BUKU AJAR

STATISTIKA

OLEH

Dr. MAKKULAU, S.Si., M.Si.

Dr. ANDI TENRI AMPA, S.Si., M.Si.

PROGRAM STUDI STATISTIKA

FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM

UNIVERSITAS HALU OLEO

KENDARI

2023

DAFTAR ISI

BAB I
PENGETAHUAN DASAR STATISTIKA

1.1 Arti dan Peran Statistika

Selama ini masih banyak orang yang menganggap bahwa kata *statistik* dan *statistika* adalah dua kata yang arti dan maknanya sama, padahal kedua kata tersebut mempunyai arti dan makna yang berbeda.

Ada beberapa pengertian yang biasa dikenal, antara lain:

- Statistik adalah kumpulan informasi baik berupa angka-angka maupun non angka mengenai suatu masalah. Dengan kata lain statistik adalah **data.**
- Statistika adalah ilmu atau metode ilmiah yang mempelajari pengumpulan, pengaturan, perhitungan, penggambaran dan penganalisaan data serta penarikan kesimpulan yang valid berdasarkan penganalisaan yang dilakukan dan pembuatan keputusan yang rasional.
- Statistika adalah ilmu yang mempelajari pengumpulan data, pengolahan data, analisa data, penarikan kesimpulan sampai pengambilan keputusan/kebijakan.
- Statistika adalah ilmu yang berhubungan dengan pengumpulan, analisis dan penafsiran data.

Dengan kata lain statistika merupakan sekumpulan konsep dan metode yang digunakan untuk mengumpulkan dan menginterpretasikan data tentang bidang kegiatan tertentu dan mengambil keputusan dalam situasi dimana ada *ketidakpastian* dan *variansi*.

Disadari atau tidak statistika sangat berperan dalam berbagai aspek kehidupan yang meliputi pengumpulan data, penyajian data, analisa data serta penafsiran data, misalnya pernyataan/masalah berikut (Makkulau & Ampa, 2012):

⊕ Ahmad ingin memperkirakan berapa rata-rata penghasilan sebuah RT di Kota Bau-Bau. (Statistika Biasa)

⊕ Siti ingin mengetahui berapa rata-rata kelahiran prematur di sebuah rumah sakit di Kota Kendari. (Statistik Biasa)

- ⊕ Ada 75% penduduk Kota Kendari memerlukan perumahan sehat. (Statistika Deskriptif)
- ⊕ Hasil jambu mente musim panen mendatang diperkirakan 25 kuintal tiap hektar. (Statistika Inferensi)
- ⊕ Faktor-faktor yang mempengaruhi pendapatan seorang PNS ditinjau dari sudut ekonomi. (Statistika Ekonomi)
- ⊕ Hubungan pendapatan dengan pengeluaran seorang petani di masa Pandemi Covid 19. (Statistika Ekonomi)
- ⊕ Seorang manajer perusahaan ingin mengetahui apakah ada keterkaitan dan seberapa besar keterkaitan antara jumlah iklan yang ditayangkan dengan peningkatan omset perusahaan. (Statistika Ekonomi atau Ekonometrika)
- ⊕ Pemda Kota Kendari ingin memprediksi pertumbuhan ekonomi pada periode 5 tahun yang akan datang. (Statistika Ekonomi, Analisis Regresi atau Time Series)
- ⊕ Badan Meteorologi dan Geofisika meramalkan akan terjadinya gempa bumi dan Tzunami di suatu tempat. (Time Series)
- ⊕ Seorang pemuda ingin mencari gadis pujaannya/*Biro Jodoh*. (Analisis Profil)
- ⊕ Seorang ahli Arkeologi ingin mengetahui umur suatu fosil manusia purba. (Analisis Multivariat)

Statistika berdasarkan fungsinya (cara pengolahan datanya) dibagi atas:

Θ <u>Statistika Deskriptif (Statistika Deduktif)</u>

Adalah cabang Statistika yang mempelajari cara-cara menggambarkan aspek-aspek yang sangat penting dari data, misalnya dengan menyusun tabel, diagram, grafik dan besaran-besaran lain. Statistika Deskriptif hanya berfungsi untuk menerangkan keadaan, gejala atau persoalan atau data digambarkan ke dalam bentuk grafik/diagram dan belum ada penarikan kesimpulan. Dengan kata lain Statistika Deskriptif adalah Analisis Data Eksplorasi, yaitu metode-metode numerik dan grafik.

Misalnya: Sekurang-kurangnya 25% anak usia sekolah di Sulawesi Tenggara masih kurang gizi.

Statistika Deskriptif mencakup ruang bahasan sebagai berikut:

a. Distribusi Frekuensi beserta bagian-bagiannya yaitu:
   - Histogram, poligon frekuensi dan ogif.
   - Ukuran Pemusatan, yaitu rata-rata, median, modus, kuartil, desil dan presentil.
   - Ukuran Penyebaran, yaitu range, deviasi rata-rata, variansi, standar deviasi dan lain-lain.
   - Ukuran Keruncingan dan Kemencengan
b. Angka Indeks.
c. Time Series (Runtum Waktu)

Di awal-awal lahirnya statistika, Statistika Deskriptif banyak digunakan. Lalu berkembang lagi metode yang lebih modern, yaitu Statistika Inferensi. Namun pada awal-awal abad XX, Prof. J.W. Tukey mempelopori kebangkitan kembali Statistika Deskriptif dengan mengembangkan banyak metode numerik/grafik yang lebih inovatif.

Θ Statistika Inferensi (Statistika Induktif)

Adalah cabang Statistika yang digunakan untuk analisis sebagian data untuk kemudian sampai pada peramalan dan penarikan kesimpulan tentang keadaan populasinya. Jadi Statistika Inferensi adalah penarikan kesimpulan berdasarkan sampel dari populasi. Dalam Statistika Inferensi ini juga biasanya memasukkan unsur peluang dalam menarik kesimpulannya.

Misalnya:
* Peramalan laju pertumbuhan ekonomi 5 tahun yang akan datang.
* Penarikan kesimpulan tentang obat baru ditemukan yang disimpulkan lebih efektif dari obat lama.

Statistika Inferensi mencakup ruang bahasan sebagai berikut:

⊕ Teori Peluang/Probabilitas	⊕ Sampling & Distribusi sampling
⊕ Model Distribusi Teoritis	⊕ Analisis Regresi & Analisis Korelasi
⊕ Estimasi Parameter & Uji Hipotesis	⊕ Uji Perbedaan Dua Mean
⊕ Analisis Variansi & Covariansi	⊕ Analisis Multivariat, dan sebagainya.

Menurut Webster, kata kerja *to infer* berarti ‘membuat derivasi (hasil runutan) sebagai suatu konsekuensi, kesimpulan, atau kemungkinan (probabilitas)’. Jika dilihat seorang perempuan yang jari manis sebelah kanannya dilingkari cincin, maka dibuat inferensi bahwa ia sudah menikah.

Dalam Statistika Inferensi diperhatikan dua macam permasalahan, yaitu perkiraan/penafsiran parameter populasi dan uji hipotesis. Persoalan dalam Statistika Inferensi adalah bagaimana menarik kesimpulan tentang sejumlah peristiwa (*events*) berdasarkan pengamatan terhadap sebagian saja dari peristiwa itu. Pada Statistika Inferensi ini biasanya memasukkan unsur peluang dalam menarik kesimpulan.

Secara umum untuk melakukan penelitian suatu masalah digunakan Statistika Deskriptif lebih dahulu kemudian Statistika Inferensi sebagai senjata terakhir untuk mengukur dan menguatkan jawaban yang ditemukan dalam Statistika Deskriptif awal.

Statistika berdasarkan bentuk parameternya dibedakan atas dua, yaitu:

Θ Statistika Parametrik

Adalah cabang Statistika yang dalam penerapannya ‘misalnya dalam Statistika Inferensi‘ didasarkan pada asumsi-asumsi bahwa parameternya mengikuti suatu bentuk distribusi tertentu. Misalnya dalam Analisis Variansi, diasumsikan bahwa sampel telah diambil dari suatu populasi berdistribusi normal.

Θ Statistika Nonparametrik

Adalah cabang Statistika yang dalam penerapannya tidak ada asumsi-asumsi tertentu tentang bentuk distribusi dari parameter tetapi hasil analisis memiliki tingkat kesahihan yang tinggi. Misalnya Uji T, Uji Kolmogorov-Smirnov, Uji Rangkaian Wald-Wolfowitz, dan sebagainya.

1.2 Macam-macam Data

Dalam menyelidiki suatu masalah selalu diperlukan data. Data dapat diartikan sebagai keterangan yang diperlukan untuk memecahkan masalah.

Berikut akan diberikan macam-macam data ditinjau dari beberapa segi:

1. Menurut Sifatnya
   a. Data Kualitatif; yaitu data yang berbentuk kategori atau atribut (datanya bukan angka).
   b. Data Kuantitatif; yaitu data yang berbentuk angka.
      Data kuantitatif dapat dibagi atas:
      1. Data Diskrit; adalah data yang diperoleh dengan cara menghitung atau membilang.
      2. Data Kontinu; adalah data yang diperoleh dengan cara mengukur.
2. Menurut Cara Memperolehnya
   a. Data Primer; adalah data yang dikumpulkan dan diolah sendiri oleh suatu organisasi serta diperoleh langsung dari obyeknya.
   b. Data Sekunder; adalah data yang diperoleh dalam bentuk yang sudah jadi, sudah dikumpulkan dan diolah oleh pihak lain.

1.3 Proses Pengumpulan Data

Ada beberapa istilah yang sering digunakan dalam hal penelitian, antara lain:

- Populasi; adalah keseluruhan atau kumpulan dari obyek yang akan diamati atau diteliti.
- Sampel; adalah bagian dari populasi.
- Parameter; ciri atau karakteristik dari populasi, misalnya:
  μ adalah mean dari populasi
  σ^2 adalah variansi dari populasi
  σ adalah simpangan baku atau deviasi standar dari populasi, dan lain-lain.
- Survey, adalah pengumpulan informasi tentang sekelompok manusia, dimana suatu hubungan langsung dengan obyek yang diperoleh seperti individu, organisasi. Masyarakat, dan sebagainya, diadakan melalui suatu cara yang sistematis seperti pengisisan daftar pernyataan, wawancara, dan lain-lain sebagainya.

- Responden, adalah sampel yang dipilih (beberapa orang: banyak pada penelitian sosial) untuk menjawab pertanyaan peneliti.

Dalam Statistika proses pengumpulan data dibagi atas:

1. Sensus; adalah cara pengumpulan data jika setiap anggota populasi diteliti satu per satu.
2. Sampling; adalah cara pengumpulan data jika hanya sebagian anggota populasi saja yang diteliti.

Beberapa metode yang dapat digunakan untuk mendapatkan data yang diperlukan dalam penelitian, yaitu dengan:

1. Mencari data yang sudah dipublikasikan oleh sumber-sumber tertentu, baik oleh Pemerintah {misalnya BPS (Badan Pusat Statistik)}, perusahaan, maupun individu.
2. Merancang suatu percobaan.
3. Melakukan survey.

Untuk memilih sampel dari suatu populasi dapat dilakukan dengan cara:

a. Cara Acak, dan
b. Cara Tidak Acak.

1.4 Skala Pengukuran

Skala pengukuran dibagi atas:

1. Skala Nominal, berupa penggolongan/klasifikasi semata; misalnya jenis kelamin, agama, golongan darah.
2. Skala Ordinal, selain penggolongan juga ada urutan/tingkatan, tapi belum ada konsep jarak; misalnya pangkat, pernyataan sikap, tingkat pendidikan.
3. Skala Interval, di sini ada penggolongan, ada urutan/tingkatan, dan ada konsep jarak tapi tidak punya 0 (nol) mutlak; misalnya suhu, tahun lahir.
4. Skala Rasiol, di sini ada penggolongan, ada urutan/tingkatan, ada konsep jarak, dan ada/punya 0 (nol) mutlak; misalnya umur, berat, tinggi.

1.5 Aturan Pembulatan

- ***Aturan 1***: Jika angka terkiri yang harus dihilangkan kurang dari 5 ($x < 5$), maka angka terkanan dari angka yang mendahuluinya tetap (tidak berubah). Contoh: 1,2 = 1.
- ***Aturan 2***: Jika angka terkiri yang harus dihilangkan lebih dari 5 ($x > 5$), maka angka terkanan dari angka yang mendahuluinya bertambah dengan satu. Contoh: 2,6 = 3 atau 3,5001 = 4.
- ***Aturan 3***: Jika angka terkiri yang harus dihilangkan sama dengan 5 ($x = 5$) atau 5 diikuti oleh bilangan nol (0) semua, maka:
  - angka terkanan dari angka yang mendahuluinya TETAP jika angka tersebut **genap**. Contoh: 2,5 = 2 atau 4,500000 = 4 .
  - angka terkanan dari angka yang mendahuluinya BERTAMBAH satu. jika angka tersebut **ganjil.** Contoh: 3,5 = 4 atau 4,500010 = 5.

BAB II
PENYAJIAN DATA

Di berbagai instansi atau perusahaan baik negeri maupun swasta, akan dijumpai berbagai gambaran data tentang instansi atau perusahaan tersebut. Gambaran tersebut meliputi tentang Pegawai, Struktur Organisasi, dan lain-lainnya.

Telah dimaklumi bahwa data yang telah dikumpulkan dari hasil observasi untuk keperluan laporan maupun analisis lebih lanjut perlu diatur, disusun dan disajikan dalam bentuk yang jelas dan baik.

Secara garis besarnya ada dua cara penyajian data yang sering dipakai, yaitu:

2.1 Penyajian Data dalam bentuk Tabel atau Daftar

- Tabel atau daftar
- Grafik atau diagram atau gambar

Ada beberapa bentuk tabel yang sudah dikenal yaitu:

a. Senarai
b. Tabel baris-kolom
c. Tabel kontingensi
d. Tabel distribusi frekuensi

2.2 Penyajian Data dalam bentuk Grafik atau Diagram atau Gambar

Ada beberapa bentuk grafik yang biasa digunakan, seperti:

a. Grafik batang (*Histogram*)
b. Grafik garis
c. Grafik lambang atau grafik simbol
d. Grafik pastel dan diagram lingkaran
e. Grafik peta atau *kartogram*
f. Grafik pencar atau diagram titik
g. Diagram Dahan Daun (*Stem-and-Leaf*)
h. Diagram Kotak (*Boxplot*).

Berikut akan diberikan contoh penyajian data dalam bentuk grafik/gambar/diagram di antaranya *Histogram*, Diagram Lingkaran, Diagram Garis, Diagram Dahan Daun (*Stem-and-Leaf*) dan Diagram Kotak (*Boxplot*).

Contoh Data Mahasiswa:

NAMA	UTS	UAS	TUGAS
DARWA	72	75	95
RAHIM	90	73	100
ARNY	88	76	95
IRIANA	88	75	95
ERFINA	79	75	95
HERMIN	91	65	95
MUS	60	75	95
SALMI	91	77	95
TATI	81	76	95
ASRI	88	75	95

PENYAJIAN DATA DALAM BENTUK GAMBAR

Diagram *Histogram* dari Data Mahasiswa:

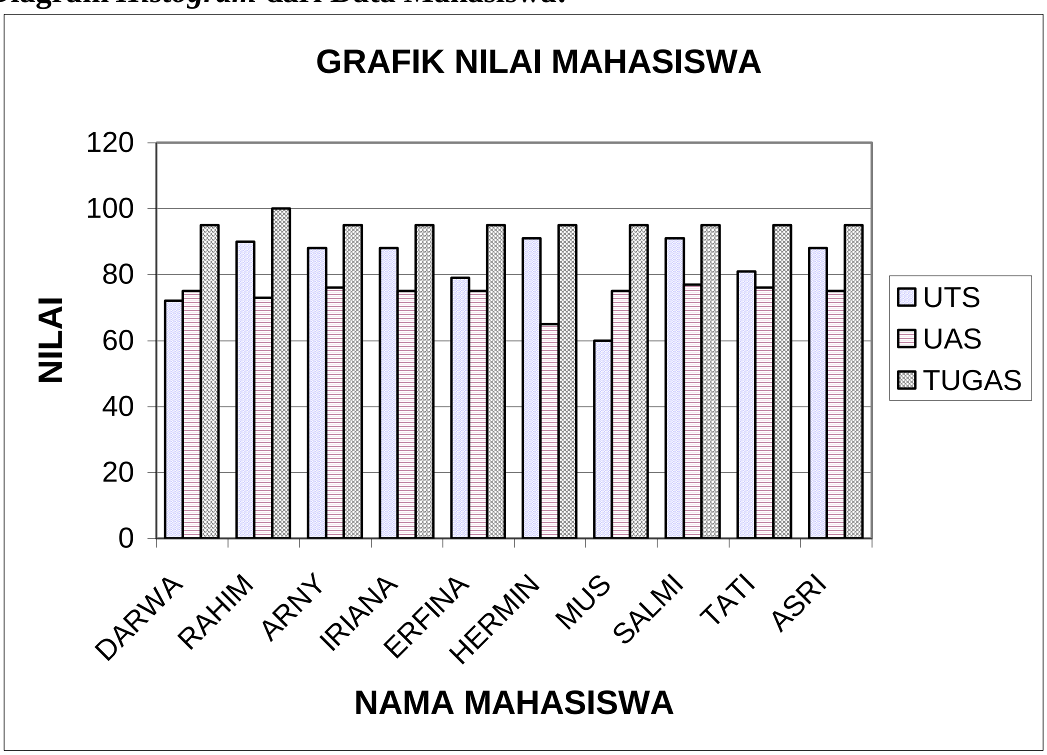


Diagram Lingkaran dari Data UTS Mahasiswa:

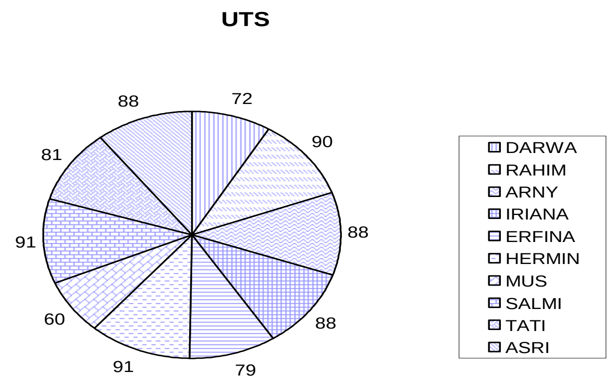


Diagram Garis dari Data Mahasiswa:

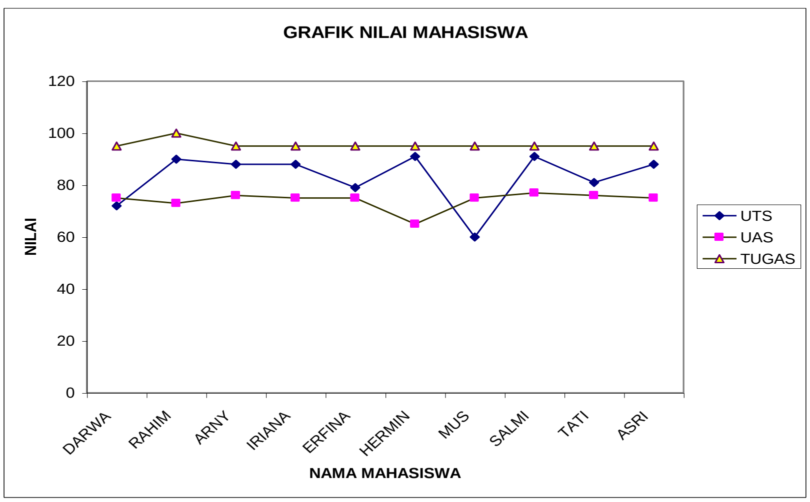


Diagram Dahan Daun dari Data UTS Mahasiswa:

Character Stem-and-Leaf Display

Stem-and-leaf of C2   N = 10

1	6	0
1	6	
2	7	2
3	7	9
4	8	1
(3)	8	888
3	9	011

BAB III
UKURAN PEMUSATAN

Disadari atau tidak, statistika telah banyak digunakan dalam kehidupan sehari-hari. Statistika sering dipakai untuk menyatakan ukuran sebagai wakil dari kumpulan data, yang melukiskan atau menggambarkan data tersebut. Statistika yang hanya menggambarkan atau melukiskan dan menganalisis data yang diberikan tanpa membuat atau menarik kesimpulan tentang data tersebut dinamakan statistika deskriptif.

Dalam menyelidiki segugus data kuantitatif, akan sangat membantu bila didefenisikan ukuran-ukuran numerik yang menjelaskan ciri-ciri data yang penting. Salah satu cara yang dapat ditempuh adalah penggunaan ukuran pemusatan.

Ukuran pemusatan dari sekumpulan data merupakan harga yang dipandang dapat menggambarkan data itu, khususnya dalam hal letaknya (lokasinya). Ada beberapa ukuran pemusatan, yaitu: rata-rata, median, kuartil, desil dan persentil, modus, rata-rata geometric, dan rata-rata harmonik.

3.1 Rata-rata

Rata-rata untuk data kuantitatif yang terdapat dalam sebuah sampel dihitung dengan jalan membagi jumlah nilai data oleh banyak data. Simbol rata-rata untuk sampel adalah $\overline{X}$, sedangkan rata-rata untuk populasi disimbolkan dengan μ .

Misalkan terdapat *n* data angka $X_1, X_2, \ldots, X_n$, rata-rata dari *n* data angka ini didefenisikan sebagai berikut:

$$\overline{X} = \frac{X_1 + X_2 + \ldots + X_n}{n} = \frac{1}{n}\sum_{i=1}^{n} X_i$$

Contoh 1:

Misalkan IPK dari 10 Mahasiswa Universitas Halu Oleo (UHO) adalah 3.4, 3.6, 2.8, 2.9, 3.0, 2.5, 2.1, 2.2, 1.9, dan 3.7, maka rata-rata dari data tersebut adalah:

$$\overline{X} = \frac{(3.4+3.6+2.8+2.9+3.6+2.5+2.1+2.2+1.9+3.7)}{10}$$
$$= 2.81$$

3.2 Median

Median dari sekumpulan data adalah harga yang di tengah apabila data telah diurutkan menurut besarnya. Jika nilai median sama dengan *Me*, maka 50 % dari data harga-harganya paling tinggi sama dengan *Me*, sedangkan 50% lagi harga-harganya paling rendah sama dengan *Me*.

Langkah-langkah menentukan median:

1. Menyusun data dari besar ke kecil atau sebaliknya.
2. Jika banyaknya data n maka:

$$\text{Median} = \begin{cases} X_{(n+1)/2} & ; \text{jika } n \text{ ganjil} \\ (X_{n/2} + X_{(n+2)/2})/2 & ; \text{jika } n \text{ genap} \end{cases}$$

Contoh 2:

Sampel dengan data: 4, 12, 5, 7, 8, 10, 10, setelah disusun menurut nilainya diperoleh: 4, 5, 7, 8, 10, 10, 12. Karena jumlah data ganjil, maka:

$$Me = X_{(n+1)/2} = X_4$$

sehingga data paling tengah adalah data ke-4 yang bernilai 8. Jadi $Me = 8$.

Contoh 3:

Diberikan sampel dengan data: 12, 7, 8, 14, 16, 19, 8, 10, setelah diurutkan menurut nilainya diperoleh: 7, 8, 8, 10, 12, 14, 16, 19. Karena jumlah data genap, maka:

$$Me = (X_{n/2} + X_{(n+2)/2})/2 = (X_4 + X_5)/2 = (10 + 12)/2 = 11$$

Jadi diperoleh $Me = 11$.

3.3 Modus

Untuk menyatakan fenomena yang paling banyak terjadi atau paling banyak terdapat digunakan ukuran modus disingkat Mo. Modus merupakan nilai yang paling sering muncul atau nilai data yang frekuensinya tertinggi. Suatu data bisa saja mempunyai lebih dari satu modus, bahkan ada data yang tidak mempunyai modus. Misalnya data IPK dari 10 mahasiswa UHO: 3.4, 3.6, 2.8, 2.9, 3.0, 2.5, 2.1, 2.2, 1.9, 3.7 tidak mempunyai modus.

Ukuran modus dalam keadaan tidak disadari sering dipakai untuk menentukan rata-rata data kualitatif. Jika didengar atau dibaca: pada umumnya kecelakaan lalu lintas karena kecerobohan pengemudi, maka ini tiada lain masing-masing merupakan modus penyebab kematian dan kecelakaan lalu lintas.

Modus untuk data kuantitatif ditentukan dengan jalan menentukan frekuensi terbanyak diantara data itu.

Contoh 6:

Terdapat sampel dengan dengan nilai data: 12, 34, 14, 34, 28, 24, 34, 28, 14, dalam tabel dapat disusun seperti dibawah ini:

X_i	f_i
12	1
14	2
28	3
34	4

Frekuensi terbanyak ialah f = 4, terjadi untuk data bernilai 34, maka modus $M_o = 34$.

3.4 Kuartil, Desil, dan Persentil

Jika sekumpulan data dibagi menjadi empat bagian yang sama banyak, sesudah disusun menurut urutan nilainya, maka bilangan pembaginya disebut **kuartil**. Ada tiga buah kuartil yaitu: kuartil pertama, kuartil kedua, dan kuartil ketiga yang masing-masing dituliskan sebagai K_1, K_2, dan K_3. Pemberian nama ini dimulai dengan kuartil paling kecil. Untuk menentukan nilai kuartil caranya adalah sebagai berikut:

1. susun data menurut urutan nilainya,
2. tentukan letak kuartil,
3. tentukan nilai kuartil.

Latak kuartil ke-*i*, diberi lambang K_i, ditentukan oleh rumus:

$$\text{Letak } K_i = \text{data ke } \frac{i(n+1)}{4} \text{ , dengan } i = 1,2,3.$$

Contoh 4:

Sampel dengan data: 75, 82, 66, 57, 64, 56, 92, 94, 86, 52, 60, 70, setelah disusun menjadi: 52, 56, 57, 60, 64, 66, 70, 75, 82, 86, 92, 95.

Letak K_1 = data ke $\frac{12+1}{4}$ = data ke-3 ¼, yaitu antara data ke-3 dan data ke-4 seperempat jauh dari data ke tiga.

Nilai K_1 = data ke-3 + ¼ (data ke-4 – data ke-3)

$K_1 = 57 + ¼ (60 – 57) = 57 ¾$

Letak K_3 = data ke- $\frac{3(12+1)}{4}$ = data ke-9 ¾ .

Dengan cara seperti di atas, nilai K_3 dapat ditentukan dengan:

K_3 = data ke-9 + ¾ (data ke-10 – data ke-9)

$K_3 = 82 + ¾ (86 – 82) = 85.$

Jika kumpulan data itu dibagi menjadi 10 bagian yang sama, maka didapat 9 pembagi dan setiap pembagi dinamakan desil. Karenanya ada sembilan buah desil, yaitu desil pertama, desil kedua, …, desil ke sembilan yang disingkat dengan D_1, D_2, …, D_9.

Desil-desil ini dapat ditentukan dengan:

1. susun data menurut urutan nilainya,
2. tentukan letak desil,
3. tentukan nilai desil.

Letak desil ke- *i*, diberi lambang D_i, ditentukan oleh rumus:

$$\text{Letak } D_i = \text{data ke } \frac{i(n+1)}{10} \text{ , dengan } i = 1, 2, \ldots, 9$$

Contoh 5:

Untuk data yang pada contoh 4, maka diperoleh:

Letak D_7 = data ke $\frac{7(n+1)}{10}$ = data ke-9,1

Nilai D_7 = data ke-9 + (0,1)(data ke-10 – data ke-9)

$D_7 = 82 + (0,1)(86 – 82) = 82,4.$

Jika sekumpulan data yang dibagi menjadi 100 bagian yang sama akan menghasilkan 99 pembagi yang berturut-turut dinamakan persentil ke-1, persentil ke-2, …, persentil ke-99. Simbol yang digunakan berturut-turut P_1, $P_{2, \ldots}$, P_{99}.

Cara perhitungan persentil sama dengan perhitungan desil, maka letak persentil P_i (i = 1, 2, …, 99) dirumuskan:

$$\text{Letak } P_i = \text{data ke-} \frac{i(n+1)}{100} \text{ , dengan } i = 1, 2, \ldots , 99.$$

3.5 Rata-rata Geometrik

Jika perbandingan tiap dua data berurutan tetap atau hampir tetap, rata-rata geometri lebih baik dipakai dari pada rata-rata hitung, apabila dikehendaki rata-ratanya. Untuk data bernilai X_1, $X_{2, \ldots,}$ X_n, maka rata-rata geometrik didefinisikan sebagai:

$$GM = \sqrt[n]{X_1 . X_2 .... X_n}$$

Untuk bilangan-bilangan bernilai besar lebih baik digunakan logaritma yaitu:

$$\text{Log}\,GM = \frac{\log X_1 + \log X_2 + ... + \log X_n}{n}$$

Contoh 7:

Misalkan diambil $X_1 = 2$, $X_2 = 4$, dan $X_3 = 8$, dengan log 2 = 0,3010, log 4 = 0,6021, dan log 8 = 0,9031.

maka, $\text{Log}\,GM = \dfrac{\log 2 + \log 4 + \log 8}{3}$

$$= \frac{0{,}3010 + 0{,}6021 + 0.9031}{3} = 0{,}6021$$

Dari daftar logaritma, diperoleh rata-rata geometri GM = 4.

3.6 Rata-rata Harmonik

Untuk data X_1, $X_{2, \ldots,}$ X_n dalam sebuah sampel berukuran *n*, maka rata-rata harmonik ditentukan oleh:

$$H = \frac{n}{\sum_{i=1}^{n} \left(\frac{1}{X_i} \right)}$$

Atau lengkapnya $H = \dfrac{n}{\frac{1}{X_1} + \frac{1}{X_2} + ... + \frac{1}{X_n}}$

Contoh 8:

Rata-rata harmonik untuk kumpulan data: 3, 5, 6, 6, 7, 10, 12, dengan *n* = 7 adalah:

$$H = \frac{7}{\frac{1}{3} + \frac{1}{5} + \frac{1}{6} + \frac{1}{6} + \frac{1}{7} + \frac{1}{10} + \frac{1}{12}} = 5{,}87$$

Penggunaan lain mengenai rata-rata harmonik adalah dalam hal sebagai berikut:

Si A bepergian pulang pergi. Waktu pergi ia melakukan kecepatan 10 km/jam sedangkan waktu kembalinya 20 km/jam. Berapakah rata-rata kecepatan pulang pergi?
Jawab: Jika digunakan rata-rata hitung biasa, diperoleh ½(10+20) km/jam = 15 km/jam.

Hal ini salah, karena jika panjang jalan 100 km, maka untuk pergi diperlukan waktu 10 jam dan kembali 5 jam. Pulang pergi perlu waktu 15 jam dan menempuh 200 km. Rata-rata kecepatan jadinya = $\frac{200}{15}$ km/jam = $13\frac{1}{3}$ km/jam.

Hasil ini tiada lain adalah rata-rata harmonic, yaitu:

$$H = \frac{2}{\frac{1}{10}+\frac{1}{20}} = \frac{40}{3} = 13\frac{1}{3}.$$

BAB IV
UKURAN PENYEBARAN

Untuk mengetahui karakteristik atau ciri-ciri yang penting dari sekelompok data, pengukuran bisa dilakukan pada aspek *central tendency* (ukuran untuk mengetahui pusat data) data, yakni mengukur rata-rata, median dan modus suatu data. Terkait dengan tujuan statistik untuk menggambarkan isi, maka itu belum cukup. Untuk itu diperlukan pengukuran tambahan untuk lebih bisa menggambarkan sekelompok data, karena adanya kemungkinan variansi data. Cara paling sederhana untuk mengukur variansi data adalah range, yang pada dasarnya menghitung selisih antara data terbesar dengan data terkecil. Selain *range*, ada dua pengukuran dispersi lain yang lazim digunakan, yakni variansi dan deviasi standar. Kedua alat pengukur ini justru paling populer dalam praktek statistik dan keduanya juga mempunyai keeratan hubungan, karena variansi adalah hasil pengkuadratan dari nilai deviasi standar.

Keunggulan dari dua alat ini adalah cakupan proses perhitungannya. Pada *range* yang diukur hanyalah ada dua titik saja dan tidak mengukur semua data yang ada, sehingga bisa terjadi dua kelompok data yang mempunyai *range* sama mempunyai interkuartil *range* yang berbeda. Data tidak akan terdeteksi jika hanya menghitung *range* saja. Hal ini berbeda jika pengukuran dilakukan dengan variansi dan deviasi standar yang mengukur variansi dengan mempertimbangkan semua isi data yang ada, sehingga walaupun *range* kedua data sama, namun akan menghasilkan variansi dan kemudian deviasi standar yang berbeda.

4.1 Rentang (*Range)*

Rentang (*range)* adalah selisih antara nilai tertinggi dan nilai yang terendah di dalam pengumpulan data. Pemakaian keterangan yang diberikan oleh *range* sebagai tambahan bagi keterangan yang telah diberikan oleh harga rata-rata mengenai sekumpulan data dengan memberi gambaran yang lebih terang mengenai kumpulan data itu.

Ada dua bentuk *range*, yang juga dipakai sebagai ukuran penyebaran di dalam hal-hal tertentu. Kedua *range* tersebut itu, yaitu 10 – 90 persentil, yaitu *range* yang merupakan selisih antara persentil ke-90 dengan persentil ke-10. Ini berarti bahwa persentil di dalam *range* tersebut termasuk 80 % dari nilai-nilai yang terdapat di dalam kumpulan data yang bersangkutan.

Contoh:

Perhatikan deretan bilangan yang berikut:

2, 102, 107, 110, 120, 139, 145, 140, 1202

Jawab:

Jika menerangkan penyebaran nilai-nilai ini dengan range sepenuhnya, maka diperoleh:

Range = data terbesar - data terkecil

= 1202 - 2 = 1200

Dengan hasil ini tidak diketahui deretan bilangan yang sebenarnya, tapi akan dibayangkan sederetan bilangan di antara 2 dan 1202, yang tersebar meluas. Padahal seperti dapat dilihat dari deretan bilangan di atas, bilangan-bilangan itu tidak tersebar hanya di antara 102 dan 150 saja, kecuali kedua nilai ekstrimnya. Jika ingin dihilangkan kedua ekstrim itu yaitu dengan memakai *range* 10 – 90 persentil, maka:

Range 10 – 90 persentil: 150 – 102 = 48

Nilai ini jauh lebih kecil dari *range* penuh, sehingga *range* 10 – 90 persentil lebih baik dipakai jika dibandingkan *range* penuh.

Dari contoh di atas jelas kelihatan bahwa kebaikan pemakaian *range* 10 – 90 persentil yaitu bahwa nilai-nilai ekstrim tidak berpengaruh di dalam penentuannya. Sedangkan jika dipakai *range* penuh sebagai ukuran penyebaran, maka nilai-nilai ekstrim saja yang menentukannya. Padahal seperti ditunjukkan oleh contoh, 80 % dari pada nilai-nilai tertumpuk di dalam *range* yang kecil saja.

Kebaikan daripada pemakaian *range* 10 – 90 persentil dipakai sebagai ukuran penyebaran yaitu mudah dimengerti dan hanya terbatas kepada hal-hal yang istimewa saja. Sedangkan kelemahannya adalah terdapat dalam defenisinya sendiri, dimana penentuannya tidak sama dari semua data itu dipertimbangkan. Ukuran lain yang sebenarnya sangat menyerupai *range* 10 – 90 persentil yaitu *interquartile range*.

4.2 Jarak Antar Kuartil (*Interquartile Range*)

Jarak Antar Kuartil (JAK) merupakan modifikasi dari *range* yang sederhana, yaitu mencoba mempersempit jarak yang akan diukur. Jika pada range sederhana, jarak kedua titik adalah data terbesar dan terkecil, maka pada JAK data yang digunakan yaitu:

$$Q = Q_3 - Q_1$$

dimana:

Q_3 = kuartil ketiga yang meliputi 75 % dan

Q_1 = kuartil pertama yang meliputi 25 %.

JAK biasa digunakan untuk mendeteksi pencilan (*outlier*) pada data. Suata data dianggap pencilan jika data tersebut lebih besar 1,5JAK dari Q_3 atau lebih kecil 1,5JAK dari Q_1.

4.3 Deviasi Rata-rata

Deviasi rata-rata adalah harga rata-rata penyimpangan tiap data terhadap meannya. Besaran ini memberikan gambaran tentang dispersi (pemencaran) data terhadap harga meannya. Misalnya adalah $X_1, X_2, ..., X_n$ dengan rata-rata $\bar{X}$, maka deviasi rata-ratanya adalah:

$$dr = \sum_{i=1}^{n} \left| X_1 - \bar{X} \right| / n$$

Contoh 1:

Diberikan data berat badan 5 calon mahasiswa UHO sebagai berikut:

No.	Berat badan (X_1) (kg)	Rata-rata ($\bar{X}$)	($X_1 - \bar{X}$)	$\left\|X_1 - \bar{X}\right\|$
1	62		-3,2	3,2
2	63	65,2	-2,2	2,2
3	65		-0,2	0,2
4	70		4,8	4,8
5	66		0,8	0,8
Jumlah	326		0	11,2

Deviasi rata-ratanya adalah = 11,2/5 = 2,24

Deviasi rata-rata data yang disusun dalam tabel distribusi frekuensi ditentukan sebagai berikut. Misalnya X_1 merupakan titik tengah interval ke-*i* dengan frekuensi f_i, rata-rata data adalah $\overline{X}$ dan banyaknya interval k, maka deviasi rata-rata adalah:

$$dr = \frac{\sum_{i=1}^{k} f_i \left| X_1 - \overline{X} \right|}{\sum_{i=1}^{k} f_i}$$

Contoh 2.

Berat	Titik Tengah	Frek	$\overline{X}$	$\left\| X_1 - \overline{X} \right\|$	$f_1 \left\| X_1 - \overline{X} \right\|$
1,45-1,95	1,7	2		1,7	3,4
1,95-2,45	2,2	1		1,2	1,2
2,45-2,95	2,7	4		0,7	2,8
2,95-3,45	3,2	15	3,4	0,2	3,0
3,45-3,95	3,7	10		0,3	3,0
3,95-4,45	4,2	5		0,8	4,0
4,45-4,95	4,7	3		1,3	3,9

Deviasi rata-ratanya adalah = 21/40 = 0,525

4.4 Variansi (Ukuran Penyebaran)

Jenis ukuran lain yang membantu menjelaskan bentuk distribusi adalah ukuran yang menunjukan tersebarnya pengamatan-pengamatan itu di sekitar rataan-rataan dan semacam ini disebut ukuran penyebaran.

Sepintas lalu jumlah simpangan pengamatan dari rataan hitung, $\sum_{i=1}^{n}(X_i - \overline{X})$, terlihat merupakan suatu ukuran yang baik untuk maksud ini, tetapi pada pemeriksaan lebih lanjut nampak nilainya sama dengan nol. hal ini dapat diatasi dengan cara mengkuadratkan simpangan itu sebelum dijumlahkan. Variansi dapat dikatakan sebagai jumlah kuadrat simpangan pengamatan dari X dibagi dengan jumlah pengamatan kurang satu. Dan disajikan dengan lambang s^2 atau dapat ditulis dengan rumus:

$$s^2 = \frac{(X_1 - \overline{X})^2 + (X_2 - \overline{X})^2 + \ldots\ldots + (X_n - \overline{X})^2}{n-1}$$

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

Contoh:

Sampel A	Sampel B
5 6 7 10	6 7 7 8

Pengukuran variansi untuk daerah A adalah:

$$s^2 = \frac{(5-7)^2+(6-7)^2+(7-7)^2(10-7)^2}{4-1} = \frac{2^2+1^2+0^2+3^2}{3} = 4{,}67$$

Untuk daerah/sampel B adalah:

$$s^2 = \frac{(6-7)^2+(7-7)^2+(7-7)^2+(8-7)^2}{4-1} = 0{,}67$$

Semakin kecil variansi sebuah data, semakin tidak bervariasi data tersebut. Sebaliknya semakin besar variansi sebuah data, maka semakin bervariasi data tersebut.

4.5 Deviasi Standar/Standar Deviasi

Untuk menghitung variansi data yang disusun dalam tabel distribusi frekuensi dapat dilakukan dengan menganggap besar tiap data sama dengan titik tengah setiap interval yang bersangkutan:

Variansi:

$$s^2 = \frac{1}{(n-1)}\sum_{i=1}^{k}(X_i-\overline{X})^2 \quad \text{atau}$$

$$s^2 = \frac{1}{(n-1)}\sum_{i=1}^{k}f_i(X_i-\overline{X})^2 = \frac{1}{(n-1)}\left(\sum_{i=1}^{n}f_iX^21-\frac{1}{n}\left(\sum_{i=1}^{n}f_iX_1\right)^2\right)$$

Deviasi standar :

$$s = \sqrt{\frac{1}{(n-1)}\sum_{i=1}^{k}(X_i-\overline{X})^2} \quad \text{atau}$$

$$s = \sqrt{\frac{1}{(n-1)}\sum_{i=1}^{k}f_i(X_i-\overline{X})^2} = \sqrt{\frac{1}{(n-1)}\left(\sum_{i=1}^{n}f_iX^2-\frac{1}{n}\left(\sum_{i=1}^{n}f_iX_1\right)^2\right)}$$

dimana: f_i = frekuensi interval ke-*i* (*i* = 1, 2, ..., *k*)
X_i = titik tengah interval ke-*i*
$\overline{X}$ = rata-rata data

n = banyaknya data.

Contoh 3.

Tentukan variansi dan deviasi standar untuk data 3, 4, 5, 6, 6, dan 7.

Jawab:

Data disusun dalam tabel sebagai berikut:

X_1	3	4	5	6	6	7	31
X_1^2	9	16	25	36	36	49	171

Variansi adalah s^2 = 1/5 x (171 – 1/6 x $(31)^2$) = 2,166

Deviasi standar adalah s = 1,47.

4.6 Deviasi Kuartil

Dapat juga digunakan sebagai ukuran penyebaran data karena apabila data semakin memencar, perbedaan harga-harga kuartilnya juga semakin membesar. Deviasi kuartil dapat didefinisikan sebagai:

$$dk = ½ \, x(K_{III}-K_1)$$

4.7 Deviasi Relatif

Ukuran penyebaran ini digunakan untuk membandingkan deviasi dua kelompok data atau lebih. Karena deviasi relatif mengabaikan suatu pengukuran, maka deviasi kelompok data dengan satuan pengukuran yang berbeda dapat dibandingkan. Deviasi relatif didefinisikan sebagai:

$$v = s/x \text{ X } 100$$

Deviasi relatif bentuk lain:

$$V_{dr} = \frac{\text{Deviasi rata - rata}}{\overline{X}} \text{ X } 100 \text{ dan} \qquad v_{dk} = \frac{K_{III} - K_I}{K_{III} + K_I}$$

4.8 Kemencengan (*skewness*) dan Kesimetrisan

Apabila suatu disribusi simetris, maka mean, median dan modus akan berhimpit (sama besar). Jika frekuensi harga-harga yang kecil lebih besar, maka mean akan semakin kecil dan bila frekuensi harga-harga yang besar lebih besar, maka mean akan semakin besar. Sehingga semakin menceng suatu distribusi jarak antara mean dan modus juga akan semakin besar.

$$km = \frac{\text{mean - modus}}{\text{deviasi standar}}$$

Dari rumus dapat diperoleh bahwa km = 0. Jika mean lebih besar dari modus, maka km bernilai positif dan dikatakan distribusi menceng positif (ke kanan). Begitu pula sebaliknya.

BAB V
DISTRIBUSI FREKUENSI

5.1 Latar Belakang

Kata data (pengamatan) dalam statistika merupakan kata yang tidak asing lagi. Dimana data tersebut diperoleh dengan cara atau metode pengumpulan yang berbeda. Setelah data tersebut dikumpulkan, kemudian disusun dengan susunan yang teratur dan sistematis, sehingga sifat-sifat data dapat dilihat dengan mudah. Setelah data primer yang dikumpulkan dengan menggunakan daftar pertanyaan dikumpulkan, perlu diklasifiaksikan dan ditabulasi agar nampak sifat-sifat data yang menonjol. Dengan persentasi data dalam bentuk tabel-tabel, maka kegiatan statistik telah cukup, karena bentuk-bentuk ini sebagai upaya persentasi data telah menunjukkan informasi penting dari data tersebut. Apabila persentasi data tersebut dipandang belum cukup memberi informasi, maka dapat dilakukan analisa terhadap data tersebut. Dalam hal ini, klasifikasi data tersebut sangat membantu untuk kegiatan analisa data.

Data yang telah dikumpulkan, baik berasal dari populasi ataupun dari sampel, untuk keperluan laporan dan atau analisis selanjutnya perlu diatur, disusun, disajikan dalam bentuk yang jelas dan baik. Garis besarnya ada dua cara penyajian data yang sering dipakai ialah *tabel* atau *daftar* dan *grafik* atau *diagram.*

Ciri-ciri penting sejumlah besar data dapat diketahui melalui pengelompokkan data ke dalam beberapa kelas dan kemudian dihitung banyaknya data yang masuk ke dalam setiap kelas. Susunan data seperti ini sering dikenal dengan distribusi frekuensi.

5.2 Distribusi Frekuensi

Distribusi frekuensi atau tabel frekuensi adalah suatu tabel dimana banyaknya kejadian/frekuensi (*cases*) didistribusikan kedalam kelompok-kelompok (kelas-kelas) yang berbeda.

Dalam daftar distribusi frekuensi, banyak obyek yang dikumpulkan dalam kelompok-kelompok berbentuk a – b, yang disebut *kelas interval*. Kedalam kelas interval a – b dimasukkan semua data yang bernilai *mulai dari a sampai dengan b.*

Urutan kelas interval disusun mulai data terkecil terus kebawah sampai nilai data terbesar. Berturut-turut mulai dari atas, diberi nama kelas interval pertama, kelas interval kedua, …, kelas interval terakhir. Ini semua ada dalam kolom kiri. Kolom kanan berisikan bilangan-bilangan yang menyatakan berapa buah data terdapat dalam tiap kelas interval. Jadi kolom ini berisikan *frekuenai*, disingkat dengan f.

Bilangan-bilangan diseelah kiri kelas interval disebut *ujung bawah* dan bilangan-bilangan disebelah kanannya disebut *ujung atas*. Selisih positif antara tiap ujung bawah berurutan disebut *panjang kelas interval* (p). Selain dari *ujung kelas interval* ada lagi yang biasa disebut *bats kelas interval*. Ini tergantung pada ketelitian data yang digunakan. Jika data yang dicatat teliti hingga satuan, maka *batas bawah kelas* sama dengan ujung bawah dikurangi 0,5. Batas atasnya didapat dari ujung atas ditambah dengan 0,5. Untuk data dicatat hingga satu desimal, bats bawah sama dengan ujung bawah dikurangi 0,05 dan batas atas sama dengan ujung atas ditambah 0,05. Kalau data hingga dua desimal, batas bawah sama dengan ujung bawah dikurangi 0,005 dan batas atas sama dengan ujung atas ditambah 0,005 dan begitu seterusnya. Untuk perhitungan nanti, dari tiap kelas interval biasa diambil sebuah nilai sebagai *wakil* kelas itu. (Sudjana, 1996).

Distribusi frekuensi dapat dibedakan menjadi dua macam, yaitu distribusi frekuensi kuantitatif (distribusi frekuensi yang kelas-kelasnya dinyatakan dalam bentuk bilangan-bilangan atau angka-angka) dan distribusi frekuensi kualitatif/kategorik (distribusi frekuensi yang kelas-kelasnya dinyatakan dalam bentuk kategori-kategori).

a. Distribusi Frekuensi Kualitatif

Contoh

Distribusi Frekuensi Kualitatif:

Tabel 1 Jumlah Mahasiswa Peserta Statistika di FMIPA:

Jurusan	Jumlah
Matematika	45
Fisika	40
Kimia	43
Biologi	35
Statistika	47

b. Distribusi Frekuensi Kuantitatif

Contoh Distribusi Frekuensi Kuantitatif:

Tabel 2 Jumlah Mahasiswa Peserta Biostatistika di STIKA:

Kelas Interval	Frekuensi (f)
25 - 27	20
28 - 30	26
31 - 33	10
34 - 36	4
Jumlah	60

Dalam menyusun distribusi frekuensi, ada beberapa syarat yang harus diperhatikan agar diperoleh distribusi frekuensi (tabel frekuensi) yang baik, yaitu:

1. Tabel frekuensi hendaknya mempunyai nomor tabel, judul tabel, dan satuan. Nomor tabel dimaksudkan untuk mempermudah membedakan suatu tabel dengan tabel yang lain.
2. Banyaknya kelas sedapat mungkin ditentukan dengan menggunakan pedoman sturges. Meskipun jumlah kelas ditentukan dengan pedoman sturges, namun sebaiknya dihindarkan penggunaan jumlah kelas yang lebih dari 20.
3. Hindarkan adanya kelas terbuka (*open end class*), karena kelas terbuka tidak ada batasnya.
4. Hindarkan adanya interval kelas yang tidak sama (*unequalclass interval*).
5. Hindarkan adanya kelas yang berulang (*overlapping class*). Dengan adanya kelas yang berulang akan menimbulkan kemungkinan suatu data dihitung atau dimasukkan secara berulang pula.
6. Sumber data hendaknya disebutkan, karena sumber data dapat digunakan untuk pengecekan kembali apabila ada keragu-raguan terhadap data (Boediono & Koster, 2002).

5.3 Distribusi Frekuensi Kualitatif dan Distribusi Frekuensi Kuantitatif

Dalam bab sebelumnya, telah dibahas mengenai macam distribusi frekuensi, yaitu distribusi frekuensi kuantitatif dan distribusi frekuensi kualitatif (kategorik). Dimana dalam menyusun distribusi frekuensi tersebut, terdapat langkah-langkah yang harus ditempuh.

Langkah-langkah menyusun distribusi frekuensi kuantitatif adalah:

1. Tentukan rentang data, dimana rentang data = data terbesar – data terkecil. Misalnya untuk data 79, 80, 70, 68, 90, 92, 80, 70, 63, 76, maka rentang data adalah 92 – 63 = 29.
2. Tentukan banyaknya interval kelas yang digunakan, biasanya dipilih $5 \leq p \leq 15$ tergantung kebutuhan. Untuk $n \geq 200$, maka digunakan Aturan Sturges yaitu banyaknya kelas (p) = 1 + 3,3 log(n), dimana n adalah banyaknya data. Untuk data yang kita punyai misalkan diambil p = 6. Ada tiga hal yang harus diperhatikan berkenaan dengan interval kelas:
   a. Interval kelas tidak boleh tumpang tindih agar satu titik data tidak masuk pada dua interval atau lebih.
   b. Tidak boleh ada jarak yang terlalu jauh antara satu interval dengan interval berikutnya sehingga tidak ada data yang jatuh pada celah tersebut.
   c. Interval yang digunakan sebaiknya mempunyai lebar yang sama.
3. Lebar interval kelas = $\frac{rentang}{banyaknya\ \ interval\ \ kelas}$, kemudian dibulatkan. Untuk data di atas, lebar interval kelas = 29/6 = 4,8 dibulatkan menjadi 5. jika datanya bulat, batas-batas interval sebaiknya pecahan (setengah). Untuk data di atas salah satu susunan intervalnya adalah : 62,5 – 67,5; 67,5 – 72,5; 72,5 – 77,5; 77,5 – 82,5; 82,5 – 87,5; 87,5 – 92,5. Ujung kiri setiap interval disebut batas bawah, ujung kanan setiap interval disebut batas atas. Titik tengah interval = (batas bawah + batas atas) / 2.
4. Hitung banyaknya observasi dalam himpunan data tersebut yang termasuk dalam tiap-tiap interval kelas. Banyaknya observasi yang termasuk dalam tiap interval kelas disebut frekuensi kelas.
5. Frekuensi relatif setiap kelas diperoleh dengan membagi frekuensi kelas bersangkutan dengan banyaknya data keseluruhan (n). Tabel distribusi frekuensi relatif dinyatakan dalam persen, dan tabel distribusi frekuensi absolut menggambarkan frekuensi absolut setiap kelas.

Cara dan syarat-syarat yang harus dipenuhi dalam menyusun distribusi frekuensi kuantitatif pada umumnya berlaku pula pada penyusunan frekuensi kualitatif. Langkah-langkahnya adalah:

1. Tentukan banyaknya kelas (kategori) yang digunakan
2. Tentukan frekuensi setiap kelas.

Dalam menyusun distribusi frekuensi kualitatif, banyaknya kelas yang dipakai tergantung pada keadaan data dan keinginan kita. Harus dihindari adanya data yang tidak termuat dan adanya kelas-kelas yang tumpang tindih. Untuk itu setiap kelas harus didefinisikan secara tepat dan jelas. Kadang-kadang dijumpai beberapa data yang tidak dapat dimasukkan kedalam kelas-kelas yang sudah ada, untuk mengatasi masalah ini harus didefinisikan suatu kelas lagi yang disebut kelas lain-lain.

5.4 Menggambar Distribusi Frekuensi

Menggambarkan distribusi frekuensi (tabel frekuensi) dalam bentuk berbagai diagram, dimaksudkan agar informasi lebih mudah dibaca. Cara menggambar distribusi frekuensi ada berbagai macam, yaitu:

1. Histogram
2. Polygon
3. Frekuensi komulatif
4. Frekuensi relatif.

5.4.1 Histogram

Histogram pada dasarnya adalah rangkaian berbagai bidang segi empat yang masing-masing bidang menunjukkan banyaknya frekuensi pada masing-masing interval kelasnya. Apabila interval kelasnya sama, maka skala tertinggi frekuensinya akan sama. Histogram distribusi frekuensi absolut digambarkan pada sistem sumbu silang X – Y. Kelas-kelas interval diletakkan pada sumbu X, sedangkan frekuensi kelas diletakkan pada sumbu Y.

Keuntungan histogram antara lain: masing-masing bidang segi empat menunjukkan kelas-kelas yang terpisah sebab bidang-bidangnya menunjukkan jumlah frekuensi yang ada pada masing-masing kelas.

Dalam menggambarkan data ini selanjutnya digunakan tepi kelas dan bukan batas kelas, karena tepi kelas (*class boundaries*) dapat berfungsi menghilangkan kesenjangan (*gap*) yang ada pada masing-masing kelas.

5.4.2 Polygon

Poligon adalah garis yang menghubungkan titik-titik tengah dari kelas-kelas suatu distribusi frekuensi atau suatu histogram. Dalam menggambarkan polygon kita harus menambahkan satu kelas pada awal dan satu kelas pada akhir, yang masing-masing tidak mempunyai frekuensi (tidak ada datanya), sehingga garis polygon masing-masing ujungnya memotong sumbu horizontal (datar).

Beberapa keuntungan dari polygon adalah lebih sederhana dibandingkan histogram, pola grafik lebih mudah memberikan informasi sedang apabila ditambah kelas-kelasnya atau datanya polygon akan menggambarkan kurva yang kontinu.

5.4.3 Frekuensi Komulatif (Ogive)

Ogive adalah polygon frekuensi untuk distribusi frekuensi komulatif. Daftar distribusi frekuensi komulatif dapat dibentuk dari daftar distribusi frekuensi biasa, dengan jalan menjumlahkan frekuensi demi frekuensi. Frekuensi komulatif dari suatu distribusi frekuensi dapat menunjukkan beberapa frekuensi yang terletak di atas atau di bawah suatu nilai tertentu dalam interval kelas. Ada dua macam frekuensi komulatif, yakni :

1. Frekuensi komulatif "kurang dari" (*less than*)
2. Frekuensi komulatif "lebih dari" (*more than*)

Frekuensi komulatif kurang dari untuk menunjukkan jumlah frekuensi yang kurang dari nilai tertentu, sedang frekuensi komulatif lebih dari untuk menunjukkan jumlah frekuensi yang lebih dari nilai-nilai tertentu. Sehingga ogive juga terbagi dua, yaitu ogive kurang dari dan ogive lebih dari.

5.4.4 Frekuensi Relatif

Disamping histogram yang menggambarkan distribusi frekuensi absolut pada masing-masing kelas, maka dapat pula digambarkan frekuensi relatifnya untuk masing-masing kelas. Frekuensi relatif masing-masing jelas dapat dihitung berdasarkan frekuensi absolut pada masing-masing kelas dibagi dengan seluruh frekuensinya. Jadi jika data dinyatakan dalam persen, maka diperoleh daftar distribusi frekuensi relatif. Frekuensi relatif disingkat f_{ref} atau f (%).

Histogram dengan nilai absolut dan frekuensi relatif pada dasarnya sama, perbedaannya terletak dari skala vertikal pada histogram dengan nilai absolut, skalanya frekuensi, sedang pada frekuensi relatif skalanya frekuensi relatif.

Keuntungan frekuensi relatif ini adalah dapat mempermudah perbandingan distribusi yang menggunakan sampel yang berbeda. Pada histogram dengan nilai relatif hubungan antara kelas adalah tetap.

5.5 Rata-rata

Rata-rata untuk data kuantitatif yang sudah dikelompokkan atau disusun dalam tabel distribusi frekuensi ditentukan sebagai berikut:

$$\overline{\mathrm{X}} = \frac{\sum_{i}^{n} f_i . X_i}{\sum_{i}^{n} f_i}$$

5.6 Median

Median untuk data kuantitatif yang sudah dikelompokkan atau disusun dalam tabel distribusi frekuensi ditentukan sebagai berikut:

$$Me = b + P\left(\frac{\frac{1}{2}n - F}{f}\right)$$

dimana:

b : batas bawah interval median

P : panjang interval

b_1 : selisih frekuensi interval modus dengan interval sebelumnya

b_2 : selisih frekuensi interval modus dengan interval sesudahnya.

5.7 Modus

Modus untuk data kuantitatif yang sudah dikelompokkan atau disusun dalam tabel distribusi frekuensi ditentukan sebagai berikut:

$$Mo = b + P\left(\frac{b_1}{b_1 + b_2}\right)$$

dimana:

b : batas bawah interval modus

P : panjang interval

b_1 : selisih frekuensi interval modus dengan interval sebelumnya

b_2 : selisih frekuensi interval modus dengan interval sesudahnya.

5.8 Kuartil, Desil, dan Persentil

Kuartil untuk data kuantitatif yang sudah dikelompokkan atau disusun dalam tabel distribusi frekuensi ditentukan sebagai berikut:

$$\text{Letak } K_i = \ K_i = b_i + P\left(\frac{\frac{i}{4}n - cfb}{f_K}\right)$$

Desil untuk data kuantitatif yang sudah dikelompokkan atau disusun dalam tabel distribusi frekuensi ditentukan sebagai berikut:

$$\text{Letak } D_i = \ D_i = b_i + P\left(\frac{\frac{i}{10}n - cfb}{f_D}\right)$$

Persentil untuk data kuantitatif yang sudah dikelompokkan atau disusun dalam tabel distribusi frekuensi ditentukan sebagai berikut:

$$\text{Letak } P_i = \ P_i = b_i + P\left(\frac{\frac{i}{100}n - cfb}{f_P}\right)$$

5.9 Deviasi Rata-rata

Deviasi rata-rata data kuantitatif yang sudah dikelompokkan atau disusun dalam tabel distribusi frekuensi ditentukan sebagai berikut. Misalnya X_1 merupakan titik tengah interval ke-*i* dengan frekuensi f_i, rata-rata data adalah $\overline{X}$ dan banyaknya interval *k*, maka deviasi rata-rata adalah:

$$\text{dr} = \frac{\sum_{i=1}^{k} \text{fi}\left|X_1 - \overline{X}\right|}{\sum_{i=1}^{k} \text{fi}}$$

Contoh 2.

Berat	Titik Tengah	Frek	$\overline{X}$	$\lvert X_1 - \overline{X} \rvert$	$f_1 \lvert X_1 - \overline{X} \rvert$
1,45-1,95	1,7	2		1,7	3,4
1,95-2,45	2,2	1		1,2	1,2
2,45-2,95	2,7	4		0,7	2,8
2,95-3,45	3,2	15	3,4	0,2	3,0
3,45-3,95	3,7	10		0,3	3,0
3,95-4,45	4,2	5		0,8	4,0
4,45-4,95	4,7	3		1,3	3,9

Deviasi rata-ratanya adalah = 21/40 = 0,525

5.10 Variansi

Variansi untuk data kuantitatif yang sudah dikelompokkan atau disusun dalam tabel distribusi frekuensi ditentukan sebagai berikut:

$$s^2 = \frac{f_1(X_1 - \overline{X})^2 + f_2(X_2 - \overline{X})^2 + \ldots\ldots + f_n(X_n - \overline{X})^2}{n-1} = \frac{\sum_{i-1}^{n} f_i(X_i - \overline{X})^2}{n-1}$$

5.11 Simpangan Baku

Simpangan baku untuk data kuantitatif yang sudah dikelompokkan atau disusun dalam tabel distribusi frekuensi ditentukan sebagai berikut:

$$s = \sqrt{\frac{f_1(X_1 - \overline{X})^2 + f_2(X_2 - \overline{X})^2 + \ldots\ldots + f_n(X_n - \overline{X})^2}{n-1}} = \sqrt{\frac{\sum_{i-1}^{n} f_i(X_i - \overline{X})^2}{n-1}}$$

BAB VI
PELUANG (PROBABILITAS)

6.1 Pendahuluan

Teori peluang sering disebut dengan teori kemungkinan yang merupakan konsep dasar dari ilmu statistika. Di dalam kehidupan sehari-hari kita sering menjumpai suatu kejadian kadang terjadi maupun tidak jadi terjadi, atau kita juga sering mendapatkan suatu pertanyaan-pertanyaan yang di dalamnya mengandung suatu ketidakpastian. Umumnya pernyataan ketidakpastian tersebut dinyatakan dengan perkataan "mungkin" atau "kemungkinan". Sebagai misal, kemungkinan seorang mahasiswa tidak lulus ujian, kemungkinan seorang pejudi memenangkan lotre, kemungkinan seorang atlit memenangkan pertandingan dan lain-lain. Terjadinya suatu peristiwa tersebut mempunyai tingkat yang berbeda-beda, ada yang kemungkinan terjadinya besar dan ada yang kemungkinan terjadinya kecil.

Dari uraian di atas, jelas kita menyatakan ketidakpastian dengan kata/pernyataan kemungkinan. Jika kita kaji lebih lanjut, istilah kemungkinan tersebut digunakan untuk dua hal yang berbeda. Pertama, ukuran ketidakpastian (dinyatakan dengan kemungkinan kecil sekali dan mungkinkah). Kedua, merupakan suatu alternatif untuk memperoleh sesuatu bila terdapat lebih dari satu jenis/kriteria/kategori.

Untuk hal yang pertama, kemungkinan yang menyangkut ketidakpastian, kita namakan peluang (atau *probability*), sedangkan yang bersifat lebih umum kita sebut dengan kemungkinan begitu saja (*possibility*). Jelasnya peluang kita gunakan untuk menunjukkan ukuran atau derajat ketidakpastian atau kepastiannya.

Kaitannya dengan statistika inferensial, teori peluang adalah bagian integral dari ilmu statistika, dan merupakan salah satu bagian terpenting dalam teori statistika inferensial. Karena statistika inferensial berkaitan dengan metode pendugaan dan penarikan kesimpulan terhadap karakteristik suatu populasi berdasarkan informasi yang diperoleh dari sampel. Dalam proses pendugaan atau penarikan kesimpulan tersebut terkandung suatu unsur "ketidakpastian", karena pada proses tersebut jarang sekali di dukung oleh informasi atau input yang sempurna. Secara statistik derajat/tingkat ketidakpastian tersebut dikuantifikasikan dengan menggunakan teori peluang.

6.2 Percobaan dan Ruang Contoh

Dalam statistika kita menggunakan istilah percobaan dan ruang contoh untuk membangkitan suatu data. Percobaan (*event*) adalah suatu proses atau kegiatan yang menghasilkan satu kejadian dari beberapa kejadian yang mungkin dihasilkan atau terjadi. Suatu percobaan dapat berupa pelemparan satu dadu. Dalam percobaan ini terdapat enam kemungkinan hasil yang terjadi yaitu kemungkinan munculnya mata dadu 1, 2, 3, 4, 5 dan 6. Jadi terjadinya kejadian munculnya mata dadu tersebut tidak dapat diduga dengan pasti. Tetapi kita mengetahui semua kemungkinan hasil untuk setiap pelemparan tersebut. Himpunan kemungkinan hasil suatu percobaan disebut dengan ruang contoh (*sample space*). Setiap kemungkinan hasil dalam suatu percobaan dalam suatu ruang contoh disebut titik contoh (*sample point*).

Contoh 1:

Pelemparan sekeping uang logam, maka ruang contohnya adalah:

S = {G, A}

Contoh 2:

Pelemparan dua keping uang logam secara bersamaan, maka ruang contohnya adalah:

S = {GG, GA, AG, AA}

Contoh 3:

Pelemparan satu dadu bermata enam, maka ruang contohnya adalah:

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

6.3 Permutasi dan Kombinasi

Nilai peluang suatu kejadian, sering kali dapat ditentukan hanya dengan menghitung jumlah kejadian yang terdapat dalam ruang sampel dari suatu percobaan tanpa harus mendaftarkan seluruh unsur/kejadian dalam ruang sampel tersebut. Untuk itu perlu adanya suatu teori yang dapat dipakai untuk menghitung nilai peluang jika seluruh unsurnya tidak diketahui. Salah satu caranya adalah dengan menggunakan permutasi dan kombinasi.

Pengetahuan tentang permutasi dan kombinasi akan memudahkan dalam perhitungan-perhitungan peluang yang dilakukan. Misalkan saja dalam percobaan pelemparan mata uang logam tiga kali. Ruang contoh yang diperoleh adalah:

S = {GGG,GGA,GAG,AGG,GAA,AGA,AAG,AAA}

Bila ingin dihitung peluang untuk mendapatkan 2A dalam percobaan tersebut, maka yang dipertimbangkan adalah:

GAA, AGA dan AAG

dan diketahui bahwa P(GAA) = P(AGA) = P(AAG)

Perhatikan susunan A, jika A nampak pada pelemparan pertama diberi catatan A_1, kedua A_2 dan ketiga A_3.dengan demikian diperoleh susunan 2A yang mungkin dari ketiga pelemparan:

A_2 $A_{3,}$ $A_1A_{3,}$ A_1A_2

Yaitu sebanyak 3 susunan berbeda berukuran 2 dari obyek berbeda berukuran 3 (A_1, A_2, dan A_3). Susunan berbeda seperti susunan tersebut dinamakan kombinasi berukuran 2 dari obyek berukuran 3, jadi :

A_1A_2, A_1A_3, A_2A_1, A_1A_3 , A_2A_3, A_3A_2

dinamakan permutasi.

Jika obyek berbeda tersebut sebanyak n sedangkan susunan beranggotakan r maka untuk permutasi dan kombinasi banyaknya berturut-turut ditulis dengan:

$${}_nP_r \text{ dan } {}_nC_r$$

(untuk kombinasi seringkali ditulis dengan $[{n \atop r}]$ sebagai pengganti ${}_nC_{r.}$)

Jika dari n unsur yang berbeda diambil r unsur (r ≤ n), maka jumlah permutasinya dinotasikan dengan ${}_nP_r$, dimana:

$${}_nP_r = n! / (n-r)! \qquad (6.1)$$

dan jika dalam menyusun unsur-unsur tersebut, ada kalanya kita hanya tertarik pada jumlah susunan yang berbeda tanpa menghiraukan urutan dalam setiap susunan maka disebut dengan kombinasi. Kombinasi r unsur yang berbeda dari n unsur yang ada, maka jumlah kombinasinya dinotasikan dengan ${}_nC_r$ dimana:

$${}_nC_r = n! / \{r! (n-r)!\} \qquad (6.2)$$

Contoh 4:

Satu set lampu hias mempunyai 9 bola lampu. Jika dipunyai 3 bola lampu berwarna merah, 4 berwarna kuning dan 2 bola berwarna hijau, tentukan jumlah susunan yang dapat dibuat untuk menempatkan ke-9 bola lampu tersebut?

Penyelesaian:

Jumlah permutasi yang mungkin dapat disusun dari ke-9 bola lampu tersebut adalah:

$$p_n^r = \frac{n!}{n_1!n_2!n_3!} = \frac{9!}{3!4!2!} = 1260$$

Banyaknya cara yang mungkin untuk menyusun 9 bola lampu tersebut adalah 1260 cara.

Contoh 5:

Dalam satu keranjang terdapat 20 buah jeruk, 8 jeruk berwarna kuning dan 12 jeruk berwarna hijau. Jika diambil 5 buah jeruk berurutan secara acak ada berapa cara untuk memperoleh 2 jeruk kuning?

Penyelesaian:

Kombinasi yang mungkin dilakukan untuk memilih 2 jeruk kuning dari 5 jeruk adalah:

$$C_n^r = \frac{n!}{(n-r)!r!} = \frac{5!}{(5-2)!2!} = 10.$$

6.4 Definisi Peluang

Telah cukup banyak usaha yang dilakukan para ahli statistika untuk mendefinisikan peluang suatu kejadian secara tepat. Dua macam pendekatan dalam menginterprestasikan peluang akan kita bahas dalam bagian ini, yaitu pendekatan peluang secara klasik dan pendekatan dengan konsep frekuensi relatif.

Pada pendekatan klasik, peluang suatu kejadian diinterpretasikan berdasarkan atas asumsi simetris dari sifat percobaan. Misalnya pada percobaan pelemparan satu mata uang, hanya ada dua kejadian yang mungkin dihasilkan, yaitu timbulnya sisi muka atau sisi belakang. Dengan asumsi simetris kita menganggap bahwa kedua permukaan tersebut mempunyai peluang yang sama untuk terjadi. Oleh karena itu, peluang timbulnya sisi muka sama dengan peluang timbulnya sisi belakang yaitu sama dengan ½. (1 dari 2 kejadian). Secara umum, jika suatu percobaan dapat menghasilkan *n* kejadian, maka dengan pendekatan klasik, peluang terjadinya salah satu kejadian tersebut adalah 1/*n*.

Misalkan diambil kejadian A dalam ruang contoh S. Misalkan kejadian A mempunyai unsur *n*A dan ruang contoh dengan unsur sebanyak *n* (atau ada *n*A pada A dan *n* pengamatan pada S). Dengan demikian:

“Peluang terjadinya A ditulis dengan P(A) adalah nisbah antara banyak unsur/pengamatan pada A dengan banyaknya unsur hasil yang mungkin (unsur dalam S) dari suatu percobaan atau:

$$P(A) = \frac{nA}{n} \qquad (6.3)$$

Beberapa sifat dari peluang klasik diantaranya :

1. Untuk setiap A dalam S, $0 < P(A) < 1$
2. $P(S) = 1$
3. Jika $A_1, A_2, \ldots, A_k$ adalah kejadian-kejadian yang mutual eksklusif dalam S, maka:

$$P(\bigcup_{i=1}^{K} A_1) = P(A_1) + P(A_2) + \ldots + P(A_k)$$

Penggunaan pendekatan klasik dalam menentukan nilai peluang sangat tergantung pada asumsi bahwa semua kejadian yang mungkin dihasilkan mempunyai peluang yang sama. Jika asumsi tersebut tidak dipenuhi, maka nilai peluang yang dihasilkan akan salah.

Pendekatan yang lain adalah pendekatan empiris (frekuensi relatif) melalui percobaan yang diulang-ulang tak terbatas. Bila peluang setiap titik contoh tidak dapat dianggap sama, maka peluang itu harus diberikan berdasarkan pengetahuan sebelumnya atau berdasarkan bukti percobaan. Hal ini didasarkan pada kenyataan bahwa kesempatan yang sama untuk terjadi sebagaimana pada pendekatan klasik tidak selamanya dapat terpenuhi. Bahkan kesempatan yang sama tersebut tidak pernah dipenuhi. Misalnya, percobaan pelemparan uang logam yang di ulang n kali. Jika dari percobaan-percobaan tersebut timbul kejadian tertentu misalnya kejadian A, sebanyak f kali, maka jika n cukup besar, nilai proporsi f/n dapat digunakan sebagai suatu pendekatan bagi nilai peluang terjadinya kejadian A. Jelasnya, peluang terjadinya kejadian A, untuk n yang cukup besar adalah :

$$P(A) = \frac{nA}{n}$$

dimana:

nA = banyaknya/frekuensi kejadian A

n = banyaknya percobaan yang dilakukan (banyak pengamatan atau total frekuensi)

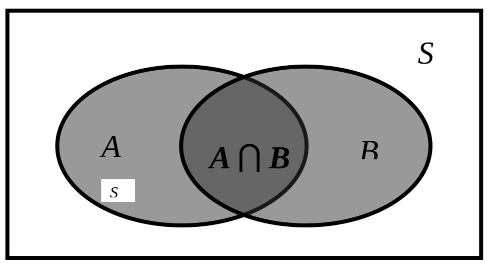
6.5 Kaidah Penjumlahan

Bila A dan B adalah dua kejadian sembarang, maka:

$$P(A \cup B) = P(A) + P(B) - P(A \cap B) \qquad (6.4)$$

<u>Bukti</u>:

Lihat diagram Venn berikut:



Gambar 6.1. Diagram venn untuk kaidah penjumlahan

$P(A \cup B)$ = jumlah peluang semua titik contoh dalam ($A \cup B$).

P(A) + P(B) = jumlah semua peluang dalam A ditambah jumlah semua peluang dalam B.

Contoh 6:

Bila peluang Amin lulus Matematika adalah 2/3, dan peluang ia lulus Biostatistika adalah 4/9, serta peluang ia lulus kedua mata kuliah tersebut adalah 1/4. Berapa peluang Amin lulus, paling tidak satu mata kuliah?

Penyelesaian:

$$P(M \cup S) = P(M) + P(S) - P(M \cap S)$$

$$= 2/3 + 4/9 - 1/4 = 31/36$$

Contoh 7:

Berapa peluang didapatnya jumlah 7 atau 11, bila dua dadu dilemparkan sekali?

Penyelesaian:

Misal A = kejadian jumlah 7 dan B = kejadian jumlah 11

Jumlah 7 dapat muncul 6 dari 36 titik contoh dan jumlah 11 dapat muncul 2 dari 36 titik contoh, Oleh karena itu,

$$P(A) = \frac{6}{36} = \frac{1}{6} \quad dan \quad P(B) = \frac{2}{36} = \frac{1}{18}$$

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

$$= \frac{1}{6} + \frac{1}{18} = \frac{2}{9}$$

Teori Komplemen

Bila A dan A' kejadian yang saling berkomplemen, maka:

$$P(A') = 1 - P(A) \qquad (6.5)$$

Bukti:

$A \cup A' = S$, dan $A \cap A' = \varnothing$,

maka: $1 = P(S) = P(A \cup A') = P(A) + P(A')$

sehingga:

$$P(A') = 1 - P(A)$$

Contoh 8:

Sekeping uang logam dilemparkan 6 kali berturut-turut. Berapa peluang paling sedikit sisi gambar muncul sekali ? .

Penyelesaian :

Ruang contoh S mempunyai $2^6 = 64$ titik contoh, karena setiap pelemparan menghasilkan 2 kemungkinan.

Misal E = munculnya sisi gambar, dan E' = bila semua pelemparan menghasilkan sisi angka, maka:

$$P(E) = 1 - P(E') = 1 - \frac{1}{64} = \frac{63}{64}$$

6.6 Kaidah Penggandaan

Bila dalam percobaan kejadian A dan B keduanya dapat terjadi sekaligus, maka:

$P(A \cap B) = P(A)\,P(B|A)$ *atau*

$P(B \cap A) = P(B)\,P(A|B)$ (6.6)

Contoh 9:

Dalam suatu kotak berisi 20 sekering, 5 diantaranya cacat. Bila 2 sekering diambil satu demi satu secara acak tanpa pengembalian. Berapa peluang kedua sekering yang diambil tersebut cacat?

Penyelesaian:

Misal: A = kejadian sekering pertama cacat

B = kejadian sekering kedua cacat

maka $A \cap B$ = kejadian A terjadi, dan kemudian kejadian B terjadi setelah A terjadi.

Jadi:

$$P(A \cap B) = \left(\frac{1}{4}\right)\left(\frac{4}{19}\right) = \frac{1}{19}$$

Definisi Kaidah Penggandaan Khusus

Bila dua kejadian A dan B bebas jika dan hanya jika

$P(A \cap B) = P(A)\ P(B)$ (6.7)

Contoh 10:

Bila pada contoh 9 sekering pertama dimasukkan kembali ke dalam kotak, maka peluang mendapatkan sekering rusak pada pengambilan kedua tetap sebesar ¼, sehingga: $P(B|A) = P(B)$ dan kedua kejadian A dan B dikatakan bebas. Jadi:

$$P(A \cap B) = \left(\frac{1}{4}\right)\left(\frac{1}{4}\right) = \frac{1}{16}$$

Contoh 11:

Pada suatu kota kecil memiliki 1 mobil pemadam kebakaran dan 1 ambulans. Peluang mobil kebakaran itu dapat digunakan pada saat diperlukan adalah 0,98, dan peluang ambulans tersedia waktu diperlukan adalah 0.92. Dalam hal terjadi kecelakaan akibat kebakaran, hitunglah peluang kedua mobil tersebut tersedia dan siap digunakan?

Penyelesaian:

Misal: A = mobil pemadam kebakaran siap digunakan

B = ambulans siap digunakan

maka:

$P(A \cap B) = P(A)\ P(B) = (0.98)\ (0.92) = 0.9016$

Definisi Kaidah Penggandaan Umum

Jika dalam suatu percobaan kejadian-kejadian $A_1, A_2, \cdots, A_k$ dapat terjadi, maka:

$$P(A_1 \cap A_2 \cap A_3 \cap \cdots \cap A_k) = P(A_1)P(A_2|A_1)P(A_3|A_1 \cap A_2) \cdots P(A_k|A_1 \cap A_2 \cap \cdots \cap A_{k-1}) \quad (6.8)$$

Jika kejadian-kejadian $A_1, A_2, \cdots, A_k$ bebas, maka:

$$P(A_1 \cap A_2 \cap A_3 \cap \cdots \cap A_k) = P(A_1)\ P(A_2)\ P(A_3) \cdots P(A_k) \quad (6.9)$$

Contoh 3.12:

Tiga kartu diambil berturut-turut dan tanpa pengembalian. Tentukan peluang bahwa kartu yang terambil pertama adalah *ace* merah, yang kedua sepuluh atau *jack*, dan yang ketiga lebih besar dari 3 tetapi kurang dari 7.

Penyelesaian:

Misal : A_1 : kartu pertama adalah *ace* merah,

A_2 : kartu kedua adalah sepuluh atau *jack*,

A_3 : kartu ketiga adalah > 3 tetapi < 7

maka:

$P(A_1) = \dfrac{2}{52}$; $P(A_2 | A_1) = \dfrac{8}{51}$

$P(A_3 | A_1 \cap A_2) = \dfrac{12}{50}$

Jadi,

$$P(A_1 \cap A_2 \cap A_3) = P(A_1)\ P(A_2 \mid A_1) P(A_3 \mid A_1 \cap A_2)$$

$$= \left(\frac{2}{52}\right)\left(\frac{8}{51}\right)\left(\frac{12}{50}\right) = \frac{8}{5525}$$

Contoh 3.13:

Satu uang logam tidak setimbang sehingga peluang munculnya sisi gambar dua kali lebih besar daripada sisi angka. Bila uang itu dilemparkan 3 kali, berapa peluang mendapatkan 2 sisi angka dan 1 sisi gambar?

Penyelesaian:

Misalkan ruang contoh untuk pelemparan pertama adalah $S_1 = \{G, A\}$. Dengan memberikan peluang w pada sisi angka dan $2w$ pada sisi gambar, didapatkan $3w = 1$ atau $w = 1/3$ sehingga $P(G)=2/3$ dan $P(A)=1/3$. Jika B adalah kejadian mendapatkan 2 sisi angka dan 1 sisi gambar pada 3 kali pelemparan uang logam, maka $B = \{AAG, AGA, GAA\}$ dan karena hasil pada setiap pelemparan itu bebas, maka:

$$P(AAG) = P(AGA) = P(GAA)$$

$$= P(A \cap A \cap G) = P(A)P(A)P(G)$$

$$= \left(\frac{1}{3}\right)\left(\frac{1}{3}\right)\left(\frac{2}{3}\right) = \frac{2}{27}$$

Jadi $P(B) = \dfrac{2}{27} + \dfrac{2}{27} + \dfrac{2}{27} = \dfrac{2}{9}$

6.7 Dalil Peluang Total

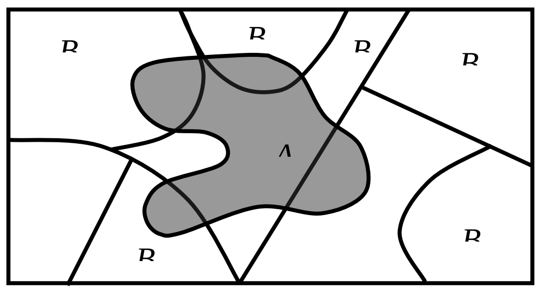
Bila kejadian-kejadian $(B_1, B_2, \cdots, B_k)$ dengan $P(B_i) \neq 0$, untuk $i = 1, 2, \cdots, k$, maka untuk setiap kejadian A yang merupakan himpunan bagian S berlaku:

$$P(A) = \sum_{i=1}^{k} P(B_i \cap A) = \sum_{i=1}^{k} P(B_i)\ P(A \mid B_i) \qquad (6.10)$$

Bukti:

Perhatikan diagram venn pada gambar 2. Kejadian A adalah union dari kejadian terpisah $B_1 \cap A,\ B_2 \cap A,\ \cdots,\ B_k \cap A$. Dengan menggunakan akibat 2 dari kaidah penjumlahan dan kaidah penggandaan, maka akan diperoleh:

$$
\begin{aligned}
P(A) &= [(B_1 \cap A) \cup (B_2 \cap A) \cup \cdots \cup (B_k \cap A)] \\
&= P(B_1 \cap A) + P(B_2 \cap A) + \cdots + P(B_k \cap A) \\
&= \sum_{i=1}^{k} P(B_i \cap A) = \sum_{i=1}^{k} P(B_i) P(A \mid B_i)
\end{aligned}
$$



Gambar 6.2. Penyekatan ruang contoh S.

Contoh 14:

Misal, S adalah ruang contoh yang menyatakan populasi sarjana di suatu kota dan diketahui sebagai berikut:

	Bekerja	Tidak bekerja	*Total*
Lelaki	460	40	500
Wanita	140	260	400
Total	600	300	900

Akan diambil secara acak seorang di antara mereka untuk ditugaskan ke kota tertentu. Bila M menyatakan lelaki yang terpilih dan E menyatakan orang yang terpilih dalam status bekerja, serta diketahui 36 orang dari status bekerja dan 12 yang tidak bekerja adalah anggota koperasi. Berapa peluang orang yang terpilih tersebut anggota koperasi?

Penyelesaian:

Misal M : lelaki yang terpilih

E : orang yang terpilih dalam status bekerja

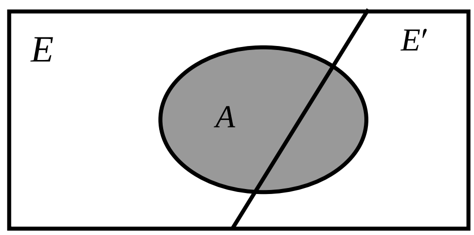
A : kejadian orang yang dipilih anggota koperasi.

Dengan melihat gambar berikut, A dapat ditulis sebagai gabungan dua kejadian yang terpisah, $E \cap A$ dan $E' \cap A$, sehingga:

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

dan menurut akibat 1 dari kaidah penjumlahan dan kaidah penggandaan maka ;

$P(A) = P[(E \cap A) + (E' \cap A)] = P(E \cap A) + P(E' \cap A) = P(E)P(A|E) + P(E')P(A|E')$



Gambar 6.3. Diagram Venn, memperlihatkan kejadian A, E, dan E′

Dari data yang ada, memungkinkan untuk menghitung:

$$P(E) = \frac{600}{900} = \frac{2}{3}, \quad dan \quad P(E') = \frac{1}{3}$$

$$P(A|E) = \frac{36}{600} = \frac{3}{50} \quad dan \quad P(A|E') = \frac{12}{300} = \frac{1}{25}$$

$$P(A) = \left(\frac{2}{3}\right)\left(\frac{3}{50}\right) + \left(\frac{1}{3}\right)\left(\frac{1}{25}\right) = \frac{4}{75}$$

6.9 Peluang Kejadian Bebas

Jika A dan B merupakan dua kejadian yang bebas sesamanya, maka peluang A dan B terjadi secara bersama-sama merupakan hasil kali peluang kejadian masing-masing.

Contoh 16:

Suatu permainan melempar dadu dan melempar sekeping uang logam dilakukan bersama-sama akan diperoleh ruang sampel sebanyak 12 (6 x 2) hasil yang mungkin. Dengan demikian peluang untuk memperoleh satu mata dadu misalkan mata dadu 5 dan sisi muka dalam satu kali pelemparan bersama-sama adalah:

$P(A \cap B) = P(A)\ P(B)$ (6.11)

P (5 dan muka) = 1/6 x ½ = 1/12

Contoh 17:

Peluang bahwa A akan hidup dalam waktu 50 tahun adalah 0,7 dan peluang bahwa B akan hidup dalam waktu 90 tahun adalah 0,1, maka peluang bahwa keduanya akan hidup dalam waktu 90 tahun adalah:

$P(A)P(B) = (0,7)(0,1) = 0,07$

Contoh 18:

Satu kotak berisi 20 buah apel, 8 apel berwarna hijau dan 12 apel berwarna merah. Diambil 3 apel secara bersamaan. Berapa peluang terambilnya 2 buah apel hijau dan 1 apel merah?

S = {MMM,MMH,MHM,HMM,MHH,HMH,HHM,HHH}

P(X = 2H1M) = 3 (8/20 x 8/20 x 12/20) = 2304/8000 = 0,288

6.10 Peluang Bersyarat

Peluang terjadinya kejadian B bila diketahui bahwa suatu kejadian lain A telah terjadi disebut peluang bersyarat (*Conditional probability*) dan dinyatakan oleh P(B/A). Suatu kejadian tidak bebas ini sama dengan pengambilan tanpa dikembalikan di mana biasa dinyatakan:

$P(B/A) = P(A \cap B)/P(A)\ ;\ P(A) > 0$ (6.12)

Bila dua kejadian A dan B bebas sesamanya, maka:

$P(A/B) = P(A)$ dan $P(B/A) = P(B)$

sehingga dapat dinyatakan:

$P(A \cap B) = P(A/B)\ P(B) = P(B/A)\ P(A) = P(A)\ P(B)$ (6.13)

Contoh 19:

Dalam satu kotak berisi 10 apel hijau dan 5 apel merah. Di mana dalam setiap kali pengambilan apel tersebut tidak dikembalikan. Berapa peluang terambilnya apel hijau jika pada pengambilan pertama apel merah sudah terpilih?

P(A) = peluang terambilnya apel merah = 5/15 = 1/3 maka peluang terambilnya apel hijau dengan syarat merah sudah terambil adalah:

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

$P(A \cap B) = P(B/A)\ P(A) = (10/14)\ (5/15 = 10/42$

Contoh 20:

Tabel mahasiswa yang mendaftar di PS Keperawatan STIKA

	Diterima (A)	Ditolak (A^c)	Total
Laki-Laki (B)	25	150	175
Perempuan (B^c)	40	200	240

Dari tabel di atas dapat menentukan:

$P(A \mid B) = P(A \cap B) / P(B) = (25/415) /(175/415) = 25/175 = 0,1428$

$P(A \mid B^c) = P(A \cap B^c) / P(B^c)= (40/415) / (240/415) = 40/240 = 0,167$

$P(A^c \mid B) = P(A^c \cap B) / P(B) = (150/415) / (175/415) = 150/175 = 0,857$

$P(A^c \mid B^c) = P(A^c \cap B^c) / P(B^c) = (200/415) / (240/415) = 200/240 = 0,833$

6.11. Distribusi Peluang Diskrit

Jika peubah X dapat menerima suatu himpunan diskrit dari nilai-nilai $X_1, X_2, \ldots, X_n$ dengan probabilitas masing-masing $p_1, p_2, \ldots, p_n$, di mana $p_1 + p_2 + \ldots + p_n = 1$, maka dapat dikatakan bahwa nilai tersebut merupakan suatu distribusi peluang diskrit.

Distribusi peluang (*probability distribution*) bagi X merupakan suatu daftar yang memuat nilai peluang bagi semua nilai variabel acak X yang mungkin terjadi. Distribusi peluang bagi variabel acak diskrit dapat disajikan dalam bentuk tabel, grafik atau rumus yang mengkaitkan nilai peluang dengan setiap nilia variabel acaknya.

Contoh tabel distribusi frekuensi relatif untuk hasil pelemparan 1 mata uang 4 kali:

X (hasil)	0	1	2	3	4
F	1	4	6	4	1
Frekuensi relatif P(X=x)	1/16	4/16	6/16	4/16	1/16

6.12 Distribusi Peluang Kontinu

Misalkan, sebuah peubah acak kontinu X mempunyai fungsi kepekatan f(x). jika X_0 merupakan nilai dari peubah acak X tersebut, maka:

P(X < X$_0$) = $\int_{-\infty}^{x_0} f(x)dx$ untuk -∞ < X < + ∞

Artinya, peluang (P) untuk X lebih kecil atau sama dengan X_0 adalah sama dengan luas daerah di bawah kurva f(x) di atas X_0 ke kiri. P(X < X_0) disebut dengan fungsi sebaran kumulatif atau fungsi sebaran peluang atau fungsi sebaran.

Beberapa sebaran kontinu yang memegang peranan penting dalam statistika:

1. Distribusi Uniform

Sebaran uniform, meskipun dalam penerapannya agak terbatas, namun mempunyai arti yang penting terutama sebagai penghampir sebaran-sebaran yang lain.

Sebuah peubah acak X yang tersebar secara uniform mempunyai fungsi kepekatan:

f(X) = $\frac{1}{b-a}$, untuk a < x < (6.15)

= 0, sebaliknya

dimana a dan b merupakan bilangan nyata (riil) tetap dengan syarat a < b.

Karena variabel acak tersebut secara uniform mempunyai kepekatan peluang yang konstan pada selang yang didefinisikan atau ditentukan, maka konstanta tersebut haruslah merupakan kebalikan dari panjang selang agar syaratnya dapat dipenuhi yaitu:

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

Contoh 21:

Suatu variabel acak X tersebar secara uniform pada selang 0 < x < 10. Berapa peluang nilai variabel X pada selang 4 < x < 9?

Penyelesaian:

P(4 < x < 9) = $\frac{9-4}{10-0}$ = $\frac{5}{10}$ = 0.5

Jadi peluang variabel acak X pada selang 0 < x < 9 adalah 0.5

2. Distribusi Eksponen

Satu variabel acak X yang menyebar menurut sebaran eksponen mempunyai fungsi kepekatan:

f(x) = $e^{-\lambda x}$, untuk x > 0

= 0, sebaliknya (6.15)

dimana parameter λ merupakan bilangan positif, nyata, dan konstan.

BAB VII

PENAKSIRAN (ESTIMASI) PARAMETER

7.1 PENAKSIRAN PARAMETER

Secara umum, parameter populasi akan diberi simbol θ (baca: theta). Jadi θ bisa merupakan rata-rata μ, simpamgan baku σ, proporsi π dan sebagainya. Jika θ, yang tidak diketahui harganya, ditaksr oleh harga $\hat{\theta}_1$ = θ, yaitu bisa menyatakan bahwa θ yang sebenarnya. Tetapi ini merupakan keinginan yang boleh dibilang ideal sifatnmya. Kenyataan yang bisa terjadi adalah:

a. menaksir θ oleh $\hat{\theta}$ terlalu tinggi, atau

b. menaksir θ oleh $\hat{\theta}$ terlalu rendah.

Beberapa definisi:

1) Penaksir $\hat{\theta}$ dikatakn *penaksir takbias* jika rata-rata semua harga $\hat{\theta}$ yang mungkin akan sama dengan θ. Dalam bahasa ekspektasi ditulis ε ($\hat{\theta}$) = 0. Penaksir yang tidak tak bias disebut *penaksit bias.*

2) Penaksir bervarians minimum ialah penaksir dengan variasi terkecil diantara semua penaksir untuk parameter yang sama. Jika $\hat{\theta}_1$ dan $\hat{\theta}_2$ dua penaksir untuk θ dimana varinasi untuk $\hat{\theta}_1$ lebih kecil dari variansi untuk $\hat{\theta}_2$, maka $\hat{\theta}_1$ merupakan penaksir betrvariasi minimum.

3) Misalkan $\hat{\theta}$ penaksir untuk θ yang dihitung berdasarkan sebuah sampel acak berukuran *n*. Jika ukuran sampel *n* makin besar mendekati ukuran populasi menyebabkan $\hat{\theta}$ mendekati θ, maka $\hat{\theta}$ disebut penaksir *konsistren.*

4) Penaksir yang takbias dan berfariasi minimum dinamakan *penaksir terbaik.*

(Sudjana, 2002).

1. Menaksir Rata-rata

Misalkan dipunyai populasi berukuran N dengan rata-rata μ dan simpangan baku σ. Dari populasi ini parameter rata-rata μ akan ditsaksir. Untuk keperluan ini, sampel sebuah sampel acak berukuran *n*, lalu hitung statistik yang perlu, ialah $\bar{x}$ dan s. Titik taksiran untuk rata-rata μ ialah $\bar{x}$. Dengan kata lain, nilai μ besarnya ditaksir oleh harga $\bar{X}$ yang didapat dari sampel.

Untuk itu dapat digunakan rumus:

$$P(\bar{X} - Z_{1/2\gamma}\frac{\sigma}{\sqrt{n}} < \mu < \bar{X} + Z_{1/2\gamma}\frac{\sigma}{\sqrt{n}}) = \gamma$$

Contoh :

Sebuah sampel acak terdiri dari 100 mahasiswa UHO telah diambil, lalu nilai-nilai IQ-nya dicatat. Didapat $\bar{X}$ = 122 dan s = 10.

a) Dikatakan: IQ rata-rata untuk mahasiswa UHO itu = 122 dalam hal ini titik taksiran telah digunakan.
b) Jika dikehendaki interval taksiran IQ rata-rata dengan koefisien kepercayaan 0,95, maka: $122 – (1{,}987)\frac{10}{\sqrt{100}} < \mu < 122 + (1{,}987)\frac{10}{\sqrt{100}}$ $atau\ 110{,}0 < \mu < 114{,}0$

Jadi didapat: 95% interval kepercayaan untuk IQ mahasiswa adalah 110,0 < μ <114,0. secara lain dapat dikatakan: kita merasa 95% yakin bahwa IQ rata-rata mahasiswa akan ada dalam interval dengan batas 110,0 dan 114,0 (Sudjana, 2002).

2. Menaksir simpangan baku

Penaksir tak bias dari varians σ^2. Akan tetapi simpangan baku s bukan penaksir takbias untuk simpangan baku σ titik taksiran s untuk σ adalah tak bias. Jika populasinya distribusi normal dengan variansi σ maka 100 γ % interval kepercayaan untuk σ^2 ditentukan dengan menggunakan distribusi chi-kuadrat dengan rumus:

$$\frac{(n-1)s^2}{X^2_{1/2(1+\gamma)}} < \sigma^2 < \frac{(n-1)s^2}{X^2_{1/2(1-\gamma)}}$$

Contoh:

Sebuah sampel acak berukuran 30 telah diambil dari sebuah populasi yang berdistribusi normal dengan simpangan baku σ. Dihasilkan dengan harga statistik s^2 = 7,8 dengan koefisioen kepercayaan 0,95 dan dk = 29, maka dari daftar Chi kuadrat didapat:

$X^2_{0,975} = 45,7 \; dan \; X^2_{0,025} = 16,0$ s

sehingga diperoleh:

$$\frac{29(7,8)}{45,7} \sigma^2 < \frac{29(7,8)}{16,0}$$

$atau \; 4,95 < \sigma^2 < 14,14$

Interval taksiran untuk simpangan baku σ adalah:

$2,23 < \sigma < 3,75$

dimana 95% percaya bahwa simpangan baku σ akan ada dalam interval yang dibatasi oleh 2,23 dan 3,75 (Sudjana, 2002).

BAB VIII
UJI HIPOTESIS

8.1 UJI HIPOTESIS

1. Uji Rata-rata μ: ***Uji Dua Pihak***

Misalkan sebuah populasi berdistrbusi normal dengan rata-rata μ dan simpangan baku σ akan diuji melalui parameter rata-rata μ.untuk ini, misalnya diambil sebuah sampel acak berukuran *n*, lalu dihitung statistik $\overline{X}$ dan s dapat dibedakan sebagai berkut:

$$\begin{cases} Ho: \mu = \mu o \\ H_1: \mu \neq \mu o \end{cases}$$ untuk pasangan hipotesis.

Dengan μ_0 sebuah harga yang diketahui, digunakan statistik:

$$\boxed{Z = \frac{\overline{X} - \mu o}{\sigma / \sqrt{n}}}$$

Statistik –z ini berdistribusi normal baku, sehingga untuk menentukan kriteria pengujian, digunakan daftar distribusi normal baku. H_0 ditolak jika $-z_{1/2(1-\alpha)} > z_{1/2(1-\alpha)}$ dengan $z_{1/2(1-\alpha)}$ didapat dari daftar normal baku dengan peluang ½(1 - α). *Contoh :* Pengusaha lampu pijar mengatakan bahwa lampunya bisa tahan pakai sekitar 800 jam.akhir-akhir ini timbul dugaan bahwa masa pakai lampu iti telah berubah. Untuk menentukan hal ini, dilakukan penelitian dngan menguji 50 lampu, ternyata rata-raanya 792 jam. Dari pengalaman, diketahui bahwa simpangan baku masa hidup lampu 60 jam diselidiki dengan taraf nyata 0,05 apakah kualitas lampuya itu saudah berubah atau belum.

Jawab: Dengan memisalkan masa hidup lampu berdistribusi normal, maka akan diuji:
H_0: μ = 800 jam, berarti lampu itu masa pakainya sekitar 800 jam dan

H_1: μ ≠ 800 jam, berarti kualitas lampu telah berubah dan bukan 800 jam lagi.

Dari pengalaman, simpangan baku σ = 60 jam (Sudjana, 2002).

2. Uji Rata-rata: Uji Satu Pihak

Perumusan yang umum untuk uji pihak kanan mengenai rata-rata μ berdarkan H_0 dean H_1 adalah:

$$\begin{Bmatrix} H_o : \mu = \mu_0 \\ H_1 : \mu \geq \mu_o \end{Bmatrix}$$

DImisalkan poopulasi berdistribusi normal dan dari padanya sebuah sampel acak berukuran ***n*** telah diambil dari sampel tersebut dihitung $\bar{x}$ dan ***s*** didapatkan hal-hal:

Jika simpangan baku σ untuk populasi diketahui, digunakan statistik z yang tertera dalam rumus (1) ialah menggunakan distribusi normal baku batas kriteria, tentunya didapat dari kriteria normal baku, tolak H_0 jika $z \geq z_{0,5 - \alpha}$ didapat dari daftar normal baku, menggunakan peluang (0,5 - α).

Contoh: Proses barang rata-rata menghasilkan 15,7 unit per jam.hasil produksi memnpunyai farians = 2,3. Metode baru diusulkan untuk mengganti yang lama jika rata-rata yang lama jika rata-rata per jam menghasilkjan paling seikit 16 buah. Untuk menentukan apakah metode diganti atau tidak, metode baru dicoba 20 kali dan ternyata rata-rata perjam menghasilkan 16,9 buah.

Pengusaha mengambil resiko 5 % untuk menggunakan metode baru apabila metode ini rata-rata menghasilkan 16 buah. Apakah keputusan sipengusaha?

Jawab: Misalkan hasil produksi berdistribusi normal, maka akan diuji pasangan hipotesis:

$H_0 : \mu = 16$, berarti rata-rata hasil metode baru paling tinggi 16. Jika ini terjadi, metode lama masih dipertahankan.

$H_1 : \mu > 16$, berarti rata-rata hasil metode baru lebih dari 16 dan karenanya metode lama dapat diganti (Sudjana, 2002).

3. Uji Simpangan Baku

Ketika menguji rata-rata μ untuk populasi normal, didapat hal dimana simpangan baku σ diketahui. Harga yang diketahui ini umumnya didapat dari pengalaman dan untuk menentukan besarnya perlu diadakan pengujian untuk ini, dimisalkan populasi berdistribusi normal dengan fariansi σ^2 dan daripadanya diambil sebuah sampel acak berukuran *n* variasni sampel.

Uji dua pihak:

Untuk ini, pasangan H_0 dan pasangan H_1 adalah:

$$\begin{cases} H_0 : \sigma^2 = \sigma_0^2 \\ H_1 : \sigma^2 \neq \sigma_0^2 \end{cases}$$

Untuk pengujian dipakai statistik chi-kuadrat: $x^2 = \dfrac{(n-1)\,s^2}{\sigma_O^2}$

Jika dalam pengujian dipakai taraf nyata α, maka kriteria pengujian adalah: terima H_0 jika $x^2_{½\alpha} < x^2 < x^2_{1-½\alpha}$ di mana $x^2_{½\alpha}$ dan $x^2_{1-½\alpha}$ didapat dari daftar distribusi chi kudrat dengan dk = (*n* – 1)dan masing-masing dengan peluang ½ σ dan (1 – ½ α), dalam hal lainya H_0 ditolak.

Contoh: Dalam bagian 4 bab ini terdapat contoh soal tentang masa hidup; lampu A, disitu diambil σ = 60 jam dengan sampel berukuran *n* = 50 didapat s = 55 jam. Jika masa lampu berdistribusi normal, benarkah σ = 60 jam dalam taraf α = 0,05.

Jawab: Untuk menyelidiki benar atau tidaknya tentang σ, maka dihadapkan dengan pengujian:

$$\begin{cases} Ho : \sigma^2 = 3.600\ jam \\ H_1 : \sigma^2 \neq 3.600\ jam \end{cases}$$

Dari rumus (2) dengan $n \neq 50$ dan s^2 = 3.025, maka:

$$x^2 = \frac{(50-1))(3.025)}{3.600} = 41{,}174$$

Dengan dk = 49 dan peluang 0,025 dan 0,975, daftar distribusi chi-kuadrat berturut turut didapat $x^2_{0,025} = 32{,}4\, dan\ x^2_{0,975} = 71{,}4$

Kriteria pengujian: terima H_0 jika x^2 antara 32,4 dan 71,4. Untuk harga-harga lainya, H_0 ditolak. Dari perhitungan didapat x^2 = 41,174 dan ini jatuh antara 32,4 dan 71,4; jadi dalam daerah penerimaan hipotesa. *Kesimpulan:* hipotesa σ = 60 jam diterima dengan menanggung risiko 5 % akan terjadinya penolakan hipotesa bahwa σ^2 = 3.600 jam.

Uji satu pihak:

Dalam kenyataanya sangat sering dikehendaki adanya variansi yang berharga kecil. Untuk ini diperlukan pengujian dan akan merupakan uji pihak kanan:

$$\begin{cases} H_0 : \sigma^2 = \sigma_0^2 \\ H_1 : \sigma^2 \geq \sigma_0^2 \end{cases}$$

Statistik yang digunakan masih tetap x^2 seperti dalam rumus (2). Kriteria pengujian dalam hal ini adalah: tolak H_0 jika $x^2_{hit} \geq x^2_{1-\alpha}$ dimana $x^2_{1-\alpha}$ didapat dari daftar chi-kuadrat dengan dk = (n – 1) dan peluang (1 - α) menyebabkan uji pihak kiri, yakni pasangan $\left\{\begin{matrix} H_0 : \sigma^2 = \sigma_0^2 \\ H_1 : \sigma^2 \leq \sigma_0^2 \end{matrix}\right\}$

Contoh: Proses pengisian sebuah minuman kedalam botol oleh mesin, paling tinggi mencapai farians 0,50 cc. Akhir-akhir ada digunakan isi botol variabillitas yang lebih besar. Diteliti 20 buah botol dan isinya ditakar ternyata sampel ini menghasilkan simpangan baku 0,90 cc dengan α = 0,05 perlukah mesin distel?

Jawab: Pengujian yang akan dilakukan adalah mengenai:

$$\begin{cases} H_0 : \sigma = 0{,}50 \\ H_1\ \sigma \geq 0{,}50 \end{cases}$$

Dengan s^2 = 0,81 dan $\boldsymbol{n}$ = 20 serta σ^2 = 0,50 dari rumus (2) didapat: $\chi^2 = \frac{(20-)(0{,}81)}{0{,}50} = 30{,}78$.

Dari daftar chi-kuadrat dengan dk = 19 dan peluang 0,95 diperoleh $x^2_{0,95} = 30{,}1$.

Karena chi-kuadrat dari penelitian lebih besar dari 30,1 maka H_0 ditolak pada taraf 5%. Ini berarti fariasi isi botol telah menjadi lebih besar, sehingga dianjurkan untuk menyetel kembali mesin agar mendapatkan pengisian yang lebih merata (Sudjana, 2002).

DAFTAR PUSTAKA

Algifari. 1997. Statistik Induktif untuk Ekonomi dan Bisnis. UPP AMP YKPN, Yogyakarta.

Bodiarto, E. 2002. Biostatistika untuk Kedokteran dan Kesehatan Masyarakat, Jakarta.

Boediono & Koster, W. 2002. Teori dan Aplikasi Statistika dan Probabilitas, Remaja Rosdakarya, Bandung.

Dajan, A. 1987. Pengantar Metode Statistik, Jilid I LP3ES, Jakarta.

Harini, S. & Kusumawati. 2007. Metode Statistika, Jakarta.

Hasan, I. 2002. Pokok-pokok materi Statistik-1 (Statistik Deskriptif) dan Statistik-2 (Statistik Inferensif), edisi kedua, Bumi Aksara, Jakarta.

Kusnandar, D. 2004. Metode Statistik dan Aplikasinya dengan Minitab dan Excel. Madyan Press, Yogyakarta.

Makkulau & Ampa, A.T. 2012. Statistika Dasar, Buku Ajar dalam Lingkungan sendiri, Kendari.

Larson, H.J. 1983. Introduction to the Theory of Statistics, Second edition, Mc. Graw-Hill Book Company, Inc. New York.

Steel, R.G.D. & Torrie, J.H. 1989. Prinsip dan Prosedur Statistika, Suatu Pendekatan Biometri 2nd ed, Penterjemah B. Sumantri, Penerbit PT. Gramedia, Jakarta.

Sudjana, 2002. Metode Statistik, Bandung.

Surjadi, P.A. 1990. Pendahuluan Teori Kemungkinan dan Statistika, cetakan keempat, ITB Bandung.

Tiro, M.A. 2004. Pengenalan Biostatistika, Makassar.

Usman, H. & Purnomo, S.A. 1995. Pengantar Statistika, Bumi Aksara, Jakarta.

Walpole, R. 1993. Pengantar Statistika, edisi ketiga, Gramedia, Jakarta.

Yitnosumarto, S. 1994. Dasar-dasar Statistika, Rajawali Press, Jakarta.