

METODE STATISTIKA

Rina Widyasari, M.Si.

Diterbitkan Oleh:
CV. MANHAJI
Medan 2019

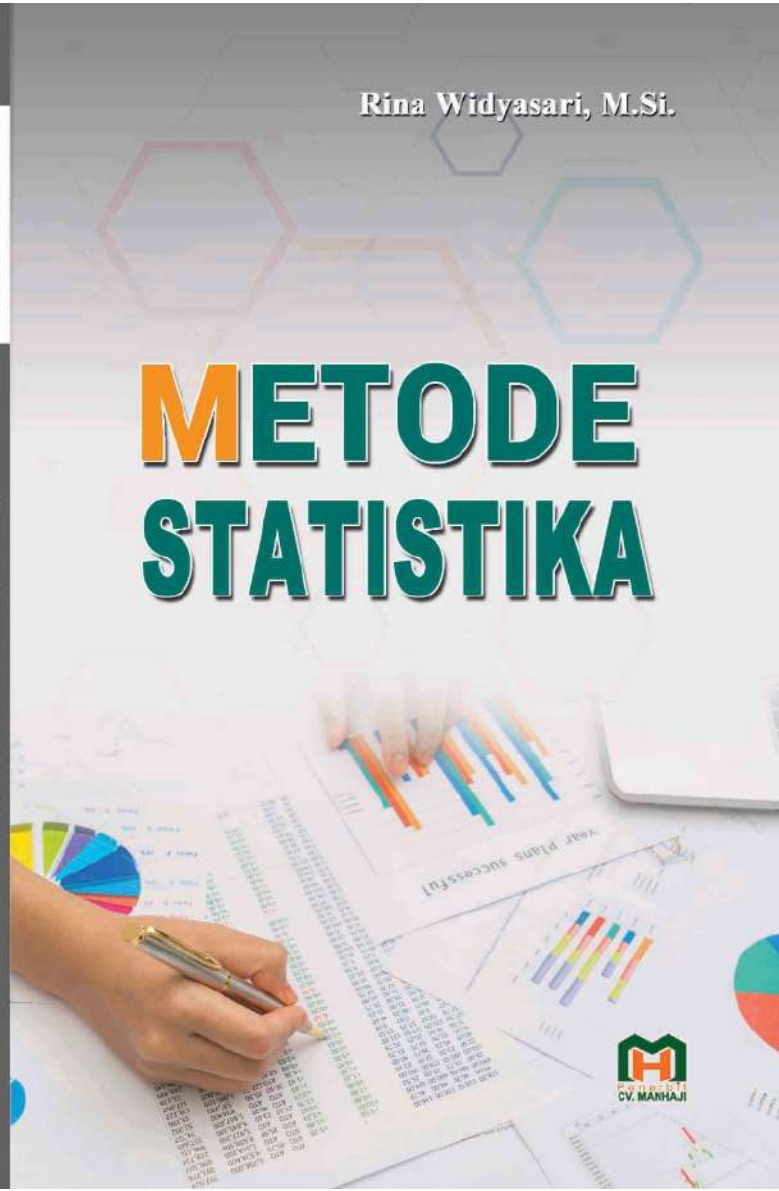


ISBN 978-602-0746-55-5

9 786020 746555

Rina Widyasari, M.Si.

METODE STATISTIKA



Rina Widyasari, M.Si.

METODE STATISTIKA



Rina Widiasari, M.Si.

METODE STATISTIKA

Ismail Husein, M.Si (Ed.)



Diterbitkan Oleh:
CV. MANHAJI Medan
2019

METODE STATISTIKA

Penulis :
Rina Widyasari, M.Si

Editor:

Copyright © 2019
Hak Cipta dilindungi Undang-Undang
All Rights Reserved

Penata Letak :
Johan Iskandar, S.Si.

Perancang Sampul :
Muhammad Hakiki, S.Kom.

Diterbitkan:
CV. Manhaji Medan
e-mail: cvmanhaji@yahoo.com
Cetakan Pertama : November 2019

ISBN: 978-602-0000-00-0

KATA PENGANTAR



Puji Syukur kehadiran Allah SWT atas Rahmat dan Karunia-Nya sehingga dapat menyempurnakan penyelesaian buku yang berjudul "*Metode Statistika*". Shalawat dan salam senantiasa tercurah kepada Muhammad SAW beserta kerabat, sahabat, para pengikutnya sampai akhir zaman, adalah sosok yang telah membawa manusia dan seisi alam dari kegelapan ke cahaya sehingga kita menjadi manusia beriman, berilmu, dan tetap beramal shaleh agar menjadi manusia yang berakhlak mulia.

Penulisan buku ini bertujuan untuk melengkapi persyaratan luaran penelitian. Buku ini juga diharapkan dapat menambah wawasan ilmu pengetahuan, khususnya matematika sains dalam instalasi nilai-nilai Islam yang terpadu dalam proses pembelajaran di lingkungan Universitas Islam Negeri Sumatera Utara.

Dalam penulisan buku ini, saya sangat menyadari bahwa masih banyak kekurangan yang perlu perbaikan di sana sini, sumbangan pemikiran yang membangun sangat penulis harapkan dari rekan-rekan sejawat terutama dari dosen-dosen senior. Juga usulan dari para pengguna buku ini terutama mahasiswa matematika.

Medan, 24 Oktober 2019
Penulis

Rina Widyasari, M.Si

DAFTAR ISI

Kata Pengantar	iii
Daftar Isi	v
BAB 1. KONSEP DASAR STATISTIKA	1
1.1 Istilah-Istilah dalam Statistika.....	1
1.2 Data dan Jenis Data	5
1.3 Populasi dan Sampel	9
1.3.1 Teknik sampling	10
1.3.2 Teknik sampling acak sederhana	11
1.3.3 Teknik sampling sistematis.....	12
1.3.4 Teknik sampling acak berstrata.....	12
1.3.5 Teknik sampling kluster.....	12
1.3.6 Teknik <i>nonprobability sampling</i>	13
BAB 2. DESKRIPSI DATA	15
2.1 Tiga Penyajian Data yang Populer	16
2.2 Ukuran Pemusatan.....	27
2.3 Ukuran Penyebaran	37
2.4 Ukuran Letak.....	44
BAB 3. TEORI PELUANG	53
3.1 Kombinatorik.....	53
3.2 Permutasi dan Kombinasi.....	56
3.3 Kejadian Bersyarat.....	58
BAB 4. DISTRIBUSI PELUANG DISKRIT	77
4.1 Pendahuluan	75
4.2 Variabel Acak	77
4.3 Distribusi Peluang Diskrit.....	84
4.3 Distribusi Poisson.....	89
4.4 Distribusi Multinom	95
4.5 Distribusi Hipergeometri	98
BAB 5. SEBARAN PELUANG KONTINU	105
5.1 Sebaran Normal.....	106

5.2 Sebaran Normal Baku	120
5.3 Sebaran t Student.....	133
5.4 Sebaran Chi-Kuadrat	138
5.5 Sebaran F	142
BAB 6 DISTRIBUSI SAMPLING.....	147
6.1 Sifat-sifat Distribusi Sampling	148
6.2 Jenis-Jenis Distribusi Sampling	152
BAB 7 PENDUGAAN PARAMETER.....	163
7.1 Metode Pendugaan Klasik	163
7.2 Pendugaan Rata-Rata.....	165
7.3 Pendugaan Beda Dua Rata-Rata Populasi.....	170
BAB 8 PENGUJIAN HIPOTESIS.....	183
8.1 Pendahuluan	183
8.2 Hipotesis Nol dan Hipotesis Alternatif	184
8.3 Dua Macam Tipe Kesalahan	186
8.4 Pengujian Hipotesis	187
8.5 Langkah Pengujian Hipotesis.....	190
8.6 Statistik Uji Untuk Pengujian Parameter Populasi	191
DAFTAR PUSTAKA	201

Bab 1

KONSEP-KONSEP DASAR STATISTIKA

1.1 ISTILAH DALAM STATISTIKA

Statistika dan statistik adalah dua kata yang memiliki perbedaan makna. Selama ini, kebanyakan orang beranggapan istilah tersebut memiliki makna yang sama. Hal tersebut terjadi karena ketidakpahaman mengenai statistika secara khusus sehingga perlu dipelajari lebih lanjut.

Dalam kehidupan sehari-hari, tanpa disadari kita telah melakukan kegiatan yang berhubungan dengan statistika, misalnya survey harga pasar, bermain melempar dadu dan koin, melakukan tes golongan darah hanya dengan setetes darah ataupun mencicipi masakan atau minuman, menduga dan menganalisis berita yang ada di televisi, bahkan membuat suatu kesimpulan atas permasalahan yang kita hadapi. Jadi, statistika itu merupakan ilmu yang mempelajari tentang pengumpulan data melalui survey atau eksperimen, penyajian data, pengolahan dan analisis data, menguji hipotesis serta penarikan kesimpulan.

Perbedaan antara statistika dan statistik berhubungan dengan populasi dan sampel. Populasi merupakan sekumpulan data yang memiliki karakteristik yang telah terdefinisi dengan baik. Sedangkan sampel merupakan bagian dari populasi yang memiliki karakteristik yang sama/hamper sama dengan populasi. Dalam populasi, terdapat ukuran data seperti rata-rata, ragam atau simpangan baku, ukuran ini dikatakan parameter sementara pada sampel ukuran ini dinamakan statistik. Untuk lebih jelasnya, simak definisi-definisi berikut.

Definisi 1.1

Populasi adalah kumpulan objek yang diamati. *Sampel* adalah himpunan bagian dari populasi. Jika sampel terdiri dari keseluruhan populasi maka disebut *sensus*.

Definisi 1.2

Pengukuran adalah jumlah atau atribut yang dihitung untuk setiap anggota populasi atau sampel. Pengukuran elemen sampel secara kolektif disebut *data sampel*.

Definisi 1.3

Parameter adalah angka yang merangkum beberapa aspek populasi secara keseluruhan. *Statistik* adalah angka yang dihitung dari data sampel dan merupakan taksiran nilai pada parameter populasi.

Definisi 1.4

Statistika adalah kumpulan metode untuk mengumpulkan, menampilkan, menganalisis, dan menarik kesimpulan dari data.

Definisi 1.5

Statistika deskriptif adalah cabang statistika yang melibatkan pengorganisasian, pemunculan, dan penggambaran data.

Definisi 1.6

Statistika inferensial adalah cabang statistika yang melibatkan pengambilan kesimpulan tentang populasi berdasarkan informasi yang terdapat dalam sampel yang diambil dari populasi tersebut.

Definisi 1.7

Data kualitatif adalah pengukuran yang tidak memiliki skala numerik alami, namun terdiri dari atribut, label, atau karakteristik nonnumerikal lainnya.

Definisi 1.8

Data kuantitatif adalah pengukuran numerik yang timbul dari skala numerik alami.

Definisi 1.9

Proporsi adalah rasio perbandingan antara jumlah suatu kelompok A dalam populasi atau sampel terhadap jumlah populasi atau sampelnya. Misal dalam populasi terdapat kelompok A dan kelompok B, jumlah populasi ada 100 sedangkan kelompok A berjumlah 58, maka proporsi kelompok A dalam populasi tersebut adalah $58/100$ atau 0,58.

Latihan 1.1

1. Jelaskan apa yang dimaksud dengan istilah populasi.
2. Jelaskan apa yang dimaksud dengan istilah sampel.
3. Jelaskan apa itu parameter.
4. Jelaskan apa itu statistik.
5. Jelaskan perbedaan antara statistika deskriptif dan statistika inferensial. Ilustrasikan dengan sebuah contoh.
6. Identifikasi ukuran-ukuran berikut ini sebagai kuantitatif atau kualitatif:
 - a. 30 pembacaan suhu tinggi dalam 30 hari terakhir.
 - b. Skor 40 siswa dalam tes bahasa Inggris.
 - c. Tipe darah 120 guru di sekolah menengah.
 - d. Empat digit terakhir nomor BPJS semua mahasiswa di kelas.
 - e. Angka pada kaus dari 53 pemain sepak bola di tim.
7. Seorang peneliti ingin memperkirakan berat rata-rata bayi baru lahir di Sumatera Utara dalam lima tahun terakhir. Dia mengambil sampel acak dari 235 bayi

yang baru lahir dan memperoleh rata-rata 3,27 kilogram.

- a. Berapakah populasi yang diamati?
 - b. Berapakah parameter yang diamati?
 - c. Berdasarkan sampel ini, apakah kita tahu berat rata-rata bayi baru lahir di Sumatera Utara? Jelaskan secara rinci.
8. Seorang peneliti ingin memperkirakan proporsi semua orang dewasa yang memiliki ponsel. Dia mengambil sampel acak dari 1.572 orang dewasa; 1.298 di antaranya memiliki ponsel, maka $1298/1572 \approx 0,83$ atau sekitar 83% memiliki ponsel.
- a. Berapakah populasi yang diamati?
 - b. Berapakah parameter yang diamati?
 - c. Statistik apa yang terlibat?
 - d. Berdasarkan sampel ini, apakah kita mengetahui proporsi semua orang dewasa yang memiliki ponsel?
9. Seorang sosiolog ingin memperkirakan proporsi semua orang dewasa di wilayah tertentu yang belum pernah menikah. Dalam sampel acak 1.320 orang dewasa, 145 tidak pernah menikah, maka $145/1320 \approx .11$ atau sekitar 11% tidak pernah menikah.
- a. Berapakah populasi yang diamati?
 - b. Berapakah parameter yang diamati?
 - c. Statistik apa yang terlibat?
 - d. Berdasarkan sampel ini, apakah kita mengetahui proporsi semua orang dewasa yang belum pernah menikah?

1.2 DATA DAN JENIS DATA

Sama halnya seperti statistika dan statistik, kebanyakan orang juga tidak paham perbedaan antara datum dan data. Data adalah catatan kumpulan fakta. Data merupakan bentuk jamak dari datum, berasal dari bahasa Latin yang berarti sesuatu yang diberikan. Dalam penggunaan sehari-hari data berarti suatu pernyataan yang diterima secara apa adanya. Pernyataan ini adalah hasil pengukuran atau pengamatan suatu variabel yang bentuknya dapat berupa angka, kata-kata ataupun citra.

Jenis-Jenis Data

1. Menurut sifatnya, data dibagi menjadi data kualitatif dan data kuantitatif, yaitu:
 - a. **Data kualitatif:** data yang tidak berbentuk angka (non-numerik).
Contoh: Kualitas produksi barang ada yang baik dan ada yang buruk.
 - b. **Data kuantitatif:** data yang berbentuk angka (numerik) yang berasal dari hasil pengamatan ataupun pengukuran.
Data numerik juga terbagi menjadi dua yaitu, diskrit dan kontinu. **Data diskrit** mempunyai nilai numerik yang berasal dari proses hitung, sedangkan **data kontinu** mempunyai nilai numerik yang berasal dari proses pengukuran.

Contoh: Panjang meja 1,5 meter, berat badan bayi 2500 gram, jumlah mahasiswa, jumlah pengunjung. Dari contoh, jumlah pengunjung dan jumlah mahasiswa merupakan data diskrit karena berasal dari proses perhitungan sedangkan panjang meja dan berat bayi merupakan data kontinu.

2. Menurut cara memperolehnya, data dibagi menjadi data primer dan data sekunder, yaitu:
 - a. **Data primer:** data yang dikumpulkan dan diolah sendiri oleh suatu organisasi/perorangan langsung dari objeknya (biasanya berdasarkan hasil pengamatan, pengukuran atau survey).
Contoh: Perusahaan telekomunikasi XYZ ingin memperoleh informasi mengenai kartu provider yang sering digunakan mahasiswa untuk internet.
 - b. **Data sekunder:** data yang diperoleh dalam bentuk jadi dan telah diolah oleh pihak lain.
Contoh: Data perekonomian Sumatera Utara tahun 2018 diperoleh dari BPS (Badan Pusat Statistika).
3. Menurut waktu pengumpulannya, data dibagi menjadi data *cross section* dan data *time series* (deret berkala), yaitu:
 - a. **Data *cross section*:** data yang dikumpulkan dalam suatu periode tertentu.
Contoh: Data sensus penduduk Indonesia tahun 2000 dan 2010 berdasarkan umur, jenis kelamin, pendidikan, pekerjaan dan lain-lain.
 - b. **Data *time series*** (deret berkala): data yang dikumpulkan dari waktu ke waktu.
Contoh: Data kurs Rupiah terhadap Dollar AS dari tanggal 1 Oktober 2018 s.d. 5 Oktober 2018.

Skala Pengukuran Data

Nilai suatu data dapat diklasifikasikan berdasarkan tingkat pengukurannya atau skala pengukurannya. Statistikawan menggunakan skala nominal dan ordinal untuk mendeskripsikan nilai data kategorik (kualitatif), dan menggunakan skala interval dan rasio untuk mendeskripsikan nilai data numerik (kuantitatif).

1. Skala Nominal

Nilai pada data berskala nominal hanya mengklasifikasi data ke dalam kategorik yang berbeda, dan tidak berarti urutan ataupun ranking.

Contoh:

Data Pribadi

Jenis Kelamin: 1. Laki-laki 2. Perempuan

Warna Rambut: 1. Hitam 2. Putih 3. Pirang

Angka 1 dan 2 hanya untuk klasifikasi ke dalam kategori, bukan berarti data berangka 1 lebih baik dibanding angka 2.

2. Skala Ordinal

Nilai pada data berskala ordinal mengklasifikasi data ke dalam kategorik yang berbeda, dan berarti urutan ataupun ranking.

Contoh 1: Nilai akhir mahasiswa

1. A 2. B 3. C 4. D 5. E

Contoh 2: Beri nilai pada masakan ini:

a. 5 b. 6 c. 7 d. 8

3. Skala Interval

Data berskala interval merupakan data dalam skala urutan yang berbeda antar pengukuran tetapi tidak termasuk nilai 0. Contoh suhu badan anak 37,8°C.

4. Skala Rasio

Data berskala rasio merupakan skala berurutan hasil pengukuran dan termasuk angka 0 di dalamnya.

Contoh 1: Tinggi badan anak 100cm.

Contoh 2: Berat badan anak 25kg.

Latihan 1.2

1. Empat jenis minuman yang berbeda habis terjual setiap harinya di suatu restoran makanan cepat saji, yaitu minuman bersoda, teh, kopi dan minuman botol.
 - a. Jelaskan mengapa tipe jenis minuman yang terjual tersebut merupakan suatu contoh data kategorik.
 - b. Jelaskan mengapa tipe jenis minuman yang terjual tersebut merupakan suatu contoh data numerik.

2. Minuman kopi yang dijual pada suatu restoran cepat saji berukuran *small*, *medium* dan *large*, jelaskan mengapa ukuran minuman kopi tersebut adalah data berskala ordinal.

3. Tentukanlah apakah variabel-variabel di bawah ini merupakan variabel kategorik atau numerik. Jika variabel tersebut adalah numerik tentukan apakah termasuk data diskrit atau kontinu.
 - a. Jumlah pengguna telepon seluler setiap rumah
 - b. Lama waktu (dalam menit) panggilan telepon terpanjang selama sebulan.
 - c. Apakah koneksi internet dalam rumah tangga berkecepatan tinggi?

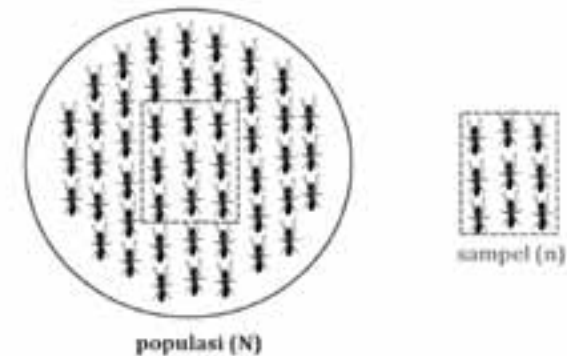
4. Toko buku kampus mengumpulkan informasi berikut ini dari mahasiswa selama seminggu kelas pertama
 - a. Jumlah waktu yang dihabiskan dalam took buku
 - b. Banyaknya buku teks yang dibeli
 - c. Jurusan Akademik
 - d. Jenis Kelamin

Kemudian klasifikasikan setiap variabel tersebut apakah termasuk data kategorik atau numerik. Jika termasuk data numerik, tentukan apakah termasuk diskrit atau kontinu, kemudian tentukan skala pengukuran datanya.

1.3 POPULASI DAN SAMPEL

Telah dijelaskan pada bagian 1.1 definisi populasi dan sampel. Bagian ini juga menjelaskan mengenai populasi dan sampel, namun lebih terperinci. *Populasi* adalah kumpulan objek yang diamati. *Sampel* adalah himpunan bagian dari populasi. Jika sampel terdiri dari keseluruhan populasi maka disebut *sensus*.

Hubungan antara populasi yang diamati dan sampel yang diambil dari populasi itu mungkin merupakan konsep yang paling penting dalam statistika. Hubungan ini digambarkan secara grafis pada Gambar 1.1



Gambar 1.1 Contoh Populasi dan Sampelnya

Gambar semut dalam lingkaran mewakili unsur populasi. Semut dalam persegi panjang mewakili unsur populasi yang dipilih secara acak dan bersama-sama membentuk sampel. Untuk setiap elemen sampel ada pengukuran yang diamati, dilambangkan dengan huruf kecil x (yang telah diberi indeks sebagai x_1, \dots, x_n), pengukuran ini secara kolektif membentuk kumpulan data sampel.

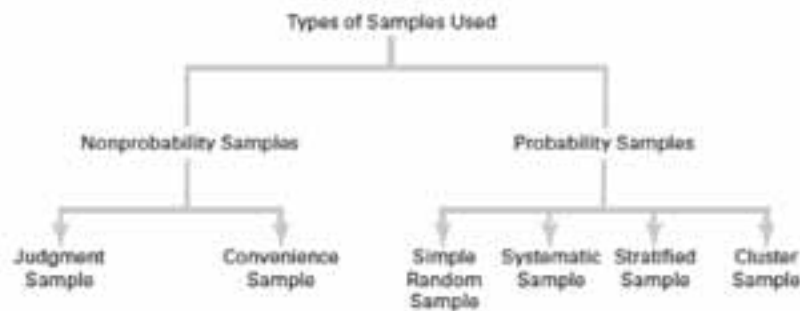
Dari data dapat dihitung berbagai statistik. Untuk mengantisipasi notasi yang akan digunakan nantinya, dapat dihitung rata-rata sampel \bar{x} dan rata-rata populasi μ .

1.3.1 Teknik sampling

Pada bagian 1.3, suatu sampel didefinisikan sebagai bagian dari populasi yang telah dipilih untuk dianalisis. Pemilihan setiap unit populasi untuk dijadikan sampel memiliki prosedur sampling khusus agar sampel tersebut benar merupakan bagian yang mewakili keadaan populasi. Hasil dari sampel kemudian digunakan untuk mengestimasi karakteristik populasi. Ada 3 alasan utama melakukan sampling, yaitu:

- a. Lebih hemat waktu dibandingkan penelitian dengan data populasi
- b. Lebih hemat biaya
- c. Hemat tenaga, dan lebih praktis.

Proses sampling bermula melalui pendefinisian *frame*, yang merupakan daftar unit-unit yang menggambarkan keadaan populasi. *Frame* merupakan sumber data seperti daftar populasi, direktori, atau peta. Sampel dipilih berdasarkan *frame*. Ketidakakuratan atau hasil sampel bias muncul jika *frame* berisi data yang tidak terdapat pada populasi. Teknik sampling dibagi dalam 2 cara, yaitu *probability sampling* (jika jumlah anggota (N) Populasi diketahui), dan *nonprobability sampling* (jika jumlah anggota populasi tidak diketahui).



Gambar 1.2 Tipe Teknik Sampling

1.3.2 Teknik sampling acak sederhana

Dalam sampel acak sederhana, setiap unit populasi memiliki kesempatan yang sama untuk dijadikan sampel. Dengan sampel acak sederhana, kita gunakan n untuk menyatakan ukuran sampel, dan N untuk menyatakan ukuran populasi.

Salah satu metode yang digunakan dalam memilih sampel acak sederhana ialah dengan menggunakan tabel bilangan acak (dalam lampiran). Tabel bilangan acak berisi barisan digit yang didaftar secara acak. Karena sistem numeric menggunakan 10 digit, maka peluang bahwa kita memilih secara acak sama dengan peluang menghasilkan angka acak lainnya. Sebagai contoh peluang 1 dari 10. Jika kita bangkitkan barisan bilangan 800 digit, maka kita menuliskan angka 80 ke dalam digit 0,80 dan menjadi 1, begitu seterusnya. Karena setiap digit atau barisan digit dalam tabel bersifat acak, tabel dapat dibaca secara horizontal maupun vertikal. Margin tabel didesain sebagai barisan angka dan kolom angka. Digit dikelompokkan ke dalam barisan 5 bilangan agar lebih mudah untuk dibaca.

Pemilihan sampel secara acak sederhana, dapat juga dilakukan dengan menentukan titik awal pada tabel bilangan acak. Salah satu metodenya ialah tutup mata, kemudian titikkan pada tabel bilangan acak dengan menggunakan pensil. Andaikan yang terpilih adalah angka pada baris 6 kolom 5 sebagai titik pertama, maka kamu dapat melanjutkan dari arah manapun, sebagai contoh kita membaca tabel dari kiri ke kanan, maka barisan tiga digit tanpa menghapusnya. Ingat bahwa, unit dalam *frame* dimulai dengan angka 0

	Column								
Row	00000	00001	11111	11112	22222	22223	33333	33334	
	12345	67890	12345	67890	12345	67890	12345	67890	
01	49280	88924	35779	00283	81163	07275	89863	02348	
02	61870	41657	07468	08612	98083	97349	20775	45091	
03	43898	65923	25078	86129	78496	97653	91550	08078	
04	62993	93912	30454	84598	56095	20664	12872	64647	
05	33850	58555	51438	85507	71865	79488	76783	31708	
Begin	06	97340	03364	88472	04334	63919	36394	11095	92470

Gambar 1.3 Contoh Tabel Bilangan Acak

1.3.3 Teknik sampling sistematis

Dalam sampling sistematis, partisi N unit populasi dalam *frame* ke dalam n kelompok yang terdiri dari k unit, dimana k merupakan bilangan bulat terdekat. Pemilihan sampel secara sistematis dilakukan dengan cara pilih unit pertama secara acak dari k unit pertama dalam *frame*. Kemudian, pilih sisa unit melalui pengambilan setiap unit ke- k dalam *frame* populasi dan seterusnya sampai jumlah sampel yang ditentukan.

$$sk = N/n$$

Contoh:

Suatu survey dilakukan pada karyawan pabrik celana jeans. Survey dilakukan dengan sampling sistematis melalui partisi *frame* 800 ke dalam 40 kelompok, maka setiap kelompok terdiri dari 20 karyawan. Kemudian kamu memilih bilangan acak dari 20 karyawan pertama, misalkan bilangan acak yang terpilih ialah 008, maka angka berikutnya adalah 028, 048, 068, 088, 108, ..., 788.

1.3.4 Teknik sampling berstrata

Dalam sampling berstrata, pertama tama kita membagi N unit pada populasi ke dalam beberapa subpopulasi atau strata. Sebuah stratum didefinisikan melalui beberapa karakteristik seperti jensi kelamin, pendidikan dan lainnya. Kemudian setiap unit sampel dari setiap stratum dipilih secara acak sederhana.

1.3.5 Teknik sampling klaster

Dalam sampling klaster, kita membagi N unit populasi ke dalam klaster. Klaster sering didesain sebagai Negara, daerah pemilihan, jenis rumah tangga, dan pendapatan. Kemudian, masing-masing unit populasi yang termasuk ke dalam klaster, dipilih secara acak sederhana.

1.3.6 Nonprobability sampling

Dalam *nonprobability sampling*, unit populasi dipilih tanpa mengetahui peluang pemilihannya. Oleh karena itu, statistika inferensial tidak dapat diaplikasikan pada data dengan sampel ini. Teknik sampling paling umum digunakan pada *nonprobability sampling* yaitu *convenience sampling* dan *judgement sampling*. Dalam *convenient sampling*, unit populasi secara mudah dan tidak sulit. Sebagai contoh, sampel ban mobil dalam suatu gudang, dipilih dengan cara memilih tumpukan ban yang paling atas dibandingkan tumpukan ban yang berada di bawah. Sedangkan dalam *judgement sampling*, kita memilih sendiri siapa unit Poyang pantas dijadikan sampel, misal kita ingin melakukan penelitian suatu kampus, maka kita unit sampel yang kita butuhkan adalah, rektor, staff pegawai dosen, dan mahasiswa.

Latihan 1.3

1. Suatu populasi terdiri dari $N = 902$, berapa angka kode yang kita tuliskan untuk:
 - a. orang pertama dalam daftar?
 - b. orang ke-40 dalam daftar?
 - c. orang terakhir dalam daftar?
2. Gunakan soal no. 1, verifikasi bahwa bilangan acak dimulai dari baris ke 5 kolom pertama (Tabel Bilangan Acak), daftarkan 20 unit sampel pertamanya.
3. Seorang peneliti ingin memperkirakan jumlah rata-rata yang dihabiskan per orang oleh pengunjung ke taman hiburan. Dia mengambil sampel acak dari empat puluh pengunjung dan memperoleh rata-rata seratus ribu rupiah per orang.

- a. Berapakah populasi yang diamati?
- b. Berapakah parameter yang diamati?
- c. Berdasarkan sampel ini, apakah kita tahu jumlah rata-rata yang dihabiskan per orang oleh pengunjung ke taman? Jelaskan secara rinci.

Bab 2

DESKRIPSI DATA

Seperti yang dijelaskan sebelumnya, statistika secara alami terbagi menjadi dua cabang, statistika deskriptif dan statistika inferensial. Titik awal untuk menangani kumpulan data adalah mengatur, menampilkan, dan meringkasnya secara efektif.

Dalam buku ini kita akan menggunakan dua format untuk menampilkan kumpulan data. Yang pertama adalah daftar data, yang merupakan daftar eksplisit semua pengukuran individual, baik sebagai tampilan dengan ruang antara pengukuran individual, atau dalam notasi set dengan pengukuran individual yang dipisahkan dengan koma.

Contoh 2.1

Data yang diperoleh dengan mengukur usia 21 mahasiswa yang dipilih secara acak yang terdaftar dalam sebuah mata kuliah mahasiswa baru di universitas dapat disajikan sebagai daftar data

18 19 18 18 19 24 19 18 19 20 18

18 22 21 20 20 18 17 18 19 17

atau dalam notasi set sebagai

{18,18,19,19,19,18,22,20,18,18,17,19,18,24,18,20,18,21,20,17,19}

Kumpulan data juga dapat disajikan dengan menggunakan tabel frekuensi data, tabel di mana setiap nilai yang berbeda x tercantum pada baris pertama dan frekuensinya f yang merupakan berapa kali nilai x muncul pada kumpulan data, adalah tercantum di bawah ini di baris kedua.

Contoh 2.2

Kumpulan data dari contoh sebelumnya ditunjukkan oleh tabel frekuensi data

x	17	18	19	20	21	22	24
f	2	8	5	3	1	1	1

Tabel frekuensi data mudah dibuat saat kumpulan data berukuran besar dan jumlah nilai yang berbeda tidak terlalu besar.

2.1 Tiga Penyajian Data yang Populer

Ada pepatah yang terkenal mengatakan bahwa "sebuah gambar bernilai seribu kata." Pepatah ini membuktikan kebenaran ketika menyajikan informasi statistik dalam kumpulan data. Ada banyak cara efektif untuk menyajikan data secara grafis. Tiga alat grafis yang diperkenalkan di bagian ini termasuk yang paling umum digunakan dan relevan dengan penyajian materi berikutnya dalam buku ini.

Diagram Batang dan Daun

Misalkan 30 mahasiswa di kelas statistika mengikuti tes dan membuat skor berikut:

86 90 40 80 83 58 25 70 68 77

73 69 73 73 100 76 70 78 100 90

87 90 83 97 69 71 92 93 95 74

Bagaimana hasil tes kelas tersebut? Sekilas sekilas angka 30 tidak segera memberikan jawaban yang jelas. Namun kumpulan data dapat direorganisasi dan ditulis ulang untuk membuat informasi yang relevan lebih terlihat. Salah satu cara untuk melakukannya adalah dengan membangun diagram batang dan daun seperti yang ditunjukkan pada Gambar 2.1 "Diagram Batang dan Daun". Angka di puluhan tempat, dari 2 sampai 9, dan tambahan angka 10, adalah "batang", dan disusun dalam urutan numerik dari atas ke bawah ke kiri garis vertikal. Jumlah di tempat unit di setiap pengukuran adalah "daun", dan ditempatkan berturut-turut di sebelah kanan batang yang sesuai, jumlah di tempat puluhan pengukuran itu. Jadi, tiga daun 9, 8, dan 9 berturut-turut mengarah ke batang 6 sesuai dengan tiga nilai ujian pada nilai 60an, 69 (pada baris pertama data), 68 (di baris ketiga), dan 69 (juga di baris ketiga). Tampilan dibuat lebih berguna untuk beberapa tujuan dengan menata ulang daun dalam urutan numerik, seperti yang ditunjukkan pada Gambar 2.2 "Diagram Batang dan Daun yang Diurutkan". Ada dua nilai sempurna; tiga siswa menghasilkan nilai di bawah 60; kebanyakan siswa mendapat nilai 70an, 80an dan 90an; dan rata-rata keseluruhan mungkin di 70-an tinggi atau rendah 80-an.

2		5									
3											
4		0									
5		8									
6		9	8	9							
7		7	3	6	0	3	3	0	1	8	4
8		6	0	3	3	7					
9		0	3	0	0	5	7	2			

10 | 0 0

Gambar 2.1 Diagram Batang dan Daun

2	5									
3										
4	0									
5	8									
6	8	9	9							
7	0	0	1	3	3	3	4	6	7	8
8	0	3	3	6	7					
9	0	0	0	2	3	5	7			
10	0	0								

Gambar 2.2 Diagram Batang dan Daun yang Diurutkan

Dalam contoh ini nilai memiliki batang alami (puluhan tempat) dan daun (yang ada). Seseorang bisa menyebarkan diagram dengan memecah setiap puluhan nomor tempat ke kategori yang lebih rendah dan lebih rendah. Misalnya, semua skor di tahun 80an dapat diwakili pada dua batang yang terpisah, lebih rendah 80an dan 80an atas:

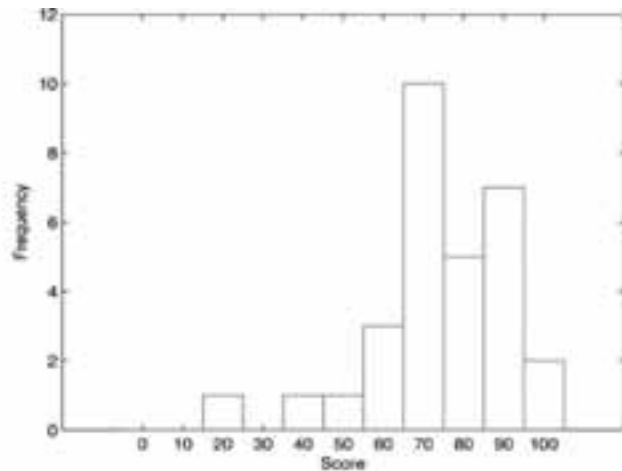
8	0	3	3
8	6	7	

Definisi batang dan daun dalam praktiknya sangat fleksibel. Tujuan umum diagram batang dan daun adalah untuk memberikan tampilan cepat bagaimana data disebarankan melintasi rentang nilai-nilai mereka; Beberapa improvisasi diperlukan untuk mendapatkan diagram yang paling sesuai dengan sasaran itu.

Perhatikan bahwa semua data asli dapat dipulihkan dari diagram batang dan daun. Ini tidak akan berlaku pada dua jenis display grafis berikut.

Histogram Frekuensi

Diagram batang dan daun tidak praktis untuk kumpulan data yang besar, jadi kita memerlukan cara yang berbeda dan murni untuk merepresentasikan data. Histogram frekuensi adalah perangkat semacam itu. Akan diilustrasikan dengan menggunakan kumpulan data yang sama dari subbagian sebelumnya. Untuk 30 skor dalam ujian, wajar untuk mengelompokkan skor pada skala sepuluh poin standar, dan menghitung jumlah skor di setiap kelompok. Jadi ada dua 100an, tujuh skor di nilai 90an, enam di nilai 80an, dan seterusnya.. Kemudian dibuat diagram yang ditunjukkan pada Gambar 2.3 "Histogram Frekuensi" dengan menggambar untuk setiap kelompok, atau kelas, sebuah batang vertikal yang panjangnya adalah jumlah pengamatan pada kelompok tersebut. Dalam contoh kita, bar berlabel 100 panjangnya 2 unit, bar berlabel 90 panjangnya 7 unit, dan seterusnya. Sementara nilai data individu hilang, kita tahu jumlahnya di setiap kelas. Nomor ini disebut frekuensi kelas, maka nama frekuensi histogram.



Gambar 2.3 Histogram Frekuensi

Prosedur yang sama dapat diterapkan pada pengumpulan data numerik. Pengamatan dikelompokkan ke dalam beberapa kelas dan frekuensi (jumlah pengamatan) masing-masing kelas dicatat.

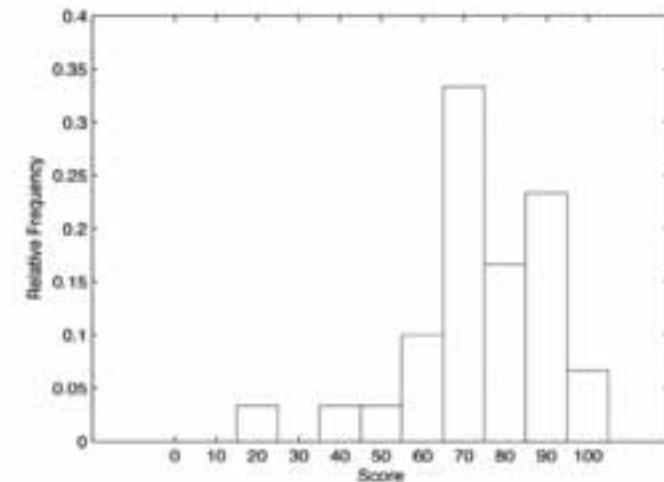
Kelas-kelas ini disusun dan ditunjukkan secara berurutan pada sumbu horizontal (disebut sumbu x), dan untuk setiap kelompok sebuah batang vertikal, yang panjangnya adalah jumlah observasi pada kelompok tersebut, ditarik. Tampilan yang dihasilkannya adalah histogram frekuensi untuk data. Kesamaan pada Gambar 2.1 "Diagram Batang dan Daun" dan Gambar 2.3 "Histogram Frekuensi" terlihat jelas, terutama jika Anda membayangkan mengubah diagram batang dan daun di sisinya dengan memutarinya seperempatnya berlawanan arah jarum jam.

Secara umum, definisi kelas dalam histogram frekuensi fleksibel. Tujuan umum histogram frekuensi sangat mirip dengan diagram batang dan daun, untuk memberikan tampilan grafis yang memberi rasa sebaran data di berbagai nilai yang muncul. Kami tidak akan membahas proses pembuatan histogram dari data karena dalam praktiknya

sebenarnya dilakukan secara otomatis dengan perangkat lunak statistik atau bahkan kalkulator ponsel.

Histogram Frekuensi Relatif

Dalam contoh nilai ujian di kelas statistika, lima mahasiswa mendapatkan nilai 80an. Angka 5 adalah frekuensi grup yang diberi label "80-an." Karena ada 30 mahasiswa di seluruh kelas statistik, proporsi yang mendapatkan nilai 80an adalah $5/30$. Angka $5/30$, yang juga bisa dinyatakan sebagai $0,1667$ atau $16,67\%$, adalah frekuensi relatif grup yang diberi label "80-an." Setiap kelompok (nilai 70an, 80an, dan seterusnya) memiliki frekuensi relatif. Dengan demikian kita dapat membuat diagram dengan menggambar untuk setiap kelompok, atau kelas, sebuah batang vertikal yang panjangnya adalah frekuensi relatif kelompok tersebut. Misalnya, bar untuk tahun 80an akan memiliki panjang $5/30$ unit, bukan 5 unit. Diagram adalah histogram frekuensi relatif untuk data, dan ditunjukkan pada Gambar 2.4 "Relative Frequency Histogram". Ini sama persis dengan histogram frekuensi kecuali sumbu vertikal dalam histogram frekuensi relatif bukan frekuensi tapi frekuensi relatif.



Gambar 2.4 Histogram Frekuensi Relatif

Prosedur yang sama dapat diterapkan pada pengumpulan data numerik. Kelas dipilih, frekuensi relatif setiap kelas dicatat, kelas disusun dan ditunjukkan secara berurutan pada sumbu horizontal, dan untuk setiap kelas, sebuah batang vertikal, yang panjangnya adalah frekuensi relatif kelas, ditarik. Tampilan yang dihasilkan adalah histogram frekuensi relatif untuk data. Titik kunci adalah bahwa sekarang jika masing-masing batang vertikal memiliki lebar 1 unit, maka luas total semua batang adalah 1 atau 100%.

Meskipun histogram pada Gambar 2.3 "Histogram Frekuensi" dan Gambar 2.4 "Relatif Frekuensi Histogram" memiliki tampilan yang sama, histogram frekuensi relatif lebih penting bagi kita, dan akan menjadi histogram frekuensi relatif yang akan digunakan berulang kali untuk mewakili data dalam buku ini.

Untuk melihat mengapa hal ini terjadi, renungkan apa yang sebenarnya Anda lihat dalam diagram yang secara cepat dan efektif mengkomunikasikan informasi kepada Anda tentang data. Ini adalah ukuran relatif dari batang. Batang yang diberi label "70-an" pada salah satu gambar menempati 1/3 dari total area semua batang, dan walaupun kita mungkin tidak memikirkannya secara sadar, kita melihat proporsi 1/3 pada gambar, yang menunjukkan bahwa sepertiga dari keseluruhan nilai di nilai 70an. Histogram frekuensi relatif penting karena pelabelan pada sumbu vertikal mencerminkan apa yang penting secara visual: ukuran relatif batang.

Latihan 2.1

1. Jelaskan satu perbedaan antara histogram frekuensi dan histogram frekuensi relatif!
2. Jelaskan satu keuntungan dari diagram batang dan daun di atas histogram frekuensi!
3. Buat diagram batang dan daun, histogram frekuensi, dan histogram frekuensi relatif untuk rangkaian data berikut. Untuk histogram menggunakan kelas 51-60, 61-70, dan seterusnya.
69 70 93 53 92 85 75 70 68 88 76 70 77 85 82 82 80 96 100 85
4. Buat diagram batang dan daun, histogram frekuensi, dan histogram frekuensi relatif untuk rangkaian data berikut. Untuk histogram menggunakan kelas 6.0-6.9, 7.0-7.9, dan seterusnya.
8.5 6.5 9.6 8.0 8.2 8.2 8.5 7.7 7.0 7.6 8.8 2.9 7.0 1.5 8.5 9.2 4.9 9.3 8.7 6.9
5. Kumpulan data berisi $n = 10$ pengamatan. Nilai x dan frekuensinya f dirangkum dalam tabel frekuensi data berikut.

x	-1	0	1	2
f	3	4	2	1

Buatlah histogram frekuensi dan histogram frekuensi relatif untuk kumpulan data tersebut.

6. Kumpulan data berisi $n = 20$ pengamatan Nilai x dan frekuensi f diringkas dalam tabel frekuensi data berikut.

x	-1	0	1	2
f	3	a	2	1

Frekuensi nilai 0 hilang. Tentukan nilai a dan kemudian sketsa histogram frekuensi dan histogram frekuensi relatif untuk kumpulan data tersebut.

7. Kumpulan data memiliki tabel sebaran frekuensi berikut:

x	1	2	3	4
f	3	a	2	1

Angka a tidak diketahui Bisakah Anda membuat histogram frekuensi? Jika ya, buatlah. Jika tidak, jelaskan mengapa tidak bisa.

8. Tabel dari beberapa frekuensi relatif yang dihitung dari kumpulan data adalah

x	1	2	3	4
f/n	0,3	p	0,2	0,1

Nilai p belum dihitung. Selesaikan tabel dan buat histogram frekuensi relatif untuk kumpulan data tersebut.

Soal Terapan

9. Nilai IQ dari sepuluh siswa yang dipilih secara acak dari sekolah dasar diberikan. 108 105 100 107 99 105 125 119 87 118

Kelompokkanlah data menjadi nilai 80an, 90an, dan seterusnya, kemudian buatlah diagram batang dan daun, histogram frekuensi, dan histogram frekuensi relatif.

10. Nilai IQ dari sepuluh siswa yang dipilih secara acak dari sekolah dasar untuk siswa berbakat secara akademis adalah sebagai berikut :

133 145 140 160 152 138 142 139 137 138.

Kelompokkanlah data menjadi ratusan dan puluhan digit biasa, kemudian susun diagram batang dan daun, histogram frekuensi, dan histogram frekuensi relatif dari data tersebut!

11. Pada suatu hari donor darah, 300 orang menyumbangkan darahnya. Jenis darah dari 300 donor ini dirangkum dalam tabel berikut:

<i>Golongan Darah</i>	A	B	AB	O
<i>Frekuensi</i>	120	32	12	136

Buat histogram frekuensi relatif untuk kumpulan data tersebut!

12. Di suatu toko perabot, penanak nasi otomatis adalah barang yang populer. Penjualan mingguan selama 20 minggu terakhir adalah sebagai berikut :

20 15 15 19 15 17 19 15 14 16 12 15 14 16 13 16 18 18 9 15

Buat histogram frekuensi relatif dengan kelas 6-10, 11-15, dan 16-20.

13. Dari satu sampel acak, masing-masing berukuran $n = 10$, diambil dari panjang dalam sentimeter dari tiga jenis ikan yang dijual pada toko ikan, dengan hasil sebagai berikut:

Sampel 1 : 108 105 100 107 99 105 125 119 87 118

Sampel 2 : 133 145 140 160 152 138 142 139 137 138

Sampel 3 : 82 60 74 79 82 83 82 80 82 80

Kelompokkan ukuran tersebut dalam ratusan dan puluhan, kemudian buatlah diagram batang dan daun, histogram frekuensi, dan histogram frekuensi relatif untuk masing-masing sampel. Bandingkan histogram dan jelaskan pola yang ditunjukkan data tersebut.

14. Di suatu toko perabot, penanak nasi otomatis adalah barang yang populer. Penjualan mingguan selama 20 minggu terakhir adalah sebagai berikut :

20 15 15 19 15 17 19 15 14 16 12 15 14 16 13 16 18 18 9 15

Dalam penjualan eceran, persediaan inventaris terlalu besar memerlukan modal yang besar, sementara biaya inventaris yang terlalu kecil akan mengakibatkan kehilangan penjualan dan kepuasan pelanggan. Dengan menggunakan histogram frekuensi relatif untuk data ini, temukan kira-kira berapa banyak penanak nasi yang harus di stok pada awal setiap minggu jika toko tersebut

- tidak kehabisan stok pada akhir minggu selama lebih dari 15% setiap minggu;
- tidak kehabisan stok pada akhir minggu selama lebih dari 5% setiap minggu

2.2 Ukuran Pemusatan

Ukuran deskriptif yang menunjukkan di mana pusat atau nilai paling khas dari variabel terletak pada kumpulan pengukuran yang diamati yang disebut ukuran pemusatan. Ukuran pemusatan sering disebut sebagai rata-rata.

Median dan rata-rata hanya berlaku untuk data kuantitatif, sedangkan modus dapat digunakan dengan data kuantitatif atau kualitatif.

Rata-rata (Mean)

Ukuran pertama dari ukuran pemusatan adalah "rata-rata" yang akrab bagi semua orang. Dalam rumus dalam definisi berikut, kami memperkenalkan notasi penjumlahan standar Σ , di mana Σ adalah huruf kapital huruf Yunani. Contoh berikut sebagai ilustrasi.

Contoh 2.3

Tentukan Σx , Σx^2 , dan $\Sigma (x-1)^2$ untuk kumpulan data

$$2 \quad 7 \quad 11$$

Penyelesaian :

$$\Sigma x = 2 + 7 + 11 = 20$$

$$\Sigma x^2 = 2^2 + 7^2 + 11^2 = 4 + 49 + 121 = 174$$

$$\Sigma (x-1)^2 = (2-1)^2 + (7-1)^2 + (11-1)^2 = 1 + 36 + 100 = 137$$

Dalam definisi kita mengikuti konvensi penggunaan huruf kecil n untuk menunjukkan jumlah pengukuran dalam sampel, yang disebut *ukuran sampel*.

Definisi 2.1

Rata-rata sampel dari satu kumpulan data sampel adalah \bar{x} didefinisikan oleh rumus :

$$\bar{x} = \frac{\sum x}{n}$$

Contoh 2.4

Nilai kuis matematika 15 mahasiswa jurusan Matematika adalah 7,8,6,4,10, 5,9,7, 3,8, 6, 5, 8, 9, dan 7. Tentukan nilai rata-ratanya.

Penyelesaian :

$$\begin{aligned} \bar{x} &= \frac{\sum x}{n} \\ &= \frac{7+8+6+4+10+5+9+7+3+8+6+5+8+9+7}{15} \\ &= \frac{102}{15} = 6,8 \end{aligned}$$

Contoh 2.5

Sampel acak dari 19 wanita dengan jumlah anak yang dilahirkan diberikan data berikut, di mana x adalah jumlah anak dan f adalah frekuensi dari nilai tersebut

x	0	1	2	3	4
f	3	6	6	3	1

Hitunglah rata-rata sampelnya!

Penyelesaian :

Setiap angka di baris pertama tabel adalah banyak data yang muncul dalam kumpulan data; Jumlah di bawah ini adalah berapa kali hal itu terjadi. Jadi nilai 0 diamati tiga kali, yaitu tiga pengukuran dalam kumpulan data adalah 0, nilai 1

diamati enam kali, dan seterusnya. Dalam konteks masalah ini berarti bahwa tiga wanita dalam sampel tidak memiliki anak, enam memiliki satu anak, dan seterusnya. Daftar eksplisit dari semua pengamatan dalam kumpulan tersebut adalah

0 0 0 1 1 1 1 1 1 2 2 2 2 2 2 3 3 3 4

Ukuran sampel dapat dibaca langsung dari tabel, tanpa daftar pertama seluruh kumpulan data, sebagai jumlah frekuensi:

$$n = 3 + 6 + 6 + 3 + 1 = 19.$$

Rata-rata sampel dapat dihitung langsung dari tabel juga, yaitu:

$$\begin{aligned} \bar{x} &= \frac{\sum x}{n} = \frac{0 \times 3 + 1 \times 6 + 2 \times 6 + 3 \times 3 + 4 \times 1}{19} = \frac{31}{19} \\ &= 1,6316 \end{aligned}$$

Dalam contoh di atas, kumpulan data digambarkan sebagai sampel. Oleh karena itu rata-ratanya itu berarti rata-rata sampel, dilambangkan dengan \bar{x} . Jika data berasal dari sensus, sehingga ada pengukuran untuk setiap elemen populasi, maka rata-ratanya dihitung dengan cara yang sama persis dengan menjumlahkan semua pengukuran dan membagi dengan berapa banyak dari jumlah itu, tapi sekarang rata-rata populasi dan dilambangkan dengan μ , huruf kecil Yunani mu.

Definisi 2.2

Rata-rata populasi dari satu kumpulan data populasi N adalah μ yang didefinisikan oleh rumus

$$\mu = \frac{\sum x}{N}$$

Rata-rata dua bilangan adalah bilangan yang berada di tengah antara keduanya. Misalnya, rata-rata angka 5 dan 17 adalah $(5 + 17) / 2 = 11$, yaitu 6 unit di atas 5 dan 6 unit di bawah 17. Dalam pengertian ini rata-rata 11 adalah "pusat" kumpulan data { 5,17}. Untuk data yang lebih besar berarti rata-rata juga dapat dianggap sebagai "pusat" data.

Median

Untuk melihat mengapa konsep rata-rata lain dibutuhkan, amati situasi berikut. Misalkan kita tertarik dengan pendapatan rata-rata tahunan karyawan di sebuah perusahaan besar. Diambil sampel acak dari tujuh karyawan, mendapatkan data sampel (dibulatkan ke ratusan ribu rupiah terdekat, dan dinyatakan dalam jutaan rupiah).

24,8 22,8 24,6 192,5 25,2 18,5 23,7

Rata-rata (dibulatkan menjadi satu desimal) adalah $\bar{x} = 47,4$, namun pernyataan "pendapatan rata-rata karyawan di perusahaan ini adalah Rp 47.400" pasti menyesatkan. Ini kira-kira dua kali lipat dari enam dari tujuh karyawan dalam sampel yang dibuat dan hal ini tidak masuk akal. Mudah untuk melihat apa yang salah: kehadiran seorang eksekutif dalam sampel, yang gajinya sangat besar dibandingkan dengan jumlah orang lain, menyebabkan pembilang dalam rumus untuk sampel yang dimaksud terlalu besar. Angka 192.5 dalam kumpulan data di atas disebut *pencilan*, angka yang jauh dari sebagian besar atau seluruh pengukuran lainnya. Sering kali pencilan adalah hasil dari semacam kesalahan, tapi tidak selalu, seperti yang terjadi di sini. Kita akan mendapatkan ukuran yang lebih baik dari "pusat" data jika kita mengatur data dalam urutan numerik,

18,5 22,8 23,7 24,6 24,8 25,2 192,5

lalu pilih bilangan tengah dalam daftar, dalam hal ini 24,6.

Hasilnya disebut *median* kumpulan data, dan memiliki sifat yang kira-kira setengah dari ukurannya lebih besar dari itu, dan kira-kira setengahnya lebih kecil. Dalam hal ini, ia menempatkan pusat data. Jika ada banyak data genap, maka akan ada dua unsur yang berada di tengah. Jadi kita mengambil rata-rata dua data yang tengah sebagai median. Dengan demikian kita memiliki definisi berikut.

Definisi 2.3

Median sampel \bar{x} dari sekumpulan data sampel, dimana banyak data ganjil, adalah pengukuran tengah saat data disusun dalam urutan numerik. Median sampel \bar{x} dari satu umpulan data sampel, dimana banyak data genap, adalah rata-rata dua data di tengah saat data disusun dalam urutan numerik.

Median populasi didefinisikan dengan cara yang sama, namun tidak akan dibahas dalam teks ini.

Median adalah nilai yang membagi pengamatan dalam kumpulan data sehingga 50% data berada di sebelah kiri dan 50% lainnya di sebelah kanannya.

Contoh 2.6

Hitung median sampel untuk data berikut :

2 -1 2 0.

Penyelesaian :

Data dalam urutan numerik adalah -1, 0, 2, 2. Dua pengukuran tengah adalah 0 dan 2, jadi

$$\bar{x} = \frac{0 + 2}{2} = 1$$

Contoh 2.7

Hitung median sampel untuk data pada Contoh 2.5 di atas!

Penyelesaian

Data dalam urutan numerik adalah

0 0 0 1 1 1 1 1 2 2 2 2 2 2 3 3 3 4.

Jumlah pengamatan adalah 19, yaitu ganjil, jadi ada satu pengukuran tengah, yaitu pengukuran kesepuluh. Karena pengukuran kesepuluh adalah 2, mediannya adalah $\bar{x} = 2$.

Penting untuk dicatat bahwa kita bisa menghitung median tanpa terlebih dahulu mencatat secara eksplisit semua pengamatan dalam kumpulan data. Kita sudah melihat di Contoh 2.5 bagaimana menemukan jumlah pengamatan langsung dari frekuensi yang tercantum dalam tabel: $n = 3 + 6 + 6 + 3 + 1 = 19$. Seperti di atas kita tahu bahwa median adalah observasi kesepuluh. Baris kedua tabel pada Contoh 2.5 menunjukkan bahwa ketika data dicantumkan dalam urutan akan ada tiga 0 diikuti oleh enam 1, maka observasi kesepuluh adalah 2. Sehingga mediannya adalah 2.

Modus

Mungkin Anda pernah mendengar pernyataan seperti "Jumlah rata-rata sepeda motor yang dimiliki oleh rumah tangga di Indonesia adalah 1,37," dan merasa bingung saat memikirkan ada sepeda motor pecahan yang sedang berjalan di jalan raya. Dalam konteks seperti ini, ukuran berikut untuk lokasi pusat mungkin lebih masuk akal.

Definisi 2.4

Modus sampel dari satu kumpulan data adalah nilai yang paling sering terjadi.

Median populasi didefinisikan dengan cara yang sama, namun tidak akan dibahas dalam teks ini.

Untuk setiap kumpulan data selalu ada satu mean dan tepat satu median. Ini tidak harus benar modus; Beberapa nilai yang berbeda bisa terjadi dengan frekuensi tertinggi, seperti yang akan kita lihat. Bahkan bisa terjadi bahwa setiap nilai terjadi dengan frekuensi yang sama, dalam hal ini konsep mode tidak masuk akal.

Contoh 2.8

Tentukan modus kumpulan data berikut. -1 0 2 0

Penyelesaian :

Nilai 0 paling sering diamati dan oleh karena itu modusnya adalah 0.

Contoh 2.9.

Tentukan modus kumpulan data pada Contoh 2.5 di atas!

Penyelesaian :

Dua nilai yang paling sering diamati dalam kumpulan data adalah 1 dan 2. Oleh karena itu, modus adalah himpunan dua nilai: {1, 2}.

Modus adalah ukuran pemusatan karena kebanyakan kumpulan data dalam kehidupan nyata memiliki pengamatan lebih banyak di dekat pusat rentang data dan lebih sedikit pengamatan pada ujung bawah dan atas. Nilai dengan frekuensi tertinggi sering berada di tengah rentang data.

Latihan 2.2

1. Untuk kumpulan data sampel {1,2,6} temukan
 - a. $\sum x$
 - b. $\sum x^2$, dan
 - c. $\sum (x-1)^2$
2. Untuk kumpulan data sampel {-1,0,1,4} temukan
 - a. $\sum x$
 - b. $\sum x^2$, dan
 - c. $\sum (x-1)^2$
3. Untuk setiap kumpulan data berikut, tentuan mean, median dan modusnya
 - a. 1 2 3 4
 - b. 4 2 4 3 1
 - c. 20 25 27 26 25 24 10
 - d. 1 0 1 0 -1 0 -1 0 1 -1 -1 1
4. Untuk kumpulan data berikut, tentuan rata-rata, median dan modusnya

<i>x</i>	3	4	5	6	7
<i>f</i>	2	4	16	20	8

5. Buat kumpulan data sampel dengan ukuran $n = 3$ yang rata-ratanya lebih dari mediannya.
6. Buat kumpulan data sampel dengan ukuran $n = 3$ yang rata-ratanya kurang dari mediannya.
7. Buat kumpulan data sampel dengan ukuran $n =$ yang rata-rata, median dan modusnya bernilai yang sama.
8. Hitunglah rata-rata dan median untuk tingkat kolesterol LDL dalam sampel sepuluh orang pasien jantung.
132 139 162 147 133 160 145 150 148 153
9. Hitunglah rata-rata, median, dan modus untuk jumlah kendaraan yang dimiliki dalam survei terhadap 52 rumah tangga.

<i>x</i>	0	1	2	3	4	5	6	7
<i>f</i>	2	12	15	11	6	3	1	2

10. Jumlah penumpang di masing-masing 120 kendaraan yang diamati secara acak pada jam sibuk pagi hari tercatat, dengan hasil sebagai berikut.

<i>x</i>	1	2	3	4	5
<i>f</i>	84	29	3	3	1

Hitunglah rata-rata, median, dan modus data di atas!

11. Seorang pemain mencata semua hasil pelemparan sepasang dadu dan mendapatkan data berikut

<i>x</i>	2	3	4	5	6	7	8	9	1	1	1
<i>f</i>	1	2	4	5	6	7	6	5	3	2	1
	0	9	0	6	8	7	7	5	9	8	1

Hitunglah rata-rata, median, dan modus data di atas!

12. Cordelia mencatat waktu perjalanan setiap harinya untuk bekerja setiap hari, sampai menit terdekat, selama dua bulan, dan mendapatkan data berikut.

<i>x</i>	26	27	28	29	30	31	32
<i>f</i>	3	4	16	12	6	2	1

- a. Berdasarkan table frekuensi di atas, apakah Anda dapat mengharapkan rata-rat dan median mempunyai nilai yang hampir sama atau berbeda, dan mengapa?
- b. Hitunglah rata-rata, median, dan modus data di atas!

dengannya. Hal ini disebut ukuran variabilitas, dan kita akan membahas tiga di antaranya.

Jangkauan (Range)

Ukuran penyebaran yang pertama akan dibahas adalah yang paling sederhana.

Definisi 2.5

Jangkauan kumpulan data adalah bilangan R yang didefinisikan oleh rumus

$$R = x_{maks} - x_{min}$$

dimana x_{maks} adalah pengukuran terbesar dalam kumpulan data dan x_{min} adalah pengukuran terkecil.

Contoh 2.10

Tentukan jangkauan setiap kumpulan data pada Tabel 2.1 "Dua Kumpulan Data".

Penyelesaian:

Untuk Data Set I data terbesar 43 dan data terkecil 38, jadi jangkauannya adalah $R = 43 - 38 = 5$.

Untuk Data Set II data terbesar 47 dan data terkecil 33, jadi jangkauannya adalah $R = 47 - 33 = 14$.

Jangkauan ini merupakan ukuran variabilitas karena menunjukkan ukuran interval dimana titik data disebarankan. Jangkauan yang lebih kecil menunjukkan variabilitas yang kurang (kurang dispersi) di antara data, sedangkan rentang yang lebih besar menunjukkan sebaliknya.

Ragam dan Simpangan Baku

Dua ukuran variabilitas lainnya yang akan kita amati lebih rumit dan juga bergantung pada apakah kumpulan data hanyalah sampel yang diambil dari populasi yang jauh lebih besar atau keseluruhan populasi itu sendiri (yaitu, sebuah sensus).

Definisi 2.6

Ragam sampel dari satu kumpulan data sampel adalah bilangan s^2 yang didefinisikan oleh rumus

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

yang oleh aljabar sama dengan rumusnya

$$s^2 = \frac{\sum x^2 - \frac{1}{n}(\sum x)^2}{n - 1}$$

Simpangan baku sampel dari satu kumpulan data sampel adalah akar kuadrat dari ragam sampel, dengan rumus

$$s = \sqrt{\frac{\sum (x - \bar{x})^2}{n - 1}} = \sqrt{\frac{\sum x^2 - \frac{1}{n}(\sum x)^2}{n - 1}}$$

Meskipun rumus pertama dalam setiap kasus terlihat lebih sederhana daripada yang kedua, namun rumus yang terakhir lebih mudah digunakan dalam perhitungan tangan, dan disebut formula jalan pintas atau rumus singkat.

Contoh 2.11

Hitunglah ragam sampel dan simpangan baku sampel dari Kumpulan Data II pada Tabel 2.1 Dua Kumpulan Data.

Penyelesaian:

Untuk menggunakan rumus yang didefinisikan (rumus pertama) dalam definisi yang pertama kita hitung untuk setiap pengamatan x adalah simpangannya $x - \bar{x}$ dari rata-rata sampel. Karena rata-rata datanya adalah $\bar{x} = 40$, maka kita dapatkan sepuluh bilangan yang ditampilkan di baris kedua tabel berikut

x	46	37	40	33	42	36	40	47	34	45
$x - \bar{x}$	6	-3	0	-7	2	-4	0	7	-6	5

Maka

$$\sum (x - \bar{x})^2 = 6^2 + (-3)^2 + 0^2 + (-7)^2 + 2^2 + (-4)^2 + 0^2 + 7^2 + (-6)^2 + 5^2 = 224$$

sehingga

$$s^2 = \frac{\sum(x-\bar{x})^2}{n-1} = \frac{224}{9} = 24,8$$

dan

$$s = \sqrt{24,8} = 4,99,$$

Untuk tkumpulan Data I, nilai ragam dan simpangan bakunya adalah $s^2 = \frac{20}{9} = 2,2$ dan $s = \sqrt{\frac{20}{9}} = 1,49$.

Contoh 2.12

Tentukan ragam sampel dan simpangan baku sampel dari sepuluh IPK mahasiswa Matematika UINSU berikut :

1,90 3,00 2,53 3,71 2,12 1,76 2,71 1,39 4,00 3,33.

Penyelesaian :

Karena

$$\sum x = 1,90 + 3,00 + 2,53 + 3,71 + 2,12 + 1,76 + 2,71 + 1,39 + 4,00 + 3,33 = 26,45$$

dan

$$\sum x^2 = 1,90^2 + 3,00^2 + 2,53^2 + 3,71^2 + 2,12^2 + 1,76^2 + 2,71^2 + 1,39^2 + 4,00^2 + 3,33^2 = 76,7321$$

dengan menggunakan rumus singkat

$$s^2 = \frac{\sum x^2 - \frac{1}{n}(\sum x)^2}{n-1} = \frac{76,7321 - \frac{1}{10}26,45^2}{9} = 0,752427$$

dan

$$s = \sqrt{0,752427} = 0,687.$$

Ragam sampel memiliki satuan yang berbeda dari data. Misalnya, jika satuan dalam kumpulan data berukuran beberapa inci, satuan baru akan berukuran inci kuadrat, atau inci persegi. Dengan demikian, hal itu terutama penting secara teoritis dan tidak akan dipertimbangkan lebih jauh dalam teks ini, kecuali secara sepintas.

Jika kumpulan data terdiri dari keseluruhan populasi, maka simpangan baku populasi, dilambangkan σ (huruf kecil huruf Yunani sigma), dan kuadratnya, ragam populasi σ^2 , didefinisikan sebagai berikut.

Definisi 2.7

Ragam populasi dan *simpangan baku populasi* dari satu kumpulan data dengan jumlah populasi N adalah σ^2 dan σ yang didefinisikan oleh rumus berikut

$$\sigma^2 = \frac{\sum(x-\mu)^2}{N} \text{ dan } \sigma = \sqrt{\frac{\sum(x-\mu)^2}{N}}$$

Perhatikan bahwa penyebut dalam pecahan adalah jumlah observasi secara keseluruhan, bukan jumlah yang dikurangi satu, seperti halnya dengan simpangan baku sampel. Karena kebanyakan kumpulan data adalah sampel, kita akan selalu bekerja dengan ragam dan simpangan baku sampel.

Latihan 2.3

1. Hitunglah jangkauan, ragam, dan simpangan baku dari kumpulan data berikut :
 - a. 2 3 2 5
 - b. 2 -3 6 0 3 1
 - c. 2 1 2 7
 - d. -1 0 1 4 1 1 1
2. Hitunglah jangkauan, ragam, dan simpangan baku dari kumpulan data berikut :

x	3	4	5	6	7
f	2	4	16	20	8

3. Tentukan jangkauan, ragam, dan simpangan baku untuk sampel dari sepuluh nilai IQ yang dipilih secara acak dari sebuah sekolah untuk siswa berbakat akademis.
132 139 162 147 133 160 145 150 148 153

4. Hitunglah simpangan baku dari data berikut

x	1	2	3	4	5	6	7	8	9	10
f	384	208	98	56	28	12	8	2	3	1

5. Sampel acak dari 49 faktor untuk perbaikan pada bengkel mobil diambil. Data disusun dalam diagram batang dan daun yang ditunjukkan. (Batangnya ribuan dolar, daunnya ratusan, jadi misalnya pengamatan terbesar adalah 3.800)

3	5	6	8								
3	0	0	1	2	4						
2	5	6	6	7	7	8	8	9	9		
2	0	0	0	0	1	2	2	4			
1	5	5	5	6	6	7	7	7	8	8	9
1	0	0	1	3	4	4	4				
0	5	6	8	8							
0	4										

Untuk data ini $\sum x = 101.100$, $\sum x^2 = 244.380.000$.

- a. Hitunglah rata-rata, median dan modusnya
 - b. Tentukan jangkauannya
 - c. Hitunglah ragam dan simpangan bakunya
6. Mulailah dengan kumpulan data berikut, sebut saja Data Set I.
5 -2 6 14 -3 0 1 4 3 2 5
 - a. Hitung simpangan baku Data Set I
 - b. Buatlah kumpulan data baru, Data Set II, dengan menambahkan 3 ke setiap bilangan di Data Set I. Hitung simpangan baku Data Set II.
 - c. Buatlah kumpulan data baru, Data Set III, dengan mengurangi 6 ke setiap bilangan di Data Set I. Hitung simpangan baku Data Set III.

- d. Bandingkan jawaban untuk bagian (a), (b), dan (c), dapatkah Anda menebak polanya? Nyatakan prinsip umum yang Anda dapatkan.

2.4 Ukuran Letak

Saat mengikuti ujian, hal yang sama pentingnya dengan skor ujian Anda adalah cara membandingkan skor Anda dengan kinerja siswa lainnya. Jika Anda menghasilkan skor 70 tapi rata-rata (apakah rata-rata, median, maupun modus) adalah 85, Anda melakukannya dengan relatif buruk. Jika Anda menghasilkan angka 70 tapi skor rata-rata hanya 55, maka Anda melakukannya dengan cukup baik. Secara umum, signifikansi satu nilai yang teramati dalam kumpulan data sangat bergantung pada bagaimana nilai tersebut dibandingkan dengan nilai pengamatan lainnya dalam kumpulan data.

Kuartil, Desil dan Persentil

Kuartil, Desil dan Persentil adalah ukuran letak yang berguna untuk membandingkan skor dalam satu kumpulan data. Untuk satu set data, Anda dapat membagi data menjadi tiga kuartil (Q_1, Q_2, Q_3), sembilan desil (D_1, D_2, \dots, D_9) dan 99 persentil (P_1, P_2, \dots, P_{99}). Kuartil bawah (Q_1) memisahkan bagian bawah 25% dari atas 75%, kuartil tengah (Q_2) adalah median dan kuartil atas (Q_3) memisahkan 25% teratas dari bawah 75%. Untuk bekerja dengan persentil, desil dan kuartil - Anda perlu belajar melakukan dua tugas yang berbeda. Pertama, Anda harus belajar bagaimana menemukan persentil yang sesuai dengan skor tertentu dan kemudian

bagaimana menemukan skor dalam sekumpulan data yang sesuai dengan persentil tertentu.

Berikut rumus untuk menghitung nilai kuartil, desil, dan persentil

$$Q_i = Tb + \frac{\frac{iN}{4} - (\sum f)_i}{f_{Q_i}}$$

$$D_i = Tb + \frac{\frac{iN}{10} - (\sum f)_i}{f_{D_i}}$$

$$P_i = Tb + \frac{\frac{iN}{100} - (\sum f)_i}{f_{P_i}}$$

Dimana

Q_i : kuartil ke- $i, i = 1, 2, 3$

D_i : desil ke- $i, i = 1, 2, \dots, 9$

P_i : Persentil ke- $i, i = 1, 2, \dots, 99$

Tb : Tepi bawah kelas dari kelas yang memuat kuartil, desil atau persentil

N : banyak pengamatan (data)

$(\sum f)_i$: frekuensi kumulatif sebelum kelas

f_{Q_i} : frekuensi kelas yang memuat Q_i

f_{D_i} : frekuensi kelas yang memuat D_i

f_{P_i} : frekuensi kelas yang memuat P_i

c : panjang kelas

Contoh 2.13

Tabel berikut menunjukkan kecepatan kendaraan di pada suatu jalan bebas hambatan pada hari-hari biasa.

Kecepatan (km/jam)	Banyak kendaraan
60-69	138
70-79	163
80-89	325
90-99	541
100-109	427
110-119	214
120-129	110
130-139	52
140-149	30
	$N=2000$

Hitunglah Q_1 , D_3 dan P_{95}

Penyelesaian :

$$Q_1 = Tb + \frac{\frac{3N}{4} - (\sum f_k)}{f_{D_1}} \cdot c = 79.5 + \frac{1500 - 3015}{325} \cdot 10 = 85.62 \text{ km/jam}$$

$$D_3 = P_{90} = Tb + \frac{\frac{3N}{10} - (\sum f_k)}{f_{D_3}} \cdot c = 79.5 + \frac{600 - 3015}{325} \cdot 10 = 88.7 \text{ km/jam}$$

$$P_{95} = Tb + \frac{\frac{95N}{100} - (\sum f_k)}{f_{P_{95}}} \cdot c = 119.5 + \frac{1900 - 18085}{110} \cdot 10 = 127.89 \text{ km/jam}$$

Nilai z

Cara lain untuk menemukan pengamatan tertentu x dalam kumpulan data adalah menghitung jaraknya dari rata-rata terhadap satuan simpangan baku.

Definisi 2.8

Nilai z dari pengamatan x adalah bilangan z yang diberikan oleh rumus komputasi

$$z = \frac{x - \bar{x}}{s} \text{ atau } z = \frac{x - \mu}{\sigma}$$

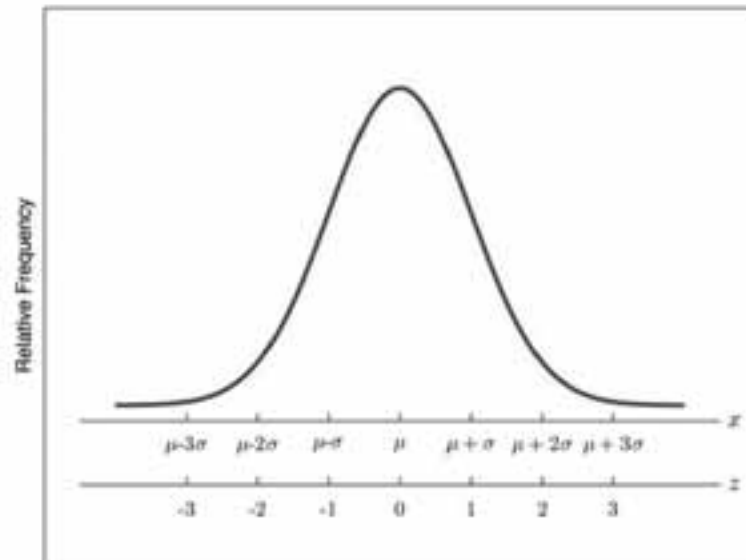
tergantung apakah kumpulan data adalah sampel atau keseluruhan populasi.

Rumus dalam definisi memungkinkan kita menghitung nilai z saat x diketahui. Jika nilai z diketahui maka x dapat ditentukan dengan menggunakan rumus berikut

$$x = \bar{x} + sz \text{ atau } x = \mu + \sigma z$$

Nilai z menunjukkan berapa jauh jarak simpangan baku setiap pengamatan x dengan rata-ratanya. Jika z negatif maka x di bawah rata-rata. Jika z adalah 0 maka x sama dengan rata-rata. Jika z positif maka x di atas rata-rata. Perhatikan gambar berikut!

Gambar 2.6 Skala x vs Nilai z



Contoh 2.14

Tentukan nilai z untuk semua sepuluh pengamatan dalam data sampel IPK mahasiswa berikut

1,90 3,00 2,53 3,71 2,12 1,76 2,71 1,39 4,00 3,33

Penyelesaian

Untuk data ini, nilai $\bar{x} = 2,645$ dan $s = 0,8674$. Nilai pengamatan yang pertama adalah $x = 1,90$, maka nilai z dari data ini adalah

$$z = \frac{1,90 - 2,645}{0,8674} = -0,8589$$

yang berarti $x = 1,90$ berada 0,8589 kali simpangan baku di bawah rata-rata kumpulan data. Untuk data kedua, yaitu $x = 3,00$, nilai z nya adalah

$$z = \frac{3,00 - 2,645}{0,8674} = 0,4093$$

yang berarti $x = 3,00$ berada 0,4093 kali simpangan baku di atas rata-rata kumpulan data. Dengan mengulangi proses yang sama untuk pengamatan yang tersisa, maka kumpulan lengkap nilai z adalah

-0,86 0,41 -0,13 1,23 -0,61 -1,02 0,07 -1,45 1,56 0,79

Contoh 2.15

Misalkan rata-rata dan simpangan baku IPK dari semua mahasiswa yang terdaftar saat ini di perguruan tinggi adalah $\mu = 2,70$ dan $\sigma = 0,50$. Nilai z dari IPK dua siswa, Anton dan Budi masing-masing adalah $z = -0,62$ dan $z = 1,28$. Berapakah IPK mereka?

Penyelesaian:

Dengan menggunakan rumus kedua tepat setelah definisi nilai z , kita menghitung IPK sebagai berikut

Anton : $x = \mu + \sigma z = 2,70 + 0,50 \cdot (-0,62) = 2,39$

Budi : $x = \mu + \sigma z = 2,70 + 0,50 \cdot (1,28) = 3,34$

Latihan 2.4

1. Tabel berikut menunjukkan sebaran frekuensi upah mingguan 65 karyawan di Perusahaan P & R.
 - a. Hitunglah rata-rata upah
 - b. Hitunglah modus upah
 - c. Hitunglah median upah
 - d. Hitunglah Q_3 dan D_8

Upah	Jumlah Karyawan
\$250.00-259.99	8
\$260.00-269.99	10
\$270.00-279.99	16
\$280.00-289.99	14
\$290.00-299.99	10
\$300.00-309.99	5
\$310.00-319.99	2
	N=65

2. Untuk populasi
 0,5 2,1 4,4 1,0 0,5 2,1 4,4 1,0
 hitung masing-masing hal berikut.
 - a. Rata-rata populasi μ .
 - b. Ragam populasi σ^2 .
 - c. Simpangan bau populasi σ .
 - d. Nilai z untuk setiap nilai dalam kumpulan data populasi.
3. Suatu pengukuran x dalam sampel dengan rata-rata $\bar{x}= 10$ dan simpangan baku $s = 3$ memiliki nilai $z = 2$. Tentukan nilai x .
4. Suatu pengukuran x dalam sampel dengan rata-rata $\bar{x}=10$ dan simpangan baku $s = 3$ memiliki nilai $z = -1$. Tentukan nilai x .
5. Suatu pengukuran x dalam populasi dengan rata-rata $\mu= 2,3$ dan simpangan baku $\sigma = 3$ memiliki nilai $z = 2$. Tentukan nilai x .

6. Suatu pengukuran x dalam populasi dengan rata-rata $\mu= 2,3$ dan simpangan baku $\sigma = 3$ memiliki nilai $z = -1,2$. Tentukan nilai x .
7. Annisa dan Fahri mengikuti mata kuliah ekonometrika yang sama, Annisa di semester berjalan, Fahri di mata kuliah pengambilan. Annisa mendapat nilai akhir 83, yang rata-rata di kelas adalah 76 dan simpangan baku 8. Fhari mendapat 79 dengan rata-rata kelas adalah 65 dan simpangan baku 12. Orang yang memiliki nilai z lebih tinggi relatif lebih baik. Apakah itu Annisa atau Fahri?

Bab 3

TEORI PELUANG

Teori Peluang dikembangkan pada abad ke XVII oleh ahli matematika dari Perancis yang bernama Pierre de Fermat dan Blaise Pascal. Awalnya teori peluang dimulai dari permainan judi atau permainan yang bersifat untung-untungan. Dalam teori peluang banyak dijumpai soal-soal yang berkaitan dengan uang logam, dadu, kartu bridge dan lain-lain. Mengapa perlu memahami teori peluang? Teori peluang dalam perencanaan ekonomi akan membantu Anda untuk menilai imbalan yang diharapkan untuk tindakan tertentu. Pada teori keputusan, peluang akan membantu dalam pengambilan keputusan yang tepat. Sedangkan pada industri asuransi akan membantu dalam meminimumkan resiko. Ilmu rekayasa memerlukan peluang dalam membuat dan memproses model yang diperlukan. Dalam bab ini akan diawali dengan kombinatorik, kemudian aksioma-aksioma peluang dan peluang bersyarat.

3.1. Kombinatorik

Dalam kehidupan sehari-hari sering dijumpai persoalan-persoalan sebagai berikut:

1. Dengan berapa cara dapat disusun n obyek menurut aturan tertentu?
2. Dengan berapa cara pengambilan sejumlah r obyek dari n obyek yang ada, bila $r < n$?

3. Dengan berapa cara sesuatu kejadian kejadian dapat terjadi?

Persoalan-persoalan di atas dapat diselesaikan dengan menggunakan kombinatorik. Ada 2 (dua) prinsip pokok yang dipakai untuk menyelesaikan persoalan kombinatorik, yaitu prinsip penjumlahan dan prinsip perkalian.

Prinsip-prinsip Berhitung

Teorema 3.1.

Prinsip dasar berhitung (Prinsip Penjumlahan) : Jika himpunan A berisi n elemen dan himpunan B berisi m elemen, maka ada mn cara yang dapat kita pilih, pertama, sebuah elemen A dan kemudian sebuah elemen B

Teorem 3.2

Perumuman Prinsip Dasar Berhitung (Prinsip Perkalian): Jika eksperimen r yang akan dilakukan sedemikian rupa sehingga yang pertama dapat menghasilkan n_1 hasil yang mungkin, dan jika untuk setiap hasil n_1 yang mungkin ada, ada kemungkinan n_2 hasil eksperimen kedua, dan jika untuk setiap kemungkinan hasil dari dua percobaan yang pertama ada n_3 kemungkinan hasil percobaan ketiga, dan seterusnya, maka ada total $n_1 n_2 n_3 \dots n_r$ kemungkinan hasil r percobaan.

Teorema 3.3

Suatu himpunan dengan banyak anggota n maka terdapat 2^n himpunan bagian.

Contoh 3.1

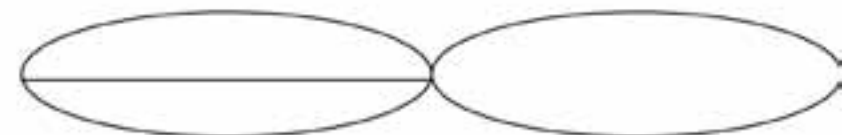
Untuk Prinsip Penjumlahan

- Suatu klub sepak bola mempunyai 40 anggota sedangkan klub bulutangkis mempunyai 20 anggota.
 - a. Jika tidak ada anggota sepak bola yang merangkap menjadi anggota bulutangkis, maka jumlah anggota kedua klub adalah $40 + 20 = 60$ anggota
Jika kedua himpunan tidak beririsan, maka jumlah anggota kedua klub ditambahkan.
 - b. Jika ada 7 anggota yang merangkap menjadi anggota kedua klub, maka dibentuk 3 himpunan yang saling lepas atau tidak beririsan, yaitu:
 - (i) Himpunan I terdiri dari pemain sepak bola saja
 - (ii) Himpunan II terdiri dari pemain bulutangkis saja
 - (iii) Himpunan III terdiri dari pemain sepak bola dan bulutangkis
 Ketiga himpunan ini saling lepas dengan masing-masing anggota $40-7, 20-7$ dan 7, dengan demikian jumlah anggota dari kedua klub adalah $33+13+7= 53$
Cara lain untuk memperoleh hasil di atas adalah dengan rumus

$$n(A \cup B) = n(A) + n(B) - n(A \cap B)$$

Contoh 3.2

- Untuk Prinsip Perkalian
Ahmad pergi dari kota A ke kota C dan harus melalui kota B. Dari kota A ke kota B ada 3 jalan alternatif dan dari kota B ke kota C ada 2 jalan alternatif. Dengan berapa banyak cara Ahmad bepergian dari kota A ke kota C?



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

Binomium Newton

$$(a + b)^n = \sum_{r=0}^n C_r^n a^{n-r} b^r$$

Soal:

1. Diketahui enam angka yaitu: 0, 1, 2, 3, 4 dan 5
 - a. Berapa banyak bilangan yang dapat dibentuk dari enam angka yang diketahui terdiri dari tiga angka (digit), bila tiap angka hanya dapat digunakan sekali
 - b. Berapa banyak daripadanya yang merupakan bilangan genap
 - c. Berapa banyak yang lebih besar dari 330
2. Dengan berapa carakah enam pohon dapat ditanam membentuk lingkaran?
3. Dari kelompok yang terdiri atas lima pria dan tiga wanita, berapa banyak panitia yang beranggotakan tiga orang dapat dibentuk:
 - a. tanpa pembatasan?
 - b. dengan dua pria dan seorang wanita?
 - c. dengan seorang wanita dan dua orang wanita bila seorang wanita tertentu harus ikut dalam panitia?
4. Tentukan koefisien x^7 dari $(2x - 3)^{10}$

3.3. Aksioma-aksioma Peluang

Pengertian Ruang Sampel dan Kejadian

Ruang Sampel adalah seluruh kemungkinan yang terjadi dalam suatu percobaan

Ruang Sampel biasanya dilambangkan dengan huruf besar " S

Contoh 3.3

Pada percobaan melempar sebuah dadu, maka ruang sampelnya ditulis:

$$S = \{ 1, 2, 3, 4, 5, 6 \}$$

Contoh 3.4

Pada percobaan melempar sebuah mata uang logam

$$S = \{ \text{Angka, Gambar} \} \text{ atau } S = \{ A, G \}$$

$$S = \{ \text{Muka, Belakang} \} \text{ atau } S = \{ M, B \}$$

Kejadian adalah bagian dari ruang sampel, biasanya untuk melambangkan suatu kejadian digunakan huruf besar.

Contoh 3.5

Pada percobaan melempar sebuah dadu.

- a. Jika A adalah kejadian muncul mata dadu bilangan genap, maka:

$$A = \{ 2, 4, 6 \}$$

- b. Jika B adalah kejadian muncul mata dadu bilangan prima, maka:

$$B = \{ 2, 3, 5 \}$$

- c. Jika C adalah kejadian muncul mata dadu yang merupakan faktor dari 12, maka:

$$C = \{ 1, 2, 3, 4, 6 \}$$

Contoh 3.6

Pada percobaan melempar dua mata uang logam.

- a. Jika P adalah kejadian kedua mata uang muncul Angka, maka:

$$P = \{ AA \}$$

- b. Jika Q adalah kejadian muncul 1 Angka dan 1 Gambar, maka:

$$Q = \{ AG, GA \}$$

Latihan 3.2

1. Jika 3 buah uang logam dilempar, tentukan:
 - a. Ruang Sampel S
 - b. Kejadian R yaitu kejadian muncul semuanya gambar
 - c. Kejadian S yaitu kejadian muncul satu angka dan dua gambar
2. Jika 2 buah dadu dilempar, yaitu dadu I dan dadu II, tentukan:
 - a. Ruang Sampel S
 - b. Kejadian A yaitu kejadian muncul jumlah kedua mata dadu sama dengan 7
 - c. Kejadian B yaitu kejadian muncul mata dadu I angka 2

Definisi 3.6 (Definisi Peluang Klasik)

Misalkan sebuah peristiwa A dapat terjadi sebanyak n kali diantara N peristiwa yang saling eksklusif dan masing-masing terjadi dengan kesempatan yang sama, maka peluang peristiwa E terjadi adalah n/N atau $P(A) = n/N$

Contoh 3.7:

Eksperimen dengan melantunkan koin Rp 1000,- sebanyak satu kali menghasilkan peristiwa-peristiwa yang terjadi :

- 1) muncul angka (G) = 1
 - 2) muncul gambar (A) = 1
- $N = 2$

$P(G) = \frac{1}{2}$; $P(A) = \frac{1}{2}$

Contoh 3.8:

Eksperimen dengan melantunkan dadu satu kali menghasilkan peristiwa, peristiwa yang terjadi :

- 1) muncul mata dadu 1 = 1
- 2) muncul mata dadu 2 = 1

- 3) muncul mata dadu 3 = 1
- 4) muncul mata dadu 4 = 1
- 5) muncul mata dadu 5 = 1
- 6) muncul mata dadu 6 = 1

$N = 6$

$P(MD1) = 1/6$; $P(MD2) = 1/6$

Contoh 3.9:

Eksperimen mengambil sebuah bola kecil dalam kotak yang berisi 2 merah, 8 hitam, 6 putih dan 4 kuning

Peristiwa yang terjadi :

- terambil bola M = 2
- terambil bola H = 8
- terambil bola P = 6
- terambil bola K = 4

$N = 20$

$P(M) = 2/20$; $P(K) = 4/20$

Sifat peluang klasik : saling eksklusif dan kesempatan yang sama

Definisi 3.7 (Definisi peluang empirik)

Peluang empirik/frekuensi relatif terjadi apabila eksperimen dilakukan berulang. Apabila kita perhatikan frekuensi absolut ($=m$) tentang terjadinya peristiwa A untuk sejumlah pengamatan ($=n$), maka peluang peristiwa itu adalah limit dari frekuensi relatif apabila jumlah pengamatan bertambah sampai tak hingga

$$P(A) = \lim_{n \rightarrow N} \frac{m}{n}$$

Contoh 3.10

Eksperimen melantunkan sebuah dadu (1000X)

Peristiwa yang muncul : - muncul mata dadu 1 hingga muncul mata dadu 6

Kejadian	M1	M2	M3	M4	M5	M6	total
m	166	169	165	167	169	164	1000

$$P(M1) = 166/1000 ; P(M6) = 164/1000$$

Latihan 3.3

Lakukan percobaan di bawah ini dengan kelompokmu !

- Melempar sebuah uang logam sebanyak: 25 kali, 30 kali, 50 kali, dan 100 kali
Kemudian hitung peluang secara frekuensi relatif munculnya gambar!
- Melempar sebuah dadu sebanyak 10 kali, kemudian hitung peluang secara frekuensi relatif
 - munculnya mata dadu bilangan prima
 - munculnya mata dadu 5
 - munculnya mata dadu 2

Definisi 3.8 (Aksioma Peluang)

Setiap kejadian di ruang sampel dikaitkan dengan bilangan antara 0 dan 1, bilangan ini disebut peluang.

- Kejadian yang tak mungkin terjadi mempunyai peluang nol
- Kejadian yang pasti terjadi mempunyai peluang satu
- Peluang dari kejadian A bernilai antara 0 dan 1
- Jika A dan B dua kejadian sehingga $A \cap B = \emptyset$, maka $P(A \cup B) = P(A) + P(B)$
- Jika A dan B dua kejadian sehingga $A \cap B \neq \emptyset$, maka $P(A \cup B) = P(A) + P(B) - P(A \cap B)$

Latihan 3.4

- Sebuah dadu dilempar 100 kali. Hasil lemparan dicatat dalam bentuk tabel sbb:

Muncul mata dadu	1	2	3	4	5	6
Frekuensi	14	17	20	18	15	16

Tentukan frekuensi relatif muncul mata dadu 3

- Tentukan frekuensi relatif muncul mata dadu 1
 - Tentukan frekuensi relatif muncul mata dadu bilangan genap
 - Tentukan frekuensi relatif muncul mata dadu bilangan prima
- Seorang dokter menggunakan obat Y untuk penyakit Z dengan peluang 0,8. Tentukan jumlah orang yang diharapkan sembuh jika ia menggunakan obat Y untuk penyakit Z pada 300 orang
 - Dua buah dadu dilantunkan secara bersama-sama. Tentukan peluang:
 - Jumlah mata dadu yang muncul 7
 - Dadu I muncul mata dadu 2 dan dadu II muncul mata dadu 3
 - Dadu I muncul mata dadu 2 atau dadu II muncul mata dadu 5

Kejadian Majemuk

Sifat 1 : Misalkan A dan B dua kejadian pada ruang sampel dengan $A \cap B = \emptyset$,

$$\text{maka : } P(A \cup B) = P(A) + P(B)$$

Sifat 2 : Misalkan A dan B dua kejadian pada ruang sampel dengan $A \cap B \neq \emptyset$,

$$\text{maka : } P(A \cup B) = P(A) + P(B) - P(A \cap B)$$

Peluang Komplemen suatu Kejadian

Sifat : Misalkan A kejadian pada ruang sampel, maka $P(A^c) = 1 - P(A)$

3.4 Kejadian Bersyarat

Misalkan sebuah dadu seimbang dilemparkan dan Anda diminta untuk menentukan peluang bahwa yang muncul mata dadu lima. Ada enam kemungkinan hasil yang sama, jadi jawaban Anda adalah 1/6. Tapi anggaplah bahwa sebelum Anda memberikan jawaban Anda, Anda diberi informasi tambahan bahwa mata dadu yang dilemparkan itu mata dadu ganjil. Karena hanya ada tiga angka ganjil yang mungkin, salah satunya adalah lima, Anda pasti akan merevisi perkiraan kemungkinan bahwa peluang munculnya mata dadu lima dari 1/6 sampai 1/3. Secara umum, peluang yang direvisi bahwa suatu kejadian A telah terjadi, dengan mempertimbangkan informasi tambahan bahwa kejadian lain B telah pasti terjadi pada percobaan ini, disebut peluang bersyarat dari A yang diberikan B dan dilambangkan dengan $P(A | B)$.

Definisi 3.9

Dua kejadian A dan B pada ruang sampel dikatakan kejadian bersyarat yaitu Kejadian B terjadi dengan syarat kejadian A terjadi lebih dahulu atau B/A, maka peluangnya adalah:

$$P(A|B) = \frac{P(A \cap B)}{P(B)}$$

atau

$$P(A \cap B) = P(A) \cdot P(B/A)$$

Kejadian Saling Bebas

Definisi 3.10

Dua kejadian A dan B pada ruang sampel dikatakan saling bebas jika

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

Contoh 3.11

Dalam sebuah penelitian tentang bahaya rokok terhadap kesehatan diketahui ada 135 orang penderita kanker. Di antara 135 orang penderita aner tersebut ada 122 orang perokok berat. Berapakah peluang yang diteliti adalah perokok berat jika diketahui pasien tersebut adalah perokok berat?

Penyelesaian

Misalkan

A = penderita kanker

B = perokok berat

$$n(A) = 135, n(A \cap B) = 122$$

$P(B/A) ?$

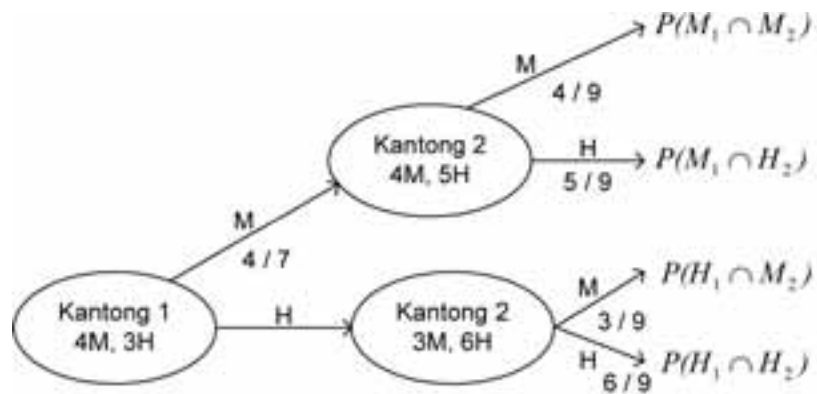
$$P(A) = \frac{n(A)}{N}, P(A \cap B) = \frac{n(A \cap B)}{N}$$

$$P(B/A) = \frac{P(A \cap B)}{P(A)} = \frac{n(A \cap B)/N}{n(A)/N} = \frac{122}{135} = 0,90$$

Contoh 3.12

Suatu kantong berisi 4 bola merah dan 3 bola hitam dan kantong kedua berisi 4 bola merah dan 5 bola hitam. Satu bola diambil dari kantong pertama dan dimasukkan tanpa melihatnya ke kantong kedua. Berapa peluangnya mengambil bola hitam dari kantong kedua?

Diagram pohonnya:



$$P(M_1 \cap H_2) + P(H_1 \cap H_2) = \frac{38}{63}$$

- H_1 = mengambil 1 bola hitam dari kantong 1
- H_2 = mengambil 1 bola hitam dari kantong 2
- M_1 = mengambil 1 bola merah dari kantong 1
- M_2 = mengambil 1 bola merah dari kantong 2

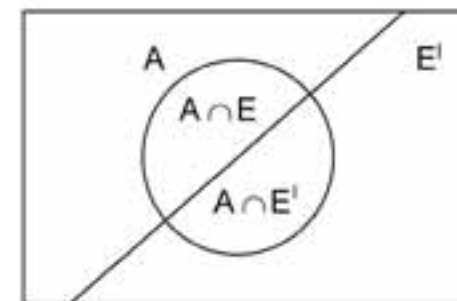
Latihan 3.5

1. Suatu pengiriman 10 pesawat TV 3 diantaranya dinyatakan cacat. Berapakah peluang sebuah hotel membeli 4 pesawat TV tersebut dan 2 TV ternyata cacat?
2. Tiga buah buku diambil secara acak dari suatu rak yang berisi empat novel, tiga buku syair dan sebuah kamus. Berapakah peluang
 - a. kamus terpilih?
 - b. dua novel dan sebuah buku syair yang terpilih?
3. Dua kartu diambil secara berturut-tan tanpa dikembalikan dari suatu kotak kartu bridge. Berapakah peluang kartu yang terpilih lebih besar dari 2 tetapi lebih kecil dari 9?

4. Bila A dan B dua kejadian yang saling asing dengan $P(A) = 0,4$ dan $P(B) = 0,5$, hitunglah:
 - a. $P(A \cup B)$
 - b. $P(A^c)$
 - c. $P(A^c \cap B)$
5. Dalam sebuah kotak berisi 15 telur 5 telur diantaranya rusak. Untuk memisahkan telur baik dan telur yang rusak dilakukan pengetesan satu persatu. Berapakah peluang diperoleh telur rusak ke 3 pada pengetesan ke 5?

Teorema Bayes

Teorema Bayes merupakan suatu formula untuk menghitung peluang bersyarat $P(A|B)$ dengan adanya informasi $P(B|A)$. Perhatikan diagram Venn berikut tentang dua kejadian A dan E yang dipartisi.



$$A = (A \cap E) \cup (A \cap E')$$

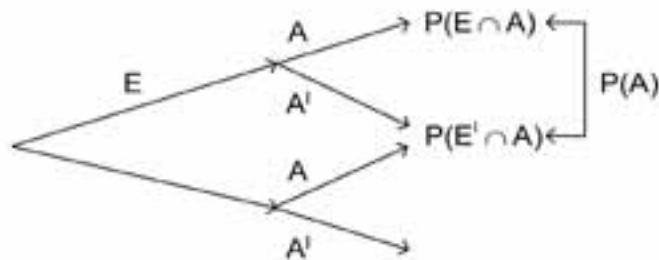
$$P(A) = P(A \cap E) + P(A \cap E')$$

Diagram Venn untuk kejadian A, E, E'

$$P(E|A) = \frac{P(A \cap E)}{P(A)} = \frac{P(A \cap E)}{P(A \cap E) + P(A \cap E')}$$

$$P(E|A) = \frac{P(E)P(A|E)}{P(E)P(A|E) + P(E^c)P(A|E^c)}$$

Diagram pohonnya



Bentuk umumnya:

B_1, B_2, \dots, B_k kejadian yang dipartisi dari ruang sampel dengan $P(B_i) \neq 0, i = 1, 2, \dots, k$ maka setiap kejadian A anggota dari ruang sampel

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

Bila $P(A) \neq 0$ maka teorema Bayes :

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

untuk $k=3$; B_1, B_2, B_3 dan A

Contoh 3.13

Misalkan suatu sekolah menerapkan suatu program pengujian narkoba yang baru. Nama siswa diambil secara acak dari catatan pendaftaran, dan siswa yang dipilih harus mengikuti tes narkoba. Diketahui bahwa 5% dari semua siswa menggunakan narkoba. Juga diketahui bahwa tingkat positif palsu tes narkoba adalah 1% (tingkat positif palsu menunjukkan penggunaan narkoba padahal sebenarnya tidak terjadi). Tingkat negatif palsu adalah 2% (negatif palsu berarti tidak menunjukkan penggunaan narkoba meskipun

hal itu terjadi). Jika seorang siswa melakukan tes narkoba dan terjadi hasilnya positif, apa kemungkinan dia benar-benar menggunakan narkoba? Kebanyakan orang akan berkata, "99 persen." Tapi jawaban itu belum tentu benar.

Penyelesaian :

Misalkan B = siswa memakai narkoba

A = hasil pengujian positif

Akan dihitung $P(B|A)$. Diketahui bahwa $P(B) = 0,05, P(\sim B) = 0,95, P(A|B) = 0,98,$ dan $P(A|\sim B) = 0,01.$

Dengan menggunakan Teorema Bayes

Pembilang = $P(A|B) \cdot P(B) = (0,98)(0,05) = 0,049$

Penyebut = $P(A|B) \cdot P(B) + P(A|\sim B) \cdot P(\sim B) = (0,98)(0,05) + (0,01)(0,95) = 0,0585$

$P(B|A) = \text{numerator/denominator} = 0,049 / 0,0585 = 0,848,$ or 84.8%

Artinya, kemungkinan seorang siswa yang dipilih secara acak yang dinyatakan positif menggunakan obat-obatan yang benar-benar menggunakan narkoba hanya sekitar 85%, bukan 99% yang beberapa orang anggap secara naif.

Hasilnya bahkan lebih dramatis dengan tes untuk penyakit, seperti HIV, di mana sebagian kecil masyarakat yang memiliki penyakit (yaitu, $P(B)$) sangat kecil. Peluang bahwa orang yang diuji secara acak (yaitu seseorang yang diuji tanpa alasan tertentu untuk berpikir bahwa dia telah terpapar) sebenarnya memiliki penyakit ini bisa serendah 50%, atau bahkan kurang.

Untuk melihat logikanya, misalkan kita memiliki populasi 10.000 orang. Diketahui bahwa 1 dari 200 orang (yaitu total 50 orang) memiliki penyakit ini. Dan misalkan tes untuk penyakit ini memiliki akurasi 98% (false positive dan false

negative rate 2%). Dari 50 penderita penyakit ini, (0.98) (50) = 49 akan positif. Dari 9.550 orang tanpa penyakit ini, (0,02) (9550) = 199 akan melakukan tes positif. Itu sebanyak 248 hasil tes positif, tapi hanya 49 dari orang-orang yang benar-benar memiliki penyakit ini. Itu hanya 19,8%. Sisanya 80,2% dari mereka yang dinyatakan positif bebas penyakit.

Latihan 3.6

- Sebuah percobaan terdiri atas pelemparan sebuah dadu kemudian mengundi sebuah koin sekali bila nomor pada dadu yang muncul adalah genap. Bila nomor pada dadu adalah ganjil, koin diundi dua kali. Dengan menggunakan notasi 4 M, misalnya untuk menunjukkan kejadian dimana dadu muncul 4 dan kemudian koin memunculkan sisi M dan 3 MB untuk menunjukkan kejadian dimana dadu memunculkan 3 diikuti oleh sisi M = muka, kemudian sisi B = belakang.
Buatlah sebuah bilangan pohon untuk menunjukkan 18 unsur dari ruang contoh S.
- Pada suatu penelitian lingkungan yg menyelidiki pengaruh gas radon pada penduduk, diambil 100 sampel (contoh) sel darah untuk dibiakan dan kemudian diamati kelainan sifat khromosomnya. Bila W_i menyatakan kelainan khromosomnya. Maka ruang contoh kejadiannya $\Omega = \{W_0, W_1, W_2, W_3, W_4, W_5\}$ dengan peluang

$$P(W_i) = \frac{C(i+1)^2}{2^{i+1}} \quad i = 0,1,2,3,4,5$$

- Tentukan nilai konstanta C
- Hitunglah peluang bahwa seorang yang dipilih secara acak mempunyai 3 kelainan.

3. A dan B adalah kejadian dengan peluang di-ketahui $P(A) = 0,7, P(B) = 0,4$ dan $P(A \cap B) = 0,2$, hitunglah peluang-peluang berikut:

- | | | |
|---------------------|-----------------------|-----------------------|
| (a) $P(A \cup B)$ | (b) $P(A^1 \cap B^1)$ | (c) $P(A^1 \cup B^1)$ |
| (d) $P(A^1 \cap B)$ | (e) $P(A \cap B^1)$ | (f) $P(A^1 \cup B)$ |
| (g) $P(A \cup B^1)$ | (h) $P(A \cap B)^1$ | (i) $P(A \cup B)^1$ |

4. Diketahui $P(A_1) = 0,5, P(A_2) = 0,4, P(A_3) = 0,4$
 $P(A_1 \cap A_2) = 0,04, P(A_1 \cap A_3) = 0,01, P(A_2 \cap A_3) = 0,2,$
 $P(A_1 \cap A_2 \cap A_3) = 0,02.$

Gunakan sistem informasi ini dan diagram venn untuk menghitung peluang berikut:

- | | |
|----------------------------------|--------------------------------------|
| (a) $P(A_1 \cup A_2 \cup A_3)$ | (b) $P(A_1^1 \cap A_2^1 \cap A_3^1)$ |
| (c) $P(A_1 \cup A_2) \cap A_3$ | (d) $P(A_1 \cup A_2) \cap A_3^1$ |
| (e) $P(A_1^1 \cap A_2 \cap A_3)$ | |

5. Diketahui $P(A) = 0,7$ dan $P(B) = 0,8$. Tunjukkan bahwa $P(A \cap B) \geq 0,5$

6. Sebuah kotak berisi 500 amplop dimana 75 diantaranya berisi \$100 tunai, 150 amplop berisi \$25, dan 275 berisi \$10. Sebuah amplop dapat dibeli dengan harga \$25 buatlah ruang contoh untuk uang yang berbeda, tentukan peluang bagi titik-titik contoh itu kemudian carilah peluang bahwa amplop pertama yang dibeli berisi kurang dari 100.

7. Dari 100 siswa yang diwisuda, 54 belajar matematika, 69 belajar sejarah, 39 belajar matematika dan sejarah. Bila seorang siswa di-pilih secara acak, hitunglah peluangnya.

- Dia belajar matematika atau sejarah
- Dia tidak belajar keduanya
- Dia belajar sejarah tapi tidak matematika

8. Peluang suatu industri akan membangun pabriknya di Bekasi 0,7 peluang, membangun pabriknya di Bandung 0,4, dan peluang membangun di Bekasi atau di Bandung atau kedua-duanya 0,5. Berapakah pabrik itu dibangun.
9. Dari pengalaman yg lalu seorang pialang saham yakin bahwa dalam keadaan ekonomi yang sekarang langganan akan menanam modalnya dalam obligasi bebas pajak dengan peluang 0,6, dalam dana bersama (Mutual Funds) dengan peluang 0,3 dan dalam keduanya dengan peluang 0,15. Pada keadaan sekarang, carilah peluang seorang langganan akan menanam modalnya.
- Dalam obligasi bebas pajak atau dana ber-sama
 - Tidak dalam salah satupun dari keduanya.
10. Di suatu penjara, ternyata $\frac{2}{3}$ dari penghuninya berumur dibawah 25 tahun. Selain itu diketahui bahwa $\frac{3}{5}$ bagian perempuan atau yang berumur 25 tahun atau lebih. Bila kita mengambil seseorang secara acak dari penjara ini berapa peluang bahwa ia berjenis kelamin perempuan dan berumur sekurang-kurangnya 25 tahun.
11. Peluang sebuah pompa bensin kedatangan 0, 1, 2, 3, 4 atau 5 atau lebih mobil selama periode 30 menit tertentu adalah 0,03; 0,18; 0,24; 0,28; 0,10 dan 0,17 hitunglah peluang bahwa dalam periode 30 menit ini
- Pompa bensin kedatangan lebih dari 2 mobil
 - Pomp bensin itu kedatangan sebanyak-banyaknya 4 mobil.
 - Pompa bensin itu kedatangan 4 atau lebih mobil.

12. Suatu contoh acak 200 orang dewasa di-klasifikasikan dibawah ini menurut jenis kelamin dan tingkat pendidikan.

Pendidikan	Laki-laki	Perempuan
SD	38	45
SM	28	50
PT	22	17

Bila seseorang diambil secara acak dari kelompok ini, hitunglah peluang bahwa.

- Yang terpilih tersebut laki-laki, bila diketahui ia berpendidikan sekolah menengah
 - Yang terpilih bukan dari perguruan tinggi, bila diketahui ia perempuan.
13. Diantara 100 siswa kelas tiga sebuah sekolah menengah atas, 42 mempelajari matematika, 68 mempelajari psikologi, 54 mempelajari sejarah, 22 mempelajari matematika dan sejarah, 10 mempelajari ketiganya, dan 8 tidak mempelajari satupun dari ketiga diatas. Bila seorang siswa mengambil secara acak hitung peluang bahwa.
- Seorang yang mempelajari psikologi mempelajari ketiganya.
 - Seorang yang tidak mempelajari psikologi mempelajari baik sejarah maupun mate-matika
14. Peluang sebuah mobil yang diisi bensin juga memerlukan pergantian oli adalah 0,25 peluang bahwa mobil itu memerlukan oli maupun penyaring oli yang baru adalah 0,14.
- Bila oli harus diganti, berapa peluang penyaring baru juga diperlukan

- (b) Bila penyaring baru juga diperlukan berapa peluang olinya juga harus diganti.
15. Peluang seorang dokter mendiagnosis suatu penyakit secara benar adalah 0,7. Bila diketahui dokter tersebut salah mendiagnosis, bahwa pasien akan menuntut kepengadilan adalah 0,9 berapakah seorang dokter tersebut salah men-diagnosis dan pasien menuntutnya?
16. Seorang pengusaha perumahan (real estate) mempunyai 8 kunci induk untuk membuka beberapa rumah baru. Suatu rumah hanya akan dapat dibuka dengan satu kunci induk tertentu. Bila 40% dari rumah biasanya tak terkunci, berapakah peluang pengusaha tersebut dapat masuk ke sebuah rumah tertentu bila dia meng-ambil tiga kunci induk secara acak sebelum meninggalkan kantornya?
17. Sebuah kota mempunyai dua mobil pemadam kebakaran yang bekerja saling bebas. Peluang suatu mobil tertentu tersedia bila diperlukan adalah 0,96
- Berapakah peluang keduanya tidak tersedia bila diperlukan
 - Berapakah peluang satu mobil tersedia bila diperlukan
18. Satu tas berisi 2 botol aspirin dan 3 botol obat masuk angin. Tas kedua berisi 3 botol aspirin, 2 botol obat masuk angin, dan 1 botol obat reumatik. Bila satu botol diambil acak dari tiap tas, cari peluangnya bahwa:
- Kedua botol berisi obat masuk angin
 - Tidak ada botol yg berisi obat masuk angin
 - Kedua botol berisi obat yang berlainan

19. Seorang pegawai mempunyai 2 mobil, satu sedan satu lagi Toyota kijang, untuk pergi bekerja dia menggunakan sedan, biasanya dia tidak kembali di rumah pukul 17.30 sebanyak 75% (75 dari 100 kali) sedangkan bila menggunakan kijang dia tiba pukul 17.30 kira-kira 60% (tapi dia merasa lebih tenang memakai kijang karena tidak terlalu khawatir diserempet mobil lain). Bila dia tiba dirumah pukul 17.30 berapa peluang dia memakai sedan.
20. Perusahaan taksi sigma di Jakarta mengklasi-fikasikan pengemudinya menjadi 3 kelas, yaitu kelas A = (baik), kelas B = (cukup), kelas C = (kurang). Klasifikasi ini berdasarkan laporan yang masuk mengenai kecelakaan-kecelakaan yang dialami para pengemudi tersebut. Dari waktu yang lampau peluang bahwa pengemudi kelas A mengalami kecelakaan satu kali dalam satu bulan = 0,02, pengemudi kelas B = 0,04 dan pengemudi C = 0,08. Berdasarkan catatan yang ada pada perusahaan, 70% pengemudi kelas A, 20% kelas B dan 10% pengemudi kelas C. Pada suatu ketika seorang pengemudi melaporkan bahwa dia mengalami kecelakaan. Berapakah peluangnya bahwa pengemudi tersebut dari kelas A.

Bab 4

DISTRIBUSI PELUANG DISKRIT

1. Pendahuluan

Distribusi peluang dapat diterapkan dalam banyak hal seperti pada kehidupan sehari-hari, kegiatan bisnis maupun pada dunia industri. Distribusi peluang berguna untuk menganalisis suatu kejadian dan memberikan keuntungan serta manfaat dalam pengaplikasiannya. Misalnya, pada suatu proses pelayanan di suatu *Bank* dapat menguji apakah dengan disediakan empat *teller*, nasabah akan menunggu lama atau kapasitas yang berlebih akan membuat boros tempat. Permasalahan ini dapat diselesaikan dengan distribusi peluang yang akan membantu Bank dalam membuat keputusan dalam menyediakan *teller*.

Distribusi peluang merupakan suatu daftar atau kumpulan dari peluang-peluang peristiwa yang mungkin terjadi. Distribusi peluang yang demikian saling berhubungan dengan semua nilai-nilai yang mungkin terjadi dan berasal dari variabel random. Variabel random adalah variabel yang nilainya merupakan suatu bilangan yang ditentukan oleh terjadinya suatu percobaan.

Fungsi distribusi peluang umumnya dibedakan menjadi distribusi peluang diskrit dan kontinyu.

2. Variabel Acak

Variabel acak merupakan parameter penting dalam sebuah distribusi peluang. Variabel acak adalah variabel yang nilainya ditentukan dari sebuah hasil percobaan. Variabel acak dinyatakan dengan huruf besar, misalnya X , sedangkan nilainya dinyatakan dengan huruf kecil misalnya x . Sebagai contoh, pada pelemparan dua koin, huruf Y menyatakan jumlah gambar yang muncul maka nilainya adalah $y = 0, 1$ dan 2 . Dari setiap nilai variabel acak yang memungkinkan akan memiliki peluang masing-masing yang disebut distribusi peluang.

Definisi 4.1

Misalkan A adalah sebuah percobaan dengan ruang sampel S . Sebuah fungsi X yang memetakan setiap anggota $a \in S$ dengan sebuah bilangan real $X(a)$ dinamakan **variabel acak**.

Contoh 4.1 :

Pada pelemparan sekeping mata uang setimbang sebanyak 3 kali, maka ruang sampelnya adalah

$$S : \{GGG, GGA, GAG, AGG, GAA, AGA, AAG, AAA\}$$

dimana $G = \text{GAMBAR}$ dan $A = \text{ANGKA}$

X : setiap sisi GAMBAR bernilai satu ($G = 1$)

$$S : \{GGG, GGA, GAG, AGG, GAA, AGA, AAG, AAA\}$$

↓	↓	↓	↓	↓	↓	↓	↓
3	2	2	2	1	1	1	0

Perhatikan bahwa $X = \{0, 1, 2, 3\}$

Nilai $x_1 = 0, x_2 = 1, x_3 = 2, x_4 = 3$

Ada dua jenis variabel acak dan dua jenis distribusi peluang: diskrit dan kontinyu.

Variabel acak diskrit, jika daerah hasil merupakan himpunan bilangan real yang terhingga.

Contoh 4.2

Pada sebuah percobaan melantunkan dua buah koin secara bersamaan, maka ruang sampel yang mungkin terjadi:

$$S = \{(A,A), (A,G), (G,A), (G,G)\}$$

Maka variabel acak X dinyatakan dengan banyaknya kemunculan angka.

- a. Untuk $a = (G,G)$ dipetakan ke nilai 0, $X(G,G) = 0$, karena titik (G,G) tidak mengandung angka sama sekali.
- b. Untuk $a = (G,A)$ dipetakan ke nilai 1, $X(G,A) = 1$, karena titik (G,A) mengandung 1 angka.

- c. Untuk $a = (A,G)$ dipetakan ke nilai 1, $X(A,G) = 1$, karena titik (A,G) mengandung 1 angka.
- d. Untuk $a = (A,A)$ dipetakan ke nilai 2, $X(A,A) = 2$, karena titik (A,A) mengandung 2 angka.

Karena daerah hasil = $\{0,1,2\}$, maka X merupakan variabel acak diskrit

Variabel acak kontinu, jika daerah hasil merupakan sebuah interval pada garis bilangan real.

Contoh 4.3

Pada sebuah percobaan memilih batu yang ada disekitar UINSU secara acak, maka ruang sampel yang mungkin terjadi (dalam gram

$$S = \{\text{batu 1, batu 2, batu 3, ..., batu 1000}\}$$

Maka peubah acak X dinyatakan dengan berat batu. Jika kita asumsikan bahwa berat batu yang terambil tidak ada yang kurang dari 10 gram dan tidak lebih dari 1000 gram maka daerah hasil = $\{x | 10 \leq x \leq 1000\}$, maka X merupakan peubah acak kontinu.

Definisi 4.2 (Distribusi Peluang Diskrit):

Misalkan X adalah **variabel acak diskrit** dengan nilai-nilainya $x_1, x_2, x_3, \dots, x_n$. Untuk setiap variabel acak tersebut memiliki nilai peluang, yaitu $P(x_i) = P(X = x_i)$. Nilai-nilai $P(x_i)$ harus memenuhi sifat-sifat sebagai berikut:

- a. $P(x_i) \geq 0$, untuk setiap i
- b. $\sum_{i=1}^n P(x_i) = 1$

Perhatikan hasil percobaan melantunkan dua buah koin. Jika kita menentukan peluang dari setiap variabel acaknya, pasangan nilai-nilai variabel acak X dengan peluang dari nilai $X, P(X = x)$ disebut distribusi peluang yang dapat digambarkan berikut.

$X = x$	0	1	2
---------	---	---	---

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

Maka ketika ditanyakan berapa peluang kemunculan angka dari hasil lantunan 2 buah koin secara bersamaan maksimal 1 adalah

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

Definisi 4.3 (Fungsi densitas):

Misalkan X adalah peubah acak kontinu yang didefinisikan atas himpunan bilang real. Sebuah fungsi disebut fungsi densitas dari variabel acak X , jika nilai-nilainya, $f(x)$, memenuhi sifat-sifat sebagai berikut:

1. $f(x) \geq 0$, untuk $x \in (-\infty, \infty)$
2. $\int_{-\infty}^{\infty} f(x) dx = 1$
3. Untuk setiap a, b dengan $-\infty < a < b < \infty$, maka:

$$P(a < X < b) = \int_a^b f(x) dx$$

Teorema 4.1:

Jika X adalah peubah acak kontinu a dan b adalah dua konstanta real dengan $a \leq b$, maka:

$$P(a \leq X \leq b) = P(a < X \leq b) = P(a \leq X < b) = P(a < X < b)$$

Contoh 4.4:

Diketahui $f(x) = \begin{cases} kx^2 & ; 0 < x < 2 \\ 0 & ; x \text{ lainnya} \end{cases}$

- a. Tentukan nilai k agar $f(x)$ merupakan fungsi densitas dari X .
- b. Hitung $P(-1 < x < 1)$

Penyelesaian:

- a. Sifat pertama dari sifat-sifat fungsi densitas dipenuhi, jika $k \geq 0$

$$\int_{-\infty}^{\infty} f(x) dx = 1 \leftrightarrow k = \frac{3}{8}$$

b. Maka peluang $-1 < X < 1$ adalah

$$P(-1 < x < 1) = \int_{-1}^1 \frac{3}{8} x^2 dx = \frac{1}{8}$$

Definisi 4.4:

Misalkan X adalah peubah acak. P didefinisikan sebagai **fungsi distribusi kumulatif** atau **fungsi distribusi** saja, dengan:

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

Maka untuk X peubah acak diskrit, fungsi distribusinya adalah:

$$F(x) = \sum_{u \leq x} p(u)$$

Contoh 4.5:

Perhatikan percobaan melantunkan 2 buah koin sekaligus. X adalah menyatakan kemunculan angka. Tentukan fungsi distribusinya.

Penyelesaian:

Untuk $x < 0$, maka $F(x) = 0$

Untuk $0 \leq x < 1$, maka $F(0) = \sum_{u \leq 0} p(u) = p(0) = \frac{1}{4}$

Untuk $1 \leq x < 2$, maka $F(1) = \sum_{u \leq 1} p(u) = p(0) + p(1) = \frac{1}{4} + \frac{1}{2} = \frac{3}{4}$

Untuk $2 \leq x$, maka

$$F(2) = \sum_{u \leq 2} p(u) = p(0) + p(1) + p(2) = \frac{1}{4} + \frac{1}{2} + \frac{1}{4} = 1$$

Jadi fungsi distribusi dari X adalah:

$$F(X) = \begin{cases} 0 & , x < 0 \\ \frac{1}{4} & , 0 \leq x < 1 \\ \frac{3}{4} & , 1 \leq x < 2 \\ 1 & , x \geq 2 \end{cases}$$

Jika kita memiliki fungsi distribusi, maka jika ditanyakan peluang;

$$P(a < X \leq b) = F_X(b) - F_X(a)$$

Contoh 4.6:

Diketahui fungsi distribusi dari peubah acak X berbentuk:

$$F_X(x) = \begin{cases} 0 & , x < 0 \\ \frac{1}{2} & , 0 \leq x < 2 \\ \frac{5}{6} & , 2 \leq x < 3 \\ 1 & , x \geq 3 \end{cases}$$

Tentukan fungsi peluangnya.

Penyelesaian:

Jika kita memperhatikan $F_X(x)$ ada 3 titik yang diskontinu yaitu, $x = 0, 2, 3$. Ketiga nilai itu merupakan nilai X yang mempunyai peluang positif.

$$p(x) = F_X(0) - F_X(0-) = \frac{1}{2}, \text{ di } x = 0$$

$$p(x) = F_X(2) - F_X(2-) = \frac{1}{3}, \text{ di } x = 2$$

$$p(x) = F_X(3) - F_X(3-) = \frac{1}{6}, \text{ di } x = 3$$

$$p(x) = 0, \text{ di } x \text{ lainnya}$$

Fungsi peluang dapat juga diperoleh dari fungsi distribusi dengan menggunakan dalil berikut.

Teorema 4.2

Jika daerah hasil dari peubah acak X terdiri dari nilai $x_1 < x_2 < x_3 < \dots < x_n$, maka:

- a. $p(x_i) = F(x_i)$ dan
- b. $p(x_i) = F(x_i) - F(x_{i-1}), i = 2, 3, \dots, n$

Definisi 4.5:

Jika X adalah peubah acak kontinu dengan fungsi densitasnya f , maka fungsi distribusinya diberikan dengan:

$$F(x) = \int_{-\infty}^x f(u)du$$

Contoh 4.7:

Misalkan fungsi densitas dari peubah acak X berbentuk:

$$f(x) = \begin{cases} \frac{3}{8}x^2 & , 0 < x < 2 \\ 0 & , x \text{ lainnya} \end{cases}$$

- a. Tentukan fungsi distribusinya dari X .
- b. Hitung $F\left(1\frac{1}{2}\right)$

Penyelesaian:

- a. Untuk $x < 0$, maka $F(x) = 0$

$$\text{Untuk } 0 \leq x < 2, \text{ maka } F(x) = \int_{-\infty}^x f(u)du = \int_0^x \frac{3}{8}u^2 du = \frac{1}{8}u^3 \Big|_0^x = \frac{1}{8}x^3$$

$$\text{Untuk } x \geq 2, \text{ maka } F(x) = \int_{-\infty}^x f(u)du = \int_0^2 \frac{3}{8}u^2 du = \frac{1}{8}u^3 \Big|_0^2 = 1$$

Jadi fungsi distribusinya dari X adalah:

$$F(x) = \begin{cases} 0 & , x < 0 \\ \frac{1}{8}x^3 & , 0 \leq x < 2 \\ 1 & , x \geq 2 \end{cases}$$

- b. Karena $1\frac{1}{2}$ terletak pada $0 \leq x < 2$, maka:

$$F\left(1\frac{1}{2}\right) = \frac{1}{8}\left(1\frac{1}{2}\right)^3 = \frac{27}{64}$$

Teorema 4.3:

Jika $f(x)$ dan $F(x)$ masing-masing merupakan nilai fungsi densitas dan nilai fungsi distribusi dari peubah acak X di x , maka:

$$P(a \leq x \leq b) = F(b) - F(a)$$

Untuk beberapa konstanta real a dan b , dengan $a \leq b$, dan:

$$f(x) = \frac{dF(x)}{dx}$$

Apabila hasil turunannya atau diferensialnya ada.

Latihan

- 1. Sebuah kotak berisi 4 bola dengan nomor 1, 2, 3, dan 4. Kemudian dua bola diambil secara acak dari kotak itu tanpa pengembalian. Jika X menunjukkan jumlah angka dari dua bola yang terambil, maka:
 - a. Tentukkan distribusi peluangnya
 - b. Hitung $P(X \leq 5)$

c. Tentukan $F(x)$

2. Misalkan peubah acak X mempunyai distribusi peluang sebagai berikut:

x	0	1	2	3	4	5
$p(x)$	k	$3k$	$3k$	k^2	$2k^2$	$6k^2 + k$

- Tentukan nilai konstanta k
 - Hitung $P(X < 4)$, $P(X \geq 4)$, dan $P(0 < X < 4)$
 - Tentukan nilai k minimum sedemikian hingga $P(X \leq k) > 0,5$
 - Tentukan fungsi distribusi dari X
3. Misalkan fungsi peluang dari peubah acak X adalah:

$$p(x) = \begin{cases} \frac{x}{10} & , x = 1,2,3,4 \\ 0 & , x \text{ lainnya} \end{cases}$$

- Tentukan $P(x = 1 \text{ atau } 2)$, $P(0,5 < X < 2,5)$
 - Tentukan fungsi distribusi dari X
4. Misalkan fungsi densitas dari peubah acak X adalah:

$$f(x) = \begin{cases} \frac{1}{4} & , -1 < x < 1 \\ 0 & , x \text{ lainnya} \end{cases}$$

- Tentukan $P(-2 < x < 0)$, $P(0 < x < 3)$
 - Tentukan fungsi distribusinya
5. Misalkan fungsi distribusi dari peubah acak X adalah:

$$F(X) = \begin{cases} 0 & , x < -1 \\ \frac{x+2}{4} & , -1 \leq x < 1 \\ 1 & , x \geq 1 \end{cases}$$

- Hitung $P(-0,5 < X < 0,5)$
- Hitung $P(X=0)$

4.3. Distribusi Peluang Diskrit

Distribusi peluang diskrit adalah suatu daftar atau distribusi di mana variabel acaknya mengasumsikan masing-masing nilainya dengan peluang tertentu. Variabel diskrit memiliki jumlah nilai kemungkinan yang terbatas atau jumlah yang tak terhingga dari nilai-nilai yang dapat dihitung. Kata "dihitung" berarti bahwa variabel acak tersebut dapat

dicacah dengan menggunakan angka 1, 2, 3, dst. Misalnya, jumlah panggilan telepon yang diterima setelah siaran TV mengudara adalah contoh variabel diskrit, karena bisa dihitung. Berikut ini akan dibahas beberapa distribusi peluang diskrit.

Distribusi Peluang Binomial

Suatu percobaan sering kali terdiri atas ulangan-ulangan dan masing-masing mempunyai dua kemungkinan hasil yang dapat diberi nama **berhasil** atau **gagal**. Misalnya saja dalam pelemparan sekeping uang logam sebanyak 5 kali, hasil setiap ulangan mungkin muncul sisi gambar atau sisi angka. Dan salah satu di antara keduanya ditentukan sebagai 'berhasil' dan yang lainnya sebagai 'gagal'. Percobaan tersebut mempunyai ciri-ciri bahwa ulangan-ulangan tersebut bersifat bebas dan peluang keberhasilan setiap ulangan tetap sama yaitu sebesar $1/2$. Percobaan semacam ini disebut **percobaan binom**.

Misalkan kita melakukan sebuah percobaan pelemparan mata uang atau dadu secara berulang-ulang sebanyak n kali. Dalam setiap pelemparan mata uang atau dadu kita selalu memiliki peluang p untuk terjadinya sebuah peristiwa katakanlah munculnya "kepala" pada mata uang atau munculnya "angka "4 pada pelemparan dadu. Dalam percobaan Bernoulli (orang yang pertama kali melakukan percobaan independensi satu peristiwa dengan peristiwa lain) peluang ini dikatakan pula sebagai peluang *sukses* dari sebuah peristiwa. Sebaliknya kita dapat menentukan peluang tidak pernah terjadinya suatu peristiwa dalam setiap percobaan atau *gagal* yaitu $q = 1 - p$. Nilai n dan p yang tidak akan pernah berubah dari satu pelemparan ke pelemparan lain ini, disebut sebagai *parameter* dari distribusi binomial.

Dalam percobaan binomial, peluang terjadinya peristiwa sukses tepat sebanyak x kali dari n percobaan dapat didekati oleh fungsi peluang :

$$f(x) = P(X = x) = \binom{n}{x} p^x q^{n-x} = \frac{n!}{x!(n-x)!} p^x q^{n-x}$$

di mana $n! = n \times (n-1) \times (n-2) \times \dots \times 1$, dan $\binom{n}{x}$ adalah koefisien binomial (lihat lampiran dalam bab ini)

Rumus di atas merupakan rumus untuk menghitung peluang terjadinya peristiwa sukses tepat sebanyak x kali dari n buah percobaan.

Secara ringkas, Percobaan Binomial adalah percobaan yang mempunyai ciri-ciri sebagai berikut:

1. Percobaan diulang n kali
2. Hasil setiap ulangan hanya dapat dikategorikan ke dalam 2 kelas;
Misal: "BERHASIL" atau "GAGAL"
("YA" atau "TIDAK"; "SUCCESS" or "FAILED")
3. Peluang keberhasilan = p dan dalam setiap ulangan nilai p tidak berubah.
Peluang gagal = $q = 1 - p$.
4. Setiap ulangan bersifat bebas satu dengan yang lain.

Contoh 4.8:

Diketahui bahwa 20% bola lampu yang diproduksi oleh sebuah mesin adalah rusak. Sebuah pemeriksaan dilakukan dengan mengambil 4 bola lampu secara acak. Dari empat bola lampu ini tentukan peluang jumlah yang rusak adalah (a) 1 bola lampu, (b) 0 bola lampu dan (c) kurang dari 2 bola lampu.

Penyelesaian:

Peluang bola lampu rusak adalah $p = 20\% = 0,2$, berarti peluang yang baik adalah $q = 1 - p = 0,80$. Misalkan X adalah variabel acak yang menunjukkan bola lampu yang rusak. Maka dengan menggunakan rumus (2) diperoleh :

$$(a) \quad P(X = 1) = \binom{4}{1} (0,2)^1 (0,8)^3 = \frac{4!}{1! \times 3!} (0,2)^1 (0,8)^3 = 0,4096$$

$$(b) \quad P(X = 0) = \binom{4}{0} (0,2)^0 (0,8)^4 = 0,4096$$

$$(c) \quad P(X < 2) = P(X = 0) + P(X = 1) = 0,4096 + 0,4096 = 0,8192$$

Contoh 4.9:

Hitunglah peluang bahwa dalam sebuah keluarga dengan 4 anak akan memiliki (a) paling sedikit satu anak laki-laki, (b) paling sedikit satu anak laki-laki dan paling sedikit satu anak perempuan. Anggaplah peluang melahirkan anak laki-laki dan perempuan adalah sama yaitu $\frac{1}{2}$.

Penyelesaian:

Misalkan X adalah variabel acak yang menunjukkan jumlah anak laki-laki. Dari persoalan di atas diketahui bahwa $n = 4$, $p = 0,5$ dan $q = 1 - 0,5 = 0,5$.

Dengan menggunakan rumus peluang binomial bisa diperoleh :

$$(a) \quad P(X \geq 1) = P(X = 1) + P(X = 2) + P(X = 3) + P(X = 4)$$

$$P(X = 1) = \binom{4}{1} (0,5)^1 (0,5)^3 = \frac{1}{4} \quad ;$$

$$P(X = 2) = \binom{4}{2} (0,5)^2 (0,5)^2 = \frac{3}{8}$$

$$P(X = 3) = \binom{4}{3} (0,5)^3 (0,5)^1 = \frac{1}{4} \quad ;$$

$$P(X = 4) = \binom{4}{4} (0,5)^4 (0,5)^0 = \frac{1}{16}$$

Jadi :

$$P(\text{paling sedikit satu anak laki-laki}) = P(X = 1) + P(X = 2) + P(X = 3) + P(X = 4) \\ = \frac{1}{4} + \frac{3}{8} + \frac{1}{4} + \frac{1}{16} = \frac{15}{16}$$

Cara lain :

$$P(X \geq 1) = 1 - P(X = 0) = 1 - \binom{4}{0}(0,5)^0(0,5)^4 = 1 - (0,5)^4 = \frac{15}{16}$$

(b) Peluang paling sedikit satu anak laki-laki dan paling sedikit satu anak perempuan

$$P(X \geq 1 \cap P(X = 4)) = 1 - P(X = 0) - P(X = 4) \\ = 1 - \frac{1}{16} - \frac{1}{16} = \frac{7}{8}$$

Beberapa sifat distribusi binomial

Beberapa sifat penting dari distribusi binomial adalah sebagai berikut :

Rata-rata	$\mu = np$
Varians	$\sigma^2 = npq$
Simpangan Baku	$\sigma = \sqrt{npq}$

Penggunaan dari sifat-sifat penting di atas dapat dilihat dalam contoh berikut ini.

Contoh 4.10.

Peluang rusaknya sebuah bola lampu adalah 0,1. Apabila diambil 400 bola lampu, berapa diharapkan dijumpai bola lampu yang rusak dan berapa variansnya.

Penyelesaian :

$$\mu = np = (400)(0,1) = 40 \text{ buah bola lampu.}$$

$$\sigma^2 = npq = (400)(0,1)(0,9) = 36$$

Untuk mempermudah perhitungan, telah tersedia tabel binomial untuk berbagai nilai n , x dan p . (lihat lampiran).

DISTRIBUSI POISSON

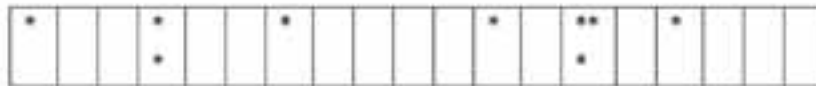
Di dunia nyata banyak peristiwa terjadi secara acak dengan perkataan lain jarang sekali kita dapat memprediksi secara tepat kapan sebuah peristiwa akan terjadi atau berapa banyak peristiwa akan terjadi secara bersamaan. Misalnya saja peristiwa terjadinya kecelakaan di jalan tol, kegagalan dalam proses produksi berteknologi tinggi seperti mobil, pesanan melalui telepon, atau kecelakaan pesawat ulang alik. Contoh lain yang dapat menggambarkan situasi ini secara tepat adalah peluruhan zat radioaktif. Tidak ada seorang pun yang dapat memperkirakan kapan atom akan luruh berikutnya atau kapan peristiwanya akan terjadi. Peristiwa-peristiwa yang terjadi secara acak dan langka dalam dimensi waktu dan ruang seperti ini, dalam jangka pendek memang sangat sulit untuk diprediksi. Namun yang mengejutkan, untuk jangka waktu yang panjang peristiwa-peristiwa ini mempunyai kesamaan bahwa kesemuanya dapat diprediksi secara akurat. Salah satu distribusi peluang yang dapat menjelaskan secara tepat peristiwa-peristiwa seperti adalah apa yang dikenal sebagai distribusi Poisson.

Distribusi Poisson merupakan distribusi peluang diskrit yang cukup memegang peranan penting dalam ilmu manajemen. Distribusi ini ditemukan oleh S.D. Poisson di awal abad ke 19. Seperti distribusi

binomial, distribusi Poisson juga termasuk ke dalam proses Bernoulli, akan tetapi tidak ada konsep yang membedakan secara jelas dalam percobaan Poisson. Olehkarenanya syarat-syarat untuk menggunakan distribusi ini tidak berbeda jauh dengan distribusi binomial, diantaranya :

- Proses yang diamati harus berbentuk "dua-peristiwa" atau proses Bernoulli
- Harus ada bilangan rata-rata dari peristiwa tertentu per pengamatan/pengukuran baik waktu maupun ruang, yang tidak berubah selama terjadinya proses
- Proses haruslah bersifat kontinu artinya tidak ada percobaan tunggal

Untuk memperjelas kita lihat contoh berikut. Misalkan dalam proses pembuatan bahan pakaian atau tekstil kadang-kadang ditemukan kain yang cacat (goresan atau sobek). Jadi dalam bahan pakaian ini hanya bisa dijumpai dua kejadian yaitu cacat atau tidak (proses Bernoulli). Sudah barang tentu yang dapat dihitung adalah jumlah cacat sedangkan yang tidak cacat adalah mustahil untuk dihitung demikian pula dengan jumlah peristiwa terjadinya cacat. Olehkarenanya tidak ada istilah peluang kain cacat untuk hal seperti ini, akan tetapi yang ada hanyalah rata-rata cacat per unit area, misalnya saja 3 cacat per meter persegi. Sebagai gambaran kita lihat contoh secara visual berikut ini.



Anggaplah tanda * merupakan cacat yang terdapat setiap meter persegi kain. Maka yang dapat kita hitung dari kain yang diproduksi adalah rata-rata cacat yang ditemukan dalam setiap ukuran luas.

Karena tidak adanya percobaan yang jelas dalam proses Poisson, maka dalam distribusi ini tidak ada parameter n dan p seperti distribusi binomial. Yang ada hanyalah satu parameter λ (baca lambda) yaitu rata-rata jumlah kejadian per satuan ukuran seperti jarak, area atau

volume. Dengan adanya nilai rata-rata kejadian ini, maka jumlah kejadian yang aktual atau sesungguhnya merupakan variabel acak, katakanlah X . Kita ambil contoh proses pembuatan kain di atas, $X = 0$, berarti tidak ada cacat, $X = 1$ terdapat 1 cacat, $X = 2$ terdapat 2 cacat dan seterusnya. Contoh distribusi peluang Poisson untuk pembuatan kain dengan $\lambda = 3$ diberikan dalam Tabel 5 berikut.

Tabel 4.1 Distribusi peluang Poisson untuk $\lambda = 3$

X	$P(X \lambda=3)$	X	$P(X \lambda=3)$
0	0.4980	7	0.0216
1	0.1494	8	0.0081
2	0.2240	9	0.0027
3	0.2240	10	0.0008
4	0.1680	11	0.0002
5	0.1008	12	0.0001
6	0.0504		

Sumber : hipotetis

Dari tabel di atas dapat kita baca bahwa peluang tidak ditemuinya cacat per meter persegi adalah 0,4980, peluang ditemuinya satu cacat adalah 0,1494 dan seterusnya.

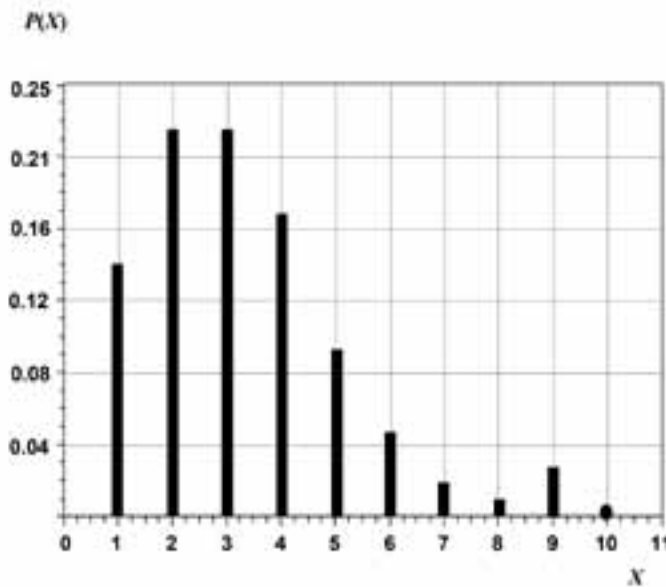
Distribusi binomial mempunyai batas atas bagi variabel acaknya artinya x tidak bisa melebihi n , sedangkan dalam distribusi Poisson nilai n tidak terhingga artinya secara teoritis x tidak mempunyai batas atas. Dalam prakteknya, seseorang biasanya mengabaikan hal yang demikian dan umumnya akan mengambil nilai X yang memiliki peluang lebih kecil dari 0,0001 seperti yang ditunjukkan dalam Tabel 4.1. Sedangkan secara visual, distribusi peluang dari peristiwa kain cacat di atas dapat dilihat dalam Gambar 1. Dari gambar ini jelas bahwa peluang terjadinya peristiwa semakin mendekati nol untuk jumlah cacat yang semakin banyak.

Dalam proses jangka panjang, distribusi Poisson dapat dijelaskan dalam bentuk fungsi distribusi peluang sebagai berikut :

$$P(X = c) = \frac{e^{-\lambda} \lambda^c}{c!}$$

di mana e adalah bilangan eksponensial 2,71828, c adalah bilangan bulat positif, sedangkan λ tidak harus bilangan bulat asalkan tidak negatif.

Untuk menghitung nilai-nilai peluang distribusi Poisson dalam rumus di atas sudah barang tentu cukup melelahkan dan rumit apalagi jika $X > 10$. Oleh karena itu untuk mempermudah perhitungan telah tersedia tabel peluang Poisson untuk berbagai nilai λ yang dapat dilihat dalam buku-buku statistika lanjutan.



Gambar 4.1 Distribusi peluang Poisson untuk $\lambda = 3$

Contoh 4.11

Berdasarkan pengalaman tahun lalu, sebuah perusahaan sewa kendaraan menerima pesanan rata-rata 6,7 kendaraan per hari. Pertanyaannya adalah berapa kendaraan harus dioperasikan?

Penyelesaian :

Misalkan variabel acak X adalah jumlah kendaraan yang dipesan dalam sehari. Maka dengan menggunakan rumus peluang distribusi Poisson diperoleh :

$$P(X = 0) = \frac{e^{-6,7} (6,7)^0}{0!} = e^{-6,7} = 0,00123$$

$$P(X = 1) = \frac{e^{-6,7} (6,7)^1}{1!} = 6,7 \times e^{-6,7} = 0,00825$$

$$P(X = 2) = \frac{e^{-6,7} (6,7)^2}{2!} = \frac{e^{-6,7} \times (6,7)^2}{2!} = 0,02763$$

Dari perhitungan di atas dapat kita lihat bahwa peluang tidak adanya orang yang menyewa kendaraan adalah 1 hari dalam 1000 kesempatan (0,001), atau peluang 1 orang menyewa adalah 8 hari dalam 1000 (0,00825), dan hanya 28 hari dari 1000 kesempatan 2 orang akan menyewa (0,002763). Perhitungan ini dapat kita lanjutkan dengan menggunakan tabel Poisson, sehingga diperoleh distribusi peluangnya sebagai berikut :

X	$P(X \lambda = 6,7)$	$P(X \leq c)$	X	$P(X \lambda = 6,7)$	$P(X \leq c)$
0	0,00123	0,00133	8	0,12397	0,76728
1	0,00825	0,00948	9	0,09229	0,85957
2	0,02763	0,03711	10	0,06183	0,92140
3	0,06170	0,09881	11	0,03765	0,95906
4	0,10335	0,20216	12	0,02103	0,98009
5	0,13849	0,34865	13	0,01084	0,99093
6	0,15465	0,49530	14	0,00519	0,99611
7	0,14802	0,64332	15	0,00232	0,99843

Apa yang dapat dimanfaatkan oleh seorang manajer dari tabel di atas.

Untuk mempermudah perhitungan selanjutnya, dalam tabel tersebut ditambahkan pula satu kolom yang memperlihatkan peluang kumulatif hingga mendekati nilai 1. Kolom ini nanti digunakan untuk

melihat peluang variabel acak X hingga mencapai sesuatu nilai atau $P(X \leq c)$.

Dengan adanya tabel di atas, maka dasar untuk pengambilan keputusan manajerial sudah lengkap. Seandainya manajer memutuskan untuk memenuhi permintaan sekitar 98% sepanjang harinya, maka dia akan membutuhkan 12 kendaraan karena $P(X \leq 12) = 0,98$.

Rata-rata dan Simpangan Baku

Seperti distribusi peluang lainnya, distribusi Poisson juga memiliki rata-rata, varians dan simpangan baku dengan mengambil bentuk :

$$\begin{aligned} \mu &= \lambda \\ \sigma^2 &= \lambda \\ \sigma &= \sqrt{\lambda} \end{aligned}$$

Pendekatan Poisson ke Binomial

Apabila jumlah pengamatan, n , sangat besar untuk tabel binomial, maka distribusi ini bisa didekati oleh distribusi Poisson jika nilai p cukup kecil. Kaidah yang umum sudah diterima dalam pendekatan ini adalah :

$$np \leq 5$$

Sebagai gambaran kita ambil contoh sebagai berikut. Misalkan kita ingin menghitung peluang menyalanya tepat 98 buah dari 100 bola lampu jika peluang setiap bola lampu menyala adalah 0,99. Dalam hal ini yang ingin dihitung adalah $P(x = 98 | n = 100, p = 0,99)$. Namun tabel binomial tidak menyediakan perhitungan hingga $n = 100$.

Secara sepiantas masalah ini tidak dapat didekati oleh Poisson karena $np = 99$ yang ternyata lebih besar dari 5. Akan tetapi jumlah bola lampu tepat menyala (sukses) sebesar 98 sebenarnya sama dengan 2 bola lampu tidak menyala (gagal) dan peluang sukses yang 0,99 adalah sama dengan peluang gagal sebesar 0,01. Dengan demikian peluangnya dapat

dinyatakan dalam bentuk $P(x = 2 | n = 100, p = 0,01)$. Karena np hanya 1, maka pendekatan Poisson bisa dilakukan.

Untuk mengubah distribusi binomial ke dalam Poisson hanya diperlukan substitusi rata-rata binomial np ke dalam rata-rata Poisson atau :

$$\mu = \lambda = n.p$$

Contoh 4.11.

Sepuluh persen peralatan yang dihasilkan dari sebuah proses produksi ditemukan cacat. Hitunglah peluang bahwa dalam 10 peralatan yang diambil secara acak dua diantaranya cacat dengan menggunakan (a) pendekatan binomial, (b) pendekatan Poisson ke Binomial.

Jawab

a) Peluang peralatan cacat adalah $p = 0,1$. Misalkan X menunjukkan jumlah peralatan yang cacat dari 10 peralatan yang dipilih. Dengan menggunakan distribusi binomial, maka :

$$P(X = 2) = \binom{10}{2} (0,1)^2 (0,9)^8 = 0,1937$$

b) Kita punya $\mu = n.p = (10)(0,1) = 1$. Berdasarkan rumus distribusi Poisson :

$$P(X = c) = \frac{\lambda^c e^{-\lambda}}{c!}$$

$$P(X = 2) = \frac{\lambda^2 e^{-\lambda}}{2!} = \frac{(1)^2 (2,71828)^{-1}}{2.1} = 0,1839$$

DISTRIBUSI MULTINOM

Merupakan perluasan dari distribusi Binomial. Misalkan sebuah eksperimen menghasilkan peristiwa-peristiwa E_1, E_2, \dots, E_k dengan peluang masing-masing adalah $P(E_1) = \pi_1, P(E_2) = \pi_2, \dots, P(E_k) = \pi_k$, dimana $\pi_1 + \pi_2 + \dots + \pi_k = 1$. Jika dalam eksperimen ini dilakukan sebanyak N kali, maka peluang akan terdapat x_1 peristiwa E_1, x_2 peristiwa E_2, \dots, x_k peristiwa E_k dapat ditentukan oleh menggunakan distribusi multinom sebagai berikut :

$$p(x_1, x_2, \dots, x_k) = \frac{N!}{x_1! x_2! \dots x_k!} \pi_1^{x_1} \pi_2^{x_2} \dots \pi_k^{x_k}$$

Dimana :

$$x_1 + x_2 + \dots + x_k = N \quad ; \quad \pi_1 + \pi_2 + \dots + \pi_k = 1 \quad ; \quad 0 < \pi_i < 1 \quad ; \quad i = 1, 2, \dots, k$$

Untuk distribusi multinom *ekspektasi* dari masing-masing peristiwa E_1, E_2, \dots, E_k adalah $N \pi_1, N \pi_2, \dots, N \pi_k$. Sedangkan variansnya adalah $N \pi_1(1 - \pi_1), N \pi_2(1 - \pi_2), \dots, N \pi_k(1 - \pi_k)$.

Contoh 4.12:

Dua dadu dilempar sebanyak 6 kali. Berapa peluang mendapatkan jumlah bilangan yang muncul sebesar 7 atau 11 sebanyak 2 kali, bilangan yang sama pada kedua sebanyak sekali dan kemungkinan lainnya 3 kali.

Penyelesaian :

Misalkan E_1 : peristiwa jumlah yang muncul 7 atau 11
 E_2 : peristiwa bilangan yang sama pada kedua dadu
 E_3 : peristiwa lainnya selain peristiwa di atas

Dari penjelasan sebelumnya (lihat pengantar peluang) kita tahu bahwa titik sampel untuk pelemparan 2 buah dadu adalah 36. Untuk peristiwa E_1 dapat ditentukan kemungkinan jumlah munculnya 7 sebanyak 6 titik dan muncul 11 sebanyak 2 titik. Maka $P(E_1) = \pi_1 = 6/36 + 2/36 = 8/36 = 2/9$. Untuk peristiwa E_2 jumlah yang mungkin adalah 6 titik. Jadi $P(E_2) = \pi_2 = 6/36 = 1/6$. Sedangkan peristiwa E_3 adalah peristiwa selain kedua peristiwa ini, sehingga $P(E_3) = \pi_3 = 1 - 2/9 - 1/6 = 11/18$. Untuk persoalan ini : $N = 6, x_1 = 2, x_2 = 1, \text{ dan } x_3 = 3$.

Gunakan rumus di atas :

$$p(x_1, x_2, \dots, x_k) = \frac{N!}{x_1! x_2! \dots x_k!} \pi_1^{x_1} \pi_2^{x_2} \dots \pi_k^{x_k}$$

$$p(x_1 = 2, x_2 = 1, x_3 = 3) = \frac{6!}{2!1!3!} (2/9)^2 (1/6)^1 (11/18)^3$$

$$p(x_1 = 2, x_2 = 1, x_3 = 3) = 0,1127$$

Contoh 4.13:

Sebuah kotak berisikan 3 barang yang dihasilkan oleh mesin A, 4 barang oleh mesin B dan 5 barang oleh mesin C. Identitas barang adalah sama kecuali berdasarkan kategori mesin. Sebuah barang diambil secara acak dari kotak tersebut kemudian dicatat identitas mesinnya dan disimpan kembali kedalam kotak. Jika 6 barang diambil dengan cara yang demikian, tentukan peluang diantara ke enam barang tersebut diperoleh 1 dari mesin A, 2 dari mesin B dan 3 dari mesin C.

Penyelesaian :

Misal :

- A : peristiwa terambilnya barang dari mesin A.
- B : peristiwa terambilnya barang dari mesin B.
- C : peristiwa terambilnya barang dari mesin C.

Jelas : $P(A) = 3/12 \quad ; \quad P(B) = 4/12 \quad ; \quad P(C) = 5/12$
 $N = 6 \quad ; \quad x_1 = 1, x_2 = 2, \text{ dan } x_3 = 3$

Dengan rumus (4) maka diperoleh :

$$p(x_1 = 1, x_2 = 2, x_3 = 3) = \frac{6!}{1!2!3!} (3/12)^1 (4/12)^2 (5/12)^3$$

$$p(x_1 = 1, x_2 = 2, x_3 = 3) = 0,1206$$

DISTRIBUSI HIPERGEOMETRIK

Ciri-ciri percobaan hipergeometri :

1. Sampel acak berukuran n diambil dari populasi berukuran N
2. k dari N objek dikategorikan sebagai **sukses** dan $N - k$ dikategorikan **gagal**.

Misalkan ada sebuah populasi berukuran N yang diantaranya terdapat k buah termasuk kategori tertentu (sukses). Dari populasi tersebut diambil sebuah sampel acak berukuran n . Berapa peluang dalam sampel tersebut terdapat x buah termasuk kategori tertentu tersebut?

Untuk menjawabnya dapat diperoleh dari *distribusi hipergeometrik* yang berbentuk :

$$P(X = x) = p(x) = \frac{\binom{k}{x} \binom{N-k}{n-x}}{\binom{N}{n}}$$

... (5)

Dengan $x = 0, 1, \dots, n$

Contoh 4.14:

Sekelompok mahasiswa terdiri dari 50 orang dan 3 diantaranya lahir pada tanggal 17 Agustus. Dari kelompok tersebut dipilih 5 orang secara acak. Berapakah peluang bahwa diantara 5 orang tersebut :

- a. tidak terdapat yang lahir pada tanggal 17 Agustus
- b. terdapat tidak lebih dari 1 orang yang lahir pada tanggal 17 Agustus

Penyelesaian :

- a. Misal X adalah banyak mahasiswa di antara $n = 5$ yang lahir pada tanggal 17 Agustus. Dari persoalan diatas dapat kita katakan : $N = 50$, $D = 3$. Maka peluang kelima mahasiswa tidak lahir pada tanggal 17 Agustus adalah :

$$P(X = x) = p(x) = \frac{\binom{k}{x} \binom{N-k}{n-x}}{\binom{N}{n}}$$

$$P(X = 0) = p(0) = \frac{\binom{3}{0} \binom{50-3}{5}}{\binom{50}{5}} = 0,724$$

- b. Tidak lebih dari 1 orang yang lahir tanggal 17 Agustus mengandung arti bahwa nilai-nilai x hanya 0 dan 1. Dari soal a) kita sudah hitung $p(0)$, jadi tinggal menghitung $p(1)$ yaitu :

$$P(X = 1) = p(1) = \frac{\binom{3}{1} \binom{50-3}{4}}{\binom{50}{5}} = 0,253$$

Jadi peluang dari kelima mahasiswa paling banyak 1 mahasiswa lahir pada tanggal 17 Agustus adalah $0,724 + 0,253 = 0,977$.

Contoh 4.15:

Sebuah panitia yang terdiri dari 5 orang diambil secara acak dari 3 perempuan dan 5 laki-laki. Tentukanlah distribusi peluang bagi banyaknya perempuan dalam panitia tersebut.

Penyelesaian:

Misal X adalah banyaknya perempuan yang duduk dalam kepanitiaan. Maka distribusi hipergeometriknya adalah untuk $X = 0, 1, 2, 3$. Nilai 0 berarti tidak ada perempuan dalam kepanitiaan tersebut.

$$P(X = 0) = p(0) = \frac{\binom{3}{0} \binom{5}{5}}{\binom{8}{5}} = \frac{1}{56}$$

$$P(X = 1) = p(1) = \frac{\binom{3}{1} \binom{5}{4}}{\binom{8}{5}} = \frac{15}{56}$$

$$P(X = 2) = p(2) = \frac{\binom{3}{2} \binom{5}{3}}{\binom{8}{5}} = \frac{30}{56}$$

$$P(X = 3) = p(3) = \frac{\binom{3}{3} \binom{5}{2}}{\binom{8}{5}} = \frac{10}{56}$$

Parameter Distribusi Hipergeometrik :

Rata-rata : $\mu = \frac{nk}{N}$

Varians: $\sigma^2 = \frac{N-n}{N-1} \cdot n \cdot \frac{k}{N} \left(1 - \frac{k}{N}\right)$

Pendekatan Binomial terhadap Hipergeometrik

Jika n relatif kecil dibandingkan dengan N ($n \ll N$), maka peluang setiap pengambilan objek akan berubah menjadi kecil sekali, maka distribusi hipergeometrik dapat didekati oleh distribusi binomial dengan mengambil $p = k/N$. Dengan demikian rata-rata dan varians hipergeometrik dapat didekati menggunakan rumus :

$$\mu = np = \frac{nk}{N}$$

$$\sigma^2 = npq = n \cdot \frac{k}{N} \left(1 - \frac{k}{N}\right)$$

Contoh 4.16:

Perusahaan Telepon melaporkan bahwa diantara 5000 pemasang tilpun baru, 4000 menggunakan telpon *tanpa kabel*. Bila 10 di antara pemasang baru tersebut diambil secara acak, berapa peluang tepat ada 3 orang yang menggunakan telpon *dengan kabel*.

Penyelesaian:

Terlihat bahwa ukuran populasi $N = 5000$ relatif sangat besar dibandingkan dengan sampel $n = 10$, maka dapat digunakan pendekatan distribusi binomial.

Peluang pemasangan menggunakan kabel dengan telepon adalah 0,2. Dengan demikian peluang tepat ada 3 orang yang menggunakan kabel adalah :

$$P(X = 3) = \binom{10}{3} (0,2)^3 (0,8)^7 = \frac{10!}{3!(7)!} 0,2^3 0,8^7 \\ = 0,2013$$

LATIHAN

- Sebuah mesin pembuat bola lampu menghasilkan 20% cacat produk. Jika 4 bola lampu diambil secara acak, hitunglah peluang
 - 2 bola lampu cacat
 - Lebih dari 1 bola lampu yang cacat
 - Kurang dari 2 bola lampu yang cacat
- Jika peluang seseorang menderita akibat penyuntikan sejenis serum adalah 0,001, tentukanlah peluang bahwa dari 2000 orang yang disuntik :
 - Tepat tiga orang menderita
 - Lebih dari dua orang menderita
- Suatu ujian terdiri atas 15 pertanyaan pilihan berganda masing-masing dengan 4 kemungkinan jawaban dan hanya satu yang benar. Berapa peluang seseorang yang menjawab secara menebak-nebak saja memperoleh 5 sampai 10 jawaban yang benar?
- Di sebuah desa di daerah Jawa Timur, secara rata-rata dilanda 6 kali puting beliung per tahun. Hitunglah peluang di suatu tahun tertentu desa tersebut akan dilanda :
 - Kurang dari 4 puting beliung
 - 6 sampai 8 puting beliung
 - Lebih dari 4 puting beliung
- Seorang sekretaris rata-rata melakukan 2 kesalahan ketik per halaman. Berapa peluang bahwa pada halaman berikutnya dia membuat :
 - 4 atau lebih kesalahan
 - Tidak satu pun kesalahan
 - Kurang dari 4 kesalahan
- Secara rata-rata 1 diantara 1000 orang membuat kesalahan menulis angka dalam membuat laporan pajak pendapatannya. Bila 10.000 formulir diambil secara acak dan diperiksa berapa peluang terdapat kurang dari 7 formulir yang mengandung kesalahan.
- Hitunglah peluang mendapat dua bilangan 1, satu bilangan 2, satu bilangan 3, dua bilangan 4, tiga bilangan 5 dan satu bilangan 6, bila sebuah dadu seimbang dilemparkan 10 kali.
- Dalam teori genetika, suatu persilangan kelinci percobaan akan menghasilkan keturunan warna merah, hitam dan putih dalam perbandingan 8 : 4 : 4. Hitunglah peluang bahwa di antara 8 keturunan semacam ini ada 5 yang berwarna merah, 2 hitam dan 1 putih.

Bab 5

SEBARAN PELUANG KONTINU

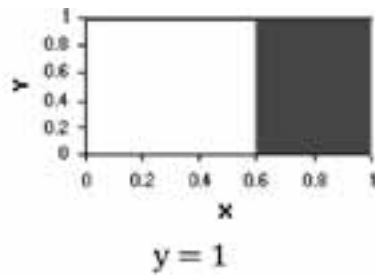
Jika variabel acak berupa variabel kontinu, sebaran peluangnya disebut sebaran peluang kontinu. Sebaran peluang kontinu berbeda dari sebaran peluang diskrit dalam beberapa cara.

- Peluang bahwa variabel acak kontinu akan mengasumsikan nilai tertentu adalah nol.
- Akibatnya, sebaran peluang kontinu tidak dapat dinyatakan dalam bentuk tabel.
- Sebagai gantinya, sebuah persamaan atau rumus digunakan untuk menggambarkan sebaran peluang kontinu.

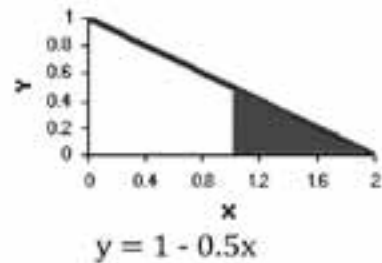
Persamaan yang digunakan untuk menggambarkan sebaran probabilitas kontinu disebut *fungsi kepadatan probabilitas (pdf)*. Semua fungsi kepadatan probabilitas memenuhi syarat-syarat berikut

- Variabel acak Y adalah fungsi X ; yaitu, $y = f(x)$.
- Nilai y lebih dari atau sama dengan nol untuk semua nilai x .
- Luas total di bawah kurva fungsi sama dengan satu.

Bagan di bawah ini menunjukkan dua sebaran probabilitas kontinyu. Bagan pertama menunjukkan fungsi kepadatan probabilitas yang digambarkan oleh persamaan $y = 1$ di atas kisaran 0 sampai 1 dan $y = 0$ di tempat lain.



Bagan berikutnya menunjukkan fungsi kepadatan probabilitas yang digambarkan oleh persamaan $y = 1 - 0,5x$ pada kisaran 0 sampai 2 dan $y = 0$ di tempat lain. Area di bawah kurva sama dengan 1 untuk kedua grafik.



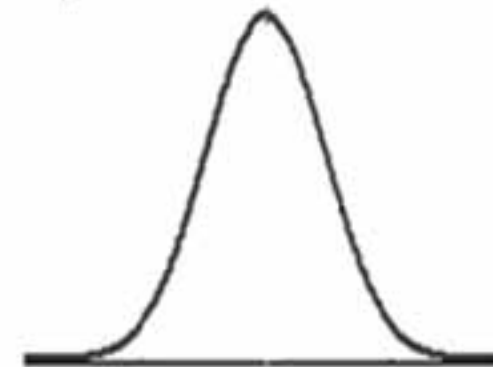
Probabilitas bahwa variabel acak kontinyu turun dalam interval antara a dan b sama dengan area di bawah kurva pdf antara a dan b . Sebagai contoh, pada bagan pertama di atas, area yang diarsir menunjukkan probabilitas bahwa variabel acak X akan turun antara 0,6 dan 1,0. Probabilitas itu adalah 0,40. Dan pada grafik kedua, area yang diarsir menunjukkan probabilitas turun antara 1,0 dan 2,0. Probabilitas itu adalah 0,25.

5.1 Sebaran Normal

Sebaran normal merupakan salah satu sebaran peluang yang populer dan banyak digunakan dalam berbagai keperluan

baik dalam ilmu sosial, ilmu alam atau teknik. Pada dasarnya sebaran ini dapat dijelaskan dari sebaran frekuensi relatif. Penjelarasannya adalah sebagai berikut. Apabila dalam sebuah sebaran frekuensi kita memperbanyak jumlah kelas intervalnya atau panjang kelasnya dipersempit, maka kita akan memperoleh titik tengah yang lebih banyak dan akhirnya akan diperoleh sebuah kurva frekuensi yang lebih halus yang dapat dianggap sebagai gambaran atau model populasi yang sedang di amati. Dengan memanfaatkan sebaran normal maka kita dapat menghitung berapa peluang variabel acak diantara nilai-nilai yang kita tetapkan. Misalnya saja ingin diketahui berapa peluang seorang karyawan yang dipilih secara acak memiliki berat badan antara 56 dan 60 kilogram atau tingginya lebih kecil dari 170 cm.

Suatu peubah acak kontinu X yang sebarannya berbentuk lonceng disebut peubah acak normal.



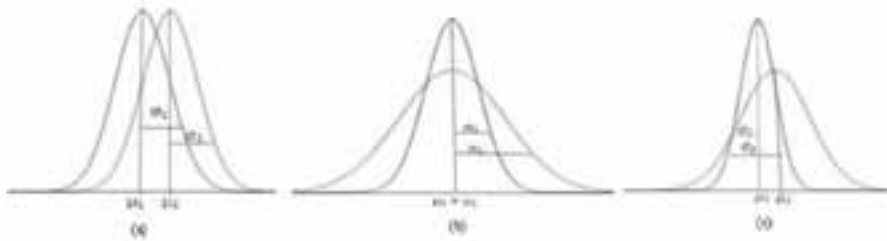
Gambar 5.1 Kurva sebaran normal

Persamaan yang memenuhi kurva tersebut dinamakan Probability Distribution Function (PDF) normal yang didefinisikan sebagai berikut:

Definisi 5.1 Probabiliti Distribution Function (PDF) Normal
 Jika X adalah peubah acak bebas, maka fungsi

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

dengan $\pi = 3,14159\dots$ dan $e = 2,71828\dots$ disebut sebagai *Probability Distribution Function* (PDF) normal atau sebaran normal dan umumnya ditulis sebagai $f(x; \mu, \sigma)$ yang berarti peubah acak X tersebaran normal dengan parameter μ dan σ



Gambar 5.2 Kurva Sebaran Normal dengan (a) $\mu_1 < \mu_2$ dan $\sigma_1 = \sigma_2$; (b) $\mu_1 = \mu_2$ dan $\sigma_1 < \sigma_2$; (c) $\mu_1 < \mu_2$ dan $\sigma_1 < \sigma_2$

Pada Gambar 5.1 melukiskan kurva normal yang berbentuk seperti lonceng. Begitu parameter μ dan σ diketahui, maka seluruh kurva normal dapat diketahui dan kurangnya dapat digambarkan. Misalkan jika diketahui $\mu = 0$ dan $\sigma = 1$, maka ordinat $f(x)$ dapat dengan mudah dihitung untuk berbagai harga x dan kurangnya dapat digambarkan.

Pada Gambar 5.2 (a) telah dilukiskan dua kurva normal yang mempunyai simpangan baku sama, tetapi rataannya berbeda. Kedua kurva bentuknya persis sama, tetapi titik tengahnya terletak di tempat yang berbeda di sepanjang sumbu datar.

Pada Gambar 5.2(b) terlukis dua kurva normal dengan rataannya yang sama tetapi simpangan bakunya berbeda. Kedua kurva mempunyai titik tengah sama tetapi kurva dengan

simpangan baku yang lebih besar tampak kurvanya lebih landai/rendah dan lebih menyebar.

Pada Gambar 5.2 (c) memperlihatkan lukisan dua kurva normal yang baik rataannya maupun simpangan bakunya berlainan. Jelas keduanya mempunyai titik tengah yang berlainan pada sumbu datar dan bentuknya pun mencerminkan dua harga simpangan baku yang berlainan. Lebih lanjut, mahasiswa dapat mengamati dengan lengkap sifat kurva normal tersebut dengan melakukan simulasi pada aplikasi Matlab (*Matriks Laboratorium*) dengan menggunakan perintah `disttool` pada *command windows* Matlab.

Dengan mengamati grafik serta memeriksa turunan pertama dan kedua dari $f(x) = f(x; \mu; \sigma)$, maka dapat diperoleh lima sifat kurva normal sebagai berikut:

Sifat 5.2 SIFAT KURVA NORMAL

- a. Titik pada sumbu datar yang memberikan maksimum kurva, terdapat pada $x = \mu$
- b. Kurva simetris terhadap garis tegak yang melalui μ
- c. Kurva mempunyai titik belok pada $x = \mu \pm \sigma$ cekung dari bawah untuk $\mu - \sigma < X < \mu + \sigma$, dan cekung dari atas untuk harga x lainnya
- d. Kedua ujung kurva normal mendekati asimtot sumbu datar, jika harga x bergerak menjauhi μ baik ke kiri maupun ke kanan;
- e. Seluruh luas di bawah kurva dan di atas sumbu datar sama dengan 1.

Contoh 5.1

Seorang mahasiswa melakukan penelitian pada kelas untuk mengetahui pengaruh penggunaan media pembelajaran BINGKAI AJAIB dalam meningkatkan prestasi belajar matematika. Hasil evaluasi pembelajaran yang diperoleh sebagai berikut:

Data prestasi belajar kelas II A

12,5 75,0 53,0 64,8 80,0 27,0 87,0 35,0 25,5 38,0 40,0 47,5 42,8 45,5 43,5
 46,0 47,0 45,0 42,0 43,0 45,0 48,0 49,5 54,0 58,0 69,5 65,5 66,0 64,0 93,5
 64,5 72 68,5 65,0 80,0 85,0 72,5 78,5

Data prestasi belajar kelas II B

22,5 35,0 36,5 32,0 77,5 59,0 60,0 54,0 54,0 59,0 52,5 54,0 59,5 50,0 46,0
 51,5 56,0 70,0 58,0 53,0 75,0 76,5 32,0 34,0 28,0 68,5 68,0 65,0 66,5 70,0
 81,5 85,0 75,5 36,0 76,5 36,0 35,0

Selidiki apakah data prestasi belajar siswa kelas tersebut tersebar normal?

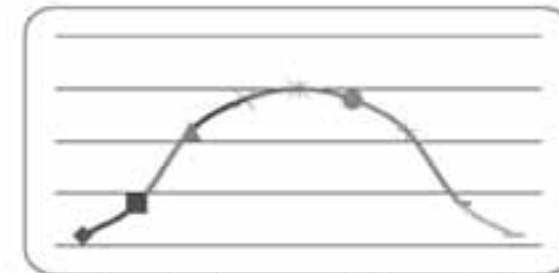
Penyelesaian:

Data prestasi belajar tersebut dapat dikatakan tersebar normal apabila data tersebut tersebar sebagaimana pada Gambar 5.1. Oleh karena itu, untuk mengetahuinya, prosedur paling sederhana yang dapat ditempuh adalah dengan membuat tabel sebaran frekuensi dan menggambar grafik sebaran frekuensinya. Apabila grafiknya menyerupai gambar kurva normal sebagaimana pada gambar 5.1, maka data tersebut dapat dinyatakan sebagai data yang tersebar normal. Karena pengamatan adalah kelas II, maka data tersebut digabungkan dan dibuat tabel sebaran frekuensi sebagai berikut:

Tabel 5.1 Sebaran frekuensi

Interval Nilai	Frekuensi
11 - 20	1
21 - 30	4
31 - 40	11
41 - 50	14
51 - 60	15
61 - 70	14
71 - 80	11
81 - 90	4
91 - 100	1
Jumlah	75

Selanjutnya nilai pada kolom frekuensi digambar grafiknya dan diperoleh sebagai berikut:



Gambar 5.3 Sebaran frekuensi data prestasi belajar siswa

Berdasarkan Gambar 5.3, maka dapat dikatakan bahwa data prestasi belajar siswa kelas menyerupai model kurva sebaran normal.

Contoh 5.2

Selidiki apakah $f(x) = n(x; \mu, \sigma)$ benar suatu fungsi peluang?

Penyelesaian:

$f(x) = n(x; \mu, \sigma)$ dapat dinyatakan sebagai fungsi peluang apabila memenuhi syarat

(i). $0 \leq f(x) \leq 1$ dan (ii). $\int_{-\infty}^{\infty} f(x) = 1$

Dengan menggunakan metode analisis real, berikut syarat (i) akan diselidiki:

$$f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} \Rightarrow \ln f(x) = \ln \left[\frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} \right]$$

$$\Rightarrow \ln f(x) = \ln \left[\frac{1}{\sqrt{2\pi\sigma^2}} \right] + \ln \left[e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} \right]$$

$$\Rightarrow \ln f(x) = \ln \left[\frac{1}{\sqrt{2\pi\sigma^2}} \right] - \frac{1}{2} \left(\frac{x-\mu}{\sigma} \right)^2$$

$$\Rightarrow \frac{1}{2} \left(\frac{x-\mu}{\sigma} \right)^2 = \ln \left(\frac{1}{\sqrt{2\pi\sigma^2}} \right) - \ln f(x)$$

$$\Rightarrow \left(\frac{x-\mu}{\sigma} \right)^2 = 2 \left[\ln \left(\frac{1}{\sqrt{2\pi\sigma^2}} \right) - \ln f(x) \right]$$

$$\Rightarrow \frac{x-\mu}{\sigma} = \sqrt{2 \left[\ln \left(\frac{1}{\sqrt{2\pi\sigma^2}} \right) - \ln f(x) \right]}$$

$$\Rightarrow x = \mu + \sigma \sqrt{2 \left[\ln \left(\frac{1}{\sqrt{2\pi\sigma^2}} \right) - \ln f(x) \right]}$$

Bentuk $\sqrt{2 \left[\ln \left(\frac{1}{\sqrt{2\pi\sigma^2}} \right) - \ln f(x) \right]}$

Terdefinisi jika $2 \left[\ln \left(\frac{1}{\sqrt{2\pi\sigma^2}} \right) - \ln f(x) \right] \geq 0$, akibatnya

$$2 \left[\ln \left(\frac{1}{\sqrt{2\pi\sigma^2}} \right) - \ln f(x) \right] \geq 0 \Rightarrow \left[\ln \left(\frac{1}{\sqrt{2\pi\sigma^2}} \right) - \ln f(x) \right] \geq 0$$

$$\Rightarrow \ln \frac{1}{\sqrt{2\pi\sigma^2}} \geq 0 \Rightarrow \ln \frac{1}{\sqrt{2\pi\sigma^2}} \geq \ln e^0$$

$$\Rightarrow \frac{1}{\sqrt{2\pi\sigma^2}} \geq e^0 \Rightarrow \frac{1}{f(x)\sqrt{2\pi\sigma^2}} - e^0 \geq 0$$

$$\Rightarrow \frac{1}{f(x)\sqrt{2\pi\sigma^2}} - \frac{e^0 f(x)\sqrt{2\pi\sigma^2}}{f(x)\sqrt{2\pi\sigma^2}} \geq 0$$

$$\Rightarrow \frac{1 - f(x)\sqrt{2\pi\sigma^2}}{f(x)\sqrt{2\pi\sigma^2}} \geq 0 \Rightarrow (1 - f(x)\sqrt{2\pi\sigma^2}) (f(x)\sqrt{2\pi\sigma^2}) > 0$$

Akar-akar karakteristik

$$f(x)\sqrt{2\pi\sigma^2} = 1 \text{ atau } f(x)\sqrt{2\pi\sigma^2} = 0$$

Diperhatikan untuk bentuk $f(x)\sqrt{2\pi\sigma^2} = 1$

$$f(x)\sqrt{2\pi\sigma^2} = 1 \Rightarrow f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} > 0 \dots\dots\dots(a)$$

$$\text{Dari bentuk } f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \Rightarrow f(x) \leq 1 \dots\dots\dots(b)$$

$$\text{Dari bentuk } f(x)\sqrt{2\pi\sigma^2} = 0 \Rightarrow f(x) = 0 \dots\dots\dots(c)$$

Dari (a), (b) dan (c) diperoleh syarat (i) $0 \leq f(x) \leq 1$ terpenuhi

Selanjutnya diperiksa apakah syarat (ii) terpenuhi, yakni

apakah $\int_{-\infty}^{\infty} f(x) dx = 1$

$$\int_{-\infty}^{\infty} f(x) dx = \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} dx = \frac{1}{\sqrt{2\pi\sigma^2}} \int_{-\infty}^{\infty} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} dx$$

Misal

$$u = \frac{1}{2} \left(\frac{x-\mu}{\sigma} \right)^2 \Rightarrow 2u = \left(\frac{x-\mu}{\sigma} \right)^2 \Rightarrow x = \mu + \sigma\sqrt{2u}$$

$$\Rightarrow dx = \frac{\sigma\sqrt{2}}{2} u^{-\frac{1}{2}} du, \text{ akibatnya}$$

$$\begin{aligned} \int_{-\infty}^{\infty} f(x) dx &= \frac{1}{\sqrt{2\pi\sigma^2}} \int_{-\infty}^{\infty} e^{-u} \left(\frac{\sigma\sqrt{2}}{2} u^{-\frac{1}{2}} du \right) = \frac{\sigma\sqrt{2}}{2\sqrt{2\pi\sigma^2}} \int_{-\infty}^{\infty} e^{-u} u^{-\frac{1}{2}} du \\ &= \frac{\sigma}{2\sqrt{\sigma^2}} \frac{1}{\sqrt{\pi}} \left[\int_{-\infty}^0 e^{-u} u^{-\frac{1}{2}} du + \int_0^{\infty} e^{-u} u^{-\frac{1}{2}} du \right] = \frac{1}{2\sqrt{\pi}} \left[\int_0^{\infty} u^{-\frac{1}{2}} e^{-u} du + \int_0^{\infty} u^{-\frac{1}{2}} e^{-u} du \right] \end{aligned}$$

$$= \frac{1}{2\sqrt{\pi}} \left[\Gamma\left(\frac{1}{2}\right) + \Gamma\left(\frac{1}{2}\right) \right] = \frac{1}{2\sqrt{\pi}} \left[\sqrt{\pi} + \sqrt{\pi} \right] = \frac{2\sqrt{\pi}}{2\sqrt{\pi}} = 1$$

Berdasarkan hasil terakhir, maka syarat (ii), yakni bahwa

$$\int_{-\infty}^{\infty} f(x) dx = 1 \text{ terpenuhi}$$

Karena syarat (i) dan (ii) terpenuhi, maka terbukti benar

$$\text{bahwa } f(x) = n(x; \mu, \sigma) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}, -\infty < x < \infty \text{ adalah}$$

suatu fungsi peluang.

Karena $f(x)$ adalah suatu fungsi peluang, maka tentunya $f(x)$ memiliki rata-rata dan simpangan baku. Rataan dan simpangan baku dari PDF normal adalah sebagai berikut:

Teorema 5.1 RATAAN DAN RAGAM SEBARAN NORMAL
 Jika X peubah acak bebas yang tersebaran normal, maka rata-rata X adalah $\mu_x = \mu$ dan dan Ragam X adalah $\sigma_x^2 = \sigma^2$.

Bukti:

Berdasarkan definisi rata-rata bahwa $\mu_x = E(x) = \frac{dM_x(t)}{dx} \Big|_{t=0}$

dengan $M_x(t)$ adalah fungsi pembangkit moment yang terdefinisi sebagai $M_x(t) = E(e^{tx})$ maka diperoleh bahwa

$$\begin{aligned} M_x(t) &= \int_{-\infty}^{\infty} e^{tx} f(x) dx = \int_{-\infty}^{\infty} e^{tx} \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} dx = \frac{1}{\sqrt{2\pi\sigma^2}} \int_{-\infty}^{\infty} e^{tx - \frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} dx \\ &= \frac{1}{\sqrt{2\pi\sigma^2}} \int_{-\infty}^{\infty} e^{\left(\frac{1}{2\sigma^2}(t\sigma^2 x - (x-\mu)^2)\right)} dx = \frac{1}{\sqrt{2\pi\sigma^2}} \int_{-\infty}^{\infty} e^{-\frac{1}{2\sigma^2}(x-\mu)^2 + tx} dx \end{aligned}$$

Sekarang perhatikan bentuk berikut:

$$\begin{aligned} [x - (\mu + t\sigma^2)]^2 - 2\mu t\sigma^2 - t^2\sigma^4 &= x^2 - 2x(\mu + t\sigma^2) + (\mu + t\sigma^2)^2 - 2\mu t\sigma^2 - t^2\sigma^4 \\ &= x^2 - 2x\mu - 2xt\sigma^2 + (\mu^2 + 2\mu t\sigma^2 + (t\sigma^2)^2) - 2\mu t\sigma^2 - t^2\sigma^4 \\ &= x^2 - 2x\mu - 2xt\sigma^2 + \mu^2 + 2\mu t\sigma^2 + t^2\sigma^4 - 2\mu t\sigma^2 - t^2\sigma^4 \\ &= x^2 - 2x\mu + \mu^2 - 2xt\sigma^2 \\ &= (x - \mu)^2 - 2\sigma^2 xt \end{aligned}$$

Jadi $(x - \mu)^2 - 2\sigma^2 xt = [x - (\mu + t\sigma^2)]^2 - 2\mu t\sigma^2 - t^2\sigma^4$ sehingga diperoleh

$$\begin{aligned} M_x(t) &= \frac{1}{\sqrt{2\pi\sigma^2}} \int_{-\infty}^{\infty} e^{\frac{1}{2\sigma^2}([x - (\mu + t\sigma^2)]^2 - 2\mu t\sigma^2 - t^2\sigma^4)} dx = \frac{1}{\sqrt{2\pi\sigma^2}} \int_{-\infty}^{\infty} e^{-\frac{1}{2\sigma^2}([x - (\mu + t\sigma^2)]^2)} e^{\frac{1}{2\sigma^2}(-2\mu t\sigma^2 - t^2\sigma^4)} dx \\ &= e^{\frac{1}{2\sigma^2}(-2\mu t\sigma^2 - t^2\sigma^4)} \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{1}{2\sigma^2}([x - (\mu + t\sigma^2)]^2)} dx \\ &= e^{\mu t - \frac{t^2\sigma^2}{2}} \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{1}{2}\left(\frac{x - (\mu + t\sigma^2)}{\sigma}\right)^2} dx = e^{\mu t - \frac{t^2\sigma^2}{2}} n(x, \mu + t\sigma^2, \sigma) \\ &= e^{\mu t - \frac{t^2\sigma^2}{2}} (1) = e^{\mu t - \frac{t^2\sigma^2}{2}} \\ \therefore M_x(t) &= e^{\mu t - \frac{t^2\sigma^2}{2}} \end{aligned}$$

sehingga diperoleh rataan sebagai berikut

$$\mu_x = E(x) = \frac{dM_x(t)}{dt} \Big|_{t=0} = (\mu + t\sigma^2) e^{\mu t + \frac{1}{2}t^2\sigma^2} \Big|_{t=0} = \mu \text{ dan}$$

$$E(x^2) = \frac{d^2 M_x(t)}{dt^2} \Big|_{t=0} = \frac{d}{dt} \left((\mu + t\sigma^2) e^{\mu t + \frac{1}{2}t^2\sigma^2} \right) \Big|_{t=0} = \left[(0 + \sigma^2) e^{\mu(0) + \frac{1}{2}(0)^2\sigma^2} + (\mu + t\sigma^2) e^{\mu(0) + \frac{1}{2}t^2\sigma^2} (\mu + t\sigma^2) \right] \Big|_{t=0}$$

$$= (0 + \sigma^2) e^{\mu(0) + \frac{1}{2}(0)^2\sigma^2} + (\mu + (0)\sigma^2) e^{\mu(0) + \frac{1}{2}(0)^2\sigma^2} (\mu + (0)\sigma^2) = \sigma^2 + \mu^2$$

sehingga diperoleh ragam sebagai

$$\sigma_x^2 = E(x^2) - [E(x)]^2 = \sigma^2 + \mu^2 - \mu^2 = \sigma^2.$$

Contoh 5.3

Kembali lihat contoh 5.2. Tentukan berapakah peluang seoprang siswa yang tdipilih secara acak, memperoleh nilai tepat 65?

Penyelesaian:

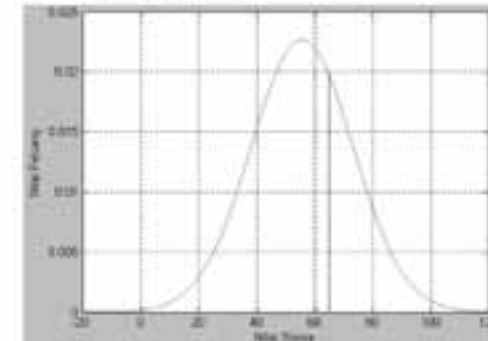
Karena data prestasi belajar tersebut tersebar normal, maka permasalahan pada contoh 5.2 dapat diselesaikan dengan menggunakan fungsi normal. Berdasarkan contoh 5.2 di atas, telah dibuktikan bahwa $f(x) = n(x; \mu, \sigma)$ adalah suatu fungsi peluang. Oleh karena untuk suatu $X = x$, maka $P(X=x) = f(x)$. Untuk itu perlu dicari nilai rataan dan simpangan baku data prestasi belajar siswa. Karena pengamatan dalam permasalahan ini adalah seluruh kelas II dan telah dimiliki data seluruh kelas II, maka rataan dan simpangan baku diberikan sebagai berikut:

$$\mu = \frac{\sum_{i=1}^{75} x_i}{75} = 55,8 \text{ dan } \sigma = \sqrt{\frac{\sum_{i=1}^{75} (x_i - 55,8)^2}{75 - 1}} = 17,6 \text{ sehingga}$$

diperoleh Peluang siswa yang terpilih secara acak akan

memperoleh nilai tes tepat 65 adalah

$$P(x = 65) = f(65) = \frac{1}{\sqrt{2\pi(17,6)^2}} e^{-\frac{1}{2} \left(\frac{65 - 55,8}{17,6} \right)^2} = 0,02 = 2\%$$



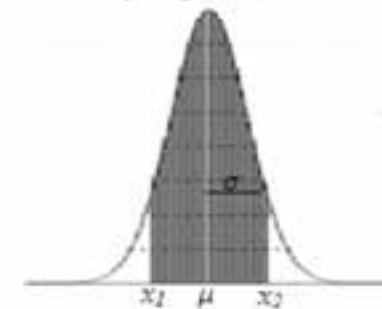
Gambar 5.4 Kurva sebaran normal dengan $\mu = 55,7$ dan $\sigma = 17,5$ serta nilai dari $P(x=65)=2\%$

5. 4 Luas di bawah Kurva Normal

Kurva setiap sebaran peluang kontinu atau fungsi padat peluang dibuat sedemikian rupa sehingga luas di bawah kurva diantara dua ordinat $x = x_1$ dan $x = x_2$ sama dengan peluang peubah acak X mendapat harga antara $x = x_1$ dan $x = x_2$. Jadi untuk kurva normal pada gambar di atas,

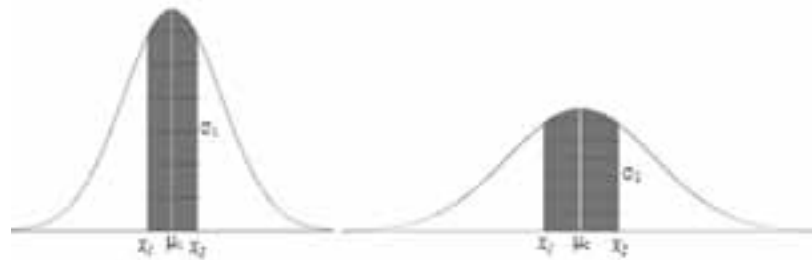
$$P(x_1 \leq X \leq x_2) = \int_{x_1}^{x_2} n(x; \mu, \sigma) dx = \int_{x_1}^{x_2} \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{1}{2} \left(\frac{x-\mu}{\sigma} \right)^2} dx$$

menyatakan luas daerah yang diarsir.



Gambar 5.5 Luas daerah di bawah kurva normal antara x_1 dan x_2

Di depan telah ditunjukkan dengan beberapa gambar, bahwa kurva normal tergantung pada rata-rata dan simpangan baku sebaran. Luas daerah di bawah kurva antara dua ordinat sembarang jelas tergantung pada harga μ dan σ . Hal ini terlihat seperti gambar berikut. Pada gambar tersebut daerah yang berpadanan dengan $P(x_1 \leq X \leq x_2)$ untuk kedua kurva dengan rata-rata dan simpangan baku yang berbeda adalah daerah yang terdapat dua arsiran.



Gambar 5.6 $P(x_1 < X < x_2)$ untuk kurva normal dengan rata-rata dan simpangan baku yang berbeda

Contoh 5.4

Lihat kembali contoh pada 5.2. Tentukan peluang seorang siswa yang terpilih secara acak akan memperoleh nilai

- a. Kurang dari 65
- b. Lebih dari 75
- c. Diantara 65 dan 75

Penyelesaian:

a. Peluang siswa yang terpilih secara acak akan memperoleh nilai tes kurang dari 65 adalah

$$\begin{aligned}
 P(X < x = 65) &= \int_{-\infty}^{65} f(x)dx = \int_{-\infty}^{55,8} f(x)dx + \int_{55,8}^{65} f(x)dx \\
 &= \int_{-\infty}^{55,8} \frac{1}{\sqrt{2\pi(17,6)^2}} e^{-\frac{1}{2}\left(\frac{x-55,8}{17,6}\right)^2} dx + \int_{55,8}^{65} \frac{1}{\sqrt{2\pi(17,6)^2}} e^{-\frac{1}{2}\left(\frac{x-55,8}{17,6}\right)^2} dx
 \end{aligned}$$

Karena $\mu = 55,8 \Rightarrow \int_{-\infty}^{55,8} f(x)dx = 0,5$ dan dengan menggunakan integral numerik metode deret Reimant untuk $\Delta x = 0.0001$, diperoleh bahwa $\int_{55,8}^{65} f(x)dx = 0,19942$ sehingga nilai

dari $P(X < x = 65) = 0,5 + 0,19942 = 0,69942 = 69,94\%$

Jadi peluang siswa yang terpilih memperoleh nilai tes kurang dari 65 adalah 69,942%.

b. Peluang siswa yang terpilih secara acak akan memperoleh nilai tes lebih dari 75 adalah $P(X \geq x=75)$. Nilai ini diberikan sebagai berikut:

$$\begin{aligned}
 P(X \geq x = 75) &= \int_{75}^{\infty} f(x)dx = \int_{55,8}^{\infty} f(x)dx - \int_{55,8}^{75} f(x)dx \\
 &= \int_{55,8}^{\infty} \frac{1}{\sqrt{2\pi(17,6)^2}} e^{-\frac{1}{2}\left(\frac{x-55,8}{17,6}\right)^2} dx - \int_{55,8}^{75} \frac{1}{\sqrt{2\pi(17,6)^2}} e^{-\frac{1}{2}\left(\frac{x-55,8}{17,6}\right)^2} dx
 \end{aligned}$$

Karena $\mu = 55,8 \Rightarrow \int_{55,8}^{\infty} f(x)dx = 0,5$ dan dengan menggunakan

integral numerik metode deret Reimant untuk $\Delta x = 0.0001$, diperoleh bahwa $\int_{55,8}^{75} f(x)dx = 0,36235$.

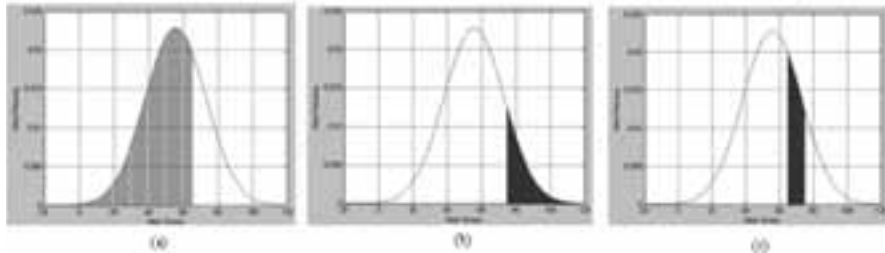
Sehingga nilai dari

$P(X > x = 75) = 0,5 - 0,36235 = 0,13765 = 13,765\%$

Jadi peluang siswa yang terpilih memperoleh nilai tes lebih dari 75 adalah 13,77%.

c. Peluang siswa yang terpilih secara acak akan memperoleh nilai tes antara 65 sampai 75 adalah

$$P(65 \leq X \leq 75) = \int_{65}^{75} \frac{1}{\sqrt{2\pi}(17.6)^2} e^{-\frac{1}{2}\left(\frac{x-55.8}{17.6}\right)^2} dx = 0,16293 = 16,3\%$$



Gambar 5.7 Kurva sebaran normal dengan $\mu = 55,7$ dan $\sigma = 17,5$ serta (a) $P(X=65)$; (b) $P(X>x=75)$; (c) $P(65<X75)$

5.2 Sebaran Normal Baku

Untuk memudahkan perhitungan dalam menentukan nilai integral fungsi padat normal, maka dibuat tabel seperti tertera pada tabel 3 lampiran 3 di belakang. Akan tetapi, tidak mungkin membuat tabel yang berlainan untuk setiap nilai μ dan σ . Seluruh nilai dalam tabel merupakan hasil transformasi setiap peubah acak X menjadi himpunan pengamatan baru satu peubah acak normal Z dengan rata-rata 0 (nol) dan variansi 1 (satu).

Teorema 5.2

Jika X peubah acak yang tersebaran normal dengan rata-rata μ dan simpangan baku σ , maka peubah acak baru $Z = \frac{X - \mu}{\sigma}$ akan tersebaran normal baku

Bukti:

$$X \sim n(x, \mu, \sigma) \Rightarrow f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}; -\infty < x < \infty$$

$z = \frac{x - \mu}{\sigma} \Rightarrow x = z\sigma + \mu \Rightarrow dx = \sigma dz$, akibatnya diperoleh

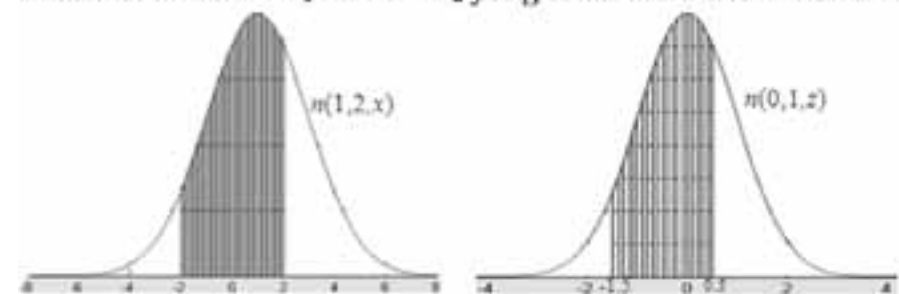
$$\int_{-\infty}^{\infty} f(x) dx = \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} dx = \int_{-\infty}^{\infty} \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}z^2} \sigma dz = \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}z^2} dz$$

$$= \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}(1)^2} e^{-\frac{1}{2}\left(\frac{z-0}{1}\right)^2} dz = \int_{-\infty}^{\infty} n(z, 0, 1) dz$$

Bentuk integral terakhir membuktikan bahwa

$$X \sim n(x, \mu, \sigma) \Rightarrow Z = \frac{X - \mu}{\sigma} \sim n(z, 0, 1)$$

Sebaran asli yang telah ditransformasikan dilukiskan pada Gambar 5.5. Karena semua nilai X antara x_1 dan x_2 mempunyai nilai Z padanan antara z_1 dan z_2 , maka luas daerah di bawah kurva X antara ordinat $x = x_1$ dan $x = x_2$ pada Gambar 5.5 sama dengan luas daerah di bawah kurva Z antara ordinat $z = z_1$ dan $z = z_2$ yang telah ditransformasikan.



Gambar 5.8 Luas daerah di bawah kurva sebaran normal dan di bawah kurva z

Sekarang banyaknya tabel kurva normal yang diperlukan telah diperkecil menjadi satu, yaitu sebaran normal baku. Tabel 3 (lihat lampiran 3) memberikan luas di bawah normal baku yang berpadanan dengan $P(0 < z < b)$. Perlu diingat, bahwa luas adalah besaran skalar, jadi selalu positif (tidak pernah negatif). Sedangkan tanda negatif pada z menunjukkan letak daerahnya saja. Misalnya daerah antara z

$= 0$ dan $z = 2,15$ adalah daerah luasan di sebelah kanan $\mu = 0$, sedangkan daerah antara $z = 0$ dan $z = -2,15$ adalah luasan di sebelah kiri $\mu = 0$. Luas kedua daerah ini sama yang besarnya akan dihitung melalui contoh-contoh berikut.

Contoh 5.5.

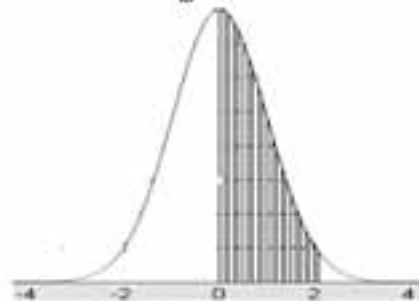
Dengan menggunakan tabel sebaran normal baku (tabel z), tentukan luas daerah

- a. Antara $z = 0$ dan $z = 2,15$
- b. Antara $z = 0$ dan $z = -1,86$
- c. Antara $z = -1,86$ dan $z = 2,15$
- d. Antara $z = 1,34$ dan $z = 2,06$
- e. Untuk $z > 1,96$
- f. Untuk $z > 1,96$

Solusi.

- a. Antara $z = 0$ dan $z = 2,15$

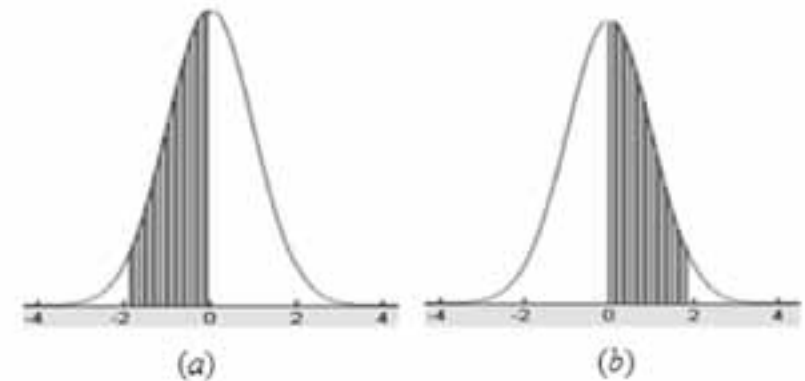
Cara menghitung luasnya adalah dengan melihat tabel z Pada kolom paling kiri cari angka 2,1 dan cari pada baris paling atas angka 05. Dari 2,1 tarik arah ke kanan dan dari 05 tarik menurun, didapat angka 4842. Angka ini merupakan angka dibelakang koma, sehingga luas daerah yang dicari adalah 0,4842. Gambarnya adalah sebagai berikut:



Gambar 5.9 Luar daerah di bawah kurva normal baku antara $z = 0$ dan $z = 2,15$

- b. Antara $z = 0$ dan $z = -1,86$

Karena z bertanda negatif, maka daerah luasannya di sebelah kiri rataan. Karena sifat simetris kurva normal, maka luas daerah antara $z = -1,86$ dan $z = 0$ sama dengan luas daerah antara $z = 0$ dan $z = 1,86$. Seperti cara mencari pada soal a, pada daftar z lihat kolom paling kiri nilai 1,8, kemudian baris paling atas lihat angka 06. Dari 1,8 tarik ke kanan dan dari .06 tarik ke bawah diperoleh angka 4686. Jadi luas daerah yang dicari adalah 0,4686

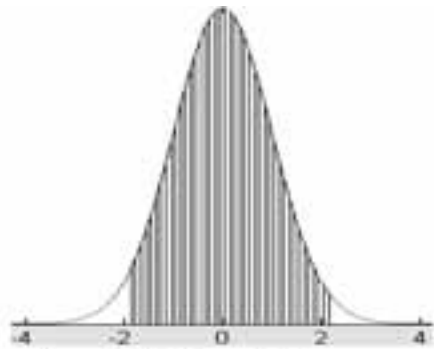


Gambar 5.10 Luar di bawah kurva normal baku antara (a) $z = -1, 86$ dan $z = 0$, (b) $z = 0$ dan $z = 1,86$

- c. Antara $z = -1,86$ dan $z = 2,15$

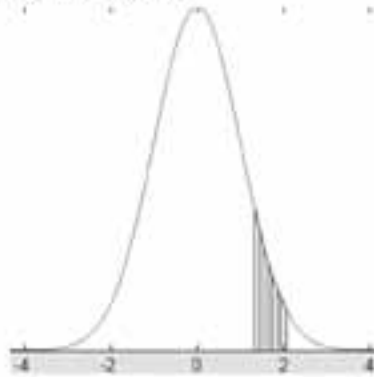
Dari grafik sebagaimana pada Gambar 4.8 di bawah, terlihat bahwa kita perlu mencari luas daerah dua kali, lalu menjumlahkannya. Mengikuti cara pada soal a dan b, maka diperoleh luas daerah antara $z = -1,86 < Z < z = 2,15$ diberikan sebagai berikut:

$$z = -1,86 < Z < z = 2,15 = (z = -1,86 < Z) + (Z < z 2,15) \\ = 0,4686 + 0,4842 = 0,9528$$



Gambar 5.8 Luar daerah di bawah kurva normal baku antara $z = -1.86$ dan $z = 2.15$

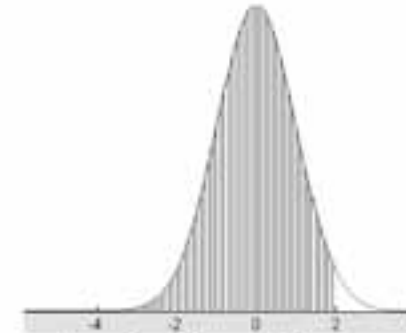
d. Antara $z = 1,34$ dan $z = 2,06$



Gambar 5.9 Luas di bawah kurva normal baku antara $z = 1.34$ dan $z = 2.06$

Untuk soal ini, kita menghitung irisan antara luas $z = 0$ sampai $z = 2,06$ dan $z = 0$ sampai $z = 1,34$. Dengan mengurangi luas $z = 0$ sampai $z = 2,06$ dan $z = 0$ sampai $z = 1,34$. dengan cara yang sama seperti soal-soal sebelumnya, maka diperoleh luas daerah yang diarsir adalah $Z_{1.36} < Z < Z_{2.06} = z_{2.06} - z_{1.34} = 0,4803 - 0,4099 = 0,0704$.

e. $z < 1,96$

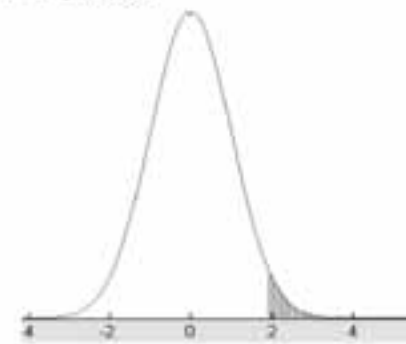


Gambar 5.10 Luas daerah di bawah kurva normal baku antara $z < 1.96$

Luasnya sama dengan dari $z = 0$ ke kiri ditambah dengan luas dari $z = 0$ sampai ke $z = 1,96$. Jadi, luasnya = $0,5000 + 0,4750 = 0,9750$

f. $z > 1,96$

Luasnya sama dengan dari $z = 0$ ke kanan ($=0,5$) dikurangi dengan luas dari $z = 0$ sampai ke $z = 1,96$. Jadi, luasnya = $0,5000 - 0,4750 = 0,0250$.



Gambar 5.11 Luas dibawah kurva normal baku dengan $z > 1,96$

Contoh 5.6

Selesaikan permasalahan pada contoh 5.3 di atas menggunakan sebaran normal baku.

Solusi:

Untuk $x = 65 \Rightarrow z = \frac{65 - 55.7}{17.5} = 0,5314$ dan untuk

$x = 75 \Rightarrow z = \frac{75 - 55.7}{17.5} = 1,1029$ sehingga diperoleh

$$a. \quad P(X < x = 65) = \int_{-\infty}^{65} f(x)dx = \int_{-\infty}^{0,53} f(z)dz = \int_{-\infty}^0 f(z)dz + \int_0^{0,53} f(z)dz$$

$$= 0,5 + P(0 < Z < 0,53) = 0,5 + 0,2019 = 0,7019 = 70,19\%$$

$$b. \quad P(X > x = 75) = \int_{75}^{\infty} f(x)dx = \int_{1,10}^{\infty} f(z)dz = \int_0^{1,10} f(z)dz - \int_0^{\infty} f(z)dz = 0,5 - 0,3643 = 0,1357 = 13,57\%$$

$$c. \quad P(65 < X < 75) = \int_{65}^{75} f(x)dx = \int_{0,53}^{1,10} f(z)dz = \int_0^{1,10} f(z)dz - \int_0^{0,53} f(z)dz = 0,3643 - 0,2019 = 0,1624 = 16,24\%$$

Contoh 5.7

Dari suatu penelitian di Rumah Sakit Sayang Ibu Selong, diperoleh data bahwa rata-rata berat bayi yang baru lahir adalah 3750 gram dengan simpangan baku 325 gram. Jika berat bayi bersebaran normal, maka tentukan ada:

- Berapa bayi yang beratnya lebih dari 4500 gram, jika ada 1000 bayi?
- Berapa bayi yang beratnya antara 3500 gram dan 4500 gram, jika semuanya ada 1000 bayi
- Berapa bayi yang beratnya kurang atau sama dengan 4000 gram, jika semuanya ada 1000 bayi?
- Berapa bayi yang beratnya 4250 gram, jika semuanya ada 5000 bayi?

Penyelesaian:

a. Dengan memakai rumus transformasi, untuk $x = 4.500$ diperoleh: $z = \frac{4.500 - 3.750}{325} = 2,31$ dengan

menggunakan tabel Z, diperoleh nilai $z_{2,31} = 0,4896$. Karena yang ditanyakan adalah berat bayi yang lebih dari 4.500 gram, pada grafik luasnya ada di sebelah kanan $z_{2,31}$, seperti contoh-contoh yang terdahulu, maka luasnya adalah $0,5000 - 0,4896 = 0,0104$. Jadi, ada $0,0104 \times 1.000 = 10,4 \approx 10$ bayi yang beratnya lebih dari 4.500 gram.

b. Dengan $x_1 = 3.500$ dan $x_2 = 4.500$ diperoleh: $z_1 = \frac{3.500 - 3.750}{325} = -0,77$ dan $z_2 = \frac{4.500 - 3.750}{325} = 2,31$.

Dengan bantuan tabel z diperoleh luas $z_{0,77} = 0,2794$ dan $z_{2,31} = 0,4896$, sehingga luas keseluruhan $0,2794 + 0,4896 = 0,7690$. Jadi banyak bayi yang beratnya antara 3.500 gram sampai dengan 4.500 gram diperkirakan ada $(0,7690)(1.000) = 769$ bayi

c. Beratnya kurang dari atau sama dengan 4.000 gram, maka beratnya harus kurang dari 4.000,5 gram. Sehingga untuk $X = 4000,5$ maka $z = \frac{4.000,5 - 3.750}{325} = 0,77$, dan $z_{0,77} = 0,2794$. Peluang

berat bayi kurang dari atau sama dengan 4.000 gram = $0,5 + 0,2794 = 0,7794$. Jadi banyaknya bayi yang beratnya kurang dari atau sama dengan 4.000 gram = $(0,7794) (1.000) = 779,4 \approx 779$ bayi

d. Berat 4.250 gram, berarti berat antara 4.249,5 dan 4.250,5 dengan $x_1 = 4.249,5$ dan $x_2 = 4.250,5$ diperoleh $z_1 = \frac{4.249,5 - 3.750}{325} = 1,53$ dan

$$z_2 = \frac{4.250,5 - 3.750}{325} = 1,54. \text{ Dengan bantuan tabel } z$$

diperoleh luas $z_{1,53} = 0,4370$ dan $z_{1,54} = 0,4382$, sehingga luas keseluruhan adalah $0,4382$, sehingga luas keseluruhan adalah $0,4382 - 0,4370 = 0,0012$. Jadi, banyaknya bayi yang beratnya 4.250 gram = $(0,0012)(5000) = 6$ bayi

APLIKASI SEBARAN NORMAL

Dengan sedikit imajinasi, berat ayam potong seperti yang dicontohkan di atas bisa dianalogikan dengan pengukuran lainnya yang berkaitan dengan keputusan manajerial seperti hasil penjualan bulanan, pengukurann daya rentang suatu material, volume minuman dalam kemasan dan lain sebagainya. Jika data ini bersebaran normal, maka dapat diketahui peluang setiap unsur data apakah termasuk ke dalam suatu nilai-nilai tertentu.

Sebaran normal dapat dikatakan sebagai sebaran yang paling banyak digunakan dalam analisis statistika lanjutan. Banyak analisis statistika lanjutan untuk keperluan penaksiran parameter maupun pengujian hipotesis yang mensyaratkan bahwa data yang dikumpulkan harus bersebaran normal. Olehkarena itu, pemahaman yang mendalam tentang sebaran ini perlu dimiliki oleh seorang manajer agar informasi yang dihasilkan dari analisis data statistika dapat digunakan secara benar dalam proses pengambilan keputusan.

PENDEKATAN NORMAL TERHADAP BINOMIAL

Telah dijelaskan sebelumnya bahwa apabila n sangat besar (di luar tabel binomial) dan p sangat kecil (seperti $np \leq 5$), maka sebaran binomial dapat didekati oleh sebaran Poisson. Akan tetapi apabila n di luar nilai tabel dan p bernilai sangat kecil atau sangat besar, maka sebaran binomial dapat

didekati oleh sebaran normal. Sebagai petunjuk dalam melakukan pendekatan normal dari binomial adalah :

$$n \geq 30$$

$$np \text{ dan } n(1 - p) \geq 5$$

Contoh 5.8 :

Anggota suatu dewan juri berisikan 55% wanita. Berapa peluang terpilihnya 50 anggota juri yang dipilih secara acak akan berisikan anggota wanita sebanyak 30 orang atau lebih. Pemilihan ini jelas merupakan proses binomial dengan $n = 50$, $p = 0,55$ dan $x \geq 30$. Tabel binomial tidak mempunyai nilai untuk $n = 50$. Pendekatan Poisson juga tidak dapat dilakukan karena $np = 27,5$. Demikian pula tehnik menggunakan $1 - p$ untuk p tidak dapat dilakukan juga karena $n(1 - p) = 23,5$. Akan tetapi kriteria untuk pendekatan normal sudah dipenuhi dimana parameter binomial untuk mendekati sebaran normal adalah :

$$\mu_B = np = 27,5$$

$$\sigma_B = \sqrt{np(1 - p)} = 3,52$$

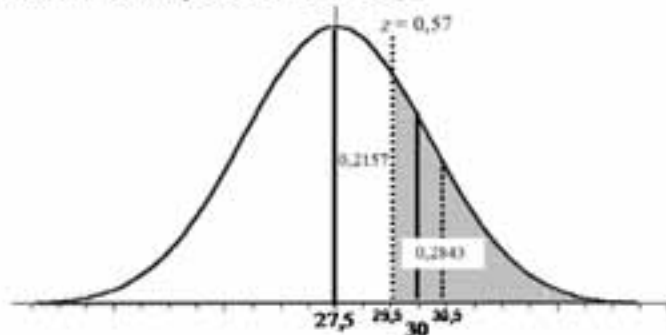
Sebelum menghitung peluang sebaran normal, terlebih dahulu perlu dihitung suatu koreksi yang memperkenankan kita melakukan pendekatan dari sebaran diskrit ke sebaran kontinu. Dalam sebaran kontinu, nilai 29 didefinisikan mengambil nilai antara "28,5 sampai 29,5", nilai 30 di antara nilai 29,5 sampai 30,5 dan seterusnya. Dengan demikian, nilai-nilai diskrit yang sama atau lebih besar dari 30 dapat diperlihatkan dalam Gambar 4.3. Akhirnya persoalan di atas dapat diselesaikan sebagaimana persoalan sebaran normal biasa yaitu :

$$z = \frac{29,5 - 27,5}{3,52} = 0,57$$

Luas area dari 0 sampai 0,57 adalah 0,2157. Jadi :

$$P(X \geq 29,5) = 0,5 - 0,2157 = 0,2843$$

Artinya peluang (pendekatan) terpilihnya anggota juri wanita lebih dari 30 orang adalah 0,2843. (Jika dihitung dengan sebaran binomial diperoleh 0,2862).



Gambar 5.12 Pendekatan normal terhadap binomial

PENDEKATAN NORMAL TERHADAP POISSON

Apabila rata-rata sebaran Poisson lebih dari 10, maka mustahil untuk menggunakan tabel peluang Poisson (meskipun sebenarnya dapat dilakukan dengan komputer). Oleh karenanya pendekatan normal kepada binomial dapat diperluas kepada sebaran Poisson (dalam hal ini $\lambda > 10$).

Contoh 5.9 :

Rata-rata jumlah kendaraan yang mengunjungi bengkel pada jam 16.00 – 17.00 di akhir pekan adalah 16. Berapa peluang bahwa kurang dari 20 kendaraan akan mengunjungi bengkel pada jam yang sama di hari Selasa mendatang.

Rata-rata sebaran Poisson λ lebih dari 10, sehingga pendekatan normal dapat dilakukan. Parameter Poisson yang ekuivalen dengan sebaran normal adalah :

$$\mu_p = \lambda = 16$$

$$\sigma_p = \sqrt{\lambda} = 4$$

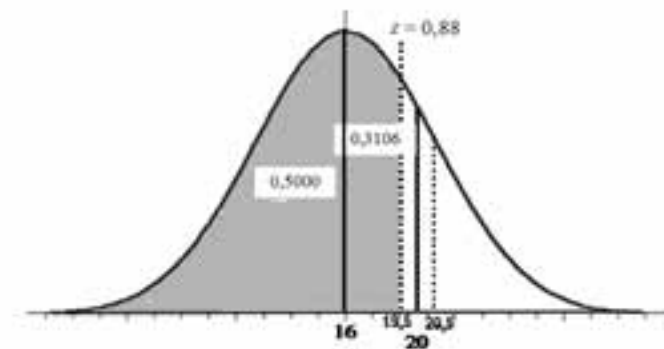
Koreksi dari sebaran diskrit ke kontinu perlu dilakukan seperti yang dicontohkan sebelumnya. Jadi dalam hal ini peluang “kurang dari 20” dapat kita didefinisikan sebagai “kurang atau sama dengan 19,5”. Luas area di bawah kurva normal lihat Gambar 4.4) dapat dihitung dengan :

$$z = \frac{19,5 - 16}{4} = 0,88$$

Dengan menggunakan tabel diperoleh luas areanya adalah 0,3106. Karena nilai z positif, maka luas area yang dicari adalah mulai dari $z = 0,88$ ke arah kiri atau :

$$P(X \leq 19,5) = 0,5 + 0,3106 = 0,8106$$

Jadi peluang (pendekatan) kendaraan yang mengunjungi bengkel di hari Selasa kurang dari 20 buah adalah 0,8106 (perhitungan secara eksak dengan menggunakan sebaran Poisson adalah 0,8122).



Gambar 5.13. Pendekatan normal terhadap Poisson

LATIHAN 5.2

1. Pengukuran tinggi yang dilakukan oleh lembaga pendidikan tentara terhadap sejumlah besar calon prajurit ternyata bersebaran normal. Anggaplah rata-rata tinggi yang diperoleh adalah 168 cm dengan simpangan baku 4 cm. Berapakah peluang seseorang yang diambil secara acak tingginya :
 - a. Kurang dari 165 cm
 - b. Lebih dari 170 cm

2. Jika rata-rata daya tahan lampu pijar adalah 3 tahun dengan simpangan baku satu tahun. Gunakanlah metode kurva normal untuk menghitung peluang satu bola lampu yang dipilih secara acak memiliki daya tahan 4 tahun atau lebih.

3. Berikut ini adalah nilai statistika dari 200 bagi mahasiswa non-eksakta :
 Jika seorang profesor menganggap bahwa nilai ujiannya bersebaran normal dan menetapkan bahwa pembagian skor mutu adalah A = 5%, B = 20%, C = 50%, D = 20% dan E = 5% (biasanya disebut penetapan nilai dengan basis 5-20-50-20-5). Tentukanlah berapa mahasiswa yang mendapat nilai A, B, C, D dan E.

Nilai	Jml. Mhs	Nilai	Jml. Mhs	Nilai	Jml. Mhs	Nilai	Jml. Mhs
9	1	8	2	7	1	6	2
9		7		5	4	4	
9	1	8	3	7	1	6	3
8		5		4	2	3	
9	1	8	4	7	1	6	4
7		4		3	1	2	
9	2	8	5	7	1	6	2
6		3		2	0	1	

9	1	8	6	7	6	6	2
5		2		1		0	
9	1	8	5	7	9	5	1
4		1		0		9	
9	1	8	6	6	7	5	1
3		0		9		7	
9	2	7	8	6	5	5	2
2		9		8		5	
9	2	7	8	6	3	5	2
0		8		7		4	
8	2	7	1	6	5	5	1
9		7	0	6		3	
8	3	7	1	6	3	5	2
8		6	6	5		2	
						5	3
						0	

5.3 Sebaran t Student

Sebaran *t* (atau sering disebut sebaran *t Student*) adalah sebaran probabilitas yang digunakan untuk memperkirakan parameter populasi bila ukuran sampel kecil dan / atau bila ragam populasi tidak diketahui.

Mengapa Menggunakan Sebaran t?

Menurut teorema batas pusat, sebaran sampling statistik (seperti mean sampel) akan mengikuti sebaran normal, selama ukuran sampel cukup besar. Oleh karena itu, ketika kita mengetahui standar deviasi populasi, kita dapat menghitung nilai *z*, dan menggunakan sebaran normal untuk mengevaluasi probabilitas dengan mean sampel.

Tapi ukuran sampel terkadang kecil, dan seringkali kita tidak tahu standar deviasi populasi. Ketika salah satu dari

masalah ini terjadi, statistikawan mengandalkan sebaran statistik t (juga dikenal sebagai nilai t).

Derajat kebebasan

Sebenarnya ada banyak sebaran t yang berbeda. Bentuk sebaran t tertentu ditentukan oleh derajat kebebasannya. Derajat kebebasan mengacu pada jumlah observasi independen dalam sekumpulan data.

Ketika memperkirakan skor rata-rata atau proporsi dari satu sampel, jumlah pengamatan independen sama dengan ukuran sampel dikurangi satu. Oleh karena itu, sebaran statistik t dari sampel berukuran 8 akan dijelaskan oleh sebaran t yang memiliki $8 - 1$ atau 7 derajat kebebasan. Demikian pula, sebaran t yang memiliki 15 derajat kebebasan akan digunakan dengan sampel berukuran 16.

Sifat dari Sebaran t

Sebaran t memiliki sifat sebagai berikut: Rata-rata sebaran sama dengan 0.

- a. Ragamnya sama dengan $v / (v - 2)$, di mana v adalah derajat kebebasan (lihat bagian terakhir) dan $v > 2$.
- b. Ragamnya selalu lebih besar dari 1, meski mendekati 1 bila ada banyak derajat kebebasan. Dengan derajat kebebasan yang tak terbatas, sebaran t sama dengan sebaran normal standar.

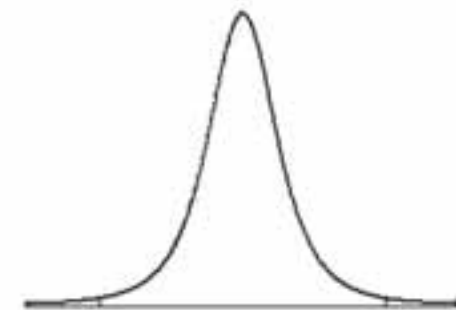
Kapan harus menggunakan sebaran t

Sebaran t dapat digunakan dengan statistik yang memiliki sebaran berbentuk lonceng (yaitu kira-kira normal). Sebaran sampling statistik harus berbentuk lonceng jika ada persyaratan berikut yang berlaku.

- a. Sebaran populasi normal.

- b. Sebaran populasi simetris, unimodal, tanpa outlier, dan ukuran sampel paling sedikit 30.
- c. Sebaran populasi agak miring, tidak rata, tanpa outlier, dan ukuran sampel paling sedikit 40.
- d. Ukuran sampel lebih besar dari 40, tanpa outlier.

Suatu peubah acak kontinu X yang tersebaran sebagaimana pada Gambar 5.14 berikut disebut sebagai sebaran Student's- t



Gambar 5.14 Kurva sebaran Student's- t

Persamaan yang memenuhi kurva tersebut diberikan diberikan sebagai berikut:

Definisi 5.2 PROBABILITY DENSITY FUNCTION (PDF) STUDENT'S t

Peubah acak X tersebaran student's t jika fungsi padatnya diberikan oleh

$$f(t) = \begin{cases} \frac{\Gamma\left(\frac{n+1}{2}\right)}{\sqrt{n\pi}\Gamma\left(\frac{n}{2}\right)\left(1+\frac{t^2}{n}\right)^{\frac{(n+1)}{2}}}, & \text{untuk } -\infty < t < \infty \\ 0 & \text{Untuk } x \text{ lainnya} \end{cases}$$

dengan n bilangan bulat positif yang disebut sebagai derajat kebebasan

Peluang $P(a \leq T \leq b)$ dicari seperti pada fungsi densitas yang lain, yaitu :

$$P(a \leq T \leq b) = \int_a^b f(t) dt.$$

Karena menghitung integralnya cukup sulit, maka dalam mencari probabilitas digunakan tabel sebaran student's t (lihat Tabel 5. lampiran 5).

Contoh 5.9

Tulislah fungsi densitas student's t , jika derajat kebebasannya $n = 4$

Penyelesaian:

Diketahui bahwa $n = 4$, maka diperoleh

$$f(t) = \frac{\Gamma\left(\frac{4+1}{2}\right)}{\Gamma\left(\frac{4}{2}\right)} \cdot \frac{1}{\sqrt{4\pi}} \cdot \frac{1}{\left(1+\frac{t^2}{4}\right)^{\frac{4+1}{2}}} = \frac{\Gamma\left(\frac{5}{2}\right)}{\Gamma(2)} \cdot \frac{1}{2\sqrt{\pi}} \cdot \frac{1}{\left(1+\frac{t^2}{4}\right)^{\frac{5}{2}}} = \frac{\Gamma\left(2+\frac{1}{2}\right)}{2\sqrt{\pi}} \cdot \frac{1}{\left(1+\frac{t^2}{4}\right)^{\frac{5}{2}}}$$

$$= \frac{\frac{3}{4}\sqrt{\pi}}{2\sqrt{\pi}} \cdot \frac{1}{\left(1+\frac{t^2}{4}\right)^{\frac{5}{2}}} = \frac{3}{8\left(1+\frac{t^2}{4}\right)^{\frac{5}{2}}}; \text{ untuk } -\infty < t < \infty$$

Contoh 5.10 Hitunglah dengan menggunakan tabel sebaran t :

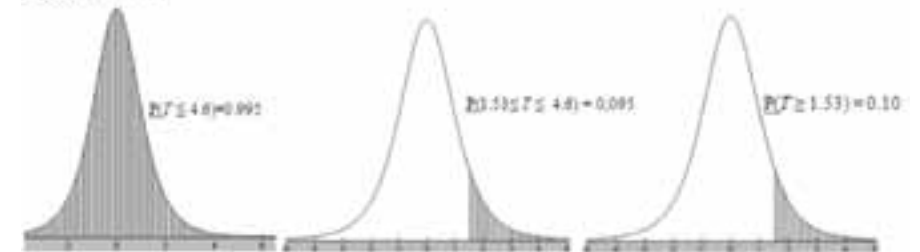
- a. $P(T \leq 4,6)$
- b. $P(1,53 \leq T \leq 4,6)$
- c. $P(T \geq 1,53)$

Penyelesaian :

- a. Untuk menjawab soal ini maka, perlu kita melihat tabel sebaran t , cari kolom paling kiri, temukan angka 4 (sesuai dengan derajat kebebasan-4), dari angka 4 ini kita tarik ke kanan, sehingga menemukan angka 4,6. Dari 4,6 ini kita tarik ke atas, maka kita temukan angka $t_{0,995}$. Jadi, $P(T \leq 4,6) = 0,995$.

- b. Seperti langkah soal a, tetapi kita temukan dahulu $P(T \leq 1,53) = 0,9$ dan $P(T \leq 4,6) = 0,995$, jadi, $P(1,53 < T < 4,6) = P(T \leq 4,6) - P(T \leq 1,53) = 0,995 - 0,90 = 0,095$
- c. $P(T \geq 1,53) = 1 - P(T < 1,53) = 1 - 0,90 = 0,10$

Jawaban tersebut digambarkan dengan grafik seperti di bawah ini



Gambar 5.15 Grafik luas daerah di bawah kurva sebaran student's t

Teorema 5.3 RATA-RATA DAN RAGAM SEBARAN STUDENT'S t

Jika X tersebaran student's t dengan derajat kebebasan n , maka rata-rata dan ragam diberikan oleh $\mu = 0$ dan

$$\sigma^2 = \frac{n}{n-2}, n > 2$$

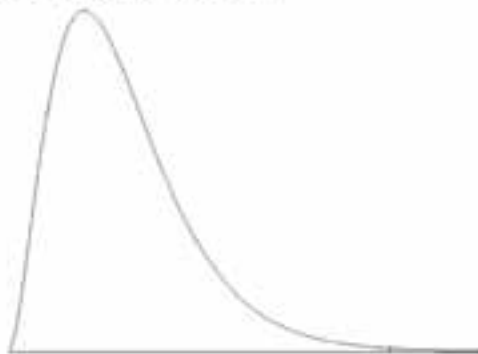
Bukti teorema ini diberikan sebagai latihan bagi mahasiswa. Cara penyelesaian dapat menyesuaikan dengan pembuktian pada teorema 5.1.

Contoh 5.11

Dari contoh 5.10, jika kita hitung rata-rata dan ragamnya, maka diperoleh $\mu = 0$ untuk $n = 4$ dan $\sigma^2 = \frac{4}{4-2} = 2$ untuk $n = 4$.

5.4 Sebaran Chi-Kuadrat

Suatu peubah acak kontinu X yang tersebaran sebagaimana pada gambar 5.16 berikut disebut sebagai sebaran eksponensial, dengan parameter β ,



Gambar 5.16 Kurva sebaran Chi-Kuadrat

Persamaan yang memenuhi kurva tersebut diberikan dengan memilih $\alpha = n/2$ dan $\beta = 2$ pada PDF gamma.

Definisi 5.3 PROBABILITY DENSITY FUNCTION (PDF) GAMMA

Peubah acak X tersebaran chi-kuadrat jika fungsi padatnya diberikan oleh

$$f(x) = \begin{cases} \frac{1}{2^{\frac{n}{2}} \Gamma\left(\frac{n}{2}\right)} x^{\frac{n}{2}-1} e^{-x/2} & x > 0 \\ 0 & x \text{ lainnya} \end{cases}$$

dengan n bilangan bulat positif yang disebut sebagai

derajat kebebasan

Sebaran Chi-Kuadrat ini juga dilambangkan dengan χ^2 .

Seperti juga pada fungsi densitas, maka probabilitasnya dicari dengan $P(a \leq X \leq b) = \int_a^b f(x) dx = \int_a^b \frac{1}{2^{\frac{n}{2}} \Gamma\left(\frac{n}{2}\right)} x^{\frac{n}{2}-1} e^{-x/2} dx$ atau dapat

dicari lewat tabel sebaran Chi-Kuadrat (lihat Tabel 4 pada Lampiran 4).

Contoh 5.12

Tuliskan fungsi densitas chi-kuadrat dengan derajat kebebasan 4.

Penyelesaian:

Diketahui bahwa derajat kebebasan 4, berarti $n = 4$, jadi fungsi densitasnya adalah:

$$f(x) = \begin{cases} \frac{1}{2^2 \Gamma(2)} x^{2-1} e^{-x/2} = \frac{1}{4} x e^{-x/2} & \text{untuk } x > 0 \\ 0 & \text{untuk } x \text{ lainnya} \end{cases}$$

Contoh 5.13

Dari fungsi densitas soal contoh 4.7 di atas, hitunglah:

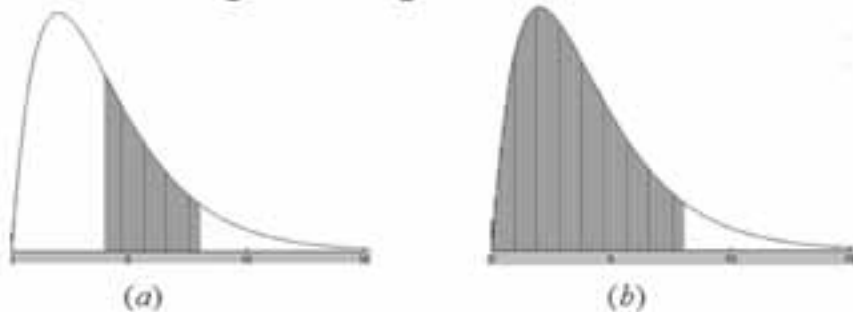
- a. $P(4 \leq X \leq 8)$
- b. $P(X \leq 8)$

Penyelesaian:

Telah diperoleh bahwa fungsi densitasnya adalah

$f(x) = \frac{1}{4} x e^{-x/2}$, untuk $x > 0$, maka diperoleh

$$\begin{aligned}
 \text{a. } P(4 \leq X \leq 8) &= \int_4^8 \frac{1}{4} x e^{-\frac{x}{4}} dx = \frac{1}{4} \int_4^8 2x d(e^{-\frac{x}{4}}) = -\frac{1}{2} \left[x e^{-\frac{x}{4}} + 2 e^{-\frac{x}{4}} \right]_4^8 \\
 &= 3e^{-2} - 5e^{-4} = 0.0578 \\
 \text{b. } P(X \leq 8) &= \int_0^8 \frac{1}{4} x e^{-\frac{x}{4}} dx = -\frac{1}{2} \left[x e^{-\frac{x}{4}} + 2 e^{-\frac{x}{4}} \right]_0^8 = 1 - 5e^{-4} = 0.9084
 \end{aligned}$$



Gambar 5.17 Luas daerah di bawah kurva chi-kuadrat antara (a) $x_1=4$ dan $x_2=8$; (b) $X < 8$

Contoh 5.14

Diketahui peubah acak X bersebaran chi-kuadrat dengan derajat kebebasan 8. Dengan menggunakan tabel sebaran chi-kuadrat, hitunglah:

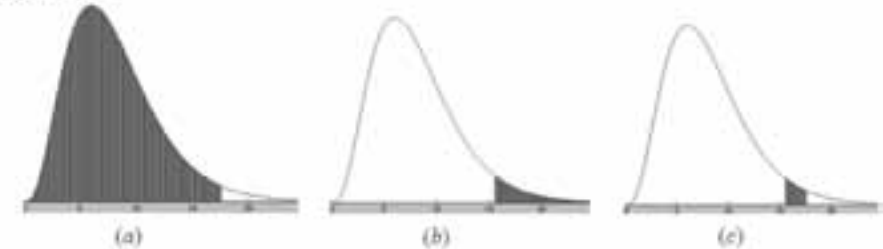
- a. $P(X < 17,5)$
- b. $P(X > 15,5)$
- c. $P(15,5 < X < 17,5)$

Penyelesaian:

- a. Pada kolom pertama (kolom derajat kebebasan) pada tabel, pilih angka 8, kemudian tarik ke kanan hingga diperoleh angka 17,5. Selanjutnya tarik ke atas dan diperoleh nilai 0.975. Jadi $P(X < 17,5) = 0.975$
- b. Dengan cara yang sama, dari angka 8, tarik kekanan hingga diperoleh nilai 15,5. Selanjutnya tarik ke atas dan diperoleh nilai 0.95. Nilai tersebut adalah $P(X < 15,5)$. Hubungan antara $P(X < 15,5)$ dengan $P(X > 15,5)$ adalah

$$\begin{aligned}
 P(X < 15,5) + P(X > 15,5) &= 1. \text{ Jadi nilai dari } P(X > 15,5) \\
 &= 1 - P(X < 15,5) = 1 - 0.95 = 0.05 \\
 \text{c. } P(15,5 < X < 17,5) &= P(X < 17,5) - P(X < 15,5) = 0.975 - 0.95 = 0.025
 \end{aligned}$$

Gambar dari kurva peluang yang dimaksud diberikan sebagai berikut:



Gambar 5.18 Luas daerah di bawah kurva chi-kuadrat dengan (a) $P(X < 17,5) = 0.975$; (b) $P(X < 15,5) = 0.95$ dan (c) $P(15,5 < X < 17,5) = 0.025$

Teorema 5.4 RATAAN DAN RAGAM SEBARAN CHI-KUADRAT

Jika X tersebaran chi-kuadrat dengan derajat kebebasan n , maka rata-rata dan ragam diberikan oleh $\mu = n$ dan $\sigma^2 = 2n$

Bukti teorema ini diberikan sebagai latihan bagi mahasiswa. Cara penyelesaian dapat disesuaikan dengan pembuktian pada Teorema 5.1.

Contoh 5.15

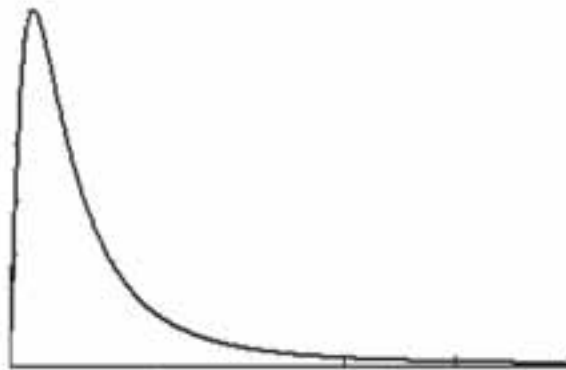
Tentukan rata-rata dan variansi dari contoh 5.14.

Penyelesaian:

Dari soal diperoleh nilai $n = 8$, maka $\mu = 8$ dan $\sigma^2 = 2 \cdot 8 = 16$

5.5 Sebaran F

Suatu peubah acak kontinu X yang tersebar sebagaimana pada gambar 5.19 berikut disebut sebagai sebaran F .



Gambar 5.19 Kurva sebaran F

Persamaan yang memenuhi kurva tersebut diberikan diberikan sebagai berikut:

Definisi 5.4: PROBABILITY DENSITY FUNCTION (PDF) F
 Peubah acak tersebar F jika fungsi padatnya diberikan

$$\text{oleh } f(F) = \begin{cases} \frac{\tau\left(\frac{n_1+n_2}{2}\right)\left(\frac{n_1}{n_2}\right)^{\frac{n_1}{2}} F^{\frac{n_1}{2}-1}}{\tau\left(\frac{n_1}{2}\right)\tau\left(\frac{n_2}{2}\right)\left(1+\frac{n_1 F}{n_2}\right)^{\frac{n_1+n_2}{2}}} & 0 < F < \infty \\ 0 & \text{x lainnya} \end{cases}$$

dengan n_1, n_2 bilangan bulat positif dimana n_1 disebut sebagai derajat kebebasan pembilang dan n_2 disebut sebagai derajat kebebasan penyebut.

Seperti juga sebaran lainnya, untuk keperluan perhitungan dengan sebaran F , daftar (tabel) sebaran F telah disediakan, seperti dalam lampiran 6. Tabel tersebut berisikan nilai-nilai F untuk peluang 0,01 dan 0,05 dengan derajat kebebasan V_1 dan V_2 peluang ini sama dengan luas daerah ujung kanan yang diarsir, sedangkan $dk = V_1$ ada pada baris paling atas dan $dk = V_2$ pada kolom paling kiri. Untuk tiap $dk = V_2$, daftar terdiri atas dua baris, yang di atas untuk peluang $P = 0,05$ dan yang bawah untuk $P = 0,01$.

Contoh 5.16

Untuk pasangan derajat kebebasan $V_1 = 24$ dan $V_2 = 8$, ditulis juga $(V_1, V_2) = (24,8)$, maka untuk $P = 0,05$ didapat $F = 3,12$, sedangkan untuk $P = 0,01$, didapat $F = 5,28$ (lihat tabel 6, lampiran 6). Ini diperoleh dengan cara mencari 24 pada baris atas dan 8 kolom kiri. Jika dari 24 turun dan dari 8 ke kanan, maka diperoleh angka-angka tersebut. Yang atas untuk $P = 0,05$ dan yang bawah untuk $P = 0,01$. Notasi lengkap untuk nilai-nilai F daftar sebaran F dengan peluang P dan $dk = (V_1, V_2)$ adalah $F_{P(V_1 \text{ dan } V_2)}$. Sehingga pada contoh ini, $F_{0,05(24,8)} = 3,12$ dan $F_{0,01(24,8)} = 5,28$.

Meskipun daftar yang diberikan hanya untuk peluang $P = 0,05$ dan $P = 0,01$, tetapi sebenarnya masih bisa didapat nilai-nilai F dengan peluang 0,99 dan 0,95.

Untuk itu diperlukan hubungan $F_{(1-P)(V_2, V_1)} = \frac{1}{F_{P(V_1, V_2)}}$. Dalam rumus di atas, perhatikan antara P dan $(1-P)$ dan pertukaran antara derajat kebebasan (V_1, V_2) menjadi (V_2, V_1) .

Contoh 5.17

Berdasarkan Tabel F, $F_{0,05(24,8)} = 3,12$, maka $F_{0,95(8,24)} = \frac{1}{F_{0,05(24,8)}} = \frac{1}{3,12} = 0,321$

Teorema 5.5 RATAAN DAN RAGAM SEBARAN F

Jika X tersebar F dengan derajat kebebasan n_1 dan n_2 , maka rata-rata dan ragam diberikan oleh $\mu = \frac{n_2}{n_2 - 2}$ dan

$$\sigma^2 = \frac{2n_2^2(n_2 + n_1 - 2)}{n_1(n_2 - 2)^2(n_2 - 4)}$$

Bukti teorema ini diberikan sebagai latihan bagi mahasiswa. Cara penyelesaian dapat menyesuaikan dengan pembuktian pada Teorema 5.1.

Latihan 5.5

1. Seorang dipilih dari 10 dosen untuk mengawasi suatu ujian penyaringan dengan cara memilih suatu gulungan kertas dari sebuah kantong berisi 10 gulungan bernomor 1 sampai 10. Jika X adalah peubah acak yang menyatakan bilangan yang terulisi dalam gulungan kertas yang diambil secara acak,
 - a. tentukan rumus sebaran peluang X
 - b. berapakah peluang mengambil bilangan lebih kecil dari 4
 - c. rata-rata peubah acak X
 - d. ragam peubah acak X
2. Dalam suatu penelitian di Kota X , diperoleh data bahwa 75% dari pencurian karena alasan profesi. Tentukan peluangnya bahwa diantara 5 kasus pencurian selanjutnya yang dilaporkan di daerah tersebut
 - a. tepat 2 karena alasan profesi
 - b. paling banyak 3 karena alasan profesi

3. Seorang petani bawang merah mengeluh karena $\frac{2}{3}$ dari panen bawang merahnya terserang sejenis virus. Tentukan peluangnya bahwa diantara 4 bawang merah yang diperiksa dari hasil panen ini
 - a. semuanya terserang virus
 - b. antara 1 sampai 3 yang terserang virus.
4. Peluang seorang mahasiswa lulus pada ujian mata kuliah Statistik matematika adalah 0,9.
 - a. Berapakah peluang tepat 5 dari 7 yang mengikuti ujian akan lulus ?
 - b. Hitung rata-ratanya
 - c. Hitung ragamnya
5. Seorang insinyur pengawas lalu lintas melaporkan bahwa 75% kendaraan yang melintasi daerah Masbagik adalah berasal dari kota Selong. Berapakah peluang bahwa kurang dari 4 dari 9 kendaraan mendatang yang melalui pemeriksaan tersebut berasal dari luar kota ?
6. Suatu penelitian tentang kesukaan warna sepeda motor di kota Sakra menunjukkan bahwa 20% lebih menyenangi warna hitam dari pada warna lainnya.
 - a. Berapakah peluang bahwa lebih dari setengah dari 20 sepeda motor yang akan dibeli warga kota Sakra akan berwarna hitam.
 - b. Hitunglah rata-ratanya!
 - c. Hitunglah ragamnya!
7. Menurut teori genetika, persilangan tertentu sejenis marmut akan menghasilkan keturunan berwarna merah, hitam, dan putih dalam perbandingan 8 : 4: 4. Tentukan peluang bahwa 5 dari 8 turunan akan berwarna 2 merah, 2 hitam, dan 1 putih.
8. Seseorang menanam 6 bibit bunga di pekarangannya, yang diambil secara acak dari sebuah kotak berisi 5 bibit

bunga matahari, dan 4 bibit bunga mawar, berapakah peluang dia menanam 2 bibit bunga matahari. Dan 4 bibit bunga mawar ?

Bab 6

DISTRIBUSI SAMPLING

Misalkan kita mengambil semua sampel ukuran n yang mungkin dari populasi tertentu. Misalkan lebih jauh bahwa kita menghitung statistik (misalnya, rata-rata, proporsi, simpangan baku) untuk setiap sampel. Distribusi peluang statistik ini disebut *distribusi sampling*. Dan simpangan baku statistik ini disebut *standard error*.

Beberapa istilah dalam distribusi sampling:

- **Sampling** : Penentuan (pemilihan) anggota sampel (unit analisis) dari populasinya dengan cara-cara tertentu. Selanjutnya terhadap unit analisis yang terpilih dilakukan pengamatan mengenai karakteristik tertentu, melalui proses kuantifikasi, baik melalui pengukuran maupun penghitungan.
- **Sampel (*sample*)** : Sebagian dari anggota populasi yang diambil dengan cara-cara tertentu.
- **Populasi (*population*)** : Keseluruhan hasil yang mungkin terjadi dari suatu proses pengukuran atau penghitungan mengenai karakteristik tertentu.
- **Sampel acak (*random sample*)** : Sebagian dari populasi yang diambil dengan cara-cara tertentu, dan setiap anggota populasi memiliki peluang yang sama untuk terambil ke dalam sampel.

- Sampling juga diartikan sebagai : pengumpulan data dengan cara memeriksa (mengukur atau menghitung) karakteristik setiap anggota sampel.
- **Sensus** : Pengumpulan data dengan cara memeriksa karakteristik seluruh anggota populasi.
- Kelebihan Sampling di banding Sensus :
 - (1) Biaya relatif lebih kecil
 - (2) Waktu lebih cepat
 - (3) Ketelitian lebih tinggi
- Sampel acak diperoleh dengan teknik sampling acak :
 - (1) Sampling Acak Sederhana (*Simple Random Sampling*)
 - (2) Sampling Acak Stratifikasi (*Stratified Random Sampling*)
 - (3) Sampling Acak Berklaster (*Clustered Random Sampling*)
 - (4) Sampling Acak Sistematis (*Systematic Random Sampling*)
- Sampel non acak (*non random sample*) diperoleh dengan cara :
 - (1) *Accidental Sampling*
 - (2) *Quota Sampling*
 - (3) *Purposive sampling*
 - (4) Snow ball sampling
 - Distribusi Sampling akan meliputi :
 - (1) Distribusi sampling rata-rata
 - (2) Distribusi sampling beda dua rata-rata
 - (3) Distribusi sampling proporsi
 - (4) Distribusi sampling beda dua proporsi

6.1. Sifat-sifat Distribusi Sampling

Sifat distribusi sampling ini disebut Central Limit Theorem (teorema limit pusat). Sifat inilah yang mendasari teori inferens. Sifat sifat tersebut adalah sebagai berikut .

Sifat 1

Apabila sampel-sampel random dengan n elemen masing – masing diambil dari satu populasi normal., yang mempunyai mean = μ varian σ^2 , distribusi sampling harga mean akan mempunyai mean sama dengan μ dan varian σ^2/n atau standar deviasi σ / \sqrt{n} . Standar deviasi distribusi sampling harga mean ini dikenal sebagai "Standar Error " (SE).

Sifat 2

Apabila populasi berdistribusi normal, distribusi sampling harga mean juga akan berdistribusi normal. Maka berlaku sifat seperti persamaan dibawah ini (z score adalah nilai deviasi relatif antara nilai sampel dan populasi = nilai distribusi normal standar).

$$z = \frac{\bar{x} - \mu}{SE}$$

Sifat 3

Walaupun populasi berdistribusi sembarang, kalau diambil sampel-sampel berulang kali secara random, distribusi harga meannya akan membentuk distribusi normal.

Contoh 6.1 :

Dipunyai populasi lima orang penderita penyakit "D" yang masa inkubasinya sebagai berikut :

No Pasien	Masa inkubasi (hari)
1	2
2	3
3	6
4	8
5	11

$\mu = 6$ hari berasal dari $2+3+6+8+11/5$

$\sigma^2 = 10,8$ hari berasal dari $\frac{\sum(x-\mu)^2}{n-1}$

$\mu = \sqrt{10,8} = 3,29$ hari

Diambil sampel dengan ukuran $n = 2$.

Dari populasi kemungkinan sampel yang terjadi $5^2 = 25$

Sampel sampel tersebut seperti tertera di dalam tabel dibawah ini.

Sampel	Pasien yang terpilih	Masa Inkubasi	Mean
1	1;1	2;2	2
2	1;2	2;3	2.5
3	1;3	2;6	4
4	1;4	2;8	5
5	1;5	2;11	6.5
6	2;1	3;2	2.5
7	2;2	3;3	3
8	2;3	3;6	4.5
9	2;4	3;8	5.5
10	2;5	3;11	7
11	3;1	6;2	4
12	3;2	6;3	4.5
13	3;3	6;6	6
14	3;4	6;8	7
15	3;5	6;11	8.5
16	4;1	8;2	5
17	4;2	8;3	5.5
18	4;3	8;6	7
19	4;4	8;8	8
20	4;5	8;11	9.5
21	5;1	11;2	6.5
22	5;2	11;3	7

Sampel	Pasien yang terpilih	Masa Inkubasi	Mean
23	5;3	11;6	8.5
24	5;4	11;8	9.5
25	5;5	11;11	11

Dari data distribusi sampling (Data pada kolom 4) didapatkan :

$$\bar{x}_x = \frac{2+2,5+4+...+11}{25} = 6,..... = \mu$$

$$\text{Varian (SE}^2) = \frac{\sum(\bar{x}-x)^2}{n-1} = 5,4 \text{ nilai ini tidak lain adalah } \sigma^2/n$$

$$= 10,8/2 = 5,4$$

$$SE = \sqrt{5,4} = 2,32 \text{ hari.}$$

Contoh 6.2:

Jika diketahui bahwa kadar hemoglobin (Hb) orang sehat (μ) = 12 g% dan (σ) = 2,5 g%. Jika seorang peneliti mengambil sebanyak 25 orang sampel diambil secara random, berapa probabilitas dari rata-rata Hb sampel tersebut.

a. Sekurang-kurangnya 13 g%

$$\mu = 12 \text{ g\%} \quad \sigma = 2,5 \text{ g\%} \quad n = 25$$

$$SE = \frac{\sigma}{\sqrt{n}} = \frac{2,5}{25} = 0,5 \text{ g\%}$$

$$Z = \frac{13-12}{0,5} = 2 \rightarrow \text{tabel} = 0,4772$$

$$\text{Jadi } p(x > 13 \text{ g\%}) = 0,5 - 0,4772 = 0,0228$$

b. Antara 11 sampai dengan 13,5 g%

$$Z_1 = \frac{11-12}{0,5} = -2 \rightarrow \text{tabel} 0,4772$$

$$Z_2 = \frac{13,5-12}{0,5} = 3 \rightarrow \text{tabel} 0,4987$$

$$\text{Jadi } p(11 \text{ g\%} < \bar{x} < 13,5 \text{ g\%}) = 0,4772 + 0,4987 = 0,9759$$

6.2 Jenis-Jenis Distribusi Sampling

Berdasarkan besaran statistik yang digunakan, dikenal beberapa jenis distribusi sampling yaitu distribusi rata-rata (distribusi rata-rata sampel), proporsi (distribusi proporsi sampel), beda dua rata-rata, dan beda dua proporsi.

6.2.1 Distribusi sampling rata-rata

Merupakan distribusi probabilitas yang dapat terjadi dari rata-rata sampelnya, yang di dasarkan pada sejumlah sample tertentu dan informasi secara global dari parameter populasinya.

Rumus : $\mu_{\bar{x}} = \mu$, jika $\frac{n}{N} \leq 0,05$ maka tidak perlu menggunakan faktor koreksi

Faktor koreksi (Fk) = $\sqrt{\frac{N-n}{N-1}}$

Angka Baku Z = $\frac{\bar{X}-\mu}{s}$

Dimana s = $\frac{\delta}{\sqrt{n}}$ → angka baku diatas menjadi Z = $\frac{\bar{X}-\mu}{\delta/\sqrt{n}}$

Jika n/N > 0,05 maka gunakan faktor koreksi → Z = $\frac{\bar{X}-\mu}{\frac{\delta}{\sqrt{n}} \sqrt{\frac{N-n}{N-1}}}$

Latihan 6.1

1. Misalkan sebuah populasi berukuran 80, variabel acak X dalam populasi tersebut berdistribusi Normal dengan rata-rata 69,7 dan varians 35. Kemudian dengan sampling tanpa pengembalian, diambil sampel acak berukuran Tentukanlah probabilitas bahwa dalam sampel tersebut rata-ratanya :
 - a. Paling sedikit 65
 - b. Antara 68 dan 71
 - c. Lebih dari 71
2. Berat kantong-kantong semen yang diisikan secara otomatis mengikuti distribusi Normal dengan rata-rata 50

Kg dan varians 10,24 Kg. Bila diambil sampel acak berukuran 25 kantong semen, maka tentukanlah peluang bahwa rata-rata berat kantong semen tersebut :

- a. Paling sedikit 48,5 Kg
- b. Antara 48,5 sampai dengan 49 Kg

3. Suatu populasi berukuran 100, variabel acak Y dalam populasi tersebut berdistribusi Normal dengan rata-rata 50 dan varians 16. Kemudian dari populasi tersebut diambil sampel acak berukuran 30, tentukanlah probabilitas bahwa rata-rata sampel adalah :
 - a. Paling besar 51
 - b. Antara 49 dan 51
 - c. Paling sedikit 48,5
4. Suatu perusahaan mempekerjakan 800 pegawai yang terdiri dari 75% pegawai pria dan sisanya pegawai wanita. Berdasarkan catatan bagian personalia, rata-rata waktu terlambat masuk kerja adalah 32 menit dan simpangan baku 9,6 menit. Suatu ketika pimpinan perusahaan melakukan sidak, dengan mengambil secara acak sampel berukuran 100 pegawai. Tentukanlah ada berapa pegawai yang terlambat masuk :
 - a. Paling sedikit 30 menit
 - b. Kurang dari 33 menit
 - c. Antara 30 hingga 33 menit

6.2.2 Distribusi sampling proporsi

Proporsi dari populasi dinyatakan dengan $P = \frac{x}{N}$ dan proporsi untuk sampel dinyatakan dengan $p = \frac{x}{n}$. Distribusi sampling proporsi adalah distribusi dari proporsi yang diperoleh dari semua sampel sama besar yang mungkin ada dari satu populasi. Distribusi sampling proporsi juga memiliki asrti penting sebagaimana halnya distribusi sampling rata-

rata. Distribusi sampling proporsi dapat digunakan untuk mengetahui persentase atau perbandingan antara dua hal yang berkomplemen (peristiwa binomial) seperti anemia dan tidak anemia, persentase ASI eksklusif dan Bukan ASI eksklusif, dan perbandingan antara pemakai dan bukan pemakai kontrasepsi tertentu.

Contoh 6.3:

Ada sebuah populasi beranggotakan 6 orang, 3 orang diantaranya anemia dan selebihnya tidak anemia. Apabila diambil sampel yang beranggotakan 3 orang, proporsi atau banyaknya sampel untuk ketiganya anemia, dua anemia dan satu tidak anemia, satu anemia dan dua tidak anemia, dan ketiganya tidak anemia dapat diketahui (pengambilan sampel tanpa pengembalian), misalnya anggota populasi adalah A, B, C untuk anemia dan X,Y, Z untuk yang tidak anemia. Banyaknya sampel yang dapat diambil adalah:

$$C_3^6 = \frac{6!}{3!(6-3)!} = 20 \text{ buah sampel, kedua puluh}$$

sampel itu adalah:

- | | | | |
|--------|---------|---------|---------|
| 1. ABC | 6. ACY | 11. BCX | 16. BYZ |
| 2. ABX | 7. ACZ | 12. BCY | 17. CXY |
| 3. ABY | 8. AXY | 13. BCZ | 18. CXZ |
| 4. ABZ | 9. AXZ | 14. BXY | 19. CYX |
| 5. ACX | 10. AYZ | 15. BXZ | 20. XYZ |

Distribusi sampling proporsi (X = Anemia, n = 3) adalah seperti disajikan dalam tabel berikut.

Disribusi Sampling Proporsi			
Sampel yang mungkin (X)	Proporsi sampel (X/n)	f	Probabilitas
X = 3 (3 (a), 0 (ba))	1	1	0,05
	0,67	9	0,45

X = 2 (2 (a), 1 (ba))	0,33	9	0,45
	0	1	0,05
X = 1 (1 (a), 2 (ba))			
X = 0 (0 (a), 3 (ba))			
Jumlah		20	1,00

Keterangan: a = anemia; ba = bukan anemia; 3(a), 0 (ba) = ABC; 2 (a), 1 (ba) = ABX, ABY, ABZ, ACX, ACY, ACZ, BCX, BCY, BCZ; 1(a), 2 (ba) = AXY, AXZ, AYZ, BXY, BXZ, BYZ, CXY, CXZ, CYZ ; 0(a), 3 (ba) = XYZ

Pada distribusi sampling proporsi, berlaku hal-hal sebagai berikut.

1. Jika pengambilan sampel dengan pengembalian atau jumlah populasi besar dibandingkan dengan jumlah sampel yakni $\frac{n}{N} \leq 5\%$, memiliki rata-rata dan standar deviasi sebagai berikut.

$$\mu_p = P$$

$$\sigma_p = \sqrt{\frac{P(1-P)}{n}} = \sqrt{\frac{PQ}{n}}$$

Keterangan:

P = Proporsi kejadian sukses

Q = Proporsi kejadian gagal (1-P)

2. Jika pengambilan sampel tanpa pengembalian atau jika jumlah populasi kecil dibandingkan dengan jumlah sampel yakni $n/N > 5\%$, memiliki rata-rata dan standar deviasi sebagai berikut.

$$\begin{aligned} \mu_p &= P \\ \sigma_p &= \sqrt{\frac{P(1-P)}{n}} = \sqrt{\frac{N-n}{N-1}} \\ &= \sqrt{\frac{PQ}{n}} \sqrt{\frac{N-n}{N-1}} \end{aligned}$$

LATIHAN 6.2

1. Ada petunjuk kuat bahwa 10% anggota masyarakat berpendidikan SD. Sebuah sampel acak terdiri dari 100 orang telah diambil. Tentukan probabilitas bahwa dalam sampel tersebut terdapat :
 - a. Kurang dari 11 orang yang berpendidikan SD
 - b. Paling sedikit 14 orang yang berpendidikan SD
 - c. Antara 12 dan 14 orang yang berpendidikan SD
2. Telah diambil 1000 sampel anak-anak masing-masing berukuran 200 anak. Dalam berapa buah sampel diharapkan dapat ditemukan :
 - a. Anak laki-laki kurang dari 42%
 - b. Anak perempuan antara 43% dan 55%
 - c. Anak laki-laki lebih dari 56%
3. Seorang pramuniaga yang telah berpengalaman bertahun-tahun dalam menjual suatu alat pembersih udara berpendapat bahwa probabilitas berhasil menjual satu set alat tersebut setelah memberikan presentasi kepada para pelanggan adalah 0,8. Jika suatu ketika ia mempresentasikan kepada 100 pelanggan, maka berapakah probabilitas bahwa :
 - a. Paling sedikit 75 pelanggan akan membeli
 - b. kurang dari 15 orang yang tidak akan membeli

4. Dari 1000 mahasiswa FE yang menempuh ujian statistik diperoleh rata-rata 58,2 dan standar deviasi 20,7. Sesuai dengan kriteria kelulusan, ternyata dari 1000 mahasiswa peserta ujian tersebut berhasil lulus sebanyak 650 orang. Jika diambil satu kelas secara acak yang terdiri dari 60 mahasiswa, maka berapakah probabilitas dari kelas tersebut akan terdapat :
 - a. Paling sedikit 30 mahasiswa berhasil lulus ujian
 - b. Kurang dari 27 mahasiswa yang tidak lulus ujian
 - c. Antara 30 dan 36 mahasiswa berhasil lulus ujian

6.2.3 Distribusi sampling beda dua rata-rata

Distribusi sampling beda dua rata-rata adalah distribusi dari perbedaan dua nilai rata-rata yang muncul dari sampel-sampel dua populasi. Misalnya, dua populasi normal N_1 dan N_2 memiliki rata-rata μ_1 dan μ_2 dan standar deviasi masing-masing σ_1 dan σ_2 . Dari kedua populasi N_1 dan N_2 tersebut di ambil sampel random yakni n_1 dan n_2 dengan rata-rata \bar{X}_1 dan \bar{X}_2 kemudian dari kedua rata-rata tersebut di hitung semua bedanya. Dari semua beda rata-rata yang diperoleh akan membentuk suatu distribusi yang disebut distribusi sampling beda rata-rata. Untuk distribusi sampling beda rata-rata pada N_1 dan N_2 cukup besar berlaku hal-hal berikut :

1. Rata-rata

$$\mu_{x_1-x_2} = \mu_1 - \mu_2$$
2. Standar Deviasi

$$\sigma_{x_1-x_2} = \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}$$
3. Untuk n_1 dan n_2 dengan $n_1, n_2 > 30$. Distribusi sampling beda rata-rata akan memberikan distribusi normal dengan variabel random standar yang rumus Z-nya adalah:

$$Z = \frac{(X1-X2) - (\mu1-\mu2)}{\sigma_{x1-x2}}$$

Contoh 6.4:

Misalkan rata-rata kadar kolesterol manajer dan karyawan masing-masing adalah 250 mg/dl dengan standar deviasi 50 mg/dl dan 180 mg/dl dengan standar deviasi 35 mg/dl. Jika di ambil sampel random manajer sebanyak 50 orang dan karyawan sebanyak 200 orang. Hitunglah :

a. Beda rata-rata kadar kolesterol sampel

$$\begin{aligned} \mu1 &= 250 & \mu2 &= 180 \\ \sigma1 &= 50 & \sigma2 &= 35 \\ n1 &= 50 & n2 &= 200 \end{aligned}$$

$$\begin{aligned} \mu_{x1-x2} &= \mu1-\mu2 \\ &= 250-180 \\ &= 70 \text{ mg/dl} \end{aligned}$$

b. Standar deviasi rata-rata kadar kolesterol sampel

$$\begin{aligned} \sigma_{x1-x2} &= \sqrt{\frac{\sigma1^2}{n1} + \frac{\sigma2^2}{n2}} \\ &= \sqrt{\frac{50^2}{50} + \frac{35^2}{200}} = 7,49 \end{aligned}$$

c. Probabilitas beda rata-rata kadar kolesterol manajer dan karyawan lebih dari 80 mg/dl.

$$\begin{aligned} Z &= \frac{(X1-X2) - (\mu1-\mu2)}{\sigma_{x1-x2}} \\ &= \frac{80-70}{7,49} = 1,34 \end{aligned}$$

Kemudian dilihat luas area pada tabel distribusi normal

(tabel-Z Z= 0,4099

$$\begin{aligned} P(\bar{X}1-\bar{X}2 > 80) &= P(Z > 1,34) \\ &= 0,5 - 0,4099 \\ &= 0,0901 \end{aligned}$$

Latihan 6.3

1. Rata-rata tinggi badan mahasiswa 167 cm dan simpangan baku 5,2 cm. Sedangkan rata-rata tinggi badan mahasiswi 156 cm dan simpangan baku 4,9 cm. Jika dari dua kelompok mahasiswa tersebut masing-masing diambil sampel acak berukuran sama yaitu 140 orang, maka berapa peluang rata-rata tinggi badan mahasiswa paling sedikit 10 cm lebihnya dari rata-rata tinggi badan mahasiswi.
2. Sebuah perusahaan kabel menghasilkan dua jenis kabel, yaitu kabel A dan kabel B. Kabel A mempunyai rata-rata kekuatan tarik 4000 kg dan simpangan 300 kg. Sedangkan kabel B mempunyai rata-rata kekuatan tarik 4500 kg dan simpangan baku 200 kg. Jika diambil sampel acak masing-masing berukuran 100 potong kabel A dan 50 potong kabel B, maka tentukan pro-babilitas bahwa kabel B memiliki rata-rata kekuatan tarik :
 - a. Paling sedikit 600 kg lebih kuat daripada kabel A
 - b. Kurang dari 450 kg lebih kuat daripada kabel A

6.2.4 Distribusi sampling beda dua proporsi

Distribusi sampling beda dua proporsi merupakan distribusi dari perbedaan dua nilai proporsi yang berasal dari sampel dua populasi. Sebagai contoh, terdapat dua populasi N1 dan N2 (dua populasi binominal), lalu di ambil sampel random yakni n1 dan n2 dengan P1 dan P2 maka beda antara kedua sampel proporsi (p1-p2) akan membentuk suatu distribusi yang disebut sebagai distribusi sampling beda proporsi. Pada distribusi sampling beda proporsi berlaku ketentuan sebagai berikut.

- a. Rata-rata

$$\mu_{p1-p2} = P1 - P2$$

b. Standar Deviasi

$$\sigma_{p1-p2} = \sqrt{\frac{P1(1-P1)}{n1} + \frac{P2(1-P2)}{n2}}$$

c. Apabila n1 dan n2 (n1, n2 ≥ 30) cukup besar, maka distribusi sampling beda prporosi akan mendekati distribusi normal dengan variabel random standar yang rumus Z-nya sebagai berikut :

$$Z = \frac{(p1-p2) - (P1-P2)}{\sigma_{p1-p2}}$$

Keterangan :

$$p1-p2 = \frac{x1}{n1} - \frac{x2}{n2}$$

Contoh 6.5:

Sebanyak 35% dari balita gizi buriy yang mendapat PMT-pemulihan berhasil mencapai status gizi baik. Mereka tahun sebelumnya pernah mendapat PMT-pemulihan namun tidak berhasil mencapai status gizi baik. Sebanyak 30% balita gizi buruk yang baru (kasus baru) berhasil mencapai status gizi baik. Jika diambil sampel random sebanyak 200 anak balita gizi buruk, baik yang belum pernah diberi PMT-pemulihan maupun yang belum (kasus baru), berapa probabilitas bahwa beda proporsi balita akan berhasil mencapai status gizi baik antara yang sudah pernah maupun baru yang mendapat PMT-Pemulihan adalah kurang dari 2,5%?

P1 = 35% = 0,35 P2 = 30% = 0,30

n1 = 200 n2 = 200

p1 - p2 = 2,5% = 0,025

$$Z = \frac{(p1-p2) - (P1-P2)}{\sigma_{p1-p2}}$$

$$Z = \frac{0,025 - (0,35 - 0,30)}{\sqrt{\frac{(0,35)(0,65)}{200} + \frac{(0,30)(0,70)}{200}}} = - 0,5345$$

Nilai - Z yang diperoleh 0,5345 (minus diabaikan) lihat pada tabel Z memperoleh luas area sebesar 0,2019 maka P (Z < - 0,53) = 0,5 - 0,20 = 0,3.

LATIHAN 5.4

1. Ada petunjuk kuat bahwa Mr. A akan mendapat suara 60% dalam pemilihan Gubernur Propinsi X periode tahun 2001-2003. Secara independen diambil dua buah sampel acak yang masing-masing berukuran 300 orang. Tentukanlah probabilitas bahwa akan terdapat perbedaan persentase suara tidak lebih dari 10% dari kedua sampel tersebut akan memilih Mr. A.
2. Dari hasil penelitian mengenai preferensi konsumen terhadap merk rokok G menunjukkan bahwa disukai 15% konsumen rokok di perkotaan dan disukai 12% konsumen rokok di pedesaan. Jika test pasar dilakukan lagi dengan mengambil dua sampel acak secara independen masing-masing berukuran 100 orang, maka tentukanlah probabilitas bahwa konsumen rokok di perkotaan paling sedikit 10 orang lebih banyak daripada konsumen rokok di pedesaan yang menyukai rokok merk G tersebut.
3. Dua sejoli A dan B yang suka judi melakukan taruhan dengan cara mengundi sekeping uang ratusan bergambar burung garuda (G) di satu sisi dan bergambar Karapan sapi (S) di sisi lainnya. Si A dan si B masing-masing diberi kesempatan melempar sebanyak 25 kali. Salah seorang akan dinyatakan menang taruhan jika memperoleh paling sedikit 5 sisi G lebih banyak dibanding lawannya. Berapakah peluang si A akan memenangkan taruhan tersebut.

Bab 7

PENDUGAAN PARAMETER

Dalam statistika, *pendugaan (estimasi)* mengacu pada proses dimana seseorang membuat kesimpulan tentang populasi, berdasarkan informasi yang diperoleh dari sampel.

Teori inferensia statistik mencakup semua metode yang digunakan dalam penarikan kesimpulan atau generalisasi mengenai suatu populasi.

Distribusi sampling memungkinkan kita mengaitkan suatu taraf kepercayaan tertentu dengan setiap kesimpulan statistik yang kita buat, sebagai suatu ukuran seberapa jauh kita menaruh kepercayaan pada ketepatan statistik dalam menduga parameter populasinya.

Inferensia statistik dapat dikelompokkan ke dalam 2 bidang utama, yaitu: **pendugaan** dan **pengujian hipotesis**.

7.1 METODE PENDUGAAN KLASIK

Sebuah nilai $\hat{\theta}$ bagi suatu statistik $\hat{\Theta}$ disebut suatu **nilai dugaan** bagi parameter populasi Θ . Statistik yang digunakan untuk memperoleh sebuah nilai dugaan disebut sebagai **penduga** atau **fungsi keputusan**.

Definisi *Ruang Keputusan*:

Himpunan semua kemungkinan nilai dugaan yang dapat diambil oleh suatu penduga.

Definisi Penduga Tak Bias:

Statistik $\hat{\theta}$ dikatakan sebagai penduga tak bias bagi parameter θ bila $\mu^{\hat{\theta}} = E(\hat{\theta}) = \theta$.

Bila $\hat{\theta}_1$ dan $\hat{\theta}_2$ keduanya adalah penduga tak bias bagi parameter populasi θ yang sama, kita akan memilih penduga yang distribusi samplingnya mempunyai ragam terkecil. Bila $\sigma^2_{\hat{\theta}_1} < \sigma^2_{\hat{\theta}_2}$, kita katakan bahwa $\hat{\theta}_1$ merupakan penduga bagi θ yang **lebih efisien** dari pada $\hat{\theta}_2$.

Definisi Penduga Paling Efisien:

Di antara semua kemungkinan penduga tak bias bagi parameter θ , yang ragamnya terkecil adalah penduga paling efisien bagi θ .

Untuk sebaran normal, dapat diperlihatkan bahwa baik \bar{X} mau-pun \tilde{X} keduanya merupakan penduga tak bias bagi rata-rata populasi μ , tetapi ragam \bar{X} lebih kecil dari pada ragam \tilde{X} . Namun sebenarnya bahkan penduga tak bias yang paling efisien pun, kecil sekali kemungkinannya menduga parameter populasi secara tepat.

Salah satu jalan pemecahannya adalah dengan menggunakan pendugaan selang bagi parameter populasi θ , yaitu suatu selang yang berbentuk $\theta_1 < \theta < \theta_2$, dengan θ_1 dan θ_2 bergantung pada nilai statistik $\hat{\theta}$ untuk suatu contoh

tertentu dan juga pada sebaran penarikan contoh bagi $\hat{\theta}$. Bila ukuran sampel naik, maka $\sigma_{\bar{x}}^2 = \sigma^2/n$ akan menurun, sehingga nilai dugaan kita kemungkinan lebih mendekati nilai parameter μ , dan ini pada akhirnya mengakibatkan selang tersebut menjadi lebih pendek.

Bila kita dapat menentukan $\hat{\theta}_1$ dan $\hat{\theta}_2$ sehingga:

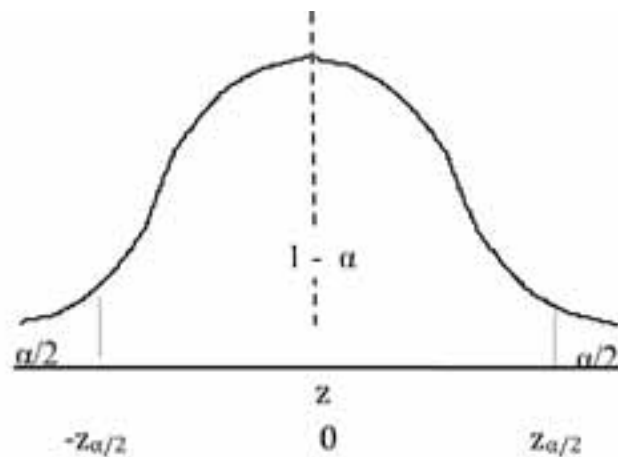
$$P(\hat{\theta}_1 < \theta < \hat{\theta}_2) = 1 - \alpha$$

Untuk $0 < \alpha < 1$, maka kita mempunyai peluang $1 - \alpha$ untuk memperoleh suatu sampel acak yang akan menghasilkan suatu selang yang mengandung θ .

- Selang $\hat{\theta}_1 < \theta < \hat{\theta}_2$, yang dihitung dari contoh terpilih, disebut **selang kepercayaan** $(1 - \alpha)$ 100%.
- Nilai $1 - \alpha$ disebut **koefisien kepercayaan** atau **derajat kepercayaan**.
- Kedua titik ujung, $\hat{\theta}_1$ dan $\hat{\theta}_2$, masing-masing disebut **batas kepercayaan** sebelah atas dan bawah.

1.2 PENDUGAAN RATA-RATA

Salah satu penduga titik bagi rata-rata populasi μ adalah statistik \bar{X} . Distribusi sampling \bar{X} berpusat di μ , dan dalam sebagian besar penerapannya, ragamnya lebih kecil dari pada ragam penduga-penduga lainnya. Gambar 1 di bawah ini menunjukkan bahwa:



Gambar 7.1. $P(-z_{\alpha/2} < Z < z_{\alpha/2}) = 1 - \alpha$

Dengan melambangkan $z_{\alpha/2}$ bagi nilai z yang luas daerah di sebelah kanan di bawah kurva normalnya adalah $\alpha/2$, kita dapat melihat bahwa:

$$P(-z_{\alpha/2} < Z < z_{\alpha/2}) = 1 - \alpha$$

sedangkan dalam hal ini $Z = \frac{\bar{X} - \mu}{\sigma / \sqrt{n}}$, sehingga dengan demikian:

$$P(-z_{\alpha/2} < \frac{\bar{X} - \mu}{\sigma / \sqrt{n}} < z_{\alpha/2}) = 1 - \alpha$$

Dengan beberapa manipulasi aljabar didapatkan

$$P(\bar{X} - z_{\alpha/2} \frac{\sigma}{\sqrt{n}} < \mu < \bar{X} + z_{\alpha/2} \frac{\sigma}{\sqrt{n}}) = 1 - \alpha$$

Selang Kepercayaan bagi μ ; σ diketahui.

Bila \bar{x} adalah rata-rata sampel acak berukuran n yang diambil dari suatu populasi dengan ragam σ^2 diketahui, maka selang kepercayaan $(1-\alpha) \times 100\%$ adalah:

$$\bar{x} - z_{\alpha/2} \frac{\sigma}{\sqrt{n}} < \mu < \bar{x} + z_{\alpha/2} \frac{\sigma}{\sqrt{n}}$$

sedangkan $z_{\alpha/2}$ adalah nilai z yang luas daerah di sebelah kanan di bawah kurva normal baku adalah $\alpha/2$.

Sampel yang berbeda akan menghasilkan nilai \bar{x} yang berbeda pula, sehingga selang kepercayaan bagi parameter μ yang diperoleh juga berbeda. Untuk asumsi σ tidak diketahui, maka kita akan mengganti σ dengan s , asalkan $n \geq 30$.

Contoh 7.1

Suatu sampel acak 36 mahasiswa tingkat akhir menghasilkan rata-rata dan simpangan baku nilai-mutu rata-rata sebesar berturut-turut 2,6 dan 0,3. Buatlah selang kepercayaan 95% dan 99% bagi rata-rata nilai-mutu rata-rata seluruh mahasiswa tingkat akhir.

Selang kepercayaan $(1 - \alpha) \times 100\%$ memberikan ukuran sejauh mana ketelitian atau akurasi nilai dugaan titiknya.

Kecil sekali kemungkinannya \bar{x} akan tepat sama dengan μ , sehingga nilai dugaan itu akan mesti mempunyai galat.

Teorema 7.1. Galat dalam Pendugaan.

Bila \bar{X} digunakan untuk menduga μ , kita percaya $(1 - \alpha) \times 100\%$ bahwa galatnya tidak akan melebihi $z_{\alpha/2} \frac{\sigma}{\sqrt{n}}$.

Teorema 7.2. Ukuran Sampel bagi Pendugaan μ .

Bila \bar{X} digunakan untuk menduga μ , kita boleh percaya $(1 - \alpha) \times 100\%$ bahwa galatnya tidak akan melebihi suatu nilai tertentu e bila ukuran contohnya diambil sebesar $n = \left(\frac{z_{\alpha/2} \sigma}{e} \right)^2$.

Sesungguhnya rumus dalam Dalil 2 hanya boleh digunakan bila kita mengetahui ragam populasi yang akan diambil sampelnya. Bila kita tidak mempunyai informasi ini, suatu sampel awal ber-ukuran $n \geq 30$ dapat diambil untuk memberikan dugaan bagi σ .

Contoh 7.2:

Seberapa besar contoh harus diambil dalam soal nomor 7.1, bila kita ingin percaya 95% bahwa nilai dugaan kita tidak menyimpang dari μ lebih dari 0,05?

Pada kondisi ragamnya tidak diketahui dan ukuran sampel kecil ($n < 30$):

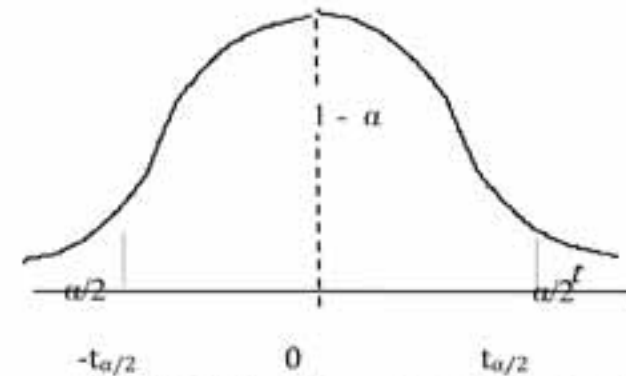
a. Jika populasi kira-kira berbentuk genta, selang kepercayaan masih bisa diperoleh meskipun σ^2 tidak diketahui dan n kecil.

b. Memanfaatkan distribusi sampling bagi T , yang dalam hal

$$\text{ini } T = \frac{\bar{X} - \mu}{S/\sqrt{n}}$$

c. Prosedur sama seperti pada sampel besar, kecuali di sini digunakan sebaran t sebagai pengganti sebaran normal baku.

Seperti halnya pada sebaran normal baku z , maka dengan meli-hat ke Gambar 7.2, maka kita dapat mengatakan bahwa:



Gambar 7.2. $P(-t_{\alpha/2} < T < t_{\alpha/2}) = 1 - \alpha$

$$P(-t_{\alpha/2} < T < t_{\alpha/2}) = 1 - \alpha$$

Dengan substitusi nilai T di atas, maka didapatkan:

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

dan melalui beberapa manipulasi aljabar, maka didapatkan:

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

Selang kepercayaan bagi μ untuk sampel berukuran kecil; σ Tidak Diketahui.

Bila \bar{X} dan s adalah rata-rata dan simpangan baku contoh berukuran $n < 30$, yang diambil dari populasi berbentuk genta yang ragamnya σ^2 tidak diketahui, maka selang kepercayaan $(1 - \alpha) \times 100\%$ bagi μ diberikan oleh rumusan:

$$\bar{X} - t_{\alpha/2} \frac{S}{\sqrt{n}} < \mu < \bar{X} + t_{\alpha/2} \frac{S}{\sqrt{n}}$$

Sedangkan dalam hal ini $t_{\alpha/2}$ adalah nilai t dengan $\nu = n - 1$ derajat bebas yang di sebelah kanannya terdapat daerah seluas $\alpha/2$.

Contoh 7.3.

Isi 7 kaleng asam sulfat adalah 9,8; 10,2; 10,4; 9,8; 10,0; 10,2; dan 9,6 liter. Tentukan selang kepercayaan 95% bagi rata-rata semua kaleng demikian ini bila isi kaleng itu menyebar normal.

1.3 PENDUGAAN BEDA DUA RATA-RATA POPULASI

Menurut teori sampling untuk beda dua rata-rata, kita dapat mengharapkan bahwa distribusi sampling bagi $\bar{X}_1 - \bar{X}_2$ kira-kira akan menyebar normal dengan rata-rata $\mu_{\bar{X}_1 - \bar{X}_2} = \mu_1 - \mu_2$ dan simpangan baku sebesar $\sigma_{\bar{X}_1 - \bar{X}_2} = \sqrt{(\sigma_1^2/n_1) + (\sigma_2^2/n_2)}$. Dengan demikian kita dapat menyatakan dengan peluang $1 - \alpha$ bahwa peubah acak normal baku:

$$Z = \frac{(\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2)}{\sqrt{(\sigma_1^2/n_1) + (\sigma_2^2/n_2)}}$$

akan jatuh antara $-z_{\alpha/2}$ dan $z_{\alpha/2}$. Melalui logika yang sama dengan pendugaan selang nilai rata-rata maka kita dapat merumuskan selang kepercayaan $(1 - \alpha) \times 100\%$ bagi $\mu_1 - \mu_2$ sebagai berikut ini.

Selang kepercayaan bagi $\mu_1 - \mu_2$; σ_1^2 dan σ_2^2 Diketahui. Bila \bar{x}_1 dan \bar{x}_2 masing-masing adalah rata-rata sampel acak bebas berukuran n_1 dan n_2 yang diambil dari populasi dengan ragam σ_1^2 dan σ_2^2 diketahui, maka selang kepercayaan $(1 - \alpha) \times 100\%$ bagi $\mu_1 - \mu_2$ adalah

$$(\bar{x}_1 - \bar{x}_2) - z_{\alpha/2} \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}} < \mu_1 - \mu_2 < (\bar{x}_1 - \bar{x}_2) + z_{\alpha/2} \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}$$

Sedangkan dalam hal ini $z_{\alpha/2}$ adalah nilai peubah normal z yang luas daerah di sebelah kanannya sebesar $\alpha/2$.

Contoh 7.4

Suatu ujian kimia diberikan pada 50 siswa perempuan dan 75 siswa laki-laki. Siswa perempuan mencapai rata-rata 76 dengan simpangan baku 6, sedangkan siswa laki-laki mendapat rata-rata 82 dengan simpangan baku 8. Tentukan selang kepercayaan 96% bagi beda $\mu_1 - \mu_2$, dalam hal ini μ_1 adalah nilai tengah skor semua siswa laki-laki dan μ_2 adalah rata-rata skor semua siswa perempuan yang mungkin mengambil ujian ini.

Jika dimisalkan σ_1^2 dan σ_2^2 tak diketahui serta n_1 dan n_2 kecil (<30). Apabila $\sigma_1^2 = \sigma_2^2 = \sigma^2$, kita mendapatkan peubah normal baku dalam bentuk:

$$Z = \frac{(\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2)}{\sqrt{\sigma^2 \left[\frac{1}{n_1} + \frac{1}{n_2} \right]}}$$

Sedangkan σ^2 harus diduga dengan cara menggabungkan kedua ragam sampel. Penduga gabungan S_p^2 diperoleh melalui rumus:

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

Dengan mensubstitusi S_p^2 bagi σ^2 dan memperhatikan bahwa pembagi dalam S_p^2 adlh $n_1 + n_2 - 2$ maka kita mendapatkan statistik:

$$T = \frac{(\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2)}{S_p \sqrt{\left(\frac{1}{n_1} + \frac{1}{n_2} \right)}}$$

yang menyebar menurut sebaran t dengan $\nu = n_1 + n_2 - 2$ derajat bebas. Selanjutnya, melalui beberapa manipulasi aljabar seperti sebelumnya, kita dapat memperoleh selang kepercayaan $(1 - \alpha) \times 100\%$ bagi $\mu_1 - \mu_2$ sebagai berikut:

Selang Kepercayaan bagi $\mu_1 - \mu_2$ untuk sampel berukuran kecil; $\sigma_1^2 = \sigma_2^2$ tetapi nilainya tidak diketahui.

Bila \bar{x}_1 dan \bar{x}_2 masing-masing adalah rata-rata sampel acak bebas berukuran kecil n_1 dan n_2 , yang diambil dari dua

populasi yang hampir normal dengan ragam sama tetapi tidak diketahui nilainya, maka selang kepercayaan $(1 - \alpha)100\%$ bagi $\mu_1 - \mu_2$ diberikan rumus:

$$(\bar{x}_1 - \bar{x}_2) - t_{\alpha/2} s_p \sqrt{\frac{1}{n_1} + \frac{1}{n_2}} < \mu_1 - \mu_2 < (\bar{x}_1 - \bar{x}_2) + t_{\alpha/2} s_p \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}$$

Dalam hal ini s_p adalah nilai dugaan gabungan bagi simpangan baku populasi, dan $t_{\alpha/2}$ adalah nilai t dengan $\nu = n_1 + n_2 - 2$ derajat bebas yang luas daerah di sebelah kanannya sebesar $\alpha/2$.

Contoh 7.5

Suatu pelajaran matematika diberikan pada 12 siswa dengan metode pengajaran yang biasa. Pelajaran yang sama diberikan pula pada 10 siswa tetapi dengan metode pengajaran yang menggunakan bahan yang telah diprogramkan. Pada akhir semester pada setiap kelas diberikan ujian yang sama. Kelas yang pertama mencapai nilai rata-rata 85 dengan simpangan baku 4, sedangkan kelas yang kedua mencapai nilai rata-rata 81 dengan simpangan baku 5. Tentukan selang kepercayaan 90% bagi selisih antara kedua rata-rata populasi, bila diasumsikan kedua populasi menyebar mendekati normal dengan ragam yang sama.

Untuk sampel yang kecil, bila kedua ragam populasi yang tidak diketahui tersebut kecil sekali kemungkinannya untuk sama, dan kita tidak mungkin mengambil sampel yang sama, maka statistik yang digunakan adalah:

$$T = \frac{(\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2)}{\sqrt{(S_1^2 / n_1) + (S_2^2 / n_2)}}$$

yang sebarannya menghampiri sebaran t dengan ν derajat bebas, sedangkan dalam hal ini:

$$\nu = \frac{\left(s_1^2/n_1 + s_2^2/n_2 \right)^2}{\left[\left(s_1^2/n_1 \right)^2 / (n_1 - 1) \right] + \left[\left(s_2^2/n_2 \right)^2 / (n_2 - 1) \right]}$$

Melalui beberapa simulasi aljabar sebagaimana sebelumnya, maka kita akan memperoleh hasil berikut ini:

Selang Kepercayaan bagi $\mu_1 - \mu_2$ untuk sampel berukuran kecil; $\sigma_1^2 \neq \sigma_2^2$ dan nilainya tidak diketahui.

Bila \bar{x}_1 dan s_1^2 , dan \bar{x}_2 dan s_2^2 masing-masing adalah rata-rata dan ragam sampel bebas berukuran kecil n_1 dan n_2 yang mendekati normal dengan ragam tidak sama dan tidak diketahui nilainya, maka selang ke-percayaan $(1 - \alpha) \times 100\%$ bagi $\mu_1 - \mu_2$ yang merupakan hampiran diberikan oleh rumus:

$$(\bar{x}_1 - \bar{x}_2) - t_{\alpha/2} \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}} < \mu_1 - \mu_2 < (\bar{x}_1 - \bar{x}_2) + t_{\alpha/2} \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}$$

sedangkan dalam hal ini $t_{\alpha/2}$ adalah nilai t dengan derajat bebas:

$$\nu = \frac{\left(s_1^2/n_1 + s_2^2/n_2 \right)^2}{\left[\left(s_1^2/n_1 \right)^2 / (n_1 - 1) \right] + \left[\left(s_2^2/n_2 \right)^2 / (n_2 - 1) \right]}$$

yang di sebelah kanannya terdapat daerah seluas $\alpha/2$.

Contoh 7.6.

Catatan selama 15 tahun terakhir menunjukkan bahwa curah hujan rata-rata di suatu daerah selama bulan Mei adalah 4,93 cm, dengan simpangan baku 1,14 cm. Di daerah lain, catatan serupa selama 10 tahun terakhir menunjukkan bahwa curah hujan rata-rata di bulan Mei adalah 2,64 cm dengan simpangan baku 0,66 cm. Tentukan selang kepercayaan 95% bagi selisih curah hujan rata-rata yang sebenarnya selama bulan Mei di kedua daerah tersebut, bila diasumsikan bahwa pengamatan-pengamatan itu berasal dari dua populasi normal dengan ragam yang berbeda.

Ilustrasi terakhir pada prosedur pendugaan bagi selisih dua rata-rata adalah bila sampelnya tidak bebas dan ragam kedua populasi tidak dapat dianggap sama. Hal ini terjadi, bila pengamatan dalam kedua sampel saling berpasang-pasangan, sehingga kedua pengamatan itu berhubungan.

Selang kepercayaan $(1 - \alpha) \times 100\%$ bagi μ_D dapat diperoleh dengan menyatakan:

$$P(-t_{\alpha/2} < T < t_{\alpha/2}) = 1 - \alpha$$

sedangkan dalam hal ini,

$$T = \frac{\bar{D} - \mu_D}{S_d / \sqrt{n}}$$

Dengan melakukan manipulasi aljabar yang diperlukan sampai menghasilkan selang kepercayaan $(1 - \alpha) \times 100\%$ bagi $\mu_1 - \mu_2 = \mu_D$ maka didapatkan:

Selang kepercayaan bagi $\mu_D = \mu_1 - \mu_2$ untuk Pengamatan Berpasangan.

Bila \bar{d} dan s_d adalah rata-rata dan simpangan baku selisih n pengamatan berpasangan, maka selang kepercayaan $(1 - \alpha) \times 100\%$ bagi $\mu_D = \mu_1 - \mu_2$ adalah:

$$\bar{d} - t_{\alpha/2} \frac{s_d}{\sqrt{n}} < \mu_D < \bar{d} + t_{\alpha/2} \frac{s_d}{\sqrt{n}},$$

Sedangkan dalam hal ini $t_{\alpha/2}$ adalah nilai t dengan $\nu = n - 1$ derajat bebas yang luas daerah di sebelah kanannya $\alpha/2$.

LATIHAN 7.1

SOAL-SOAL ESTIMASI RATA-RATA

1. Telah diambil secara acak sampel yang terdiri dari 100 orang mahasiswa sebuah universitas di Jakarta. Melalui test IQ terhadap 100 mahasiswa tersebut diperoleh rata-rata IQ sebesar 112 dan varians 100. Dengan menggunakan tingkat keyakinan (*confidence level*) sebesar 95%, tentukan interval keyakinan untuk nilai rata-rata IQ seluruh mahasiswa universitas tersebut.
2. Seorang petani tomat ingin mengetahui rata-rata berat buah tomat hasil kebunnya. Untuk itu, diambil sampel secara acak 10 buah tomat dengan berat masing-masing (dalam gram) : 142, 157, 138, 175, 152, 149, 148, 200, 182, dan 164. Jika petani tersebut merasa yakin 95% bahwa rata-rata berat buah tomat akan tercakup dalam interval estimasi, maka tentukanlah interval estimasi rata-rata berat tomat hasil kebun petani tersebut.

3. Sebuah populasi yang berdistribusi Normal terdiri dari 1000 data mempunyai simpangan baku 5,75. Kemudian diambil sampel acak berukuran 80 data dengan rata-rata 68,6. Tentukan inter-val keyakinan 95% untuk rata-rata populasi tersebut.
4. Untuk menjaga cash-flow yang aman, sebuah perusahaan pemasok bahan bangunan ingin mengestimasi rata-rata saldo kredit dari para krediturnya. Untuk itu, diambil sampel acak yang terdiri dari 25 kreditur, dan diperoleh rata-rata saldo kredit sebesar US\$ 3200 dengan standar deviasi US\$ 350. Tentukan interval estimasi rata-rata saldo kredit para piutang perusahaan tersebut, sehingga pimpinan perusahaan akan merasa yakin sebesar 90% terhadap kebenaran estimasi tersebut.
5. Untuk meningkatkan mutu pelayanan, manajemen Cafe MaFia melakukan penelitian terhadap para pelanggannya selama tiga bulan. Dari hasil observasi secara acak terhadap 100 pelanggan diperoleh data sebagai berikut :
 - a. Rata-rata waktu santap para pelanggan adalah 45 menit dengan standar deviasi 10 menit.
 - b. Rata-rata pembelian para pelanggan Rp 125.000,- dengan standar deviasi Rp 25.000,-

Jika manajemen Cafe tersebut merasa yakin sebesar 95% bahwa hasil penelitiannya akan tercakup dalam interval estimasi (interval konfidens), maka tentukan interval konfidens untuk :

 - a. Rata-rata waktu santap para pelanggannya
 - b. Rata-rata pembelian para pelanggannya

LATIHAN 7.2**SOAL-SOAL ESTIMASI PROPORSI**

1. Dari hasil survey yang dilakukan suatu *research agency* mengenai kebiasaan ibu rumah tangga menyaksikan tayangan iklan di TV Swasta. Ternyata diperoleh hasil bahwa 76 orang dari 180 orang ibu rumah tangga yang dipilih secara acak, biasa menyaksikan tayangan iklan paling sedikit 2 jam per minggu. Jika peneliti tersebut menggunakan taraf kefidens sebesar 90%, maka tentukan interval estimasi seluruh ibu rumah tangga yang biasa menyaksikan tayangan iklan paling sedikit 2 jam per minggu.
2. Sebuah sampel acak yang terdiri dari 100 buruh tani, ternyata sebanyak 64 buruh tani tersebut juga sebagai pemilik tanah. Tentukanlah interval keyakinan sebesar 95% untuk mengestimasi proporsi buruh tani yang juga sebagai pemilik tanah.
3. Dari hasil pemeriksaan mutu terhadap sebuah sampel acak ban mobil yang diproduksi PT. BB, ternyata sebanyak 20% tidak memenuhi standar mutu. Tentukanlah interval kefidens sebesar 90% untuk proporsi ban yang tidak memenuhi standar mutu, jika digunakan ukuran sampel :
 - a. $n = 10$
 - b. $n = 25$
 - c. $n = 100$
4. Seperempat dari 300 konsumen yang diwawancarai secara acak menyatakan tidak suka sabun mandi merk "X". Jika digunakan taraf kefidens 99%, tentukanlah

interval estimasi seluruh kon-sumen yang tidak menyukai sabun merk "X" tersebut.

LATIHAN 7.3**SOAL-SOAL ESTIMASI BEDA DUA RATA-RATA**

1. Sampel acak yang terdiri dari 100 orang buruh perusahaan A telah diperiksa ternyata rata-rata waktu menyelesaikan pekerjaannya per unit barang adalah 12 menit dengan standar deviasi 2 menit. Sedangkan dari perusahaan B yang sejenis diambil sampel acak berukuran 50, setelah diperiksa ternyata rata-rata menyelesaikan pekerjaan yang sama adalah 11 menit dengan standar deviasi 3 menit. Tentukanlah interval keyakinan sebesar 99% untuk mengestimasi beda rata-rata waktu penyelesaian pekerjaan semua buruh di perusahaan A dan perusahaan B.
2. Berdasarkan hasil survey sampel mengenai rata-rata pendapatan keluarga per tahun di dua desa yang berbeda, yaitu Desa A dan Desa B. Dari Desa A diambil secara acak sampel berukuran 100 dan diperoleh rata-rata pendapatan Rp 5,9 juta dengan varians Rp 0,81 juta. Sedangkan dari Desa B diambil secara acak sampel berukuran 120 dan diperoleh rata-rata pendapatan Rp 5,8 juta dengan varians Rp 0,64 juta. Jika digunakan taraf kefidens 95%, tentukan interval estimasi perbedaan rata-rata pendapatan keluarga per tahun di kedua desa tersebut.

LATIHAN 7.4**SOAL-SOAL ESTIMASI BEDA DUA PROPORSI**

1. Dua sampel acak masing-masing terdiri 700 mahasiswa dan 500 mahasiswi yang mengunjungi suatu bazar buku murah. Ternyata setelah kedua sampel tersebut diperiksa, terdapat 400 mahasiswa dan 325 mahasiswi yang merasa puas dengan adanya bazar tersebut. Tentukan interval konfidens sebesar 98% untuk mengestimasi perbedaan proporsi mahasiswa dan mahasiswi yang merasa puas terhadap bazar buku murah tersebut.
2. Untuk mengetahui perbedaan proporsi ketaatan pemilik mobil melunasi PKB di Kota A dan Kota B, diambil secara acak sampel di Kota A sebanyak 100 mobil dan ternyata 72 mobil telah melunasi PKB. Sedangkan di Kota B dari sampel acak sebanyak 100 mobil, ternyata 66 mobil yang sudah melunasi pajaknya. Tentukanlah interval konfidens sebesar 90% untuk mengestimasi beda proporsi pemilik mobil yang taat melunasi pajak di kedua kota tersebut.

LATIHAN 7.5**MENENTUKAN n DARI ESTIMASI PARAMETER**

1. Depkes dan Depdiknas bekerjasama untuk mengadakan penelitian mengenai persentase murid SD yang sakit gigi. Supaya dengan taraf konfidens 95% diperoleh perbedaan antara persentase sebenarnya dengan persentase dugaan tidak lebih dari 4%, maka harus berapa murid SD yang dijadikan sampel.

2. Mr. X akan dinyatakan menang dalam pemilihan gubernur, jika ia berhasil mengumpulkan suara paling sedikit 51%. Dari pemilihan sebelumnya ia mendapatkan suara 55%. Untuk menjajagi pencalonan Mr. X agar terpilih lagi menjadi gubernur yang ketujuh kalinya, maka diambil sampel acak berukuran n pemilih. Agar Mr. X merasa yakin 95% akan terpilih lagi menjadi gubernur untuk ketujuh kalinya, maka berapakah n tersebut.
3. Berapakah ukuran sampel yang diperlukan untuk mengestimasi rata-rata pengeluaran konsumsi per bulan penduduk Desa Z, jika dengan taraf konfidens 95% diinginkan kekeliruan estimasi tidak lebih dari Rp 25.000. Berdasarkan hasil survey tahun sebelumnya ditunjukkan bahwa rata-rata pengeluaran konsumsi per bulan penduduk Desa Z tersebut adalah Rp 700.000 dengan standar deviasi Rp 200.000.
4. Dua buah populasi berdistribusi Normal dengan simpangan baku sama besar yaitu 8,4. Selanjutnya dari masing-masing populasi diambil secara acak sampel yang berukuran sama yang digunakan untuk mengestimasi beda rata-rata populasi. Jika taraf konfidens 95% dan kekeliruan estimasi tidak lebih dari 2,5, maka tentukanlah masing-masing ukuran kedua sampel tersebut.

Bab 8

PENGUJIAN HIPOTESIS

Hipotesis statistika adalah asumsi tentang parameter populasi. Asumsi ini mungkin benar atau mungkin tidak benar. Uji hipotesis mengacu pada prosedur formal yang digunakan oleh ahli statistik untuk menerima atau menolak hipotesis statistik.

8.1 Pendahuluan

Hipotesis Penelitian = Hipotesis Konseptual adalah pernyataan yang merupakan jawaban sementara terhadap suatu masalah yang masih harus diuji kebenarannya secara empirik, apakah pernyataan tersebut dapat diterima atau harus ditolak. Hipotesis penelitian didasarkan pada teori tertentu dan berbentuk verbal/kalimat sehingga sukar untuk diuji secara langsung atas dasar fakta empirik. Sehingga untuk mengujinya hipotesis penelitian harus diterjemahkan terlebih dahulu ke dalam hipotesis statistik yang merupakan hipotesis operasional. **Hipotesis statistika** merupakan pernyataan khusus mengenai populasi atau sampel.

Contoh 8.1:

- Nilai rata-rata mata kuliah statistik kelas A adalah 65.
- Lebih dari 30% mahasiswa yang mengambil mata kuliah statistik mendapat nilai A.
- Rata-rata pendapatan keluarga di Bandung Rp 1.000.000,00 / bulan

- Motivasi belajar mahasiswa Politeknik Piksi Ganesha tinggi.

Setiap hipotesis bisa benar atau salah, sehingga perlu diadakan penelitian sebelum hipotesis itu diterima atau ditolak. Langkah atau prosedur untuk menentukan apakah menerima atau menolak hipotesis dinamakan **pengujian hipotesis**. Pengujian hipotesis adalah suatu cara untuk membuktikan atau menguatkan suatu dugaan atau anggapan tentang parameter populasi yang tidak diketahui, berdasarkan informasi dari sampel yang diambil dari populasi yang bersangkutan, sehingga dari pengujian ini dapat ditarik sebuah kesimpulan.

8.2 Hipotesis Nol dan Hipotesis Alternatif

Dalam bahasa statistik, apa yang diasumsikan (dihipotesiskan) dinyatakan sebagai **Hipotesis Nol (H_0)** atau **Hipotesis Alternatif (H_1)**. Hipotesis Nol (H_0) diuji berhadapan dengan hipotesis tandingannya/lawannya yakni Hipotesis Alternatif (H_1). Nilai Hipotesis Nol (H_0) ditulis dalam bentuk suatu tanda sama ($=, \geq, \leq$) dan H_1 sebagai lawannya.

- ~ Jika H_0 berisi tanda $=$, H_1 bisa bertanda $\neq, >$ atau $<$.
- ~ Jika H_0 berisi tanda \geq , H_1 akan bertanda $<$.
- ~ Jika H_0 berisi tanda \leq , H_1 akan bertanda $>$.

Berikut adalah contoh tentang pasangan tanda untuk H_0 dan H_1

1. H_0 : Nilai rata-rata mata kuliah statistik kelas A adalah 65.
 H_1 : Nilai rata-rata mata kuliah statistik kelas A tidak sama dengan 65.
 Atau ditulis dengan $\rightarrow H_0 : \mu = 65$
 $H_1 : \mu \neq 65$

2. H_0 : Lebih dari 30% mahasiswa yang mengambil mata kuliah statistik mendapat nilai A.
 H_1 : Kurang dari 30% mahasiswa yang mengambil mata kuliah statistik mendapat nilai A.
 Atau ditulis dengan $\rightarrow H_0 : \pi \geq 30\%$ atau
 $H_0 : \pi = 30\%$
 $H_1 : \pi < 30\%$
 $H_1 : \pi < 30\%$
3. H_0 : Rata-rata pendapatan keluarga di Bandung lebih kecil dari Rp 1.000.000,00 /bulan
 H_1 : Rata-rata pendapatan keluarga di Bandung lebih besar dari Rp 1.000.000,00 /bulan
 Atau ditulis dengan $\rightarrow H_0 : \mu \leq 1.000.000$ atau
 $H_0 : \mu = 1.000.000$
 $H_1 : \mu > 1.000.000$
 $H_1 : \mu > 1.000.000$
4. H_0 : Motivasi belajar mahasiswa Politeknik Piksi Ganesha tinggi
 H_1 : Motivasi belajar mahasiswa Politeknik Piksi Ganesha rendah.
5. dsb.

Teori pengujian hipotesis akan memutuskan apakah H_0 diterima atau ditolak. Keputusan menolak atau menerima H_0 didasarkan pada uji statistik yang diperoleh dari data sampel, setelah dibandingkan dengan nilai kritis dari distribusi statistik yang bersangkutan yang terdapat dalam tabel yang dibuat statistisi.

Catatan :

- a. Penerimaan suatu hipotesis terjadi karena TIDAK CUKUP BUKTI untuk MENOLAK hipotesis tersebut dan BUKAN karena HIPOTESIS ITU BENAR dan
2. Penolakan suatu hipotesis terjadi karena TIDAK CUKUP BUKTI untuk MENERIMA hipotesis tersebut dan BUKAN karena HIPOTESIS ITU SALAH.

8.3 Dua Macam Tipe Kesalahan

Dalam melakukan pengujian hipotesis, ada dua macam kesalahan yang dapat terjadi, dikenal dengan:

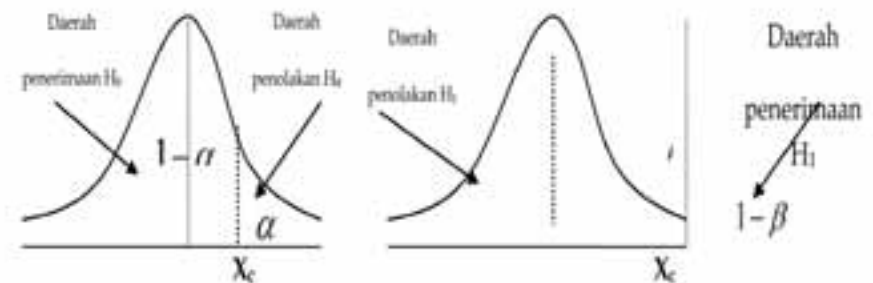
- a. **Kesalahan jenis I** adalah kesalahan akibat menolak Hipotesis Nol (H_0), padahal Hipotesis Nol benar, sehingga sesungguhnya harus diterima. Dengan kata lain kita menolak hipotesis yang seharusnya diterima.
- b. **Kesalahan jenis II** adalah kesalahan akibat menerima Hipotesis Nol (H_0), padahal Hipotesis Nol salah, sehingga sesungguhnya harus ditolak. Dengan kata lain kita menerima hipotesis yang seharusnya ditolak.

Keputusan	Keadaan yang sesungguhnya	
	H_0 benar	H_0 salah
Menerima H_0	Keputusan benar. (Probabilitas = $1 - \alpha$)	Kesalahan Jenis II (Probabilitas = β)
Menolak H_0	Kesalahan Jenis I. (Probabilitas = α)	Keputusan benar (Probabilitas = $1 - \beta$)

- a. Probabilitas melakukan kesalahan jenis I disimbolkan dengan α disebut sebagai **taraf nyata (tingkat signifikansi)**,

sedangkan probabilitas melakukan kesalahan jenis II disimbolkan dengan β .

- b. $1 - \alpha$ menyatakan **derajat kepercayaan**, sedangkan $1 - \beta$ disebut sebagai **kuasa uji**.
- c. Oleh karena α menyatakan probabilitas menolak H_0 padahal H_0 benar, maka nilai α harus sekecil mungkin. Begitu juga dengan β yang menyatakan probabilitas menerima H_0 padahal H_0 salah, maka nilai β juga harus sekecil mungkin.



- Dalam praktek pengujian hipotesis, nilai-nilai α yang biasa digunakan adalah misalnya seperti $\alpha = 0.05$ atau $\alpha = 0.01$, dsb.
- Bila dipakai taraf signifikansi 5% (atau $\alpha = 0.05$), artinya kira-kira sebanyak 5 dari setiap 100 kasus bahwa kita akan menolak hipotesis nol (H_0), padahal H_0 itu benar, sehingga seharusnya diterima. Atau dengan kata lain kira-kira 95% percaya bahwa kita telah membuat kesimpulan yang benar.

8.4 Pengujian Hipotesis

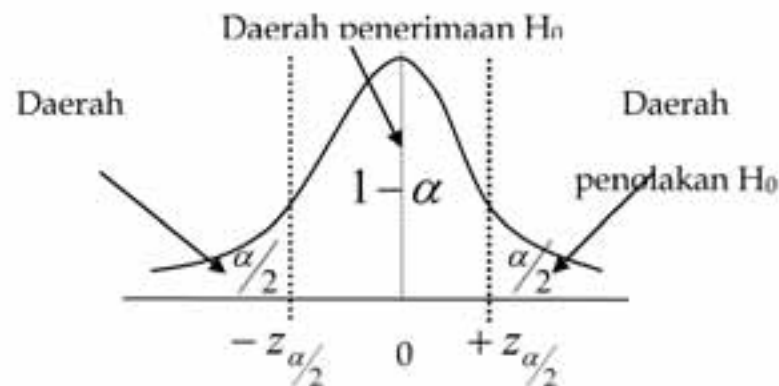
Pengujian hipotesis ada 2 bentuk, yakni :

- a. **Uji hipotesis dua arah**

- b. Uji hipotesis satu arah, yang dibedakan menjadi 2 yaitu *Uji hipotesis satu arah atas (kanan)* dan *Uji hipotesis satu arah bawah (kiri)*.

8.4.1 Uji hipotesis dua arah

- Jika H_0 berisi tanda sama dengan ($=$), H_1 berisi tanda tidak sama dengan (\neq), maka disebut uji dua arah.
 $H_0 : \theta = \theta_0$ melawan $H_1 : \theta \neq \theta_0$
- Uji dua arah ditandai dengan adanya dua daerah penolakan hipotesis nol (H_0) yang juga bergantung pada nilai kritis tertentu.
- Daerah penolakan H_0 ada dua, yaitu luas daerah di bagian paling kiri dan luas daerah di bagian paling kanan, masing-masing besarnya $\frac{\alpha}{2}$, dimana α telah ditentukan sebelumnya.
- Daerah penerimaan H_0 ditunjukkan oleh daerah $1 - \alpha$.
- Nilai kritis ada dua, yakni $-z_{\alpha/2}$ dan $+z_{\alpha/2}$, yang diperoleh dari tabel untuk nilai α yang telah ditentukan sebelumnya.



8.4.2 Uji Hipotesis Satu Arah

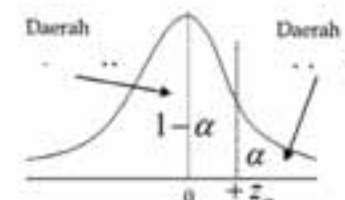
Uji satu arah ditandai dengan adanya satu daerah penolakan hipotesis nol (H_0) yang juga bergantung pada nilai kritis tertentu.

1. *Uji hipotesis satu arah atas (kanan)*

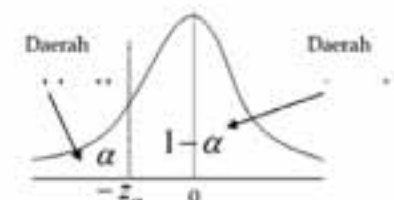
- Jika Hipotesis Nol (H_0) berisi tanda \leq dan Hipotesis alternatif (H_1) berisi tanda $>$, maka dinamakan uji searah atas (kanan).
 $H_0 : \theta \leq \theta_0$ melawan $H_1 : \theta > \theta_0$
- Daerah penolakan H_0 berada pada luas daerah di bagian paling kanan sebesar α .
- Daerah penerimaan H_0 ditunjukkan oleh daerah $1 - \alpha$.
- Nilai kritisnya adalah $+z_\alpha$ yang diperoleh dari tabel untuk nilai α yang telah ditentukan sebelumnya.

2. *Uji hipotesis satu arah bawah (kiri)*

- Jika Hipotesis Nol (H_0) berisi tanda \geq dan hipotesis alternatif (H_1) berisi tanda $<$, maka dinamakan uji searah bawah (kiri).
 $H_0 : \theta \geq \theta_0$ melawan $H_1 : \theta < \theta_0$
- Daerah penolakan H_0 berada pada luas daerah di bagian paling kiri sebesar α .
- Daerah penerimaan H_0 ditunjukkan oleh daerah $1 - \alpha$.
- Nilai kritisnya adalah $-z_\alpha$ yang diperoleh dari tabel untuk nilai α yang telah ditentukan sebelumnya.



1. Uji searah atas

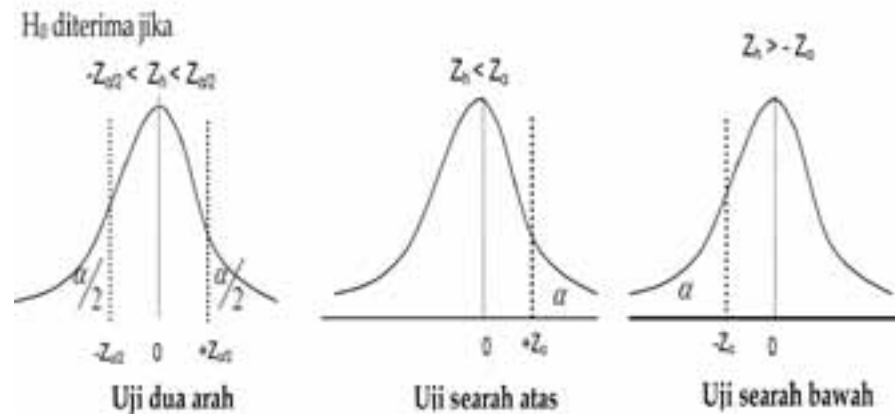


2. Uji searah bawah

8.5 Langkah Pengujian Hipotesis

Langkah-langkah Pengujian Hipotesis Secara Umum adalah :

1. Menetapkan rumusan hipotesis dengan tepat, baik Hipotesis Nol (H_0) dan Hipotesis Alternatif (H_1), apakah termasuk uji satu arah atau dua arah.
2. Tetapkanlah taraf nyata α yang diinginkan, sehingga dengan memakai nilai α tersebut dapat diperoleh nilai kritis dari tabel. Dengan demikian dapat digambarkan daerah penolakan H_0 dan daerah penerimaan H_0 .
3. Tetapkanlah statistik uji (Z_h) yang cocok untuk menguji Hipotesis Nol (H_0).
4. Hitunglah nilai statistik uji (Z_h) berdasarkan data dan informasi yang diketahui baik dari populasi maupun dari sampel yang diambil dari populasi tersebut.
5. Simpulkan; tolak H_0 , bila nilai statistik uji Z_h jatuh atau terletak di daerah penolakan H_0 , dan terima H_0 bila nilai statistik uji jatuh atau terletak di daerah penerimaan H_0 .



8.6 Statistik Uji Untuk Pengujian Parameter Populasi

Rumus-rumus uji perhitungan statistik

No	Uji Hipotesis	H_0	Nilai Statistik Uji	H_1	Wilayah Kritis
1.	Uji Rata-rata	$\mu = \mu_0$	σ diketahui $z_h = \frac{\bar{x} - \mu_0}{\sigma / \sqrt{n}}$	$\mu \neq \mu_0$ $\mu < \mu_0$ $\mu > \mu_0$	$z_h < -z_{\alpha/2}$ dan $z_h > z_{\alpha/2}$ $z_h < -z_{\alpha}$ $z_h > z_{\alpha}$
			σ tidak diketahui (ditaksir oleh s) $t_h = \frac{\bar{x} - \mu_0}{s / \sqrt{n}}$ <p>dengan</p> $s = \sqrt{\frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n-1}}$	$\mu \neq \mu_0$ $\mu < \mu_0$ $\mu > \mu_0$	$t_h < -t_{(1-\alpha/2; dk)}$ dan $t_h > t_{(1-\alpha/2; dk)}$ $t_h < -t_{(1-\alpha; dk)}$ $t_h > t_{(1-\alpha; dk)}$ $dk = n - 1$
No	Uji Hipotesis	H_0	Nilai Statistik Uji	H_1	Wilayah Kritis
			$\sigma_1 = \sigma_2 = \sigma$		

2.	Uji Beda Dua Rata-Rata	$ \mu_1 - \mu_2 = d_0$	<p>diketahui</p> $z_b = \frac{ \bar{x}_1 - \bar{x}_2 - d_0}{\sigma \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}$	$ \mu_1 - \mu_2 \neq d_0$ $z_h < -z_{\alpha/2}$ dan $z_h > z_{\alpha/2}$ $ \mu_1 - \mu_2 < d_0$ $z_h < -z_{\alpha}$ dan $z_h > z_{\alpha}$ $ \mu_1 - \mu_2 > d_0$
			<p>$\sigma_1 \neq \sigma_2$ tapi diketahui</p> $z_b = \frac{ \bar{x}_1 - \bar{x}_2 - d_0}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}$	$ \mu_1 - \mu_2 \neq d_0$ $z_h < -z_{\alpha/2}$ dan $z_h > z_{\alpha/2}$ $ \mu_1 - \mu_2 < d_0$ $z_h < -z_{\alpha}$ dan $z_h > z_{\alpha}$ $ \mu_1 - \mu_2 > d_0$
			<p>$\sigma_1 = \sigma_2 = \sigma$ tapi tidak diketahui</p> $t_h = \frac{ \bar{x}_1 - \bar{x}_2 - d_0}{s_{gab} \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}$ <p>dengan</p> $s_{gab} = \sqrt{\frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2}}$	$ \mu_1 - \mu_2 \neq d_0$ $t_h < -t_{(1-\alpha/2; dk)}$ dan $t_h > t_{(1-\alpha/2; dk)}$ $ \mu_1 - \mu_2 < d_0$ $t_h < -t_{(1-\alpha; dk)}$ $ \mu_1 - \mu_2 > d_0$

			$ \mu_2 > d_0$ $t_h > t_{(1-\alpha; dk)}$ $dk = n_1 + n_2 - 2$
			<p>$\sigma_1 \neq \sigma_2$ dan tidak diketahui</p> $t_h = \frac{(\bar{x}_1 - \bar{x}_2) - d_0}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}$
3.	Uji Kesamaan Dua Rata-rata	$\mu_1 = \mu_2$ Keterangan: Nilai-nilai statistik uji untuk uji kesamaan dua rata-rata dapat dihitung dengan menggunakan rumus yang sama seperti pada uji beda dua rata-rata namun d_0 bernilai nol ($d_0 = 0$). Begitupun dalam menentukan wilayah kritisnya. $\mu_1 - \mu_2 = 0$	

4.	Uji Proporsi	$\pi = \pi_0$	$z_h = \frac{\hat{\pi} - \pi_0}{\sqrt{\frac{\pi_0(1-\pi_0)}{n}}}$ dengan $\rightarrow \hat{\pi} = \frac{x}{n}$	$\pi \neq \pi_0$ $\pi < \pi_0$ $\pi > \pi_0$	$z_h < -z_{\alpha/2}$ dan $z_h > z_{\alpha/2}$ $z_h < -z_{\alpha}$ $z_h > z_{\alpha}$
5.	Uji Beda Dua Proporsi	$ \pi_1 - \pi_2 = d_0$	$z_h = \frac{ \hat{\pi}_1 - \hat{\pi}_2 - d_0}{\sqrt{\frac{\hat{\pi}_1(1-\hat{\pi}_1)}{n_1} + \frac{\hat{\pi}_2(1-\hat{\pi}_2)}{n_2}}}$	$ \mu_1 - \mu_2 \neq d_0$ $ \mu_1 - \mu_2 < d_0$ $ \mu_1 - \mu_2 > d_0$	$z_h < -z_{\alpha/2}$ dan $z_h > z_{\alpha/2}$ $z_h < -z_{\alpha}$ $z_h > z_{\alpha}$
6.	Uji Kesamaan Dua Proporsi	$\pi_1 = \pi_2$ Atau dapat	$z_h = \frac{ \hat{\pi}_1 - \hat{\pi}_2 }{\sqrt{\frac{\hat{\pi}_1(1-\hat{\pi}_1)}{n_1} + \frac{\hat{\pi}_2(1-\hat{\pi}_2)}{n_2}}}$	$ \mu_1 - \mu_2 \neq d_0$	$z_h < -z_{\alpha/2}$ dan z_h

		ditulis		$ \mu_1 - \mu_2 < d_0$	$> z_{\alpha/2}$
		$\pi_1 - \pi_2 = 0$		$ \mu_1 - \mu_2 > d_0$	$z_h < -z_{\alpha}$ $z_h > z_{\alpha}$

Contoh 8.1

Pengusaha lampu pijar A mengatakan bahwa lampunya bisa tahan pakai sekitar 800 jam. Akhir-akhir ini timbul dugaan bahwa masa pakai lampu itu telah berubah. Untuk menentukan hal ini, dilakukan penelitian dengan jalan menguji 50 lampu. Ternyata rata-ratanya 792 jam. Dari pengalaman, diketahui bahwa simpangan baku masa hidup lampu adalah 60 jam. Selidiki dengan taraf nyata 0.05 apakah kualitas lampu itu sudah berubah atau belum ?

Jawab :

Diketahui $n = 50$, $\bar{x} = 792$ jam, $\mu_0 = 800$ jam, $\sigma = 60$ jam, dan $\alpha = 0.05$

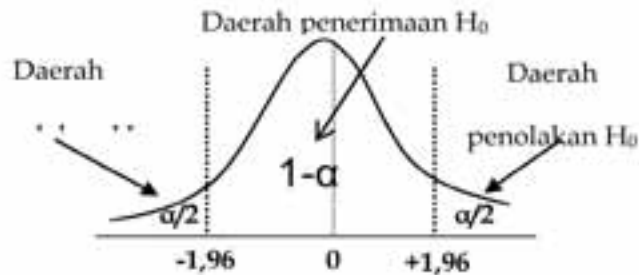
a. Dengan memisalkan masa hidup lampu berdistribusi normal, kita akan menguji
 $H_0: \mu = 800$ {Masa pakai lampu 800 jam yang berarti kualitas lampu tidak berubah}

$H_1: \mu \neq 800$ {Masa pakai lampu bukan 800 jam yang berarti kualitas lampu sudah berubah}

Termasuk uji dua arah.

b. Taraf nyata yang diinginkan $\alpha = 0.05$, sehingga $\alpha/2 = 0.025$

Nilai $z_{\alpha/2} = 1.96$ (dari tabel normal z)



c. Statistik Uji : Karena σ diketahui, maka digunakan

$$z_h = \frac{\bar{x} - \mu_0}{\sigma / \sqrt{n}} = \frac{792 - 800}{60 / \sqrt{50}} = -0.94$$

d. Kriteria Uji : Terima H_0 jika Z_h berada diantara daerah $-z_{\alpha/2} < z_h < z_{\alpha/2}$ atau $-1.96 < z_h < 1.96$.

Tolak H_0 jika Z_h berada di daerah lainnya.

e. Dari penelitian diperoleh bahwa $Z_h = -0.94$, dan ini terletak di daerah penerimaan H_0 . Dengan demikian H_0 diterima.

Kesimpulan : 95% kita percaya bahwa memang masa pakai lampu masih di sekitar 800 jam, jadi kualitas lampu belum/tidak berubah.

Contoh 8.2

Untuk contoh 1, tentang masa pakai lampu, misalkan simpangan baku populasi (σ) tidak diketahui, dan dari sampel didapatkan $s = 55$ jam. Jika diketahui $n = 50$, $\bar{x} = 792$ jam, $\mu_0 = 800$ jam, dan $\alpha = 0.05$ apakah kualitas lampu sudah berubah ?

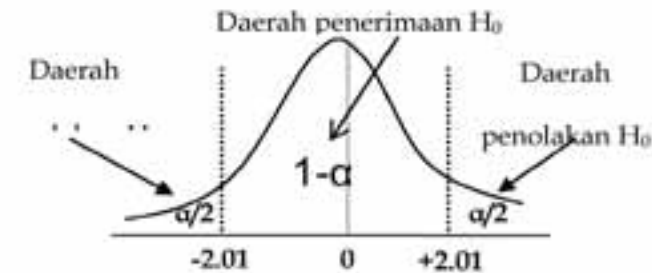
Jawab :

- a. $H_0: \mu = 800$ {Masa pakai lampu 800 jam yang berarti kualitas lampu tidak berubah}
- $H_1: \mu \neq 800$ {Masa pakai lampu bukan 800 jam yang berarti kualitas lampu sudah berubah}

Termasuk uji dua arah.

- b. Taraf nyata yang diinginkan $\alpha = 0.05$, sehingga $\alpha/2 = 0.025$

Nilai $t_{(1-\alpha/2;dk)} = 2.01 \rightarrow$ dari tabel distribusi t dengan $\alpha = 0.05$ dan $dk = n - 1 = 50 - 1 = 49$



c. Statistik Uji : Karena σ tidak diketahui, maka digunakan

$$t_h = \frac{\bar{x} - \mu_0}{s / \sqrt{n}} = \frac{792 - 800}{55 / \sqrt{50}} = -1.029$$

d. Kriteria Uji : Terima H_0 jika t_h berada di daerah $-t_{(1-\alpha/2;dk)} < t_h < t_{(1-\alpha/2;dk)}$ atau terletak diantara $-2.01 < t_h < 2.01$. Tolak H_0 jika t_h berada di daerah lainnya.

e. Dari penelitian diperoleh bahwa $t_h = -1.029$, dan ini terletak di daerah penerimaan H_0 . Dengan demikian H_0 diterima.

Kesimpulan : 95% kita percaya bahwa memang masa pakai lampu masih di sekitar 800 jam, jadi kualitas lampu belum/tidak berubah.

Contoh 8.3

Pimpinan perusahaan mengatakan bahwa paling banyak 60% dari pegawainya termasuk golongan A. Untuk meneliti kebenarannya, sampel acak yang terdiri dari 1500 orang telah diambil dan ternyata 1025 orang termasuk golongan A. Dengan menggunakan $\alpha = 0.01$ selidiki apakah pernyataan tersebut dapat diterima atau tidak ?

Jawab :

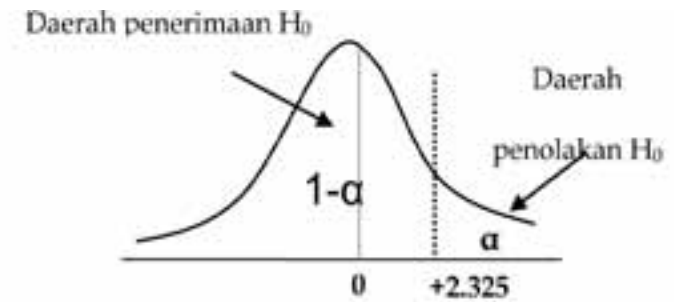
Diketahui $\pi_0 = 0.6$ $n = 1500$ $x = 1025$

$$\hat{\pi} = \frac{x}{n} = \frac{1025}{1500} = 0.68$$

- a. $H_0: \pi \leq 0.6$ {Paling banyak 60% pegawai termasuk golongan A yang berarti pernyataan dapat diterima}
- $H_1: \pi > 0.6$ {Lebih dari 60% pegawai termasuk golongan A yang berarti pernyataan tidak dapat diterima }

Termasuk uji satu arah atas (kanan).

- b. Taraf nyata yang diinginkan $\alpha = 0.01 \rightarrow$ Nilai $z_\alpha = 2.325$ (dari tabel normal z)



c. Statistik Uji :

$$z_h = \frac{\hat{\pi} - \pi_0}{\sqrt{\frac{\pi_0(1 - \pi_0)}{n}}} = \frac{0.68 - 0.6}{\sqrt{\frac{0.6 \times 0.4}{1500}}} = 6.32$$

- d. Kriteria Uji : Terima H_0 jika Z_h berada di daerah $z_h < z_\alpha$ atau $z_h < 2.325$. Tolak H_0 jika Z_h berada di daerah lainnya.
- e. Dari penelitian diperoleh bahwa $Z_h = 6.32$, dan ini terletak di daerah penolakan H_0 . Dengan demikian H_0 ditolak.
Kesimpulan : 95% kita percaya bahwa persentase pegawai yang termasuk golongan A lebih dari 60%, maka pernyataan pimpinan perusahaan tidak dapat diterima.

DAFTAR PUSTAKA

- ANDERSON, T. W., An Introduction to Multivariate Statistical Analysis New York: J. Wiley & Sons, 1958.
- APOSTOL, T., Mathematical Analysis, 2nd ed. Reading, MA: Addison-Wesley, 1974.
- BARNDORFF-NIELSEN, O. E., AND D. P. Cox, Asymptotic Techniques for Use in Statistics New York: Chapman and Hall, 1989.
- BILLINGSLEY, P., Probability and Measure, 3rd ed. New York: J. Wiley & Sons, 1979, 1995.
- BIRKHOFF, G., AND S. MacLANE, A Survey of Modern Algebra, rev. ed. New York: Macmillan, 1953.
- BIRKHOFF, G., AND S. MacLANE, A Survey of Modern Algebra, 3rd ed. New York MacMillan, 1965.
- BREIMAN, L., Probability Reading, MA: Addison-Wesley, 1968.
- CHUNG, K. L., A Course in Probability Theory New York: Academic Press, 1974.
- DEMPSTER, A. P., Elements of Continuous Multivariate Analysis Reading, MA: Addison-Wesley, 1969.
- DIEUDONNE, J., Foundation of Modern Analysis, v. 1 , Pure and Applied Math. Series, Volume 10 New York: Academic Press, 1960.
- DUNFORD, N., AND J. T. SCHWARTZ, Linear Operators, Volume 1 , Interscience New York: J. Wiley & Sons, 1964.
- FELLER, W., An Introduction to Probability Theory and Its Applications, Vol. II, 2nd ed. New York: J. Wiley & Sons, 1971.
- GRIMMETT, G. R., AND D. R. STIRSAKER, Probability and Random Processes Oxford: Clarendon Press, 1992.

- HALMOS, P. R., *An Introduction to Hilben Space and the Theory of Spectral Multiplicity*, 2nd ed. New York: Chelsea, 1951.
- HAMMERSLEY, J., "An extension of the Slutsky-Fr6chet theorem," *Acta Mathematica*, 87, 243-247 (1952).
- HoEFFDJNG, W., "Probability inequalities for sums of bounded random variables," *J. Amer. Statist.A"oc.*, 58, 13-30 (1 963).
- LoEvE, M., *Probability Theory*, Vol. I, 4th ed. Berlin: Springer, 1977.
- RAo, C. R., *Linear Statistical Inference and Its Applications*, 2nd ed. New York: J. Wiley & Sons, 1973.
- ROCKAFELLAR, R. T., *Convex Analysis* Princeton, NJ: Princeton UniverSity Press, 1970.
- ROYDEN, H. L., *Real Analysis*, 2nd ed. New York: MacMillan, 1968.
- RUDIN, W., *Functional Analysis*, 2nd ed. New York: McGraw-Hill, 1991.
- SKOROKHOD, A. V., "Limit theorems for stochastic processes," *Th. Prob. Applic.*, 1, 261-290 (1956)