan bahwa $H_0 : \mu = 60$ tidak ditolak. Mengapa?

Bagaimana menghitung nilai p secara manual? Menghitung nilai p secara manual dapat kita ikuti seperti berikut ini:

$$p = P\left[\bar{X} \geq 62 \mid H_0 : \mu = 60\right]$$

$$== P\left[\frac{\sqrt{n}\left(\bar{X} - \mu\right)}{\sigma} \geq \frac{\sqrt{n}\left(62 - \mu\right)}{\sigma} \mid H_0 : \mu = 60\right]$$

$$= P\left[Z \geq \frac{5\left(62 - 60\right)}{10}\right] = P\left[Z \geq 1\right]$$

$$= 1 - P\left[Z < 1\right] = 1 - 0.8413 = 0.1587$$

Perbedaan nilai p ini terletak pada masalah pembulatan bilangan.

Latihan 8.2

8.2.1 Misal $X_1, X_2, ..., X_n$ sampel acak berukuran 16 dari sebaran normal, $X_i \sim N(\mu, 1)$, dan kita ingin menguji $H_0 : \mu = 20$ pada taraf signifikan $\alpha = 0.05,$ berdasarkan pada purata sampel $\bar{X}$.

(a) Tentukan daerah kritis yang berbentuk $A = \{\bar{x} : -\infty < \bar{x} \leq a\}$ dan $B = \{\bar{x} : b \leq \bar{x} < \infty\}$.

(b) Tentukan peluang galat jenis II, yaitu $\beta,$ untuk masing-masing daerah kritis dalam (a) jika hipotesis alternatifnya $H_1 : \mu = 21.$ Daerah kritis yang manakah yang memberikan hasil yang lebih baik.

(c) Ulangi (b) untuk hipotesis alternatif $H_1 : \mu = 19.$

(d) Berapakah taraf signifikan pengujian ini untuk daerah kritis $A \cup B$?

8.2.2. Misal suatu kotak berisi empat kelereng, sebanyak θ kelereng putih dan $4 - \theta$ kelereng hitam. Untuk pengujian $H_0 : \theta = 2$ lawan $H_1 : \theta \neq 2$ seperti berikut: Pengambilan dua kelereng dengan pengembalian dan menolak H_0 jika kedua kelereng berwarna sama, tidak menolak H_0 untuk yang lainnya.

(a) Hitung peluang galat jenis I.
(b) Hitung peluang galat jenis II, untuk semua hasil yang mungkin.
(c) Ulangi (a) dan (b) untuk pengambilan tanpa pengembalian.

8.2.3. Misal $X_1, X_2, ..., X_n$ sampel acak berukuran 20 dari sebaran normal, $X_i \sim N(\mu, \sigma^2)$, anggap $\bar{x} = 11$ dan $s^2 = 16$.

(a) Ujilah hipotesis $H_0 : \mu \geq 12$ lawan $H_1 : \mu < 12$, jika σ^2 tidak diketahui dan taraf signifikan $\alpha = 0.01$.
(b) Berapakah β jika kenyataannya $\mu = 10.5$?

8.2.4. Misal sekeping uang logam bias dengan peluang muncul muka adalah $p = 0.20$, 0.30, atau 0.80. Misal uang tersebut dilempar berulang-ulang, dan X menyatakan banyak percobaan yang diperlukan sampai mendapatkan muka. Untuk pengujian $H_0 : p = 0.80$ anggap menolak H_0 jika $X > 3$, dan tidak menolak jika yang lainnya.

(a) Berapakah peluang galat jenis I?
(b) Berapakah peluang galat jenis II, untuk masing-masing dua nilai p yang lainnya?

8.2.5. Misal peubah acak X mempunyai pdf $f(x) = \theta x^{\theta-1} I 0 < x < 1)$, di mana $\theta \in \Omega = \{\theta : \theta = 1, 2\}$. Untuk pengujian hipotesis $H_0 : \theta = 1$ lawan $H_0 : \theta = 2$, menggunakan sampel acak X_1, X_2 definisikan daerah kritis $C = \left\{(x_1, x_2) : \frac{3}{4} x_1 \leq x_2\right\}$. Tentukan fungsi kuasa untuk uji ini.

8.2.6. Misal X_1, X_2 sampel acak berukuran 2 dari sebaran dengan pdf $f(x) = \frac{1}{\theta} e^{-\frac{x}{\theta}} I(0 < x < \infty)$. Kita menolak $H_0 : \theta = 2$ dan menerima $H_1 : \theta = 1$ jika nilai

$$\frac{f(x_1;2)\,f(x_2;2)}{f(x_1;1)\,f(x_2;1)} \le \frac{1}{2}.$$

Tentukan α dan β uji ini.

8.2.7. Berikut ini adalah nilai Metode Statistika Semester 2 Tahun 2007 off G mahasiswa Jurusan Matematika FMIPA Universtas Negeri Malang:

85	90	82	89	62	95	70	77	89	86
90	76	70	87	83	66	76	89	78	83

Misal nilai-nilai tersebut dianggap berasal dari sebaran normal, $X_i \sim N(\mu, 100)$.

(a) Tentukan nilai p untuk pengujian $H_0 : \mu \le 65$ lawan $H_1 : \mu > 65$.

(b) Berdasarkan nilai p pada soal a, dan menggunakan taraf signifikan $\alpha = 0.1$, putuskanlah ditolak ataukah tidak ditolak $H_0 : \mu \le 65$.

(c) Ulangi soal (a) dan (b) untuk $H_0 : \mu = 65$ lawan $H_1 : \mu \ne 65$.

8.3 Uji-uji untuk Sebaran Normal dan Sebaran Binomial

Pada subpokok bahasan sebelumnya kita telah membahas pengujian hipotesis untuk purata dari sampel acak bersebaran normal, dalam hal varians σ^2 diketahui. Dalam kenyataannya varians ini jarang sekali diketahui. Oleh karena itu perlu adanya suatu uji yang lainnya. Statistik yang digunakan untuk pengujian purata μ jika varians σ^2 tidak diketahui adalah statistik T_0, di mana

$$T_0 = \frac{\sqrt{n}(\bar{X} - \mu_0)}{S}.$$

Teorema 8.3.1 Misal $x_1, x_2, ..., x_n$ adalah nilai pengamatan sampel acak $X_1, X_2, ..., X_n$ dari sebaran normal, $X_i \sim N(\mu, \sigma^2)$, di mana σ^2 tidak diketahui, dan misal

$$t_0 = \frac{\sqrt{n}(\bar{x} - \mu_0)}{s}.$$

1. Untuk pengujian $H_0 : \mu \le \mu_0$ lawan $H_1 : \mu > \mu_0$ dengan taraf signifikan α tolak H_0 jika $t_0 \ge t_{1-\alpha}(n-1)$.
2. Untuk pengujian $H_0 : \mu \ge \mu_0$ lawan $H_1 : \mu < \mu_0$ dengan taraf signifikan α tolak H_0 jika $t_0 \le t_{\alpha}(n-1)$.
3. Untuk pengujian $H_0 : \mu = \mu_0$ lawan $H_1 : \mu \ne \mu_0$ dengan taraf signifikan α tolak H_0 jika $t_0 \le t_{\alpha/2}(n-1)$ atau $t_0 \ge t_{1-\alpha/2}(n-1)$.

Fungsi kuasa untuk pengujian $H_0 : \mu \le \mu_0$ lawan $H_1 : \mu > \mu_0$ dengan taraf signifikan α adalah

$$\begin{aligned}\pi(\mu) &= P\left[\frac{\sqrt{n}\left(\bar{X}-\mu_0\right)}{S} \ge t_{1-\alpha}(\nu) \mid \mu\right] \\ &= P\left[\frac{\sqrt{n}\left(\bar{X}-\mu+\mu-\mu_0\right)}{S} \ge t_{1-\alpha}(\nu)\right] \\ &= P\left[\frac{Z+\delta}{\sqrt{V/\nu}} \ge t_{1-\alpha}(\nu)\right],\end{aligned}$$

di mana $\nu = n-1$, $\delta = \sqrt{n}(\mu-\mu_0)/\sigma$, dengan Z dan V bebas, $Z \sim N(0,1)$ dan $V = \frac{(n-1)S^2}{\sigma^2} \sim \chi^2(v)$. Seperti halnya pada z_α, nilai dari t_α dapat diperoleh dari sifat $t_\alpha = -t_{1-\alpha}$.

Contoh 8.3.1 Diberikan data nilai Statistika Matematis semester 2 tahun akademik 2006-2007 kelas G mahasiswa Jurusan Matematika FMIPA Universitas Negeri Malang:

68	75.8	77.8	73	70	50	91.8	84.8	74.8
76	61.8	57	59	54.8	74.8	83.8	75	

Purata data tersebut adalah $\bar{x} = 71.07$ dan simpangan bakunya adalah $s = 11.37$. Anggap data tersebut dari sebaran normal, $X_i \sim N(65, \sigma^2)$, di mana σ^2 tidak diketahui. Kita ingin menguji hipotesis $H_0 : \mu \le 65$ lawan $H_1 : \mu > 65$ dengan taraf signifikan $\alpha = 0.05$. Kita hitung nilai t_0 seperti berikut ini

$$t_0 = \frac{\sqrt{n}(\bar{x}-\mu_0)}{s} = \frac{\sqrt{17}(71.07-65)}{11.37} = 2.2011.$$

Dari Minitab $t_{1-\alpha}(n-1) = t_{0.95}(16) = 1.7459.$ Karena $t_0 > t_{1-\alpha}(n-1),$ maka $H_0 : \mu \leq 65$ ditolak atau terima $H_1 : \mu > 65.$

Pengujian hipotesis di atas jika kita lakukan pengolahan menggunakan Minitab, kita peroleh seperti berikut ini:

One-Sample T: C1

Test of mu = 65 vs mu > 65

Variable	N	Mean	StDev	SE Mean
C1	17	71.07	11.37	2.76

Variable	95,0% Lower Bound	T	P
C1	66.26	2.20	0.021

Kita lihat nilai p kurang dari nilai α sehingga hipotesis $H_0 : \mu \leq 65$ ditolak. Hasil ini sama dengan hasil sebelumnya.

Kita telah membahas uji untuk purata, jika varians tidak diketahui. Pembahasan selanjutnya adalah tentang uji untuk varians. Perhatikan hipotesis $H_0 : \sigma^2 = \sigma_0^2$ lawan $H_1 : \sigma^2 > \sigma_0^2.$ Untuk keperluan uji varians ini kita gunakan statistik $V_0,$

$$V_0 = (n-1)S^2 / \sigma_0^2,$$

di mana $V_0 \sim \chi^2(n-1)$ jika H_0 benar.

Teorema 8.3.2 Misal $x_1, x_2, ..., x_n$ adalah nilai pengamatan sampel acak $X_1, X_2, ..., X_n$ dari sebaran normal, $X_i \sim N(\mu, \sigma^2)$, dan misal

$$v_0 = (n-1)s^2 / \sigma_0^2.$$

1. Untuk pengujian $H_0 : \sigma^2 \leq \sigma_0^2$ lawan $H_0 : \sigma^2 > \sigma_0^2$ dengan taraf signifikan α adalah menolak H_0 jika $v_0 \geq \chi^2_{1-\alpha}(n-1).$ Fungsi kuasa dari uji ini adalah

$$\pi(\sigma^2) = 1 - H\left[\left(\sigma_0^2 / \sigma^2\right)\chi^2_{1-\alpha}(n-1); n-1\right].$$

2. Untuk pengujian $H_0 : \sigma^2 \geq \sigma_0^2$ lawan $H_0 : \sigma^2 < \sigma_0^2$ dengan taraf signifikan α adalah menolak H_0 jika $v_0 \leq \chi_\alpha^2(n-1)$. Fungsi kuasa dari uji ini adalah
$$\pi(\sigma^2) = H\left[\left(\sigma_0^2 / \sigma^2\right)\chi_\alpha^2(n-1); n-1\right].$$
3. Untuk pengujian $H_0 : \sigma^2 = \sigma_0^2$ lawan $H_0 : \sigma^2 \neq \sigma_0^2$ dengan taraf signifikan α adalah menolak H_0 jika $v_0 \leq \chi_{\alpha/2}^2(n-1)$ atau $v_0 \geq \chi_{1-\alpha/2}^2(n-1)$.

Bukti

Kita akan menurunkan fungsi kuasa bagian 1, dan meninggalkan yang lain sebagai latihan.
$$\begin{aligned}\pi(\sigma^2) &= P\left[V_0 \geq \chi_{1-\alpha}^2(n-1) \mid \sigma^2\right] \\ &= P\left[(n-1)S^2/\sigma^2 \geq \left(\sigma_0^2/\sigma^2\right)\chi_{1-\alpha}^2(n-1)\right] \\ &= 1 - H\left[\left(\sigma_0^2/\sigma^2\right)\chi_{1-\alpha}^2(n-1); n-1\right].\end{aligned}$$
Secara khusus,
$$\pi(\sigma_0^2) = 1 - H\left[\chi_{1-\alpha}^2(n-1); n-1\right] = 1-(1-\alpha) = \alpha.$$
Karena $\pi(\sigma^2)$ merupakan fungsi naik, maka ukuran dari daerah kritis ini adalah α.

Contoh untuk permasalahan ini kita lihat kembali Contoh 8.3.1 tentang nilai Statistika Matematis 2 sebelumnya, yang kita misalkan sebagai X, $X \sim N(65, \sigma^2)$, σ^2 tidak diketahui. Kita akan menguji $H_0 : \sigma^2 = 100$ lawan $H_1 : \sigma^2 \neq 100$ dengan taraf signifikan $\alpha = 0.05$. Dari hasil sebelumnya diperoleh $s = 11.37$, sehingga $s^2 = 129.2769$. Oleh karena itu
$$v_0 = (n-1)s^2/\sigma_0^2 = (25-1)\times 129.2769/100 = 31.026456.$$
Daerah kritis dari pengujian ini adalah
$$v_0 < \chi_{\alpha/2}^2(n-1) = 12.4012$$
atau
$$v_0 > \chi_{1-\alpha/2}^2(n-1) = 39.3641.$$

Karena nilai hitung tidak jatuh pada daerah kritis, maka $H_0 : \sigma^2 = 100$ tidak ditolak.

Pengujian hipotesis yang telah kita bahas sejauh ini adalah uji hipotesis untuk satu sampel. Pembahasan berikutnya untuk pengujian hipotesis dua sampel. Terdapat dua macam pengujian hipotesis untuk dua sampel ini, yaitu dua sampel yang tidak berpasangan dan dua sampel yang berpasangan. Pembahasan akan diawali untuk dua sampel yang tidak berpasangan.

Kita konstruksi pengujian hipotesis untuk dua varians, yaitu pengujian $H_0 : \dfrac{\sigma_2^2}{\sigma_1^2} = d_0$, berdasarkan statistik F_0, di mana

$$F_0 = \frac{S_1^2}{S_2^2} d_0,$$

dan $F_0 \sim F(n_1, n_2)$, jika H_0 benar.

Pada prakteknya nilai dari d_0 ini diambil 1. Hal ini cukup beralasan karena jika H_0 tidak ditolak berkibat bahwa $\sigma_2^2 = \sigma_1^2$.

Teorema 8.3.3 Misal $x_1, x_2, ..., x_{n_1}$ dan $y_1, y_2, ..., y_{n_2}$ adalah berturut-turut merupakan nilai-nilai pengamatan dari dua sampel acak bebas dari sebaran $X_i \sim N(\mu_1, \sigma_1^2)$ dan $Y_j \sim N(\mu_2, \sigma_2^2)$, dan misal

$$f_0 = \frac{s_1^2}{s_2^2} d_0.$$

1. Untuk pengujian $H_0 : \dfrac{\sigma_2^2}{\sigma_1^2} \le d_0$ lawan $H_1 : \dfrac{\sigma_2^2}{\sigma_1^2} > d_0$ dengan taraf signifikan α, tolak H_0 jika $f_0 \le f_\alpha(n_1 - 1, n_2 - 1)$. Jika nilai $f_\alpha(n_1 - 1, n_2 - 1)$ ini kita cari di tabel sebaran F, maka kita tentukan dari hubungan

$$f_\alpha(n_1 - 1, n_2 - 1) = \frac{1}{f_{1-\alpha}(n_2 - 1, n_1 - 1)}.$$

2. Untuk pengujian $H_0 : \dfrac{\sigma_2^2}{\sigma_1^2} \geq d_0$ lawan $H_1 : \dfrac{\sigma_2^2}{\sigma_1^2} < d_0$ dengan taraf signifikan α, tolak H_0 jika $f_0 \geq f_{1-\alpha}(n_1 - 1, n_2 - 1)$.
3. Untuk pengujian $H_0 : \dfrac{\sigma_2^2}{\sigma_1^2} = d_0$ lawan $H_1 : \dfrac{\sigma_2^2}{\sigma_1^2} \neq d_0$ dengan taraf signifikan α, tolak H_0 jika $f_0 \leq f_{\alpha/2}(n_1 - 1, n_2 - 1)$ atau $f_0 \geq f_{1-\alpha/2}(n_1 - 1, n_2 - 1)$.

Contoh 8.3.2 Diberikan data 1 dan data 2 yang merupakan nilai Statistika Matematis semester 2 kelas G mahasiswa Jurusan Matematika FMIPA Universitas Negeri Malang berturut-turut tahun akademik 2005-2006 dan tahun akademik 2006-2007 seperti berikut ini:

Data 1

33	63	90	56	58	80	43	55	49	50
81	61	64	47	63	78	61	62	62	63
65	68	70	70	58					

Data 2

68	75.8	77.8	73	70	50	91.8	84.8	7.8
76	61.8	57	59	54.8	74.8	83.8	75	

Kita akan menguji $H_0 : \dfrac{\sigma_2^2}{\sigma_1^2} = 1$ lawan $H_1 : \dfrac{\sigma_2^2}{\sigma_1^2} \neq 1$ dengan taraf signifikan $\alpha = 0.1$. Dari hasil perhitungan diperoleh $s_1^2 = 156.7504$ dan $s_2^2 = 129.2769$. Oleh karena itu menggunakan statistik F kita peroleh

$$f_0 = \frac{s_1^2}{s_2^2} = \frac{156.7504}{129.2769} = 1.2125.$$

Daerah kritis dari pengujian ini adalah

$$f_0 < f_{\alpha/2}(n_1 - 1, n_2 - 1) = 0.4789$$

atau

$$f_0 > f_{1-\alpha/2}(n_1 - 1, n_2 - 1) = \ 2.2354.$$

Karena f_0 tidak masuk pada daerah kritis, maka H_0 tidak ditolak. Jadi tidak ada perbedaan varians populasi pertama dengan populasi kedua, atau dikatakan kedua ***populasi homogen***.

Untuk pengujian beda purata dari dua populasi dalam hal kedua varians sama tetapi tidak diketahui, yaitu pengujian $H_0 : \mu_2 - \mu_1 = d_0$ kita gunakan statistik T_0, di mana

$$T_0 = \frac{\bar{Y} - \bar{X} - d_0}{S_p \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}},$$

dengan

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

dan $T_0 \sim t(n_1 + n_2 - 2)$.

Pada prakteknya nilai dari d_0 ini diambil 0. Seperti halnya pada uji varians, jika H_0 tidak ditolak berakibat bahwa $\mu_2 = \mu_1$.

Teorema 8.3.4 Misal $x_1, x_2, ..., x_{n_1}$ dan $y_1, y_2, ..., y_{n_2}$ berturut-turut merupakan nilai-nilai pengamatan dari dua sampel acak bebas dari sebaran σ $X_i \sim N(\mu_1, \sigma_1^2)$ dan $Y_j \sim N(\mu_2, \sigma_2^2)$, di mana $\sigma_1^2 = \sigma_2^2 = \sigma$ tidak diketahui dan misal

$$t_0 = \frac{\bar{y} - \bar{x} - d_0}{s_p \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}.$$

1. Untuk pengujian $H_0 : \mu_2 - \mu_1 \le d_0$ lawan $H_1 : \mu_2 - \mu_1 > d_0$ dengan taraf signifikan α, tolak H_0 jika $t_0 \ge t_{1-\alpha}(n_1 + n_2 - 2)$.
2. Untuk pengujian $H_0 : \mu_2 - \mu_1 \ge d_0$ lawan $H_1 : \mu_2 - \mu_1 < d_0$ dengan taraf signifikan α, tolak H_0 jika $t_0 \le t_{\alpha}(n_1 + n_2 - 2)$.
3. Untuk pengujian $H_0 : \mu_2 - \mu_1 = d_0$ lawan $H_1 : \mu_2 - \mu_1 \ne d_0$ dengan taraf signifikan α, tolak H_0 jika $t_0 \le t_{\alpha/2}(n_1 + n_2 - 2)$ atau $t_0 \ge t_{1-\alpha/2}(n_1 + n_2 - 2)$.

Kembali ke Contoh 8.3.2, misal kita ingin $H_0 : \mu_2 - \mu_1 = 0$ lawan $H_1 : \mu_2 - \mu_1 \neq 0$ dengan taraf signifikan $\alpha = 0.05.$ Dari hasil pengolahan data diperoleh $\bar{x} = 62.00,$ $\bar{y} = 71.07,$ $s_1^2 = 156.7504,$ dan $s_2^2 = 129.2769.$ Sehingga varians kedua sampel ini adalah

$$s_p^2 = \frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2} = \frac{24 \times 156.7504 + 16 \times 129.2769}{25 + 17 - 2} = 145.761.$$

Oleh karena itu

$$t_0 = \frac{\bar{y} - \bar{x}}{s_p \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} = 2.3898.$$

Daerah kritis pengujian ini adalah

$$t_0 < t_{\alpha/2}(n_1 + n_2 - 2) = -2.0244$$

atau

$$t_0 > t_{1-\alpha/2}(n_1 + n_2 - 2) = 2.0244.$$

Karena $t_0 = 2.3898$ terletak pada daerah kritis, maka $H_0 : \mu_2 - \mu_1 = 0$ ditolak. Jadi terdapat perbedaan purata populasi pertama dengan populasi kedua.

Jika kita lakukan dengan pengolahan menggunakan Minitab maka diperoleh seperti berikut ini:

Two-Sample T-Test and CI: C2, C1

Two-sample T for C2 vs C1

	N	Mean	StDev	SE Mean
C2	17	71.1	11.4	2.8
C1	25	62.0	12.5	2.5

Difference = mu C2 - mu C1
Estimate for difference: 9.07
95% CI for difference: (1.40, 16.74)
T-Test of difference = 0 (vs not =):
T-Value = 2.39 P-Value = 0.022 DF = 40
Both use Pooled StDev = 12.1

Terlihat bahwa nilai $p = 0.022$ lebih kecil dari taraf signifikan $\alpha = 0.05.$ Jadi $H_0 : \mu_2 - \mu_1 = 0$ ditolak. Hasil ini sama dengan hasil sebelumnya.

Pembahasan pada subpokok bahasan ini diakhiri dengan pembahasan tentang dua sampel yang berpasangan. Dasar teori yang digunakan pada dua sampel berpasangan ini seperti pada satu sampel, yaitu dengan memandang dua sampel yang berpasangan tersebut sebagai satu sampel. Banyak sekali kejadian yang dapat dipandang sebagai dua sampel yang berpasangan. Sebagai contoh dari suatu sampel terdapat dua nilai, seperti nilai Matematika dan Bahasa Inggris, yang diperoleh dari masing-masing anggota sampelnya. Kita dapat membandingkan purata kedua nilai Matematika dan Bahasa Inggris menggunakan teori sampel yang berpasangan.

Untuk pengujian $H_0 : \mu_2 - \mu_1 = d_0$ yang merupakan beda purata dua sampel yang berpasangan dari sebaran normal kita gunakan statistik T_0, di mana

$$T_0 = \frac{\bar{D} - d_0}{S_D / \sqrt{n}},$$

di mana $T_0 \sim t(n-1)$.

Teorema 8.3.5 Misal pasangan nilai $(x_1, y_1), (x_2, y_2), \ldots, (x_n, y_n)$ adalah pasangan nilai pengamatan dari peubah-peubah acak dari suatu sebaran, $(X_i, Y_i) \sim f(x_i, y_i)$, dan bedanya adalah $D_i = Y_i - X_i$, di mana $D_i \sim N(\mu_{D_i}, \sigma_{D_i}^2)$. Misal $\bar{d}$ dan s_D^2 masing-masing menyatakan purata dan varians sampel berpasangan berdasarkan $d_i = y_i - x_i$, untuk $i = 1, \ldots, n$, dan

$$T = \frac{\bar{D} - d_0}{s_D / \sqrt{n}}.$$

1. Untuk pengujian $H_0 : \mu_2 - \mu_1 \le d_0$ lawan $H_1 : \mu_2 - \mu_1 > d_0$ dengan taraf signifikan α, tolak H_0 jika $t_0 \ge t_{1-\alpha}(n-1)$.
2. Untuk pengujian $H_0 : \mu_2 - \mu_1 \ge d_0$ lawan $H_1 : \mu_2 - \mu_1 < d_0$ dengan taraf signifikan α, tolak H_0 jika $t_0 \le t_{\alpha}(n-1)$.

3. Untuk pengujian $H_0 : \mu_2 - \mu_1 = d_0$ lawan $H_1 : \mu_2 - \mu_1 \neq d_0$ dengan taraf signifikan α, tolak H_0 jika $t_0 \leq t_{\alpha/2}(n-1)$ atau $t_0 \geq t_{1-\alpha/2}(n-1)$.

Contoh 8.3.3 Diberikan data 1 dan data 2 yang masing-masing merupakan nilai tes 1 dan tes 2 matakuliah Statistika Matematis semester 2 kelas G mahasiswa Jurusan Matematika FMIPA Universitas Negeri Malang tahun akademik 2006-2007 seperti berikut ini:

Data 1

75	84	74	85	100	25	99	49	84	100
89	80	35	89	79	84	55			

Data 2

60	100	65	65	65	20	85	80	100	50
65	30	95	55	45	65	95			

Kita akan menguji $H_0 : \mu_2 - \mu_1 = 0$ lawan $H_1 : \mu_2 - \mu_1 \neq 0$ dengan taraf signifikan $\alpha = 0.05$.

Dari hasil pengolahan data menggunakan Minitab diperoleh seperti berikut ini:

Results for: tes 1 tes 2 kelas g 2006. MTW

Paired T-Test and CI: Tes 2, Tes 1

Paired T for Tes 2 - Tes 1

	N	Mean	StDev	SE Mean
Tes 2	17	67.06	23.46	5.69
Tes 1	17	75.65	22.04	5.35
Difference	17	-8.59	31.44	7.63

95% CI for mean difference: (-24.76, 7.58)

T-Test of mean difference = 0 (vs not = 0):

T-Value = -1.13 P-Value = 0.277

Jadi H_0 tidak ditolak. Ini berarti tidak ada perbedaan yang signifikan antara nilai purata tes 1 dan tes 2. Mengapa?

Teknik yang digunakan untuk menentukan selang kepercayaan pada sebaran Binomial dapat dimodifikasi untuk pengujian hipotesis.

Teorema 8.3.5 Misal $X_1, X_2, ..., X_n$ sampel acak dari sebaran Binomial, $X_i \sim BIN(1, p)$.

Misal

$$z_0 = \frac{\sqrt{n}(\bar{p} - p_0)}{\sqrt{p_0(1-p_0)}}.$$

Untuk n yang cukup besar, suatu pendekatan pengujian:

1. $H_0 : p \leq p_0$ lawan $H_1 : p > p_0$ dengan taraf signifikan α, tolak H_0 jika $z_0 \geq z_{1-\alpha}$.
2. $H_0 : p \geq p_0$ lawan $H_1 : p < p_0$ dengan taraf signifikan α, tolak H_0 jika $z_0 \leq z_\alpha$.
3. $H_0 : p = p_0$ lawan $H_1 : p \neq p_0$ dengan taraf signifikan α, tolak H_0 jika $z_0 \leq z_{\alpha/2}$ atau $z_0 \geq z_{1-\alpha\backslash 2}$.

Contoh 8.3.4 Misal pelemparan satu uang logam sebanyak 30 kali, dan misal $x = 12$ menyatakan banyak muka yang muncul pada pelemparan tersebut. Misal p merupakan nilai peluang kejadian muncul muka. Untuk pengujian $H_0 : p \geq 0,5$ lawan $H_1 : p < 0,5$ dengan taraf signifikan $\alpha = 0.05$ diperoleh

$$z_0 = \frac{12 - 30 \times 0.5}{\sqrt{30 \times 0.5(1-0.5)}} = \text{-1.0954}.$$

Daerah kritis untuk pengujian ini adalah $z_0 \leq z_\alpha = z_{\ 0.05} =$ -1.6449. Jadi tidak tolak H_0. Ini berarti uang logam masih seimbang.

Latihan 8.3

8.3.1. Untuk pengujian $H_0 : \sigma^2 \geq \sigma_0^2$ lawan $H_0 : \sigma^2 < \sigma_0^2$ dengan taraf signifikan α adalah menolak H_0 jika $v_0 \leq \chi_\alpha^2(n-1)$. Buktikan bahwa fungsi kuasa dari uji ini adalah

$$\pi(\sigma^2) = H\left[\left(\sigma_0^2 / \sigma_1^2\right)\chi_\alpha^2(n-1); n-1\right].$$

8.3.2. Perhatikan kembali Contoh 8.3.1. Tentukan daerah kritis dan putuskanlah ditolak tidaknya H_0 untuk pengujian hipotesis:

(a) $H_0 : \mu \geq 65$ lawan $H_1 : \mu < 65$ dengan taraf signifikan $\alpha = 0.05$.

(b) $H_0 : \mu = 65$ lawan $H_1 : \mu \neq 65$ dengan taraf signifikan $\alpha = 0.05$.

(c) $H_0 : \mu = 65$ lawan $H_1 : \mu \neq 65$ dengan taraf signifikan $\alpha = 0.1$.

8.3.3. Perhatikan kembali Contoh 8.3.1, dan misal nilai Statistika Matematis sebagai X, dan anggap $X \sim N(65, \sigma^2)$. σ^2 tidak diketahui. Tentukan daerah kritis dan putuskanlah ditolak tidaknya H_0 untuk pengujian hipotesis:

(a) $H_0 : \sigma^2 \leq 100$ lawan $H_1 : \sigma^2 > 100$ dengan taraf signifikan $\alpha = 0.05$.

(b) $H_0 : \sigma^2 \geq 100$ lawan $H_1 : \sigma^2 < 100$ dengan taraf signifikan $\alpha = 0.05$.

(c) $H_0 : \sigma^2 = 100$ lawan $H_1 : \sigma^2 \neq 100$ dengan taraf signifikan $\alpha = 0.1$.

8.3.4. Perhatikan kembali Contoh 8.3.2. Tentukan daerah kritis dan putuskanlah ditolak tidaknya H_0 untuk pengujian hipotesis:

(a) $H_0 : \dfrac{\sigma_2^2}{\sigma_1^2} \leq 1$ lawan $H_1 : \dfrac{\sigma_2^2}{\sigma_1^2} > 1$ dengan taraf signifikan $\alpha = 0.05$.

(b) $H_0 : \dfrac{\sigma_2^2}{\sigma_1^2} \geq 1$ lawan $H_1 : \dfrac{\sigma_2^2}{\sigma_1^2} < 1$ dengan taraf signifikan $\alpha = 0.05$.

(c) $H_0 : \dfrac{\sigma_2^2}{\sigma_1^2} = 1$ lawan $H_1 : \dfrac{\sigma_2^2}{\sigma_1^2} \neq 1$ dengan taraf signifikan $\alpha = 0.2$.

8.3.5. Perhatikan kembali Contoh 8.3.2. Tentukan daerah kritis dan putuskanlah ditolak tidaknya H_0 untuk pengujian hipotesis:

(a) $H_0 : \mu_2 - \mu_1 \leq 0$ lawan $H_1 : \mu_2 - \mu_1 > 0$ dengan taraf signifikan $\alpha = 0.05$.

(b) $H_0 : \mu_2 - \mu_1 \geq 0$ lawan $H_1 : \mu_2 - \mu_1 < 0$ dengan taraf signifikan $\alpha = 0.05$.

(c) $H_0 : \mu_2 - \mu_1 = 0$ lawan $H_1 : \mu_2 - \mu_1 \neq 0$ dengan taraf signifikan $\alpha = 0.1$.

8.3.6. Perhatikan kembali Contoh 8.3.3. Gunakan statistik $T_0 = \dfrac{\bar{d} - d_0}{s_D / \sqrt{n}}$ untuk menentukan daerah kritis dan keputusan ditolak tidaknya H_0 untuk pengujian:

(a) $H_0 : \mu_2 - \mu_1 \leq 0$ lawan $H_1 : \mu_2 - \mu_1 > 0$ dengan taraf signifikan $\alpha = 0.05$.

(b) $H_0 : \mu_2 - \mu_1 \geq 0$ lawan $H_1 : \mu_2 - \mu_1 < 0$ dengan taraf signifikan $\alpha = 0.05$.

(c) $H_0 : \mu_2 - \mu_1 = 0$ lawan $H_1 : \mu_2 - \mu_1 \neq 0$ dengan taraf signifikan $\alpha = 0.1$.

8.3.7. Perhatikan kembali Contoh 8.3.4. Tenentukan daerah kritis dan keputusan ditolak tidaknya H_0 untuk pengujian:

(a) $H_0 : p \leq p_0$ lawan $H_1 : p > p_0$ dengan taraf signifikan $\alpha = 0.05$.

(b) $H_0 : p = p_0$ lawan $H_1 : p \neq p_0$ dengan taraf signifikan $\alpha = 0.05$.

(c) Ulangi (b) untuk taraf signifikan $\alpha = 0.02$.

8.4 Daerah Kritis Terbaik

Dalam subpokok bahasan sebelumnya kita telah menggunakan suatu daerah kritis. Untuk suatu pengujian hipotesis dengan taraf signifikan tertentu dimungkinkan terdapat dua atau lebih daerah kritis. Daerah kritis mana yang lebih baik dengan taraf signifikan tertentu dilihat dari nilai peluang galat jenis kedua. Daerah kritis yang memberikan nilai peluang yang lebih kecil itulah yang lebih baik.

Dalam subpokok bahasan ini kita akan membahas daerah kritis paling baik yang disebut sebagai daerah kritis terbaik. Untuk tujuan tersebut berikut ini akan kita kaji suatu definisi yang memungkinkan kita untuk memilih daerah kritis yang terbaik.

Misal peubah acak X mempunyai pdf $f(x;\theta)$ di mana $\theta \in \Omega$. Kita anggap bahwa $\theta \in \omega_0$ atau $\theta \in \omega_1$ di mana $\omega_0 \cup \omega_1 = \Omega$. Kita beri label hipotesis nol lawan hipotesis alternatif sebagai

$$H_0 : \theta \in \omega_0 \text{ lawan } H_1 : \theta \in \omega_1.$$

Uji H_0 lawan H_1 dengan sampel acak $X_1, X_2, ..., X_n$ dari sebaran dari X. Dalam permasalahan ini kita akan sering menggunakan vector acak $\mathbf{X}^t = (X_1, X_2, ..., X_n)$ untuk menyatakan sampel acak dan $\mathbf{x}^t = (x_1, x_2, ..., x_n)$ untuk menyatakan nilai-nilai dari sampel. Suatu uji H_0 lawan H_1 berdasarkan daerah kritis C yang berkorespondensi dengan kaidah:

Menolak H_0 (menerima H_1), jika $\mathbf{X} \in C$,

Mempertahankan H_0 (menolak H_1), jika $\mathbf{X} \in C^C$.

Taraf signifikannya kita tulis sebagai

$$\alpha = \max_{\theta \in \omega_0} P_\theta(\mathbf{X} \in C),$$

dan fungsi kuasa dari uji ini kita tulis sebagai

$$\pi_C(\theta) = P_\theta(\mathbf{X} \in C),\ \theta \in \omega_1.$$

Kita mulai dengan pengujian hipotesis sederhana H_0 lawan hipotesis alternatif sederhana H_1. Misal $f(x;\theta)$ menyatakan pdf dari peubah acak X di mana $\theta \in \Omega = \{\theta', \theta''\}$. Oleh karena itu $\omega_0 = \{\theta'\}$ dan $\omega_1 = \{\theta''\}$. Kita akan mendefinisikan ***daerah kritis terbaik (best critical region)*** untuk pengujian hipotesis sederhana H_0 lawan hipotesis alternatif sederhana H_1.

Definisi 8.4.1 Misal C menyatakan himpunan bagian dari ruang sampel. Maka kita katakan C adalah daerah kritis terbaik ukuran α hipotesis sederhana $H_0 : \theta = \theta'$ lawan hipotesis alternatif sederhana $H_1 : \theta = \theta''$ jika,

(a) $P_{\theta'}(\mathbf{X} \in C) = \alpha$,

(b) untuk setiap himpunan bagian A dari ruang sampel,

$$P_{\theta'}(\mathbf{X} \in A) = \alpha \Rightarrow P_{\theta''}(\mathbf{X} \in C) \ge P_{\theta''}(\mathbf{X} \in A).$$

Contoh 8.4.1 Misal peubah acak X bersebaran Binomial dengan parameter $n = 5$ dan $p = \theta$. Misal $f(x;\theta)$ menyatakan pdf dari X dan misal $H_0 : \theta = \dfrac{1}{2}$ dan $H_1 : \theta = \dfrac{3}{4}$. Berikut ini adalah tabel dari nilai $f(x; \frac{1}{2})$, $f(x; \frac{3}{4})$, dan $\dfrac{f(x; \frac{1}{2})}{f(x; \frac{3}{4})}$.

TABEL 8.4.1

x	0	1	2	3	4	5
$f(x; \frac{1}{2})$	$\frac{1}{32}$	$\frac{5}{32}$	$\frac{10}{32}$	$\frac{10}{32}$	$\frac{5}{32}$	$\frac{1}{32}$
$f(x; \frac{3}{4})$	$\frac{1}{1024}$	$\frac{15}{1024}$	$\frac{90}{1024}$	$\frac{270}{1024}$	$\frac{405}{1024}$	$\frac{243}{1024}$
$\dfrac{f(x; \frac{1}{2})}{f(x; \frac{3}{4})}$	32	$\frac{32}{3}$	$\frac{32}{9}$	$\frac{32}{27}$	$\frac{32}{81}$	$\frac{32}{243}$

Kita akan menggunakan satu nilai peubah acak X untuk menguji hipotesis $H_0 : \theta = \frac{1}{2}$ lawan $H_1 : \theta = \frac{3}{4}$, dan kita ambil ukuran atau taraf signifikan $\alpha = \frac{1}{32}$. Dengan taraf signifikan ini kita peroleh daerah kritis $A_1 = \{x : x = 0\}$, $A_2 = \{x : x = 5\}$, karena

$$\pi_{\{\theta=1/2\}}(\mathbf{X} \in A_1) = \pi_{\{\theta=1/2\}}(\mathbf{X} \in A_2) = \frac{1}{32}.$$

Di pihak lain

$$\pi_{\{\theta=3/4\}}(\mathbf{X} \in A_1) = \frac{1}{1024}$$

dan

$$\pi_{\{\theta=3/4\}}(\mathbf{X} \in A_1) = \frac{243}{1024}.$$

Menggunakan Definisi 8.4.1 kita peroleh bahwa A_2 daerah kritis terbaik ukuran $\alpha = \frac{1}{32}$. Tentunya Anda dapat memilih ukuran yang lainnya, untuk menentukan masing-masing daerah kritisnya dan menentukan daerah kritis terbaik.

Prosedur dalam menentukan daerah kritis terbaik untuk contoh tersebut dapat pula kita peroleh dari nisbah pdf $\frac{f(x;\frac{1}{2})}{f(x;\frac{3}{4})}$, yaitu dengan melihat nilai nisbah yang paling kecil.

Teorema berikut ini memberikan suatu prosedur untuk mendapatkan daerah kritis terbaik.

Teorema 8.4.1 Neyman-Pearson. Misal $X_1, X_2, ..., X_n$ adalah sampel acak dari sebaran dengan pdf $f(x;\theta)$. Maka fungsi kemungkinan dari $X_1, X_2, ..., X_n$ adalah

$$L(\theta;\mathbf{x}) = \Pi_{i=1}^{n} f(x;\theta), \text{ untuk } \mathbf{x}^t = (x_1, x_2, ..., x_n).$$

Misal θ' dan θ'' nilai tertentu dari θ yang berbeda sehingga $\Omega = \{\theta : \theta = \theta', \theta''\}$, dan misal k bilangan positif. Misal C himpunan bagian dari ruang sampel sehingga:

(a) $\frac{L(\theta';\mathbf{x})}{L(\theta'';\mathbf{x})} \le k$, untuk masing-masing $\mathbf{x} \in C$.

(b) $\frac{L(\theta';\mathbf{x})}{L(\theta'';\mathbf{x})} \ge k$, untuk masing-masing $\mathbf{x} \in C^C$.

(c) $\alpha = P_{H_0}\left[\mathbf{X} \in C\right]$.

Maka C adalah daerah kritis terbaik ukuran α untuk pengujian hipotesis sederhana $H_0 : \theta = \theta'$ lawan hipotesis alternatif sederhana $H_1 : \theta = \theta''$.

Bukti

Kita buktikan untuk sampel acak kontinu. Jika hanya C daerah kritis ukuran α, maka bukti selesai. Misal ada daerah kritis ukuran α yang lain, katakanlah A. Untuk memudahkan penulisan, kita nyatakan $\int \underset{R}{\cdots} \int \int L(\theta; x_1, x_2, ..., x_n) dx_1 dx_2 ... dx_n$ sebagai $\int_R L(\theta)$. Kita akan menunjukkan bahwa

$$\int_C L(\theta'') - \int_A L(\theta'') \ge 0.$$

Perhatikan persamaan himpunan berikut ini

$$C = (C \cap A) \cup (C \cap A^C)$$

dan

$$A = (A \cap C) \cup (A \cap C^C).$$

Maka

$$\int_C L(\theta'') - \int_A L(\theta'') = \int_{C\cap A} L(\theta'') + \int_{C\cap A^C} L(\theta'') - \int_{A\cap C} L(\theta'') - \int_{A\cap C^C} L(\theta'')$$
$$= \int_{C\cap A^C} L(\theta'') - \int_{A\cap C^C} L(\theta''). \tag{8.4.1}$$

Meskipun demikian, oleh hipotesis teorema, $L(\theta'') \geq (1/k)L(\theta')$ pada masing-masing titik dari C, dan oleh karena itu pada masing-masing titik dari $C \cap A^C$. Jadi

$$\int_{C\cap A^C} L(\theta'') \geq \frac{1}{k}\int_{C\cap A^C} L(\theta').$$

Tetapi $L(\theta'') \leq (1/k)L(\theta')$ pada masing-masing titik dari C^C, dan oleh karena itu pada masing-masing titik dari $A \cap C^C$. Jadi

$$\int_{A\cap C^C} L(\theta'') \leq \frac{1}{k}\int_{A\cap C^C} L(\theta').$$

Akibatnya ketaksamaan menjadi

$$\int_{C\cap A^C} L(\theta'') - \int_{A\cap C^C} L(\theta'') \geq \frac{1}{k}\left[\int_{C\cap A^C} L(\theta') - \int_{A\cap C^C} L(\theta')\right],$$

dan dari persamaan (8.4.1) kita peroleh

$$\int_C L(\theta'') - \int_A L(\theta'') \geq \frac{1}{k}\left[\int_{C\cap A^C} L(\theta') - \int_{A\cap C^C} L(\theta')\right]. \tag{8.4.2}$$

Karena

$$\int_{C\cap A^C} L(\theta') - \int_{A\cap C^C} L(\theta') = \int_{C\cap A^C} L(\theta') + \int_{C\cap A} L(\theta') - \int_{A\cap C} L(\theta') - \int_{A\cap C^C} L(\theta')$$
$$= \int_C L(\theta') - \int_A L(\theta') = \alpha - \alpha = 0$$

Jadi

$$\int_C L(\theta'') - \int_A L(\theta'') \geq 0.$$

Untuk peubah acak diskret tinggal mengganti integral dengan jumlah.

Teorema di atas menyatakan bahwa kondisi (a), (b), dan (c) merupakan syarat cukup untuk daerah kritis terbaik C. Meskipun demikian, kondisi tersebut juga merupakan syarat perlu. Kita diskusikan hal ini secara singkat. Anggap ada daerah kritis A yang berukuran α yang tidak memenuhi (a) dan (b) dan adalah sebagai terkuasa (*powerful*) pada $\theta = \theta''$ seperti C, yang memenuhi (a), (b), dan (c). Maka persamaan (8.4.1) akan sama dengan nol, karena kuasa θ'' menggunakan A adalah sama menggunakan C. Sebetulnya dalam kasus kontinu, A dan C merupakan daerah yang sama, yaitu mereka berbeda hanya oleh himpunan yang mempunyai peluang sama dengan nol. Meskipun demikian, dalam kasus diskret, jika $P_{H_0}\left[L(\theta') = kL(\theta'')\right]$ adalah positif, A dan C dapat berbeda, tetapi masing-masing memerlukan kondisi (a), (b), dan (c) untuk menjadi daerah kritis terbaik ukuran α.

Akan kita lihat bahwa suatu uji mempunyai sifat bahwa kuasanya tidak pernah di bawah taraf signifikannya. Di pihak lain, peluang galat menolak H_0 (taraf signifikan) adalah lebih besar dari pada peluang kebenaran menolak H_0 (kuasa uji). Hal ini ditunjukkan oleh akibat berikut ini.

Akibat 8.4.1 Seperti dalam Teorema 8.4.1, misal C merupakan daerah kritis terbaik dari $H_0 : \theta = \theta'$ lawan $H_1 : \theta = \theta''$, dengan taraf signifikan α. Misal $\pi_C(\theta'') = P_{\theta''}\left[\mathbf{X} \in C\right]$ menyatakan kuasa uji ini. Maka $\alpha \le \pi_C(\theta'')$.

Bukti

Dengan tidak mengurangi keumuman bukti anggap data dari percobaan Bernoulli dengan peluang sukses α. Jika percobaan berakhir sukses, kita menolak H_0. Taraf signifikan dari uji ini adalah α. Karena kuasa dari uji adalah peluang menolak H_0 jika H_1 benar, maka kuasa uji ini juga merupakan α. Karena C daerah kritis terbaik berukuran α, maka kuasa uji haruslah lebih besar atau sama dengan α. Jadi $\alpha \le \pi_C(\theta'')$.

Aspek lain dari Teorema 8.4.1 adalah jika kita ambil C sehingga untuk semua $\mathbf{x}$ anggota C memenuhi

$$\frac{L(\theta';\mathbf{x})}{L(\theta'';\mathbf{x})} \le k, \ k > 0,$$

maka C adalah daerah kritis terbaik. Ketaksamaan ini sering dinyatakan dalam bentuk

$$u_1(\mathbf{x};\theta',\theta'') \le c_1,$$

atau

$$u_2(\mathbf{x};\theta',\theta'') \ge c_2,$$

di mana c_1 dan c_2 adalah konstanta. Anggap bahwa pertama $u_1 \le c_1$ berlaku. Karena θ' dan θ'' konstanta, maka $u_1(\mathbf{X};\theta',\theta'')$ adalah statistik, dan jika pdf dari statistik ini dapat ditentukan asalkan H_0 benar, maka taraf signifikan dari uji H_0 lawan H_1 dapat ditentukan dari sebaran ini, yaitu,

$$\alpha = P_{H_0}\left[u_1(\mathbf{X},\theta',\theta'') \le c_1\right].$$

Lebih lanjut, uji berdasarkan statistik ini, kita akan menolak H_0 jika $u_1(\mathbf{x}) \le c_1$.

Contoh 8.4.2 Misal $\mathbf{X}^t = (X_1, X_2, ..., X_n)$ sampel acak dari sebaran eksponensial, $X_i \sim EXP(\theta)$. Kita ingin menguji hipotesis sederhana $H_0 : \theta = \theta'$ lawan hipotesis alternatif $H_1 : \theta = \theta''$ di mana $\theta' < \theta''$ dengan taraf signifikan α. Kita ingat bahwa jika $X \sim EXP(\theta)$, maka

$$f(x;\theta) = \theta^{-1}\exp\{-x/\theta\}\, I(\theta > 0).$$

Oleh karena itu untuk $\mathbf{x} \in C$ kita punya

$$\frac{L(\theta';\mathbf{x})}{L(\theta'';\mathbf{x})} = \frac{(\theta')^{-n}\exp\left(-\sum x_i/\theta'\right)}{(\theta'')^{-n}\exp\left(-\sum x_i/\theta''\right)} \le k.$$

Sehingga

$$\frac{\exp\left(-\sum x_i/\theta'\right)}{\exp\left(-\sum x_i/\theta''\right)} \le k\left(\frac{\theta'}{\theta''}\right)^n$$

$$\exp\left(\sum x_i\left(\frac{1}{\theta''} - \frac{1}{\theta'}\right)\right) \le k\left(\frac{\theta'}{\theta''}\right)^n$$

$$\sum x_i\left(\frac{1}{\theta''} - \frac{1}{\theta'}\right) \le \ln\left(k\left(\frac{\theta'}{\theta''}\right)^n\right).$$

Karena $\theta' < \theta''$, maka $\frac{1}{\theta''} - \frac{1}{\theta'} < 0$. Oleh karena itu

$$\sum x_i \geq \frac{1}{1/\theta'' - 1/\theta'} \ln\left(k\left(\frac{\theta'}{\theta''}\right)^n \right).$$

Jadi

$$C = \left\{(x_1, x_2, ..., x_n) : \sum x_i \geq c\right\},$$

di mana

$$c = \frac{1}{1/\theta'' - 1/\theta'} \ln\left(k\left(\frac{\theta'}{\theta''}\right)^n \right).$$

Karena taraf signifikan yang kita pilih adalah α, maka

$$\alpha = P_{H_0}\left[\sum X_i \geq c\right].$$

Pada bab sebelumnya kita telah mempunyai bentuk sebaran

$$2\sum_{i=1}^{n} \frac{X_i}{\theta} \sim \chi^2\,(2n).$$

Sehingga di bawah hipotesis sederhana $H_0 : \theta = \theta'$,

$$P\left(2\sum X_i / \theta' < 2c/\theta'\right) = 1 - P\left(2\sum X_i / \theta' \geq 2c/\theta'\right) = 1 - \alpha.$$

Jadi

$$2c/\theta' = \chi^2_{1-\alpha}(2n),$$

atau

$$c = \frac{\theta' \chi^2_{1-\alpha}(2n)}{2}.$$

Sebagai contoh jika X menyatakan waktu hidup suatu lampu pijar dalam bulan dan misal $n = 15$, $\theta' = 12$ dan $\alpha = 0.05$, maka

$$c = \frac{12 \times 43.7730}{2} = 262.638.$$

Jadi diperoleh daerah kritis terbaik

$$C = \left\{(x_1, x_2, ..., x_n) : \sum x_i \geq 262.638\right\}.$$

Ini berarti kita akan menolak pernyataan bahwa purata waktu hidup suatu lampu $\theta' = 12$ jika jumlah waktu hidup masing-masing anggota sampel tersebut lebih dari 262.638 bulan, atau purata sampelnya lebih dari 17.5092 bulan.

Misal kita mempunyai sejumlah berhingga parameter yang muncul dalam pdf. Kita akan melihat bahwa hipotesis sederhana H_0 dan H_1 tidak memerlukan parameter untuk hipotesis tentang parameter-parameter dari sebaran. Misal H_0 adalah hipotesis sederhana dengan pdf bersama $g(x_1, x_2, ..., x_n)$, dan H_1 adalah hipotesis alternatif sederhana dengan pdf $h(x_1, x_2, ..., x_n)$, maka C adalah daerah kritis terbaik ukuran α untuk pengujian H_0 lawan H_1 jika, $k > 0$:

1. $\dfrac{g(x_1, x_2, ..., x_n)}{h(x_1, x_2, ..., x_n)} \le k$ untuk $(x_1, x_2, ..., x_n) \in C.$
2. $\dfrac{g(x_1, x_2, ..., x_n)}{h(x_1, x_2, ..., x_n)} \ge k$ untuk $(x_1, x_2, ..., x_n) \in C^C.$
3. $\alpha = P_{H_0}\left[(X_1, X_2, ..., X_n) \in C\right].$

Contoh 8.4.3 Misal $X_1, X_2, ..., X_n$ sampel acak dari sebaran dengan pdf $f(x)$ yang positif hanya pada bilangan bulat positif. Kita ingin menguji hipotesis sederhana

$$H_0 : f(x) = \frac{e^{-1}}{x!} I(x = 0, 1, 2, ...)$$

lawan

$$H_1 : f(x) = \left(\frac{1}{2}\right)^{x+1} I(x = 0, 1, 2, ...).$$

Di sini

$$\frac{g(x_1, x_2, ..., x_n)}{h(x_1, x_2, ..., x_n)} = \frac{e^{-n}/(x_1! x_2! ... x_n!)}{\left(\frac{1}{2}\right)^n \left(\frac{1}{2}\right)^{x_1 + x_2 + ... + x_n}}$$

$$= \frac{\left(2e^{-1}\right)^n 2^{\sum_{i=1}^n x_i}}{\prod_{i=1}^n (x_i!)}.$$

Jika $k > 0$, himpunan titik-titik $(x_1, x_2, ..., x_n)$ sehingga

$$\left(\sum_{i=1}^{n} x_i\right)\ln 2-\ln\left[\prod_{i=1}^{n}(x_i!)\right]\le \ln k-n\ln\left(2e^{-1}\right)=c$$

adalah daerah kritis terbaik C. Dalam hal $k=1$ dan $n=1$. Maka ketaksamaan dapat ditulis sebagai $2^{x_1}/x_1!\le e/2$. Ketaksamaan ini dipenuhi oleh semua titik-titik dalam himpunan

$$C=\{x_1 : x_1=0,3,4,5,...\}.$$

Oleh karena itu kuasa uji jika H_0 benar adalah

$$P_{H_0}(X_1\in C)=1-P_{H_0}(X_1=1,2)\approx 0.448,$$

yang juga merupakan taraf signifikan. Kuasa uji jika H_1 benar adalah

$$P_{H_1}(X_1\in C)=1-P_{H_1}(X_1=1,2)=1-\left(\tfrac{1}{4}+\tfrac{1}{8}\right)=0.625.$$

Latihan 8.4

8.4.1. Misal $X_1,X_2,...,X_n$ sampel acak dari sebaran dengan pdf $f(x;\theta)=\theta x^{\theta-1}I(\theta<x<1)$.

(a) Jika $n=1$, maka tentukan daerah kritis terbaik untuk pengujian $H_0:\theta=1$ lawan $H_1:\theta=2$ dengan taraf signifikan $\alpha=0.1$.

(b) Hitung nilai kuasa uji pada (a) untuk hipotesis alternatif $H_1:\theta=2$.

(c) Turunkan daerah kritis terbaik untuk pengujian hipotesis pada (a) berdasarkan sampel acak berukuran n.

8.4.2. Misal peubah acak X bersebaran Poisson dengan parameter θ.

(a) Turunkan daerah kritis terbaik untuk pengujian $H_0:\theta=\theta_0$ lawan $H_1:\theta=\theta_1$ di mana $\theta_1>\theta_0$ berdasarkan satu nilai pengamatan.

(b) Ulangi (a) berdasarkan n nilai sampel pengamatan.

8.4.3. Misal $X_1,X_2,...,X_n$ sampel acak dari sebaran dengan pdf

$$f(x;\theta)=\frac{1}{\sqrt{2\pi}}\exp\left(-\frac{(x-\theta)^2}{2}\right),\ -\infty<x<\infty.$$

Tunjukkan bahwa daerah kritis untuk pengujian $H_0 : \theta = \theta' = 0$ lawan $H_1 : \theta = \theta'' = 1$ mempunyai bentuk $C = \{(x_1, x_2, ..., x_n) : \bar{x} \geq c\}$. Tentukan c dan kuasa uji jika H_1 benar, untuk $n = 25$ dan taraf signifikan $\alpha = 0.05$.

8.4.4. Misal peubah acak X mempunyai pdf $f(x) = (1/\theta)e^{-x/\theta} I(0 < x < \infty)$. Perhatikan hipotesis sederhana $H_0 : \theta = \theta' = 2$ dan hipotesis alternatif sederhana $H_1 : \theta = \theta'' = 4$. Misal X_1, X_2 sampel acak berukuran 2 dari sebaran ini. Tunjukkan bahwa daerah kkritis terbaik uji H_0 lawan H_1 ditentukan oleh statistik $X_1 + X_2$.

8.4.5. Ulangi Latihan 8.4.2 jika $H_1 : \theta = \theta'' = 6$. Buatlah bentuk umum untuk $H_1 : \theta = \theta'' > 2$.

8.4.6. Misal $X_1, X_2, ..., X_{10}$ adalah sampel acak berukuran 10 dari sebaran normal, $X_i \sim N(0, \sigma^2)$. Tentukan daerah kritis terbaik ukuran $\alpha = 0.05$ untuk pengujian $H_0 : \sigma^2 = 1$ lawan $H_1 : \sigma^2 = 2$. Apakah daerah kritis ukuran $\alpha = 0.05$. untuk pengujian $H_0 : \sigma^2 = 1$ lawan $H_1 : \sigma^2 = 4$? Lawan $H_1 : \sigma^2 > 1$?

8.4.7. Misal $X_1, X_2, ..., X_n$ adalah sampel acak dari sebaran normal, $X_i \sim N(\theta, 100)$. Tunjukkan bahwa

$$C = \{(x_1, x_2, ..., x_n) : c \leq \bar{x}\}$$

adalah daerah kritis terbaik untuk pengujian $H_0 : \theta = 75$ lawan $H_1 : \theta = 78$. Tentukan n dan c sehingga

$$P[(X_1, X_2, ..., X_n) \in C \mid H_0] = P(\bar{X} \geq c \mid H_0) = 0.05$$

dan

$$P[(X_1, X_2, ..., X_n) \in C \mid H_1] = P(\bar{X} \geq c \mid H_1) = 0.90.$$

8.4.8. Misal $X_1, X_2, ..., X_n$ adalah sampel acak dari sebaran dengan pdf $f(x;p) = p^x(1-p)^{1-x}\, I(x=0,1).$ Tunjukkan bahwa

$$C = \left\{ \left(x_1, x_2, ..., x_n \right) : \sum_{i=1}^{n} x_i \le c \right\}$$

adalah daerah kritis terbaik untuk pengujian $H_0 : p = \frac{1}{2}$ lawan $H_1 : p = \frac{1}{3}$. Gunakan CLT untuk menentukan n dan c sehingga $P\left(\sum_{i=1}^{n} X_i \le c \,|\, H_0 \right) = 0.10$ dan $P\left(\sum_{i=1}^{n} X_i \le c \,|\, H_1 \right) = 0.80.$

8.4.9. Misal $X_1, X_2, ..., X_n$ sampel acak dari sebaran tertentu. Perhatikan hipotesis $H_0: X \sim UNIF(0,1)$ lawan $H_1: X \sim EXP(1)$.

(a) Tentukan daerah kritis terbaik untuk pengujian hipotesis di atas.

(b) Tentukan k, jika taraf signifikan yang digunakan $\alpha = 0.05$ dan ukuran sampel $n = 25$.

8.4.10. Misal peubah acak X tipe diskret.

(a) Berdasarkan satu nilai pengamatan X, turunkan daerah kritis terbaik untuk pengujian hipotesis $H_0: X \sim GEO(0.05)$ lawan hipotesis $H_1: X \sim POI(0.95)$ dengan taraf signifikan $\alpha = 0.0975$.

(b) Tentukan kuasa dari uji pada (a) untuk $H_1: X \sim POI(0.95)$.

8.5 Uji Paling Kuasa Seragam

Daerah kritis terbaik yang telah kita bahas di atas untuk pengujian hipotesis sederhana berhadapan dengan hipotesis sederhana. Pembahasan selanjutnya, kita akan menentukan daerah kritis terbaik untuk pengujian hipotesis sederhana berhadapan dengan hipotesis komposit atau hipotesis komposit berhadapan dengan hipotesis komposit. Jika terdapat daerah kritis terbaik tersebut berlaku untuk semua nilai parameter dalam hipotesis yang dimaksud, maka daerah kritis yang demikian disebut sebagai ***daerah kritis paling kuasa seragam (uniformly most powerful critical region).*** Selanjutnya paling kuasa seragam ditulis sebagai UMP.

Definisi 8.5.1 Daerah kritis C adalah daerah kritis UMP ukuran α untuk pengujian hipotesis sederhana H_0 lawan hipotesis alternatif komposit H_1 jika C adalah daerah kritis terbaik ukuran α untuk pengujian hipotesis sederhana H_0 lawan masing-masing hipotesis alternatif sederhana H_1. Suatu uji yang didefinisikan oleh daerah kritis C ini disebut ***uji UMP***, dengan taraf signifikan α, untuk pengujian hipotesis sederhana H_0 lawan hipotesis alternatif komposit H_1.

Prosedur yang digunakan untuk menentukan daerah kritis terbaik, yaitu Teorema Neyman-Pearson dapat digunakan untuk menentukan daerah kritis UMP, dengan memandang terlebih dahulu masing-masing hipotesisnya sebagai hipotesis sederhana. Selanjutnya dari hipotesis sederhana tersebut diberlakukan untuk hipotesis komposit.

Contoh 8.5.1 Kita lihat kembali Contoh 8.4.2 hipotesis untuk pengujian $H_0 : \theta = \theta'$ lawan $H_1 : \theta = \theta''$ di mana $\theta' < \theta''$ dengan taraf signifikan α kita telah memperoleh daerah kritis paling baik

$$C = \left\{(x_1, ..., x_n) : \sum x_i \geq c\right\}.$$

Daerah kritis ini tidak bergantung pada nilai tertentu θ'', tetapi berlaku untuk semua θ'', di mana $\theta' < \theta''$. Oleh karena itu C adalah daerah kritis UMP untuk pengujian hipotesis sederhana $H_0 : \theta = \theta'$ lawan hipotesis alternatif komposit $H_1 : \theta > \theta'$.

Contoh 8.5.2 Misal $X_1, X_2, ..., X_n$ sampel acak dari sebaran normal, $X_i \sim N(0, \theta)$. Kita akan menentukan daerah kritis UMP untuk pengujian hipotesis sederhana $H_0 : \theta = \theta'$ lawan hipotesis komposit $H_1 : \theta > \theta'$. Untuk menentukan daerah kritis UMP ini, bila ada, mula-mula kita pandang H_1 sebagai hipotesis sederhana, yaitu $H_1 : \theta = \theta''$ di mana $\theta'' > \theta'$, kemudian kita terapkan Teorema Neyman-Pearson. Karena $X_1, X_2, ..., X_n$ sampel acak dari sebaran normal, $X_i \sim N(0, \theta)$, maka

$$f(x;\theta)=\frac{1}{\sqrt{2\pi\theta}}\exp\left\{-\frac{x^2}{2\theta}\right\},$$

dan

$$L(\theta;\mathbf{x})=\left(\frac{1}{2\pi\theta}\right)^{n/2}\exp\left\{-\frac{\sum x_i^2}{2\theta}\right\}.$$

Menggunakan Teorema Neyman-Pearson kita akan menolak H_0 jika

$$\frac{L(\theta';\mathbf{x})}{L(\theta'';\mathbf{x})}=\left(\frac{\theta''}{\theta'}\right)^{n/2}\exp\left\{-\left(\frac{\theta''-\theta'}{2\theta'\theta''}\right)\sum x_i^2\right\}\le k.$$

Sehingga

$$\sum x_i^2\ge\frac{2\theta'\theta''}{\theta''-\theta'}\left(\frac{n}{2}\ln\left(\frac{\theta''}{\theta'}\right)-\ln k\right)=k_1.$$

Jadi

$$C=\left\{(x_1,\ldots,x_n):\sum x_i^2\ge k_1\right\}.$$

Karena daerah kritisnya mempunyai bentuk $\sum x_i^2\ge k_1$ yang berlaku untuk semua $\theta''>\theta'$, di mana dengan diberlakukan oleh kendala α yang ditentukan, maka daerah kritis ini merupakan daerah kritis UMP untuk pengujian hipotesis sederhana $H_0:\theta=\theta'$ lawan hipotesis komposit $H_1:\theta>\theta'$.

Contoh berikut ini merupakan suatu contoh bahwa tidak semua pengujian hipotesis mempunyai daerah kritis UMP.

Contoh 8.5.3 Misal $X_1,X_2,\ldots,X_n$ sampel acak dari sebaran normal, $X_i\sim N(\theta,1)$. Kita akan menunjukkan bahwa tidak ada daerah kritis UMP untuk pengujian hipotesis sederhana $H_0:\theta=\theta'$ lawan hipotesis komposit $H_1:\theta\ne\theta'$. Seperti halnya pada contoh sebelumnya mula-mula kita pandang H_1 sebagai hipotesis sederhana, yaitu $H_1:\theta=\theta''$ di mana $\theta''\ne\theta'$, kemudian kita terapkan Teorema Neyman-Pearson. Karena $X_1,X_2,\ldots,X_n$ sampel acak dari sebaran normal, $X_i\sim N(\theta,1)$, maka

$$f(x;\theta) = \frac{1}{\sqrt{2\pi}} \exp\left\{ -\frac{(x-\theta)^2}{2} \right\},$$

dan

$$L(\theta;\mathbf{x}) = \left(\frac{1}{2\pi}\right)^{n/2} \exp\left\{ -\frac{\sum (x_i-\theta)^2}{2} \right\}.$$

Seperti prosedur pada contoh sebelumnya kita akan menolak H_0 jika

$$\exp\left\{ -(\theta''-\theta')\sum x_i + \frac{n}{2}\left((\theta'')^2 - (\theta')^2\right) \right\} \le k,$$

atau ekivalen dengan

$$(\theta''-\theta')\sum x_i \ge \frac{n}{2}\left((\theta'')^2 - (\theta')^2\right) - \ln k.$$

Ketaksamaan terakhir ekivalen dengan

$$\sum x_i \ge \frac{n}{2}(\theta''+\theta') - \frac{\ln k}{(\theta''-\theta')},$$

jika $\theta'' > \theta',$ dan ekivalen dengan

$$\sum x_i \le \frac{n}{2}(\theta''+\theta') - \frac{\ln k}{(\theta''-\theta')},$$

jika $\theta'' < \theta'.$ Oleh karena itu daerah kritis terbaik untuk pengujian $H_0 : \theta = \theta'$ lawan $H_1 : \theta = \theta''$ asalkan $\theta'' > \theta',$ tidak sama dengan daerah kritis terbaik untuk pengujian $H_0 : \theta = \theta'$ lawan $H_1 : \theta = \theta''$ asalkan $\theta'' < \theta'.$ Sebagai contoh, daerah kritis terbaik untuk pengujian $H_0 : \theta = \theta'$ lawan $H_1 : \theta = \theta' + 1,$ bukan merupakan dengan daerah kritis terbaik untuk pengujian $H_0 : \theta = \theta'$ lawan $H_1 : \theta = \theta' - 1.$ Jadi tidak ada daerah kritis UMP untuk pengujian hipotesis sederhana $H_0 : \theta = \theta'$ lawan hipotesis komposit $H_1 : \theta \ne \theta'.$

Misal $X_1, X_2, ..., X_n$ sampel acak dari sebaran dengan pdf $f(x;\theta),$ $\theta \in \Omega.$ Anggap bahwa $Y = u(X_1, X_2, ..., X_n)$ adalah statistik cukup untuk $\theta.$ Menggunakan teorema faktorisasi, pdf bersama dari $X_1, X_2, ..., X_n$ boleh ditulis sebagai

$$L(\theta; x_1, x_2, ..., x_n) = k_1\left[u(x_1, x_2, ..., x_n); \theta\right]k_2(x_1, x_2, ..., x_n),$$

di mana $k_2(x_1, x_2, ..., x_n)$ tidak bergantung pada θ. Konsekuensinya, nisbah

$$\frac{L(\theta'; x_1, x_2, ..., x_n)}{L(\theta''; x_1, x_2, ..., x_n)} = \frac{k_1\left[u(x_1, x_2, ..., x_n); \theta'\right]}{k_1\left[u(x_1, x_2, ..., x_n); \theta''\right]}$$

bergantung pada $x_1, x_2, ..., x_n$ hanya melalui $u(x_1, x_2, ..., x_n)$. Oleh karena itu jika ada statistik cukup $Y = u\left(X_1, X_2, ..., X_n\right)$ untuk θ dan jika uji terbaik atau uji paling kuasa seragam diinginkan, maka tidak diperlukan uji yang berdasarkan pada sebarang statistik lain dari pada statistik cukup.

Dalam contoh-contoh di atas, kita telah memperkenalkan uji paling kuasa seragam. Kita akan membahas pengujian untuk hipotesis komposit H_0 lawan hipotesis komposit H_1 atau pengujian satu sisi dari bentuk

$$H_0 : \theta \leq \theta' \text{ lawan } H_1 : \theta > \theta', \tag{8.5.1}$$

atau

$$H_0 : \theta \geq \theta' \text{ lawan } H_1 : \theta < \theta'.$$

Dalam pembahasan sebelumnya taraf signifikan dari uji untuk hipotesis ini didefinisikan sebagai $\alpha = \max\limits_{\theta \leq \theta'} \pi(\theta)$.

Definisi 8.5.2 Kita katakan bahwa kemungkinan $L(\theta; \mathbf{x})$ mempunyai ***nisbah kemungkinan monoton*** dalam statistik $y = u(\mathbf{x})$, jika untuk $\theta_1 < \theta_2$, nisbah

$$\frac{L(\theta_1, \mathbf{x})}{L(\theta_2, \mathbf{x})} \tag{8.5.2}$$

adalah fungsi monoton dari $y = u(\mathbf{x})$.

Anggap fungsi kemungkinan $L(\theta; \mathbf{x})$ mempunyai nisbah kemungkinan monotan turun dalam statistik $y = u(\mathbf{x})$. Maka nisbah pada persamaan (8.5.2) sama dengan $g(y)$, di mana $g(y)$ adalah fungsi turun. Misal α menyatakan taraf signifikan. Maka kita nyatakan bahwa uji berikut adalah UMP pada taraf α untuk pengujian hipotesis (8.5.1):

$$\text{Tolak } H_0 \text{ jika } Y \geq c_Y, \tag{8.5.3}$$

di mana c_Y ditentukan oleh $\alpha = P_{\theta'}\left[Y \geq c_Y\right].$

Untuk menunjukkan kebenaran pernyataan tersebut, pertama kita pandang hipotesis nol sederhana $H_0': \theta = \theta'.$ Misal $\theta'' > \theta'$ dan misal C daerah kritis paling kuasa untuk θ' lawan $\theta''.$ Oleh Teorema Neyman-Pearson, C didefinisikan sebagai

$$\frac{L(\theta'; \mathbf{X})}{L(\theta''; \mathbf{X})} \leq k, \text{ jika dan hanya jika } \mathbf{X} \in C,$$

di mana k ditentukan oleh $\alpha = P_{\theta'}\left[\mathbf{X} \in C\right].$ Tetapi dengan Definisi 8.5.2, karena $\theta'' > \theta',$ maka

$$\frac{L(\theta'; \mathbf{X})}{L(\theta''; \mathbf{X})} = g(Y) \leq k \Leftrightarrow Y = g^{-1}(k),$$

di mana $g^{-1}(k)$ memenuhi $\alpha = P_{\theta'}\left[Y \geq g^{-1}(k)\right],$ yaitu $c_Y = g^{-1}(k).$ Oleh karena itu uji Neyman-Pearson ekivalen dengan (8.5.3). Lebih lanjut, uji adalah UMP untuk θ' lawan $\theta'' > \theta'$ karena uji hanya bergantung pada $\theta'' > \theta'$ dan $g^{-1}(k)$ ditentukan tunggal di bawah $\theta'.$

Misal $\pi_Y(\theta)$ menyatakan fungsi kuasa dari uji (8.5.3). Untuk yang terakhir, kita perlu menunjukkan bahwa $\alpha = \max\limits_{\theta \leq \theta'} \pi_Y(\theta),$ yaitu jika kita dapat menunjukkan bahwa $\pi_Y(\theta)$ adalah fungsi tidak turun. Untuk melihat ini, misal $\theta_1 < \theta_2.$ Kita catat bahwa karena $\theta_1 < \theta_2,$ uji (8.5.3) adalah uji paling kuasa untuk pengujian θ_1 lawan θ_2 dengan taraf $\pi_Y(\theta_1).$ Dengan Akibat 8.5.1, kuasa dari uji pada θ_2 haruslah tidak di bawah taraf signifikan, yaitu $\pi_Y(\theta_2) \geq \pi_Y(\theta_1).$ Jadi $\pi_Y(\theta)$ adalah fungsi tidak turun.

Contoh 8.5.4 Misal $X_1, X_2, ..., X_n$ sampel acak dari sebaran Bernoulli dengan parameter $p = \theta,$ di mana $0 < \theta < 1.$ Misal $\theta' < \theta''.$ Perhatikan nisbah kemungkinan,

$$\frac{L(\theta', \mathbf{x})}{L(\theta'', \mathbf{x})} = \frac{(\theta')^{\sum_{i=1}^{n} x_i}(1-\theta')^{n-\sum_{i=1}^{n} x_i}}{(\theta'')^{\sum_{i=1}^{n} x_i}(1-\theta'')^{n-\sum_{i=1}^{n} x_i}} = \left[\frac{\theta'(1-\theta'')}{\theta''(1-\theta')}\right]^{\sum_{i=1}^{n} x_i}\left(\frac{1-\theta'}{1-\theta''}\right)^{n}.$$

Karena $\theta'/\theta'' < 1$ dan $(1-\theta'')/(1-\theta') < 1,$ maka $\theta'(1-\theta'')/\theta''(1-\theta') < 1.$ Oleh karena itu nisbah adalah fungsi turun dari $y = \sum_{i=1}^{n} x_i.$ Jadi nisbah kemungkinan monoton dalam $Y = \sum_{i=1}^{n} X_i.$

Kita perhatikan hipotesis

$$H_0 : \theta \le \theta' \text{ lawan } H_1 : \theta > \theta'.$$

Dengan pembahasan di atas, UMP taraf α untuk pengujian H_0 lawan H_1 diberikan oleh

$$\text{Tolak } H_0 \text{ jika } Y = \sum_{i=1}^{n} X_i \ge c,$$

di mana c suatu konstanta sehingga $\alpha = P_{\theta'}\left[Y \ge c\right].$

Latihan 8.5

8.5.1. Misal $X_1, X_2, ..., X_n$ sampel acak dari sebaran uniform, $X \sim UNIF(0, \theta)$. Tentukan daerah kritis UMP untuk pengujian $H_0 : \theta \ge \theta_0$ lawan $H_1 : \theta < \theta_0.$

8.5.2. Misal $X_1, X_2, ..., X_{10}$ sampel acak dari sebaran dengan pdf $f(x;\theta) = \theta^x (1-\theta)^{1-x} I(x = 0,1).$

(a) Tentukan daerah kritis UMP untuk pengujian $H_0 : \theta = 1/4$ lawan $H_1 : \theta < 1/4$ dengan taraf signifikan $\alpha = 0.05.$

(b) Tentukan fungsi kuasa $\pi(\theta)$ pada (a) untuk $0 < \theta \le 1/4.$

8.5.3. Misal X peubah acak dari suatu sebaran dengan pdf $f(x;\theta) = (1/\theta) I(0 < x < \theta).$ Misal $Y_1 \le Y_2 \le Y_3 \le Y_4$ merupakan statistik urutan sampel acak dari sebaran ini dengan ukuran $n = 4.$ Kita akan menolak $H_0 : \theta = 1$ lawan $H_1 : \theta \ne 1,$ jika salah satu $y_4 \le 1/2$ atau $y_4 \ge 1.$ Tentukan fungsi kuasa $\pi(\theta)$ untuk $0 < \theta$ untuk pengujian ini.

8.5.4. Perhatikan sebaran-sebaran $N(\mu_1, 400)$ dan $N(\mu_2, 225)$. Misal $\theta = \mu_1 - \mu_2$. Kita akan menolak $H_0: \theta = 0$ dan menerima $H_1: \theta > 0$ jika dan hanya jika $\bar{x} - \bar{y} \geq c$, di mana $\bar{x}$ dan $\bar{y}$ purata pengamatan dari dua sampel bebas, masing-masing berukuran n, dari dua sebaran di atas. Tentukan n dan c sehingga $\pi(0) = 0.05$ dan $\pi(10) = 0.90$.

8.5.5. Misal $X_1, X_2, ..., X_n$ sampel acak dari sebaran dengan pdf $f(x;\theta) = \theta x^{\theta-1} I(0 < x < 1, \theta > 0)$. Tentukan statistik cukup untuk θ dan tunjukkan bahwa uji UMP dari $H_0: \theta = 6$ lawan $H_1: \theta < 6$ didasarkan pada statistik ini.

8.6 Uji Nisbah Kemungkinan

Dalam Subpokok Bahasan 8.4 kita telah membahas daerah terbaik untuk hipotesis sederhana lawan hipotesis sederhana. Dalam Subpokok Bahasan 8.5 kita perluas teori ini untuk daerah kritis paling kuasa seragam untuk hipotesis satu sisi dan keluarga sebaran yang mempunyai nisbah kemungkinan monoton, dan uji paling kuasa seragam ini tidak selalu ada. Sekarang kita akan membahas permasalahan yang lebih umum. Anggap peubah acak X mempunyai pdf $f(x;\boldsymbol{\theta})$ di mana $\boldsymbol{\theta}$ adalah vektor dari parameter-parameter dalam Ω. Misal $\omega \subset \Omega$ dan perhatikan hipotesis-hipotesis

$$H_0: \boldsymbol{\theta} \in \omega \text{ lawan } H_1: \boldsymbol{\theta} \in \Omega \cap \omega^C.$$

Metode untuk uji ini disebut sebagai ***uji nisbah kemungkinan.***

Contoh 8.6.1 Misal peubah acak $X \sim N(\theta_1, \theta_2)$ dan misal ruang parameter $\Omega = \{(\theta_1, \theta_2): -\infty < \theta_1 < \infty, 0 < \theta_2 < \infty\}$. Kita ingin menguji

$$H_0: \boldsymbol{\theta} \in \omega \text{ lawan } H_1: \boldsymbol{\theta} \in \Omega \cap \omega^C,$$

di mana $\omega = \{(\theta_1, \theta_2): \theta_1 = 0, 0 < \theta_2 < \infty\}$. Misal $X_1, X_2, ..., X_n$ sampel acak berukuran $n > 1$ dari sebaran ini. Maka pdf bersama dari $X_1, X_2, ..., X_n$ untuk masing-masing titik dalam Ω adalah

$$L(\theta_1,\theta_2;x_1,x_2,...,x_n)=\left(\frac{1}{2\pi\theta_2}\right)^{n/2}\exp\left[-\frac{\sum_{i=1}^{n}(x_i-\theta_1)^2}{2\theta_2}\right]=L(\Omega).$$

Untuk masing-masing $(\theta_1,\theta_2)\in\omega,$ pdf bersama dari $X_1,X_2,...,X_n$ adalah

$$L(0,\theta_2;x_1,x_2,...,x_n)=\left(\frac{1}{2\pi\theta_2}\right)^{n/2}\exp\left[-\frac{\sum_{i=1}^{n}x_i^2}{2\theta_2}\right]=L(\omega).$$

Kita akan menentukan maksimum dari $L(\omega)$ dalam $\omega,$ yaitu, maksimum dari $L(\omega)$ terhadap $\theta_2,$ dan maksimum $L(\Omega)$ dalam $\Omega,$ yaitu, maksimum $L(\Omega)$ terhadap θ_1 dan $\theta_2.$ Nisbah dari maksimum ini akan diambil sebagai kriteria untuk uji H_0 lawan $H_1.$ Misal maksimum dari $L(\omega)$ dalam ω dinyatakan sebagai $L(\hat{\omega})$ dan misal maksimum dari $L(\Omega)$ dinyatakan sebagai $L(\hat{\Omega}).$ Maka kriteria untuk uji H_0 lawan H_1 adalah nisbah kemungkinan

$$\lambda(x_1,x_2,...,x_n)=\lambda=\frac{L(\hat{\omega})}{L(\hat{\Omega})}.$$

Karena $L(\omega)$ dan $L(\Omega)$ adalah pdf, maka $\lambda\geq 0,$ dan karena ω himpunan bagian dari $\Omega,$ maka $\lambda\leq 1.$ Maksimum $L(\hat{\omega})$ dari $L(\omega)$ diperoleh dari

$$\frac{d\ln L(\omega)}{d\theta_2}=-\frac{n}{2\theta_2}+\frac{\sum_{i=1}^{n}x_i}{2\theta_2^2}=0.$$

Penyelesaian persamaan di atas adalah

$$\hat{\theta}_2=\frac{\sum_{i=1}^{n}x_i^2}{n}.$$

Oleh karena itu maksimum $L(\omega)$ adalah

$$L(\hat{\omega}) = \left(\frac{1}{2\pi \sum_{i=1}^{n} x_i^2 / n} \right)^{n/2} \exp\left(-\frac{\sum_{i=1}^{n} x_i^2}{2\sum_{i=1}^{n} x_i^2 / n} \right) = \left(\frac{ne^{-1}}{2\pi \sum_{i=1}^{n} x_i^2} \right)^{n/2} .$$

Di pihak lain maksimum $L(\hat{\Omega})$ dari $L(\Omega)$, dari Contoh 7.3.4 diperoleh

$$\hat{\theta}_1 = \overline{x}_n \text{ dan } \hat{\theta}_2 = \frac{\sum_{i=1}^{n} \left(x_i - \overline{x}_n \right)^2}{n}.$$

Oleh karena itu

$$L(\hat{\Omega}) = \left[\frac{1}{2\pi \sum_{i=1}^{n} (x_i - \overline{x})^2 / n} \right]^{n/2} \exp\left[-\frac{\sum_{i=1}^{n} (x_i - \overline{x})^2}{2\sum_{i=1}^{n} (x_i - \overline{x})^2 / n} \right]$$

$$= \left[\frac{ne^{-1}}{2\pi \sum_{i=1}^{n} (x_i - \overline{x})^2} \right]^{n/2} .$$

Jadi

$$\lambda = \left[\frac{\sum_{i=1}^{n} (x_i - \overline{x})^2}{\sum_{i=1}^{n} x_i^2} \right]^{n/2} .$$

Karena

$$\sum_{i=1}^{n} x_i^2 = \sum_{i=1}^{n} (x_i - \overline{x})^2 + n\overline{x}^2,$$

maka

$$\lambda = \frac{1}{\left\{ 1 + \left[n\overline{x}^2 / \sum_{i=1}^{n} (x_i - \overline{x})^2 \right] \right\}^{n/2}}.$$

Jika nilai pengamatan $\bar{x} = 0,$ maka percobaan menuju ke penerimaan $H_0,$ dan $\lambda = 1.$ Di pihak lain, jika $\bar{x}$ dan $n\bar{x}^2 \sum_{i=1}^{n} (x_i - \bar{x})^2$ berbeda dari nol, maka percobaan menuju ke negasi dari $H_0,$ dan $\lambda < 1.$ Jika λ ini digunakan sebagai kriteria, maka secara intuisi daerah kritis untuk pengujian H_0 didefinisikan oleh $0 < \lambda < \lambda_0,$ di mana λ_0 adalah pecahan sejati positif. Jadi kita menolak H_0 jika $\lambda < \lambda_0.$ Suatu uji dengan daerah kritis $\lambda < \lambda_0$ disebut sebagai ***uji nisbah kemungkinan.*** Dalam contoh ini $\lambda < \lambda_0$ jika dan hanya jika

$$\frac{\sqrt{n}\,|\,\bar{x}\,|}{\sqrt{\sum_{i=1}^{n} (x_i - \bar{x})^2 / (n-1)}} \geq \sqrt{(n-1)(\lambda_0^{-2/n} - 1)} = c.$$

Jika H_0 benar maka statistik

$$t(X_1, X_2, ..., X_n) = \frac{\sqrt{n}(\bar{X} - 0)}{\sqrt{\sum_{i=1}^{n} (X_i - \bar{X})^2 / (n-1)}}$$

bersebaran t dengan derajat bebas $(n-1).$ Untuk bilangan bulat positif n dan taraf signifikan α diketahui, menggunakan tabel sebaran t atau dengan perangkat lunak statistik, nilai c dapat ditentukan sehingga

$$\alpha = P\left[|\,t(X_1, X_2, ..., X_n)\,| \geq c; H_0\right].$$

Contoh 8.6.2 Misal peubah-peubah acak X dan Y bebas berturut-turut mempunyai $N(\theta_1, \theta_3)$ dan $N(\theta_2, \theta_3),$ di mana θ_3 tidak diketahui. Maka $\Omega = \left\{\left(\theta_1, \theta_2, \theta_3\right) : -\infty < \theta_1 < \infty, -\infty < \theta_2 < \infty, 0 < \theta_3 < \infty\right\}.$ Misal $X_1, X_2, ..., X_n$ dan $Y_1, Y_2, ..., Y_m$ merupakan sampel-sampel acak dari sebaran-sebaran ini. Hipotesis $H_0 : \boldsymbol{\theta} \in \omega,$ di mana

$$\omega = \left\{\left((\theta_1, \theta_2, \theta_3\right) : -\infty < \theta_1 = \theta_2 < \infty, 0 < \theta_3 < \infty\right\},$$

akan diuji lawan hipotesis alternatif. Di sini $X_1, X_2, ..., X_n,$ $Y_1, Y_2, ..., Y_m$ adalah $n + m > 2$ peubah acak bebas dengan fungsi kemungkinan

$$L(\omega) = \left(\frac{1}{2\pi\theta_3}\right)^{(n+m)/2} \exp\left[-\frac{\sum_{i=1}^{n}(x_i-\theta_1)^2 + \sum_{i=1}^{m}(y_i-\theta_1)^2}{2\theta_3}\right],$$

dan

$$L(\Omega) = \left(\frac{1}{2\pi\theta_3}\right)^{(n+m)/2} \exp\left[-\frac{\sum_{i=1}^{n}(x_i-\theta_1)^2 + \sum_{i=1}^{m}(y_i-\theta_3)^2}{2\theta_3}\right].$$

Jika

$$\frac{\partial \ln L(\omega)}{\partial \theta_1} \text{ dan } \frac{\partial \ln L(\omega)}{\partial \theta_3}$$

kedua-duanya sama dengan nol, maka (Latihan 8.6.1)

$$\sum_{i=1}^{n}(x_i-\theta_1) + \sum_{i=1}^{m}(y_i-\theta_1) = 0,$$

$$-(n+m) + \frac{1}{\theta_3}\left[\sum_{i=1}^{n}(x_i-\theta_1)^2 + \sum_{i=1}^{m}(y_i-\theta_1)^2\right] = 0. \tag{8.6.1}$$

Penyelesaian persamaan (8.6.1) untuk θ_1 dan θ_3 berturut-turut adalah

$$\hat{\theta}_1 = \frac{\sum_{i=1}^{n} x_i + \sum_{i=1}^{m} y_i}{n+m}$$

dan

$$\hat{\theta}_3 = \frac{\sum_{i=1}^{n}(x_i-\hat{\theta}_1)^2 + \sum_{i=1}^{m}(y_i-\hat{\theta}_1)^2}{n+m}.$$

Oleh karena itu maksimum dari $L(\omega)$ adalah

$$L(\hat{\omega}) = \left(\frac{e^{-1}}{2\pi\hat{\theta}_3}\right)^{(n+m)/2}.$$

Secara sama jika

$$\frac{\partial \ln L(\Omega)}{\partial \theta_1}, \quad \frac{\partial \ln L(\Omega)}{\partial \theta_2}, \text{ dan } \frac{\partial \ln L(\Omega)}{\partial \theta_3}$$

ketiga-tiganya sama dengan nol, maka diperoleh persamaan (Latihan 8.6.2)

$$\sum_{i=1}^{n}(x_i-\theta_1)=0,$$

$$\sum_{i=1}^{m}(y_i-\theta_1)=0,$$

(8.6.2)

$$-(n+m)+\frac{1}{\theta_3}\left[\sum_{i=1}^{n}(x_i-\theta_1)^2+\sum_{i=1}^{m}(y_i-\theta_2)^2\right]=0.$$

Penyelesaian untuk $\theta_1, \theta_2,$ dan θ_3 berturut-turut adalah

$$\hat{\theta}_1' = \frac{\sum_{i=1}^{n} x_i}{n},$$

$$\hat{\theta}_2' = \frac{\sum_{i=1}^{m} y_i}{n},$$

dan

$$\hat{\theta}_3' = \frac{\sum_{i=1}^{n}(x_i-\hat{\theta}_1')^2+\sum_{i=1}^{m}(y_i-\hat{\theta}_2')^2}{n+m}.$$

Oleh karena itu maksimum dari $L(\Omega)$ adalah

$$L(\hat{\Omega})=\left(\frac{e^{-1}}{2\pi\hat{\theta}_3}\right)^{(n+m)/2}.$$

Sehingga

$$\lambda(x_1,x_2,...,x_n,y_1,y_2,...,y_m)=\lambda=\frac{L(\hat{\omega})}{L(\hat{\Omega})}=\left(\frac{\hat{\theta}_3'}{\hat{\theta}_3}\right)^{(n+m)/2}.$$

Peubah acak yang didefinisikan oleh $\lambda^{2/(n+m)}$ adalah

$$\frac{\sum_{i=1}^{n}(X_i-\bar{X})^2+\sum_{i=1}^{m}(Y_i-\bar{Y})^2}{\sum_{i=1}^{n}\left\{X_i-\left[\left(n\bar{X}+m\bar{Y}\right)/(n+m)\right]\right\}^2+\sum_{i=1}^{m}\left\{Y_i-\left[\left(n\bar{X}+m\bar{Y}\right)/(n+m)\right]\right\}^2}.$$

Kita tulis

$$\sum_{i=1}^{n}\left(X_i-\frac{n\bar{X}+m\bar{Y}}{n+m}\right)^2=\sum_{i=1}^{n}\left[\left(X_i-\bar{X}\right)+\left(\bar{X}-\frac{n\bar{X}+m\bar{Y}}{n+m}\right)\right]^2$$

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

dan

$$\sum_{i=1}^{m}\left(Y_i-\frac{n\bar{X}+m\bar{Y}}{n+m}\right)^2=\sum_{i=1}^{m}\left[\left(Y_i-\bar{Y}\right)+\left(\bar{Y}-\frac{n\bar{X}+m\bar{Y}}{n+m}\right)\right]^2$$

$$=\sum_{i=1}^{m}(Y_i-\bar{Y})^2+m\left(\bar{Y}-\frac{n\bar{X}+m\bar{Y}}{n+m}\right)^2.$$

Demikian juga

$$n\left(\bar{X}-\frac{n\bar{X}+m\bar{Y}}{n+m}\right)^2=\frac{m^2n}{(n+m)^2}\left(\bar{X}-\bar{Y}\right)^2$$

dan

$$m\left(\bar{Y}-\frac{n\bar{X}+m\bar{Y}}{n+m}\right)^2=\frac{n^2m}{(n+m)^2}\left(\bar{X}-\bar{Y}\right)^2.$$

Oleh karena itu peubah acak yang didefinisikan oleh $\lambda^{2/(n+m)}$ dapat ditulis sebagai

$$\frac{\sum_{i=1}^{n}(X_i-\bar{X})^2+\sum_{i=1}^{m}(Y_i-\bar{Y})^2}{\sum_{i=1}^{n}(X_i-\bar{X})^2+\sum_{i=1}^{m}(Y_i-\bar{Y})^2+\left[nm/(n+m)\right]\left(\bar{X}-\bar{Y}\right)^2}$$

$$=\frac{1}{1+\dfrac{\left[nm/(n+m)\right]\left(\bar{X}-\bar{Y}\right)^2}{\sum_{i=1}^{n}(X_i-\bar{X})^2+\sum_{i=1}^{m}(Y_i-\bar{Y})^2}}.$$

Jika H_0 benar, maka statistik

$$T=\frac{\sqrt{\frac{nm}{n+m}}\left(\bar{X}-\bar{Y}\right)}{\sqrt{\frac{\sum_{i=1}^{n}(X_i-\bar{X})^2+\sum_{i=1}^{m}(Y_i-\bar{Y})^2}{n+m-2}}}$$

bersebaran t dengan derajat bebas $n+m-2$. Jadi peubah acak yang didefinisikan oleh $\lambda^{2/(n+m)}$ adalah

$$\frac{n+m-2}{(n+m-2)+T^2}.$$

Seperti halnya pada contoh sebelumnya, kita menolak H_0 jika $\lambda<\lambda_0<1.$ Taraf signifikan dari uji ini adalah

$$\alpha=P\left[\lambda(X_1,X_2,...,X_n,Y_1,Y_2,...,Y_m)\leq\lambda_0;H_0\right].$$

Meskipun demikian, $\lambda(X_1,X_2,...,X_n,Y_1,Y_2,...,Y_m)\leq\lambda_0$ ekivalen dengan $|T|\geq c,$ sehingga

$$\alpha=P\left(|T|\geq c;H_0\right).$$

Untuk $m,$ $n,$ dan α diketahui nilai c dapat diperoleh dari tabel sebaran $t,$ atau dari perangkat lunak statistik.

Dalam dua contoh di atas kita menentukan uji nisbah kemungkinan berdasarkan statistik yang jika H_0 benar dia bersebaran $t.$

Definisi 8.6.1 Misal peubah acak $W\sim N(\delta,1),$ peubah acak $V\sim\chi^2,$ W dan V bebas. Maka

$$T=\frac{W}{\sqrt{V/r}}$$

dikatakan ***bersebaran*** t ***takterpusat*** dengan derajat bebas r dan parameter takterpusatnya $\delta.$ Jika $\delta=0,$ maka dikatakan T mempunyai ***sebaran terpusat*** $t.$

Kita ingin menguji statistik dalam contoh-contoh di atas berdasarkan Definisi 8.6.1. Dalam Contoh 8.6.1 kita mempunyai

$$t(X_1, X_2, ..., X_n) = \frac{\sqrt{n}\bar{X}}{\sqrt{\sum_{i=1}^{n}(X_i - \bar{X})^2/(n-1)}}$$

$$= \frac{\sqrt{n}\bar{X}/\sigma}{\sqrt{\sum_{i=1}^{n}(X_i - \bar{X})^2/\left[\sigma^2(n-1)\right]}}.$$

Di sini $W_1 = \frac{\sqrt{n}\bar{X}}{\sigma} \sim N(\frac{\sqrt{n}\theta_1}{\sigma}, 1)$ dan $V_1 = \sum_{i=1}^{n}\frac{(X_i - \bar{X})^2}{\sigma^2} \sim \chi^2(n-1)$, W_1 dan V_1 bebas. Jadi, jika $\theta_1 \neq 0$, maka $t(X_1, X_2, ..., X_n)$ bersebaran t takterpusat dengan derajat bebas $n-1$ dan parameter terpusatnya $\delta_1 = \sqrt{n}\theta_1/\sigma$. Dalam Contoh 8.6.2 kita mempunyai

$$T = \frac{W_2}{\sqrt{V_2/(n+m-2)}},$$

di mana

$$W_2 = \sqrt{\frac{nm}{n+m}}\left(\bar{X} - \bar{Y}\right)/\sigma$$

dan

$$V_2 = \left[\sum_{i=1}^{n}(X_i - \bar{X})^2 + \sum_{i=1}^{m}(Y_i - \bar{Y})^2\right]/\sigma^2.$$

Di sini

$$W_2 \sim N\left(\sqrt{\frac{nm}{n+m}}\frac{\theta_1 - \theta_2}{\sigma}, 1\right)$$

dan

$$V_2 \sim \chi^2(n+m-2),$$

W_2 dan V_2 bebas. Jadi, jika $\theta_1 \neq \theta_2$, maka T bersebaran t takterpusat dengan derajat bebas $n+m-2$ dan parameter takterpusatnya $\delta_2 = \sqrt{nm/(n+m)}(\theta_1 - \theta_2)/\sigma$.

Contoh 8.6.3 Misal peubah-peubah acak X dan Y bebas berturut-turut mempunyai $N(\theta_1, \theta_3)$ dan $N(\theta_2, \theta_4)$, di mana θ_3 tidak diketahui. Maka $\Omega = \left\{\left(\theta_1, \theta_2, \theta_3\right) : -\infty < \theta_1, \theta_2 < \infty,\ 0 < \theta_3, \theta_4 < \infty\right\}$. Misal

$X_1, X_2, ..., X_n$ dan $Y_1, Y_2, ..., Y_m$ merupakan sampel-sampel acak dari sebaran-sebaran ini. Hipotesis $H_0 : \boldsymbol{\theta} \in \omega,$ di mana

$$\omega = \{((\theta_1, \theta_2, \theta_3) : -\infty < \theta_1, \theta_2 < \infty, 0 < \theta_3 = \theta_4 < \infty\},$$

akan diuji lawan hipotesis alternatif. Ditinggalkan sebagai latihan untuk menunjukkan bahwa statistik yang didefinisikan oleh $\lambda = L(\hat{\omega}) / L(\hat{\Omega})$ adalah fungsi dari statistik

$$W = \frac{\sum_{i=1}^{n} (X_i - \bar{X})^2 / (n-1)}{\sum_{i=1}^{m} (Y_i - \bar{Y})^2 / (m-1)}.$$

Jika $\theta_3 = \theta_4,$ statistik $W \sim F(n-1, m-1)$. Hipotesis $H_0 : \boldsymbol{\theta} \in \omega,$ ditolak jika $W \le c_1$ atau $W \ge c_2$. Konstanta c_1 dan c_2 dipilih sehingga, jika $\theta_3 = \theta_4,$

$$P(W \le c_1) = P(W \ge c_2) = \alpha / 2,$$

di mana α menyatakan taraf signifikan uji ini.

Latihan 8.6

8.6.1. Selidiki kebenaran persamaan (8.6.1).

8.6.2. Selidiki kebenaran persamaan (8.6.2).

8.6.3. Dalam Contoh 8.6.1 misal $n = 10,$ dan misal nilai pengamatan $\bar{x} = 0.6$ dan $\sum_{i=1}^{n} (x_i - \bar{x})^2 = 3.6.$ Jika uji yang telah diturunkan pada contoh ini digunakan, apakah menolak atau menerima $H_0 : \theta_1 = 0$ pada taraf signifikan 5 persen?

8.6.4. Dalam Contoh 8.6.2 misal $n = m = 8,$ $\bar{x} = 75.2,$ $\bar{y} = 78.6,$ $\sum_{i=1}^{8} (x_i - \bar{x})^2 = 71.2,$ dan $\sum_{i=1}^{8} (y_i - \bar{y})^2 = 54.8.$ Jika uji yang telah diturunkan dalam contoh ini digunakan, apakah menolak atau menerima $H_0:\theta_1=\theta_2$ pada taraf signifikan 5 persen?

8.6.5. Misal $X_1, X_2, ..., X_n$ sampel acak dari sebaran normal $N(\theta, 1)$. Tunjukkan bahwa prinsip nisbah kemungkinan untuk perngujian $H_0:\theta=\theta'$, di mana θ' tertentu, lawan $H_1:\theta\neq\theta'$, menolak H_0 jika $|\bar{x}-\theta'|\geq c$. Apakah ketaksamaan ini merupakan daerah kritis UMP untuk pengujian $H_0:\theta=\theta'$ lawan $H_1:\theta\neq\theta'$?

8.6.6. Misal $X_1, X_2, ..., X_n$ sampel acak dari sebaran normal $N(\theta_1, \theta_2)$. Tunjukkan bahwa prinsip nisbah kemungkinan untuk perngujian $H_0:\theta_2=\theta_2'$, di mana θ_2' tertentu dan θ_1 tidak diketahui, lawan $H_1:\theta_2\neq\theta_2'$, menolak H_0 jika $\sum_{i=1}^{n}(x_i-\bar{x})^2\leq c_1$ atau $\sum_{i=1}^{n}(x_i-\bar{x})^2\geq c_2$, di man $c_1<c_2$.

8.6.7. Misal peubah-peubah acak X dan Y bebas berturut-turut mempunyai $N(\theta_1,\theta_3)$ dan $N(\theta_2,\theta_4)$, di mana θ_3 tidak diketahui.

(a) Tunjukkan bahwa nisbah kemungkinan untuk pengujian hipotesis $H_0:\boldsymbol{\theta}\in\omega$, di mana $\omega=\left\{\left((\theta_1,\theta_2,\theta_3\right):-\infty<\theta_1,\theta_2<\infty, 0<\theta_3=\theta_4<\infty\right\}$, lawan hipotesis alternatif diberikan oleh

$$\frac{\left[\sum_{i=1}^{n}(x_i-\bar{x})^2/n\right]^{n/2}\left[\sum_{i=1}^{m}(y_i-\bar{y})^2/m\right]^{m/2}}{\left\{\left[\sum_{i=1}^{n}(x_i-u)^2+\sum_{i=1}^{m}(y_i-u)^2\right]/(n+m)\right\}^{(n+m)/2}},$$

di mana $u=(n\bar{x}+m\bar{y})/(n+m)$.

(b) Tunjukkan bahwa uji nisbah kemungkinan untuk pengujian $H_0:\theta_3=\theta_4$, di mana θ_1 dan θ_2 tidak diketahui, lawan $H_1:\theta_3\neq\theta_4$, diperoleh berdasarkan peubah acak

$$W = \frac{\sum_{i=1}^{n}(X_i - \bar{X})^2/(n-1)}{\sum_{i=1}^{m}(Y_i - \bar{Y})^2/(m-1)}.$$

DAFTAR PUSTAKA

Anton, H; terjemahan Silaban, Pantur. 1988. *Aljabar Linear Elementer*. Edisi Kelima. Jakarta: Erlangga.

Bain, L.J. dan Engelhardt, M. 1992. *Introduction to Probability and Mathematical Statistics.* Edisi Kedua. Belmont, California: Duxbury Press.

Bartle, R.G. dan Sherbert, D.R. 1982. *Introduction to Real Analysis.* New York: John Wiley

Bhat, B. R. 1981. M*odern Probability Theory an Introductory Text Book.* New Delhi: Wiley Eastern Limited.

Dudewicz, E. D. dan Mishra, S. N.; terjemahan Sembiring, R. K. 1995. *Statistika Matematika Modern.* Bandung: Penerbit ITB.

Fisz, M. 1963. *Probability Theory and Mathematical Statistics.* Edisi ketiga. New York: John Wiley.

Grahramani, S. 2005. *Fundamentals of Probability with Stochastic Processes*. Edisi Ketiga. New Jersey: Pearson Prentice Hall.

Hogg, R. V., McKean J.W., dan Craig, A. T. 2005. *Introduction to Probability and Mathematical Statistics.* Edisi Keenam. New Jersey: Pearson Prentice Hall.

http://en.wikipedia.org/wiki/Countable_set#Definition, diakses 16 Juni 2008.

http://en.wikipedia.org/wiki/Indicator_function, diakses 26 Juni 2008.

Ross, M. S. 1997. *Introduction to Probability Models.* San Diego: Academic Press.

Smith, G. 1985. *Statistical Reasoning.* Boston: Allyn and Bacon.

Susiswo. 2009. *Statistika Matematis*. Malang: UM Press.

Susiswo. 2009. *Teori Peluang*. Malang: UM Press.

Thomas, JR. G. B. 1962. *Calculus and Analytic Geometry.* Edisi Ketiga. Massachusetts: Addison-Wesley.

TABEL SEBARAN POISSON, NORMAL, χ^2, t, DAN F

1. Sebaran Poisson

$$P(X \le x) = \sum_{w=0}^{x} e^{-m} \frac{m^w}{w!}$$

x	$m = E(X)$											
	0.5	1.0	1.5	2.0	3.0	4.0	5.0	6.0	7.0	8.0	9.0	10.0
0	.607	.368	.223	.135	.050	.018	.007	.002	.001	.000	.000	.000
1	.910	.736	.558	.406	.199	.092	.040	.017	.007	.003	.001	.000
2	.986	.920	.809	.677	.423	.238	.125	.062	.030	.014	.006	.003
3	.998	.981	.934	.857	.647	.433	.265	.151	.082	.042	.021	.010
4	1.00	.996	.981	.947	.815	.629	.440	.285	.173	.100	.055	.029
5	1.00	.999	.996	.983	.916	.785	.616	.446	.301	.191	.116	.067
6	1.00	1.00	.999	.995	.966	.889	.762	.606	.450	.313	.207	.130
7	1.00	1.00	1.00	.999	.988	.949	.867	.744	.599	.453	.324	.220
8	1.00	1.00	1.00	1.00	.996	.979	.932	.847	.729	593	.456	.333
9	1.00	1.00	1.00	1.00	.999	.992	.965	.916	.830	.717	.587	.458
10	1.00	1.00	1.00	1.00	1.00	.997	.986	.957	.901	.816	.706	.583
11	1.00	1.00	1.00	1.00	1.00	.999	.995	.980	.947	.888	.803	.697
12	1.00	1.00	1.00	1.00	1.00	1.00	.995	.991	.973	.936	.876	.792
13	1.00	1.00	1.00	1.00	1.00	1.00	.999	.996	.987	.966	.926	.864
14	1.00	1.00	1.00	1.00	1.00	1.00	1.00	.999	.994	.983	.959	.917
15	1.00	1.00	1.00	1.00	1.00	1.00	1.00	.999	.998	.992	.978	.951
16	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	.999	.996	.989	.973
17	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	.998	.995	.986
18	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	.999	.998	.993
19	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	.999	.997
20	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	.998
21	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	.999
22	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00

2. Sebaran Normal

$$P(X \le x) = \Phi(x) = \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}} e^{-w^2/2}\, dw$$

Catatan bahwa peluang pada tabel hanya untuk nilai $x \ge 0$. Untuk memperoleh peluang untuk $x < 0$, gunakan kesamaan $\Phi(-x) = 1 - \Phi(x)$.

x	**0.00**	**0.01**	**0.02**	**0.03**	**0.04**	**0.05**	**0.06**	**0.07**	**0.08**	**0.09**
0.0	.5000	.5040	.5080	.5120	.5160	.5199	.5239	.5279	.5319	.5359
0.1	.5398	.5438	.5478	.5517	.5557	.5596	.5636	.5675	.5714	.5753
0.2	.5793	.5832	.5871	.5910	.5948	.5987	.6026	.6064	.6103	.6141
0.3	.6179	.6217	.6255	.6293	.6331	.6368	.6406	.6443	.6480	.6517
0.4	.6554	.6591	.6628	.6664	.6700	.6736	.6772	.6808	.6844	.6879
0.5	.6915	.6950	.6985	.7019	.7054	.7088	.7123	.7157	.7190	.7224
0.6	.7257	.7291	.7324	.7357	.7389	.7422	.7454	.7486	.7517	.7549
0.7	.7580	.7611	.7642	.7673	.7704	.7734	.7764	.7794	.7823	.7852
0.8	.7881	.7910	.7939	.7967	.7995	.8023	.8051	.8078	.8106	.8133
0.9	.8159	.8186	.8212	.8238	.8264	.8289	.8315	.8340	.8365	.8389
1.0	.8413	.8438	.8461	.8485	.8508	.8531	.8554	.8577	.8599	.8621
1.1	.8643	.8665	.8686	.8708	.8729	.8749	.8770	.8790	.8810	.8830
1.2	.8849	.8869	.8888	.8907	.8925	.8944	.8962	.8980	.8997	.9015
1.3	.9032	.9049	.9066	.9082	.9099	.9115	.9131	.9147	.9162	.9177
1.4	.9192	.9207	.9222	.9236	.9251	.9265	.9279	.9292	.9306	.9319
1.5	.9332	.9345	.9357	.9370	.9382	.9394	.9406	.9418	.9429	.9441
1.6	.9452	.9463	.9474	.9484	.9495	.9505	.9515	.9525	.9535	.9545
1.7	.9554	.9564	.9573	.9582	.9591	.9599	.9608	.9616	.9625	.9633
1.8	.9641	.9649	.9656	.9664	.9671	.9678	.9686	.9693	.9699	.9706
1.9	.9713	.9719	.9726	.9732	.9738	.9744	.9750	.9756	.9761	.9767
2.0	.9772	.9778	.9783	.9788	.9793	.9798	.9803	.9808	.9812	.9817
2.1	.9821	.9826	.9830	.9834	.9838	.9842	.9846	.9850	.9854	.9857
2.2	.9861	.9864	.9868	.9871	.9875	.9878	.9881	.9884	.9887	.9890
2.3	.9893	.9896	.9898	.9901	.9904	.9906	.9909	.9911	.9913	.9916
2.4	.9918	.9920	.9922	.9925	.9927	.9929	.9931	.9932	.9934	.9936
2.5	.9938	.9940	.9941	.9943	.9945	.9946	.9948	.9949	.9951	.9952
2.6	.9953	.9955	.9956	.9957	.9959	.9960	.9961	.9962	.9963	.9964
2.7	.9965	.9966	.9967	.9968	.9969	.9970	.9971	.9972	.9973	.9974
2.8	.9974	.9975	.9976	.9977	.9977	.9978	.9979	.9979	.9980	.9981
2.9	.9981	.9982	.9982	.9983	.9984	.9984	.9985	.9985	.9986	.9986
3.0	.9987	.9987	.9987	.9988	.9988	.9989	.9989	.9989	.9990	.9990
3.1	.9990	.9991	.9991	.9991	.9992	.9992	.9992	.9992	.9993	.9993
3.2	.9993	.9993	.9994	.9994	.9994	.9994	.9994	.9995	.9995	.9995
3.3	.9995	.9995	.9995	.9996	.9996	.9996	.9996	.9996	.9996	.9997
3.4	.9997	.9997	.9997	.9997	.9997	.9997	.9997	.9997	.9997	.9998
3.5	.9998	.9998	.9998	.9998	.9998	.9998	.9998	.9998	.9998	.9998

3. Sebaran Khi Kuadrat

$$P(X \le x) = \int_0^x \frac{1}{\Gamma(r/2)2^{r/2}} w^{r/2-1} e^{-w/2}\, dw$$

r	$P(X \le x)$							
	0.010	**0.025**	**0.050**	**0.100**	**0.900**	**0.950**	**0.975**	**0.990**
1	0.000	0.001	0.455	0.016	2.706	3.841	5.024	6.635
2	0.020	0.051	1.386	0.211	4.605	5.991	7.378	9.210
3	0.115	0.216	2.366	0.584	6.251	7.815	9.348	11.345
4	0.297	0.484	3.357	1.064	7.779	9.488	11.143	13.277
5	0.554	0.831	4.351	1.610	9.236	11.070	12.833	15.086
6	0.872	1.237	5.348	2.204	10.645	12.592	14.449	16.812
7	1.239	1.690	6.346	2.833	12.017	14.067	16.013	18.475
8	1.646	2.180	7.344	3.490	13.362	15.507	17.535	20.090
9	2.088	2.700	8.343	4.168	14.684	16.919	19.023	21.666
10	2.558	3.247	9.342	4.865	15.987	18.307	20.483	23.209
11	3.053	3.816	10.341	5.578	17.275	19.675	21.920	24.725
12	3.571	4.404	11.340	6.304	18.549	21.026	23.337	26.217
13	4.107	5.009	12.340	7.042	19.812	22.362	24.736	27.688
14	4.660	5.629	13.339	7.790	21.064	23.685	26.119	29.141
15	5.229	6.262	14.339	8.547	22.307	24.994	27.488	30.578
16	5.812	6.908	15.338	9.312	23.542	26.296	28.845	32.000
17	6.408	7.564	16.338	10.085	24.769	27.587	30.191	33.409
18	7.015	8.231	17.338	10.865	25.989	28.869	31.526	34.805
19	7.633	8.907	18.338	11.651	27.204	30.144	32.852	36.191
20	8.260	9.591	19.337	12.443	28.412	31.410	34.170	37.566
21	8.897	10.283	20.337	13.240	29.615	32.671	35.479	38.932
22	9.542	10.982	21.337	14.041	30.813	33.924	36.781	40.289
23	10.196	11.689	22.337	14.848	32.007	35.172	38.076	41.638
24	10.856	12.401	23.337	15.659	33.196	36.415	39.364	42.980
25	11.524	13.120	24.337	16.473	34.382	37.652	40.646	44.314
26	12.198	13.844	25.336	17.292	35.563	38.885	41.923	45.642
27	12.879	14.573	26.336	18.114	36.741	40.113	43.195	46.963
28	13.565	15.308	27.336	18.939	37.916	41.337	44.461	48.278
29	14.256	16.047	28.336	19.768	39.087	42.557	45.722	49.588
30	14.953	16.791	29.336	20.599	40.256	43.773	46.979	50.892

4. Sebaran t

$$P(X \le x) = \int_{-\infty}^{x} \frac{\Gamma\left[(r+1)/2\right]}{\sqrt{\pi r}\,\Gamma(r/2)(1+w^2/r)^{(r-1)/2}}\,dw$$

r	$P(X \le x)$					
	0.900	**0.950**	**0.975**	**0.990**	**0.995**	**0.999**
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	1.829	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
∞	1.282	1.645	1.960	2.326	2.576	3.090

5. **Sebaran** F

$$P(X \le x) = \int_0^x \frac{\Gamma\left[(r_1 + r_2)/2\right](r_1/r_2)^{r_1/2} w^{r_1/2-1}}{\Gamma(r_1/2)\Gamma(r_2/2)(1 + r_1 w/r_2)^{(r_1+r_2)/2}} dw$$

$P(X \le x)$	r_1					
	r_1	**1**	**2**	**3**	**4**	**5**
0.950	1	161.450	199.500	215.710	224.580	230.160
0.975	1	647.790	799.500	864.160	899.580	921.850
0.990	1	4052.180	4999.500	5403.350	5624.580	5763.650
0.950	2	18.510	19.000	19.160	19.250	19.300
0.975	2	38.510	39.000	39.170	39.250	39.300
0.990	2	98.500	99.000	99.170	99.250	99.300
0.950	3	10.130	9.550	9.280	9.120	9.010
0.975	3	17.440	16.040	15.440	15.100	14.880
0.990	3	34.120	30.820	29.460	28.710	28.240
0.950	4	7.710	6.940	6.590	6.390	6.260
0.975	4	12.220	10.650	9.980	9.600	9.360
0.990	4	21.200	18.000	16.690	15.980	15.520
0.950	5	6.610	5.790	5.410	5.190	5.050
0.975	5	10.010	8.430	7.760	7.390	7.150
0.990	5	16.260	13.270	12.060	11.390	10.970
0.950	6	5.990	5.140	4.760	4.530	4.390
0.975	6	8.810	7.260	6.600	6.230	5.990
0.990	6	13.750	10.920	9.780	9.150	8.750
0.950	7	5.590	4.740	4.350	4.120	3.970
0.975	7	8.070	6.540	5.890	5.520	5.290
0.990	7	12.250	9.550	8.450	7.850	7.460
0.950	8	5.320	4.460	4.070	3.840	3.690
0.975	8	7.570	6.060	5.420	5.050	4.820
0.990	8	11.260	8.650	7.590	7.010	6.630
0.950	9	5.120	4.260	3.860	3.630	3.480
0.975	9	7.210	5.710	5.080	4.720	4.480
0.990	9	10.560	8.020	6.990	6.420	6.060
0.950	10	4.960	4.100	3.710	3.480	3.330
0.975	10	6.940	5.460	4.830	4.470	4.240
0.990	10	10.040	7.560	6.550	5.990	5.640

$P(X \le x)$	r_1					
	r_1	**6**	**7**	**8**	**9**	**10**
0.950	1	233.990	236.770	238.880	240.540	241.880
0.975	1	921.850	948.220	956.660	963.280	968.630
0.990	1	5763.650	5920.360	5981.070	6022.470	6055.850
0.950	2	19.330	19.350	19.370	19.380	19.400
0.975	2	39.330	39.360	39.370	39.390	39.400
0.990	2	99.330	99.360	99.370	99.390	99.400
0.950	3	8.940	8.890	8.850	8.810	8.790
0.975	3	14.730	14.620	14.540	14.470	14.420
0.990	3	27.910	27.670	27.490	27.350	27.230
0.950	4	6.160	6.090	6.040	6.000	5.960
0.975	4	9.200	9.070	8.980	8.900	8.840
0.990	4	15.210	14.980	14.800	14.660	14.550
0.950	5	4.950	4.880	4.820	4.770	4.740
0.975	5	6.980	6.850	6.760	6.680	6.620
0.990	5	10.670	10.460	10.290	10.160	10.050
0.950	6	4.280	4.210	4.150	4.100	4.060
0.975	6	5.820	5.700	5.600	5.520	5.460
0.990	6	8.470	8.260	8.100	7.980	7.870
0.950	7	3.870	3.790	3.730	3.680	3.640
0.975	7	5.120	4.990	4.900	4.820	4.760
0.990	7	7.190	6.990	6.840	6.720	6.620
0.950	8	3.5080	3.500	3.440	3.390	3.350
0.975	8	4.650	4.530	4.430	4.030	3.960
0.990	8	6.370	6.180	6.030	5.910	5.81
0.950	9	3.370	3.290	3.230	3.180	3.140
0.975	9	4.320	4.200	4.100	4.030	3.960
0.990	9	5.800	5.610	5.470	5.350	5.260
0.950	10	3.220	3.140	3.070	3.020	2.980
0.975	10	4.070	3.950	3.850	3.780	3.720
0.990	10	5.390	5.200	5.060	4.940	4.850

INDEKS

ISBN 9789794959480
9 789794 959480
Anggota IKAPI No. 059/JTI/89