

Yusuf Nalim  
Salafudin

STAIN PRESS  
PEKALONGAN



# STATISTIKA Deskriptif



Yusuf Nalim  
Salafudin Turmudi

# STATISTIKA Deskriptif

## UNDANG-UNDANG HAK CIPTA NO. 19 TAHUN 2002

### Pasal 2

- (1). Hak Cipta merupakan hak eksklusif bagi Pencipta dan Pemegang Hak Cipta untuk mengumumkan atau memperbanyak ciptaannya, yang timbul secara otomatis setelah suatu ciptaan dilahirkan tanpa mengurangi pembatasan menurut peraturan perundang-undangan yang berlaku.

### Pasal 72

- (1). Barangsiapa dengan sengaja dan tanpa hak melakukan perbuatan sebagaimana dimaksud dalam Pasal 2 ayat (1) dipidana dengan pidana penjara paling singkat 1 (satu) bulan dan/atau denda paling sedikit Rp. 1.000.000,- (satu juta rupiah), atau pidana penjara paling lama 7 (tujuh) tahun dan denda paling banyak Rp. 5.000.000.000,- (lima miliar rupiah).
- (2). Barangsiapa dengan sengaja menyiarkan, memamerkan, mengedarkan, atau menjual kepada umum suatu ciptaan atau barang hasil pelanggaran Hak Cipta atau Hak Terkait sebagai mana dimaksud pada Pasal 2 ayat (1), dipidana dengan pidana penjara paling lama 5 (lima) tahun dan/atau denda paling banyak Rp. 500.000.000,- (lima ratus juta rupiah).

# STATISTIKA DESKRIPTIF

**Penulis:**

Yusuf Nalim  
Salafudin Turmudi

**Editor & Layouter:**

Abu Fahmi

**Desain Cover:**

Gus Fadhel

Diterbitkan oleh:

**STAIN Pekalongan Press**

Jl. Kusumabangsa No. 9 Pekalongan 51114  
Telp. [0285] 412575, Fax. [0285] 423418  
e-mail: p3mstainpkl@yahoo.co.id

ISBN: 978-979-3968-60-5

Hak Cipta dilindungi Undang-Undang

No. 19 Th. 2002

*All right reserved*

Cetakan pertama, Desember 2012

## KATA PENGANTAR

Segala puji dan syukur kami haturkan kehadiran Allah SWT, karena atas pertolongan-Nya dan segala nikmat yang telah diberikan sehingga kami berdua (penulis) bisa menyelesaikan penyusunan buku yang berjudul "Statistika Deskriptif". Buku ini hadir ke hadapan pembaca dengan beberapa materi yang dikemas secara jelas dan lugas disertai dengan contoh-contoh yang bersifat membimbing serta cuplikan data yang sebisa mungkin diambil dari kondisi riil di lapangan dan bersifat *up to date* dengan bantuan pencarian materi baru melalui Internet.

Buku ini sengaja kami susun berdua – karena diawali dari diskusi yang cukup intens sebelumnya – dengan tujuan untuk mengkolaborasikan berbagai macam pengetahuan, informasi dan pengalaman kami (baik sebagai pengajar maupun peneliti) agar bisa tersampaikan melalui tulisan yang dibukukan. Beberapa contoh, *statement*, dan informasi yang dikemukakan dalam buku ini juga merupakan bagian dari hasil penelitian penulis.

Kandungan materi dalam buku ini sebenarnya merupakan materi dasar dari ilmu statistika. Kajian dalam statistika deskriptif hanya bertujuan memberikan gambaran ringkas (ringkasan) tentang data (populasi/sampel) dan pengamatan yang telah dilakukan terhadap suatu peristiwa. Lingkup kajian utama dari statistika deskriptif adalah bagaimana melakukan proses pengumpulan data dalam penelitian, pengolahan data, penyajian data dalam bentuk tabel, grafik serta, mengetahui karakteristik khas dari masing-masing data dalam bentuk rata-rata, median, modus, simpangan, kuartil dan lain-lain.

Untuk mengetahui perkembangan suatu peristiwa kuantitatif maka diperlukan pengetahuan mengenai angka indeks.

Pembahasan buku ini meliputi 12 Bab, dimulai dari Konsep Dasar Statistika (Bab 1) dan diakhiri dengan ulasan mengenai Angka Indeks (Bab 12). Penulis sengaja membagi beberapa materi bahasan dalam bab tersendiri dengan harapan agar penjelasannya lebih rinci dan mudah dipahami. Seperti materi tentang Populasi dan Sampel (Bab 2) serta materi tentang Data (Bab 3). Dalam banyak literatur, materi tentang populasi, sampel dan data seringkali hanya dijadikan subbab.

Semoga dengan kehadiran buku ini bisa menjadikan nilai tambah bagi khazanah keilmuan statistika dan sekaligus bisa menjadi rujukan bagi pembaca dari semua kalangan (mahasiswa, pengajar, praktisi, peneliti dan lain-lain) yang ingin mengetahui lebih dalam tentang dasar-dasar statistika.

Sungguhpun kami telah bekerja sebaik mungkin, namun namanya manusia pasti tidak lepas dari salah dan lupa. Sebagaimana juga tentu buku ini masih jauh dari kesempurnaan. Untuk itu kritik dan saran yang konstruktif demi perbaikan kedepan senantiasa penulis harapkan.

Ucapan terimakasih juga tidak lupa kami sampaikan kepada semua pihak yang telah membantu penulis. Semoga Allah membalasnya dengan kebaikan yang berlipat. Amin

Pekalongan, 13-11-2012|11.23

Penulis

## DAFTAR ISI

KATA PENGANTAR \_\_ v

DAFTAR ISI \_\_ vii

### I. KONSEP DASAR STATISTIKA

- A. Sejarah dan Perkembangan Statistika \_\_ 1
- B. Pengertian Statistik dan Statistika \_\_ 5
- C. Penggolongan Statistika \_\_ 10
- D. Pemilihan Metode Analisis (Parametrik atau Nonparametrik) \_\_ 13
- E. Ciri Khas Statistika \_\_ 15
- F. Fungsi dan Kegunaan Statistika \_\_ 16

### II. POPULASI DAN SAMPEL

- A. Pendahuluan \_\_ 31
- B. Populasi \_\_ 32
- C. Sampel \_\_ 34
- D. Mengapa Sampel? \_\_ 36
- E. Cara Menentukan Jumlah Sampel \_\_ 38

### III. BERBAGAI HAL TENTANG DATA

- A. Pendahuluan \_\_ 41
- B. Sumber Data Statistik \_\_ 41
- C. Penggolongan Data \_\_ 42
- D. Validitas Data Statistik \_\_ 44
- E. Variabel \_\_ 47
- F. Skala Pengukuran \_\_ 49
- G. Angka Eksak dan Angka Pendekatan (Aproksimatif) \_\_ 57

- H. Angka Pendekatan sebagai Interval \_\_ 58
- I. Pembulatan Bilangan \_\_ 58

#### IV. PENGUMPULAN DATA

- A. Pendahuluan \_\_ 61
- B. Tahapan Pengumpulan Data \_\_ 61
- C. Cara Mengumpulkan Data \_\_ 62
- D. Teknik Pengumpulan Data \_\_ 64

#### V. PENYAJIAN DATA

- A. Pendahuluan \_\_ 79
- B. Penyajian Data dalam Bentuk Tabel (Daftar) \_\_ 79
- C. Penyajian Data dalam Bentuk Diagram/Grafik \_\_ 85

#### VI. TABEL DISTRIBUSI FREKUENSI

- A. Pendahuluan \_\_ 91
- B. Pengertian Tabel Distribusi Frekuensi \_\_ 92
- C. Tabel Distribusi Frekuensi Tunggal \_\_ 94
- D. Tabel Distribusi Frekuensi Kelompok \_\_ 95
- E. Istilah-istilah dalam Tabel Distribusi Frekuensi Kelompok \_\_ 97
- F. Cara Membuat Tabel Distribusi Frekuensi Kelompok \_\_ 99
- G. Tabel Distribusi Frekuensi Kumulatif \_\_ 102
- H. Tabel Distribusi Relatif \_\_ 103
- I. Grafik Distribusi Frekuensi 104

#### VII. UKURAN PEMUSATAN DATA

- A. Pendahuluan \_\_ 109
- B. Modus (Mode) \_\_ 111
- C. Median \_\_ 114
- D. Rata-rata Hitung (*arithmetic mean*) \_\_ 116

- E. Rata-rata Gabungan (weighted mean) \_\_ 122
- F. Bentuk Distribusi Data (Hubungan Mean, Median dan Modus) \_\_ 126

#### VIII. UKURAN LETAK

- A. Kuartil \_\_ 129
- B. Desil \_\_ 132
- C. Persentil \_\_ 134

#### IX. UKURAN PENYEBARAN DATA

- A. Pendahuluan \_\_ 137
- B. Jangkauan (*Range*) \_\_ 140
- C. Simpangan Kuartil (*Interquartile Range*) \_\_ 142
- D. Simpangan Persentil \_\_ 143
- E. Simpangan rata-rata \_\_ 143
- F. Variansi (*Variance*) \_\_ 145
- G. Simpangan Baku (Standar Deviasi) \_\_ 147
- H. Koefisien Variasi \_\_ 150

#### X. SKEWNESS DAN KURTOSIS

- A. Pendahuluan \_\_ 155
- B. Skewness \_\_ 156
- C. Menghitung Indeks Kemiringan (Skewness Index) \_\_ 157
- D. Kurtosis (keruncingan) \_\_ 162
- E. Perhitungan Kurtosis \_\_ 162

#### XI. TEKNIK SAMPLING

- A. Pendahuluan \_\_ 165
- B. Teknik Penentuan Ukuran Sampel \_\_ 168
- C. Teknik Pengambilan Sampel \_\_ 172

#### XII. ANGKA INDEKS

- A. Pendahuluan \_\_ 183

B. Komponen Penyusunan Angka Indeks \_\_ 184  
C. Indeks Harga Relatif Sederhana dan Agregatif \_\_ 184

DAFTAR PUSTAKA \_\_ 189

BIOGRAFI SINGKAT PENULIS \_\_ 191

*Bab 1*

KONSEP DASAR STATISTIKA

**A. Sejarah dan Perkembangan Statistika**

Kata statistika berawal dari bahasa Latin modern *statisticum collegium* (dewan negara) dan bahasa Italia *statista* yang berarti "negarawan" atau "politikus". Dalam bahasa Romawi statistika berasal dari kata *states* yang berarti negara atau negarawan. Penggunaan istilah statistika pertama kali ketika Gottfried Achenwall<sup>1</sup> (1719–1772) menggunakannya dalam bahasa Jerman pada kegiatan analisis data kenegaraan, dengan mengartikannya sebagai "ilmu tentang negara (*state*)". Kemudian Dr. E.A.W. Zimmermann memperkenalkan istilah statistika di Inggris dan selanjutnya dipopulerkan oleh Sir John Sinclair<sup>2</sup> di Statistika Akuntansi Skotlandia (1791–1799).

Pada awalnya statistika hanya digunakan untuk mengurus data yang dipakai lembaga-lembaga administratif dan pemerintahan yang hanya berisi data jumlah penduduk menurut umur, jenis kelamin dan pekerjaan yang ditunjukkan

---

<sup>1</sup> Gottfried Achenwall (20 Oktober 1719 - 1 Mei 1772) merupakan seorang filsuf Jerman, sejarawan, ekonom, dan statistisi. Dia diakui sebagai salah satu diantara para penemu statistika.

<sup>2</sup> Sir John Sinclair merupakan salah satu politikus Skotlandia. Beliau juga merupakan seorang penulis buku mengenai keuangan dan pertanian. Menurut [en.wikipedia.org](http://en.wikipedia.org) beliau adalah orang pertama yang menggunakan kata statistika dalam bahasa Inggris dalam karyanya yang terkenal: *Statistical Account of Scotland*, yang terbit dalam 21 jilid.

untuk penarikan pajak dan wajib militer. Pengumpulan data terus berlanjut, khususnya melalui sensus yang dilakukan secara teratur untuk memberi informasi kependudukan yang berubah setiap saat.

Pada tahun 1733, de Moivre mempublikasikan penemuan kurva normal (*normal curve*). Kurva ini telah terbukti sangat penting dalam pengembangan statistika. Namun saat itu De Moivre sama sekali tidak tahu bagaimana menerapkan penemuannya tersebut pada data hasil percobaan, dan karyanya ini tetap tidak diketahui sampai Karl Pearson menemukannya di suatu perpustakaan pada tahun 1924. Walaupun demikian, hasil yang sama dikembangkan kemudian oleh dua astronom matematik, Laplace, 1749-1855 dan Gauss, 1777-1855, secara terpisah.

Pada abad ke-19 Charles Lyell telah mengemukakan argumentasi yang bersifat statistik terhadap suatu masalah geologi. Dalam periode 1830-1833, diterbitkan 3 jilid *Principles of Geology* karya Lyell, yang mengurutkan batuan zaman Tertier, serta sekaligus memberi nama pada masing-masing batuan. Bersama dengan M. Deshayesn (seorang ahli biologi dari Prancis), mereka mengidentifikasi dan mendaftarkan spesies-spesies fosil yang terdapat dalam satu atau lebih strata, dan meramalkan proporsi jenis-jenis yang masih hidup di bagian-bagian laut tertentu. Berdasarkan proporsi-proporsi tersebut mereka memberi nama Pleistosen, Pliosen, Miosen, dan Eosen. Argumentasi Lyell sesungguhnya bersifat statistika. Sayangnya setelah ditetapkan dan diterimanya nama-nama tersebut, metodenya segera dilupakan orang. Hal ini terjadi baik di bidang ilmu-ilmu biologi maupun fisika.

Abad Pada awal abad 20, teori peluang mulai banyak digunakan dalam pengembangan statistika. Perkembangan statistika inferensial (salah satu cabang statistika yang pada saat ini sangat luas digunakan untuk mendukung metode ilmiah) diawali oleh Ronald Fisher (peletak dasar statistika inferensial), Karl Pearson (metode regresi linear), dan William Sealey Gosset dengan nama samaran "*student*" (meneliti problem sampel berukuran kecil).

Saat ini penggunaan statistika telah menyentuh semua bidang ilmu pengetahuan, mulai dari bidang eksak (astronomi, geologi, teknik) hingga bidang sosial (psikologi, komunikasi, linguistik). Metodologi statistika banyak mempengaruhi perkembangan bidang-bidang ekonomi, biologi dan cabang-cabang terapan lainnya. Akibatnya lahir ilmu-ilmu gabungan seperti ekonometrika (ekonomi dan statistika), biometrika (biologi dan statistika), dan psikometrika (psikologi dan statistika).

Di negara-negara maju, statistika telah sejak lama berkembang dengan pesat. Bahkan kemajuan suatu bangsa sangat ditentukan oleh sejauh mana negara itu dapat menerapkan ilmu statistika dalam memecahkan masalah-masalah pembangunan dan perencanaan pembangunan. Di Indonesia sekarang ini sudah memiliki lembaga resmi milik pemerintah yang mengurus tentang data-data statistik yaitu BPS (Badan Pusat Statistik).

Perkembangan statistika di dunia tergolong cukup pesat. Hal ini bisa dilihat dari karya-karya para pengembangannya. Tabel 1.1 berikut akan memberikan gambaran secara rinci dan runtut tentang perkembangan statistika.

Tabel 1.1.  
Perkembangan Statistika di Eropa

Tahun	Peristiwa	Tokoh
1532	Data pertama tentang kematian (dalam mingguan) di London	Sir W. Petty
1539	Awal dari pengumpulan data pada pembaptisan, pernikahan, dan kematian di Perancis	
1608	Permulaan pendaftaran jemaah Gereja di Swedia	
1662	Pertama kali diterbitkan studi demografi berdasarkan angka kematian	J. Graunt
1693	Publikasi "Estimasi tingkat kematian manusia yang diambil dari tabel kelahiran dan kematian di kota Breslaw untuk menentukan angka harapan hidup"	E. Halley
1713	Publikasi <i>Ars Conjectandi</i>	J. Bernoulli
1714	Publikasi <i>Libellus de Ratiocinus in Ludo Aleae</i>	C. Huygens
1714	Publikasi <i>The Doctrine of Chances</i>	A. De Moivre
1735	Awal pengumpulan data demografi di Norwegia	
1763	Publikasi "An essay towards solving a problem in the Doctrine of Chances"	Rev. Bayes
1809	Publikasi "Theoria Motus Corporum Coelestium"	C.F. Gauss
1812	Publikasi "Théorie analytique des probabilités"	P.S. Laplace
1834	Pendirian "The Statistical Society" di London	
1839	Pendirian "The American Statistical Association" (Boston)	
1889	Publikasi <i>Natural Inheritance</i>	F. Galton
1900	Pengembangan uji <i>Chi square</i>	K. Pearson
1901	Publikasi isu awal tentang	F. Galton et

	Biometrika	al.
1903	Pengembangan <i>Principal Component Analysis</i> (analisis komponen utama)	K. Pearson
1908	Publikasi "The probable error of a mean"	"Student"
1910	Publikasi "An introduction to the theory of statistics"	G.U. Yule
1933	Publikasi "On the empirical determination of a distribution"	A.N. Kolmogorov
1935	Publikasi "The Design of Experiments"	R.A. Fisher
1936	Publikasi "Relations between two sets of variables"	H. Hotelling
1972	Publikasi "Regression models and life tables"	D.R. Cox
1972	Publikasi "Generalized linear models"	J.A. Nelder and R.W.M. Wedderburn
1979	Publikasi "Bootstrap methods: another look at the jackknife"	B. Efron

Sumber: <http://www.uib.no/People/ngbnk/kurs/notes/node4.html>

## B. Pengertian Statistik dan Statistika

Kita mungkin lebih sering mendengar orang mengatakan statistik daripada statistika. Sebab secara pengucapan memang lebih mudah dan cepat mengatakan "statistik". Memang sebagian orang tidak terlalu mempedulikan apakah antara statistik dan statistika berbeda atau tidak. Namun, untuk meluruskan pemahaman, perlu diketahui betul bahwa statistik dan statistika merupakan dua hal yang sangat berbeda. Statistik berasal dari bahasa Inggris "statistic" yang memiliki beberapa pengertian.

Jika dikaji secara rinci, paling tidak ada tiga pengertian statistik. Pengertian statistik yang paling sederhana bisa diartikan sebagai *data*.

Contoh 1.1 (Statistik)

- a. Nilai Ekspor Indonesia bulan Agustus 2011 mencapai US\$18,81 miliar sedangkan impor sebesar US\$ 15,05 miliar (www.bps.go.id)
- b. Jumlah wisatawan mancanegara bulan Agustus 2011 naik 5,89% dan jumlah penumpang angkutan udara domestik Agustus 2011 turun 24,9% (www.bps.go.id)
- c. Aset industri perbankan syariah per September 2011 telah mencapai Rp. 126 triliun tumbuh 46,55% dibandingkan dengan periode yang sama tahun sebelumnya Rp. 85,97 triliun (www.bisnis.com)

Pengertian statistik yang kedua adalah kumpulan data dalam bentuk angka maupun bukan angka yang disusun dalam tabel (daftar), sering pula disertai dengan diagram (grafik) dan keterangan lain yang menggambarkan suatu persoalan.

Contoh 1.2

- a. Daftar Mahasiswa Asing Jurusan Administrasi Pendidikan UPI tahun 2007

Nama Mahasiswa	Angkatan	Asal Negara	Keterangan
Gayos Aliong	1999	Malaysia	Sudah lulus
Affendy Abdul Sani	2000	Malaysia	Cuti
Jehrobiah Jehdorloh	2000	Thailand	Aktif (pernah cuti selama satu tahun)
Sharulizam Othman	2000	Malaysia	Sudah lulus
Pressely Mambang	2000	Malaysia	Sudah lulus

Sumber: www.imagegoogle.co.id

- b. Harga sembako bulan Januari 2009 di Kab. Kepulauan Mentawai

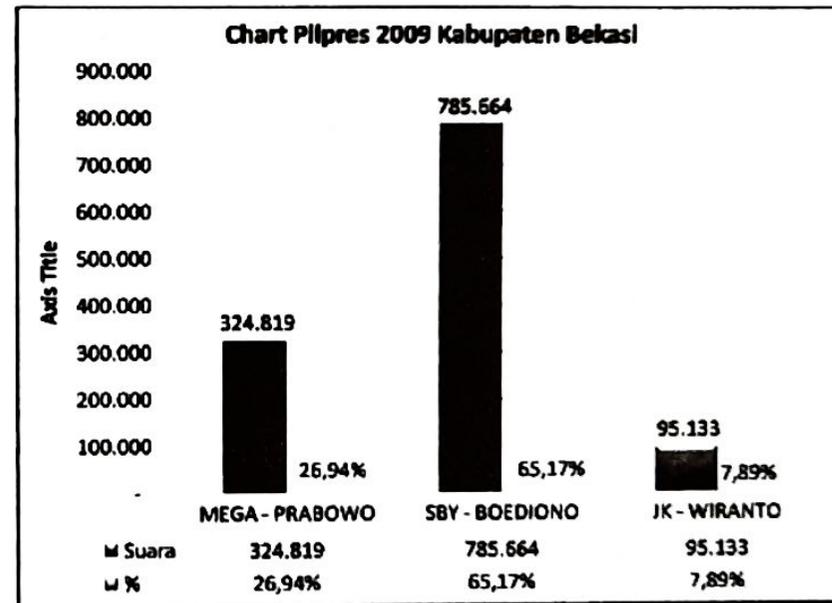
**Harga Sembako**  
**Kab. Kep. Mentawai**

Barang	Satuan	Harga Belanja	Harga Jual
Cabe Merah	1 Kg	Rp. 25.000	Rp. 30.000
Kel	1 Kg	Rp. 10.000	Rp. 12.000
Bawang Merah	1 Kg	Rp. 10.000	Rp. 12.000
Minyak Goreng	1 Kg	Rp. 7.500	Rp. 9.000
Telur	1 Kg	Rp. 2.500	Rp. 3.000
Telur	1 Kg	Rp. 2.500	Rp. 3.000
Gula Pasir	1 Kg	Rp. 6.500	Rp. 7.500
Manis	1 Kg	Rp. 14.000	Rp. 15.000
Tomat	1 Kg	Rp. 6.000	Rp. 8.000
Kentang	1 Kg	Rp. 6.000	Rp. 8.000

Sumber: www.imagesgoogle.co.id

grafik: etnis-sunda.com

- c. Hasil pemilu pilpres 2009 di Kabupaten Bekasi



Sumber: www.imagesgoogle.co.id

d. Hasil Undian Piala Dunia Afrika Selatan 2010

Groups			
<b>Group A</b> South Africa Mexico Uruguay France	<b>Group B</b> Argentina Nigeria Korea Republic Greece	<b>Group C</b> England USA Algeria Slovenia	<b>Group D</b> Germany Australia Serbia Ghana
<b>Group E</b> Netherlands Denmark Japan Cameroon	<b>Group F</b> Italy Paraguay New Zealand Slovakia	<b>Group G</b> Brazil Korea DPR Côte d'Ivoire Portugal	<b>Group H</b> Spain Switzerland Honduras Chile

Sumber: www.images.google.co.id

e. Diagram peta (kartogram) juga merupakan contoh "statistik"



Sumber: [http://media-cache-lt0.pinterest.com/upload/-311874342914962845\\_aOuruoT8.jpg](http://media-cache-lt0.pinterest.com/upload/-311874342914962845_aOuruoT8.jpg)

Sedangkan pengertian statistik yang ketiga adalah informasi (ukuran/nilai) yang diperoleh dari sampel dan digunakan untuk menduga sesuatu pada populasi.

Contoh 1.3

- Rata-rata pendapatan orang tua sebagian mahasiswa Rp. 2000.000,00 perbulan yang akan digunakan untuk menduga rata-rata pendapatan orang tua seluruh mahasiswa
- Nilai tengah harga beras berada pada posisi Rp. 6000,00 pada tahun 2009
- Nilai tukar rupiah terhadap dollar pada bulan Desember 2009 yang paling sering terjadi (modus) adalah sebesar Rp. 9.300,00

Adapun statistika berasal dari bahasa Inggris "statistics". Dalam buku *Statistics for Business and Economics, Eleventh Edition* yang ditulis oleh David R. Anderson, Dennis J. Sweeney, dan Thomas A. Williams disebutkan bahwa:

"Statistics is defined as the art and science of collecting, analyzing, presenting, and interpreting data. Particularly in business and economics, the information provided by collecting, analyzing, presenting, and interpreting data gives managers and decision makers a better understanding of the business and economic environment and thus enables them to make more informed and better decisions. In this text, we emphasize the use of statistics for business and economic decision making (Anderson et. al, 2011)"

Statistika didefinisikan sebagai seni dan ilmu yang mempelajari cara-cara pengumpulan, pengolahan, penyajian, penganalisaan data, dan pembuatan kesimpulan serta pengujian terhadap kesimpulan yang diperoleh.

Sedangkan statistika dalam bidang ekonomi dan bisnis adalah informasi yang diperoleh dengan cara mengumpulkan,

menyajikan, menganalisis, dan menginterpretasikan data yang akan menjadikan para manajer dan pembuat keputusan memiliki pemahaman yang lebih baik tentang lingkungan bisnis dan ekonomi sehingga mampu membuat keputusan yang lebih baik.

Secara singkat, statistika dapat diartikan sebagai *pengetahuan yang berhubungan dengan statistik*; atau statistika adalah *ilmu yang mempelajari statistik*. Sehingga kita mengenal ilmu statistika, bukan ilmu statistik. Saat ini kita mengenal mata kuliah statistika ekonomi, bukan mata kuliah statistik ekonomi dan lain-lain. Dengan demikian pengertian statistika jauh lebih luas daripada statistik.

Tujuan dari statistika adalah untuk mengembangkan dan menerapkan metodologi untuk mengekstraksi pengetahuan yang berguna dari percobaan dan data. Selain peran fundamental dalam analisis data, penalaran statistik juga sangat berguna dalam pengumpulan data (desain eksperimen dan survei) dan juga dalam membimbing inferensi ilmiah yang benar (Fisher, 1990).

Adapun kegiatan utama dalam statistika meliputi: desain eksperimen dan survei untuk menguji hipotesis, eksplorasi dan visualisasi data sampel, ringkasan deskripsi data sampel, pemodelan stokastik ketidakpastian (misalnya melempar koin), peramalan didasarkan pada model yang sesuai, pengujian hipotesis dan inferensi statistika, pengembangan teori dan metode statistik baru.

### C. Penggolongan Statistika

Pada sisi lain, statistika dapat pula dimaknai sebagai alat analisis dan alat untuk membuat keputusan. Sebagai alat analisis, ilmu statistika meliputi dua kelompok bidang kerja,

yaitu statistika deskriptif dan statistika inferensial yang masing-masing berkaitan dengan model analisis yang berbeda.

#### a. Statistika Deskriptif vs Inferensial

Statistika deskriptif adalah statistika yang digunakan untuk menggambarkan atau menganalisis data hasil penelitian tetapi tidak untuk mengambil kesimpulan yang lebih luas terhadap ciri-ciri populasi (generalisasi/inferensi). Ruang lingkup dari statistika deskriptif meliputi: konsep dasar statistika, distribusi frekuensi, pengukuran nilai pusat (*central tendency*), pengukuran penyebaran (*dispersion*), kemiringan (*skewness*) dan keruncingan (*kurtosis*), penyajian data dalam bentuk diagram grafik (diagram batang, diagram garis, batang histogram, *polygon*, *ogive*), angka indeks, dan *time series* atau deret waktu. Secara garis besar kegiatan statistik deskriptif meliputi:

- a. *Mengumpulkan data*: di dalamnya termasuk melakukan kegiatan *interview* atau wawancara kepada pihak-pihak yang diperlukan informasinya, membagikan kuesioner kepada responden, melakukan pengukuran terhadap variabel penelitian dengan menggunakan alat ukur (seperti mengukur tinggi badan, kadar gula, kecepatan reaksi, curah hujan dan lain-lain), melakukan pengamatan terhadap obyek penelitian (misalkan mengamati sikap konsumen, perilaku siswa dll).
- b. *Mengolah data*: pengolahan data merupakan kegiatan meringkas data baik dengan cara mengurutkan, membulatkan data, menyortir data, mengelompokkan, maupun memberi skor/kode (*koding*).

- c. *Menyajikan data*: data bisa disajikan dalam bentuk tulisan, tabel/daftar, grafik/diagram ataupun gambar yang menarik. Pembahasan mengenai penyajian data secara lengkap akan dijelaskan pada bab selanjutnya.
- d. *Analisis sederhana*: analisis di sini berupa analisis data sederhana yakni mencari ukuran-ukuran atau nilai-nilai yang bisa menggambarkan keadaan suatu kelompok data secara cepat, misalnya mencari rata-rata (*mean*), modus, median, variansi (ragam) atau standar deviasi (simpangan baku) tanpa disertai tindak lanjut, seperti penyimpulan dan pengujian terhadap suatu hipotesis.

Statistika inferensial bertujuan untuk melakukan pengujian dan menarik kesimpulan tentang ciri-ciri populasi yang dinyatakan dengan parameter populasi melalui perhitungan ciri-ciri yang ada pada sampel. Di dalamnya berisi estimasi parameter, uji hipotesis, prediksi dan perhitungan derajat asosiasi antar variabel. Adapun ruang lingkup dari statistika inferensial meliputi: peluang (probabilitas), distribusi data, estimasi parameter, uji hipotesis, analisis variansi, analisis regresi, dan analisis korelasi.

#### b. Statistika Parametrik vs Nonparametrik

Statistika parametrik adalah salah satu cabang analisis statistika yang digunakan untuk data-data yang memiliki distribusi (sebaran) normal. Jika data tidak menyebar normal maka analisisnya menggunakan statistik nonparametrik. Sebenarnya, meskipun data yang akan dianalisis tidak berdistribusi normal, kita masih memiliki peluang untuk tetap melakukan analisis parametrik dengan cara mentransformasi data terlebih dahulu.. Transformasi data perlu dilakukan agar data mengikuti sebaran normal. Transformasi dapat

dilakukan dengan mengubah data ke dalam bentuk logaritma natural, eksponensial, menggunakan operasi matematika (membagi, menambah, atau mengali dengan bilangan tertentu), dan mengubah skala data dari nominal menjadi interval. Statistika parametrik lebih banyak digunakan untuk menganalisis data yang berskala interval dan rasio dengan dilandasi asumsi tertentu seperti normalitas dan homogenitas varians. Contoh metode statistika parametrik diantaranya adalah analisis regresi linier sederhana maupun berganda, regresi logistik, uji-z (1 atau 2 sampel), uji-t (1 atau 2 sampel), korelasi pearson (*product moment*), rancangan percobaan (analisis varian dua arah), dan lain-lain.

Analisis statistik yang tidak memerlukan adanya syarat bentuk distribusi parameter populasi disebut statistika non-parametrik. Statistika nonparametrik disebut juga statistika bebas distribusi (*free distribution*). Statistika non-parametrik dapat digunakan untuk menganalisis data yang berskala nominal dan ordinal. Contoh metode statistik nonparametrik diantaranya adalah Chi-square test, Median test, Friedman test, korelasi rank Spearman, korelasi Kendall, dan lain-lain.

#### D. Pemilihan Metode Analisis (parametrik atau nonparametrik)

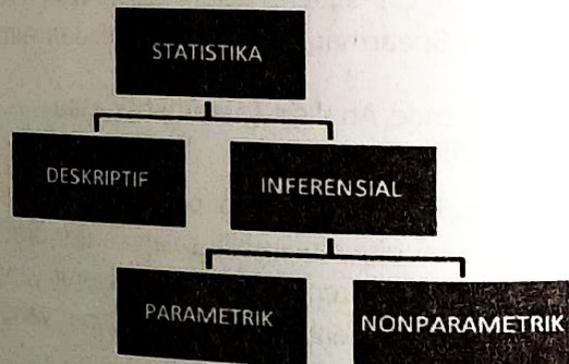
Kapan metode statistik nonparametrik digunakan? Metode pengujian ini digunakan bila salah satu syarat dalam statistik parametrik tidak terpenuhi. Syarat-syarat yang perlu diperhatikan untuk menentukan statistik apa yang akan digunakan dalam analisis, yaitu:

1. Jika distribusi data tidak diketahui maka statistik yang sesuai adalah statistik nonparametrik. Jika distribusi data

diketahui, maka kita harus melihat jenis distribusi data tersebut.

2. Jika data tidak berdistribusi normal, maka statistik yang sesuai adalah statistik nonparametrik. Jika data berdistribusi normal, maka statistik yang sesuai adalah statistik parametrik.
3. Jika sampel tidak ditarik secara random, maka statistik yang sesuai adalah statistik nonparametrik. Jika sampel ditarik secara random, maka statistik yang sesuai adalah statistik parametrik.
4. Jika varians kelompok tidak sama, maka statistik yang sesuai adalah statistik nonparametrik. Jika varians kelompok sama, maka statistik yang sesuai adalah statistik parametrik.

Jika skala pengukuran data nominal dan ordinal, maka statistik yang sesuai adalah statistik nonparametrik. Jika skala pengukuran data interval dan rasio, maka statistik yang sesuai adalah statistik parametrik.



Gambar 1.1 Bagan Penggolongan Statistika

Meski secara garis besar, kegiatan statistika terbagi dua (deskriptif dan inferensial), namun antara keduanya tidak

dapat dipisahkan, karena tanpa statistika deskriptif maka statistika inferensial tidak bisa dijalankan. Karena pada umumnya statistika deskriptif senantiasa mengawali/mendahului tahapan statistika inferensial. Sebelum dilakukan penarikan kesimpulan mengenai suatu keadaan yang sedang diteliti, maka datanya harus diuraikan dulu dalam statistika deskriptif sehingga diperoleh kesimpulan yang akurat guna memperoleh manfaat secara maksimal. Jadi, antara statistika deskriptif dan statistika inferensial diibaratkan seperti mata uang logam yang mempunyai dua sisi dimana antara dua sisi tersebut tidak bisa saling dipisahkan.

### E. Ciri Khas Statistika

Statistika merupakan salah satu bagian ilmu pengetahuan yang mempunyai ciri khas. Ciri tersebut ada 3 yaitu *kuantitatif, obyektif dan universal*.

#### Kuantitatif

Statistika selalu bekerja dengan angka (bilangan). Maka untuk melaksanakan tugasnya statistika selalu memerlukan bahan keterangan yang bersifat kuantitatif. Jika statistika akan digunakan untuk menganalisis data kualitatif (yaitu data yang tidak berwujud angka/bilangan), maka data kualitatif tersebut terlebih dahulu harus dikuantitatifkan. Proses transformasi dari data kualitatif menjadi data kuantitatif disebut proses *kuantifikasi*. Proses tersebut dilakukan dengan memberi skor kepada nilai kualitatif.

#### Obyektif

Statistika bersifat obyektif, artinya statistika bekerja menurut obyeknya, atau bekerja menurut apa adanya. Kesimpulan dan ramalan yang dihasilkan dari proses statistika,

semata-mata berdasarkan pada angka yang diolah, bukan didasarkan pada subyektifitas atau pengaruh eksternal. Dengan kata lain statistika bekerja secara jujur, tanpa dipengaruhi oleh *interest* tertentu. Dengan sifat demikian, statistika dapat digunakan sebagai "alat penilai kenyataan yang terjadi".

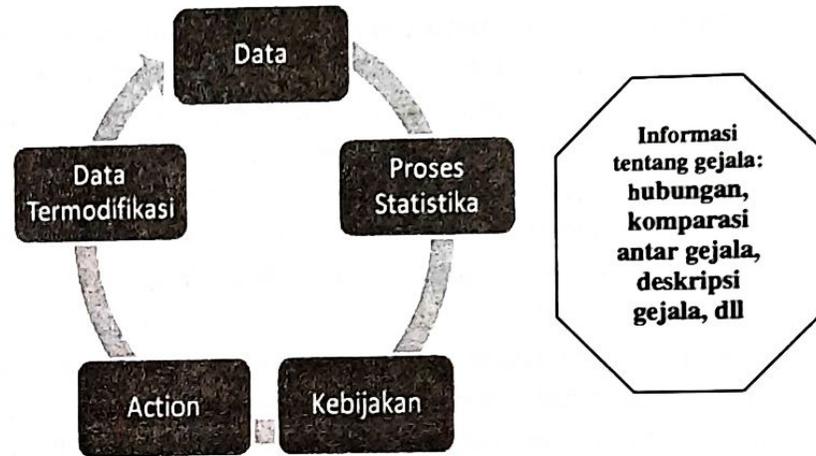
**Universal**

Statistika mempunyai ruang lingkup yang luas. Statistika dapat dipergunakan dalam hampir semua cabang kehidupan. Beberapa yang dapat disebutkan misalnya dalam bidang ekonomi terdapat statistik perdagangan, statistik ekspor-impor, statistik pertanian dan sebagainya. Dalam bidang kependudukan terdapat statistik kelahiran dan kematian, statistik nikah-talak-cerai-rujuk (NTCR), statistik kedatangan dan kepergian (imigrasi dan emigrasi). Dalam bidang hukum kita mengenal statistik kriminalitas dan statistik kecelakaan lalu-lintas. Dalam bidang pendidikan terdapat statistik pendidikan.

**F. Fungsi dan Kegunaan Statistika**

Pada hakikatnya statistika merupakan alat bantu yang dapat digunakan pada banyak bidang kehidupan. Statistika dapat digunakan dalam dunia pendidikan, dalam penelitian, dalam bisnis dan perbankan, dalam administrasi dan manajemen, dan dapat pula digunakan oleh para ilmuwan pada umumnya, bahkan masyarakat pada umumnya pun secara sadar atau tidak sadar menggunakan statistika.

Secara diagramik, posisi proses statistik dapat dilihat pada diagram berikut:



Gambar 1.2. Posisi Statistik

Kegunaan statistika pada tiap aspek kehidupan diantaranya akan diuraikan sebagai berikut

**1. Dalam Dunia Pendidikan**

Dalam bidang pendidikan, statistika digunakan sebagai alat bantu untuk mengolah, menganalisis dan menyimpulkan hasil yang telah dicapai dalam kegiatan penilaian hasil proses belajar mengajar. Dalam penilaian tersebut, umumnya digunakan bilangan (data kuantitatif). Dengan data kuantitatif tersebut dapat dilihat:

- a. Gambaran (deskripsi), baik gambaran secara khusus maupun umum tentang suatu gejala, keadaan atau peristiwa. Gambaran tersebut dapat berupa tabel, grafik, diagram atau ringkasan data numerik.
- b. Perkembangan atau pasang surut mengenai suatu gejala, keadaan atau peristiwa dari waktu ke waktu. Dengan demikian dapat dilihat *trend* (kecenderungan) yang terjadi.

- c. Komparasi antara suatu gejala yang satu dengan gejala lainnya. Apakah suatu gejala dengan gejala lainnya terdapat perbedaan atau tidak. Bila terdapat perbedaan, apakah perbedaan tersebut cukup signifikan atau hanya kebetulan saja. Dengan melihat ada atau tidak adanya perbedaan tersebut, diambil pilihan tindakan atau perlakuan (*treatment*) yang tepat sesuai dengan keadaan yang diharapkan.
- d. Korelasi antara suatu gejala dengan gejala lainnya. Apakah suatu gejala berhubungan dengan gejala lainnya. Jika berhubungan apakah hubungan tersebut signifikan atau tidak, kuat atau lemah. Dengan melihat ada atau tidak adanya hubungan, dapat dilakukan suatu tindakan yang tepat.
- e. Pola hubungan antara suatu gejala dengan gejala lainnya. Ini dilakukan dengan analisis regresi. Dengan melihat pola hubungan tersebut, dapat dilihat bobot pengaruh suatu faktor (gejala) terhadap suatu faktor lainnya. Dapat pula dilihat perkiraan kejadian bila dilakukan suatu perlakuan (*treatment*). Bila faktornya waktu, dapat pula diramalkan hal-hal yang mungkin bakal terjadi dimasa yang akan datang. Dengan demikian dapat dilakukan langkah antisipasi yang harus diambil.
- f. Kesimpulan dari data yang telah dianalisis. Dari kesimpulan tersebut, selanjutnya dapat diambil keputusan dan langkah-langkah yang dapat diambil.

## 2. Dalam Penelitian

Dalam konteks penelitian, kedudukan statistika adalah sebagai alat bantu. Urgensi penggunaan statistika tergantung pada jenis penelitian, dan akurasi pemakaiannya tergantung

pada pemakainya. Maka anggapan yang menyatakan bahwa, statistika sebagai alat yang paling tepat, maupun anggapan bahwa tanpa statistika maka penelitian tidak dapat dipertanggungjawabkan, harus dibuang jauh-jauh. Sebaliknya anggapan bahwa statistika tidak banyak memberi manfaat dalam pengembangan ilmu pengetahuan, juga tidak dapat diterima.

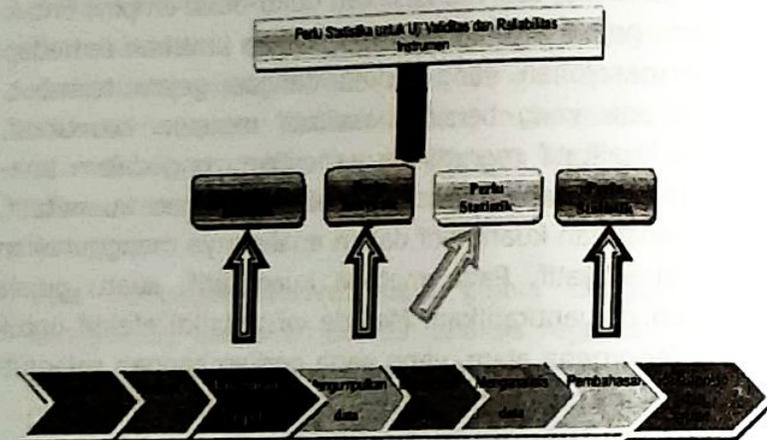
Penelitian merupakan upaya sistematis dalam menemukan, menganalisis, dan menafsirkan bukti-bukti empiris untuk memahami gejala atau untuk menemukan jawaban terhadap suatu permasalahan yang terkait dengan gejala tersebut. Penelitian ada yang bersifat kualitatif maupun kuantitatif. Penelitian kualitatif merupakan penelitian yang dalam analisisnya menggunakan metode kualitatif, bukan kuantitatif. Sedang penelitian kuantitatif dalam analisisnya menggunakan metode kuantitatif. Pada metode kuantitatif, suatu gejala diukur dan dikuantitatifkan. Metode ini terbukti efektif untuk meneliti fenomena alam, yang pada perkembangan selanjutnya digunakan pula untuk meneliti gejala-gejala sosial.

Penelitian kuantitatif menggunakan teknik statistika. Statistika bekerja dengan angka-angka, oleh karenanya akan memaksa seorang pemakai statistika untuk terlibat dengan permainan angka-angka. Meskipun demikian anggapan bahwa statistika merupakan suatu yang sulit adalah tidak benar. Sebab pada dasarnya statistika adalah ilmu yang mudah dipelajari. Terlebih saat ini banyak alat bantu yang dapat menolong, mulai dari kalkulator hingga *software* yang siap pakai, antara lain SPSS, minitab, microstat, bahkan excel pun menyediakan menu-menu statistika yang cukup lengkap.

Penelitian dilakukan dengan cara sistematis. Suharsimi Arikunto (1998;16-17) menyatakan ada sebelas langkah dalam penelitian, yaitu: 1) Memilih masalah, 2) Studi Pendahuluan,

3) Merumuskan masalah, 4) Merumuskan hipotesis, 5) Memilih pendekatan, 6) Menentukan variabel dan sumber data, 7) Menemukan dan menyusun instrumen penelitian, 8) Mengumpulkan data, 10) Menarik kesimpulan, 11) Menulis laporan.

Statistik digunakan dalam langkah-langkah penelitian. Menurut Sugiyono (1999,13) kesertaan statistik dapat dinyatakan dalam gambar 1.3. berikut:



Gambar 1.3. Proses Penelitian dan statistik yang diperlukan

Dari diagram tersebut terlihat bahwa peranan statistik dalam penelitian adalah sebagai :

1. Alat untuk menentukan besarnya anggota sampel yang diambil dari populasi, agar jumlah sampel yang diperlukan dapat dipertanggungjawabkan.
2. Alat-alat untuk menguji validitas dan reliabilitas instrumen.
3. Teknik-teknik untuk menyajikan data.
4. Alat untuk analisis data seperti menguji hipotesis penelitian yang diajukan.

Uraian masing-masing adalah sebagai berikut:

### 1). Penentuan Sampel

Untuk menentukan sampel yang akan digunakan dalam penelitian terdapat berbagai teknik sampling. Secara garis besar ada dua teknik sampling yaitu *probability sampling* dan *non probability sampling*. Peranan statistik adalah pada penentuan jumlah sampel harus dipenuhi agar dari penelitian terhadap sampel dapat dilakukan generalisasi terhadap populasi darimana sampel diambil. Statistik juga mengatur bagaimana pengambilan sampel dilakukan agar dapat dipertanggungjawabkan secara ilmiah.

### 2). Pengujian Validitas dan Reliabilitas Instrumen

Dalam penelitian kuantitatif, kualitas pengumpulan data-nya sangat ditentukan oleh kualitas instrumen atau alat pengumpulan data yang digunakan. Instrumen tersebut dikatakan berkualitas dan dapat dipertanggungjawabkan bila sudah dibuktikan validitas dan reliabilitasnya.

### 3). Penyajian data

Agar mudah dicerna, data hasil penelitian perlu disajikan. Dengan penyajian data, data lebih komunikatif sehingga mudah dibaca dan dipahami. Teknik-teknik penyajian itu antara lain tabel, grafik, diagram dengan berbagai bentuknya.

### 4). Analisis Data

Data yang telah terkumpul selanjutnya diolah, dianalisis. Analisis dilakukan untuk menguji hipotesis yang telah disusun, apakah hipotesis tersebut benar atau palsu. Dalam hal untuk melihat perbedaan antar gejala, digunakan teknik analisis komparasional. Sedang bila dimaksudkan untuk melihat hubungan antar gejala, digunakan teknik analisis korelasional. Bila ingin melihat lebih jauh pola hubungan antar gejala tersebut, digunakan analisis regresi.

Bagi peneliti di laboratorium, metode statistika memberikan peralatan yang berguna bagi perencanaan eksperimen dan evaluasi hasil eksperimen itu sendiri. Dalam merencanakan eksperimen laboratorium, peneliti harus memperhitungkan kemungkinan adanya kesalahan eksperimen (*experimental error*). Metode statistika memberikan teknik pengawasan serta pengulangan kesalahan-kesalahan tersebut di samping teknik penentuan kombinasi faktor-faktor yang diuji secara laboratoris. Mungkin, kontribusi terbesar metode statistika modern pada dunia penelitian yang bersifat eksperimen ialah perkembangan cara eksperimen dalam laboratorium dengan kondisi-kondisi yang terkontrol secara cermat kearah eksperimen yang bersifat lapangan (*field experiment*) dimana kondisi-kondisi yang terkontrol sedikit demi sedikit ditinggalkan agar penelitian dapat diselenggarakan dalam kondisi-kondisi yang lebih mendekati kenyataan.

Dalam bidang pertanian, perkembangan penelitian yang terjadi merupakan konsekuensi dari perkembangan metode statistika modern sejak tahun 1925. Demikian pula dalam bidang teknologi modern, metode statistika juga secara intensif digunakan dalam berbagai riset di pabrik-pabrik kertas, tekstil, bahan farmasi, gelas, karet, besi-baja dan cabang industri kimia serta metalurgi lainnya. Selain riset teknis di atas, riset di bidang pendidikan, ekonomi, kesehatan umum, keamanan jalan, psikologi, sosiologi, antropologi juga membutuhkan metode statistika sebagai peralatannya.

### 3. Dalam Bidang Manajemen dan Administrasi

Statistika digunakan para manajer dan administrator untuk pengumpulan data, penyajian data, pengolahan data, bahan evaluasi dan sebagai dasar pengambilan kebijakan

maupun untuk memecahkan masalah (*problem solving*). Bagi manajer profesional, pengambilan keputusan dengan berbasis data merupakan suatu keniscayaan. Dengan demikian kebijakan dan pemecahan masalah yang dilakukan akan akurat dan mantap.

### 4. Dalam Bidang Bisnis dan Perbankan

Dalam dunia bisnis dan perbankan kegunaan statistika sangat banyak. Bisnis dan perbankan antara lain mencakup permasalahan *accounting, economics, finance, management* dan *marketing*. Bagaimana statistika dapat digunakan dalam persoalan bisnis dapat, diilustrasikan dalam paparan berikut.

#### a. Akuntansi

Banyak keputusan yang dibuat oleh akuntan tentang status keuangan, likuiditas, dan persediaan didasarkan atas analisis rasio keuangan. Akuntan menggunakan metode statistik berupa data rasio keuangan untuk menganalisis kondisi keuangan perusahaan. Analisis keuangan dapat menunjukkan sehat tidaknya rasio keuangan perusahaan. Manajer, karyawan, dan investor perusahaan sangat berkepentingan dengan hasil analisis ini, sebab perusahaan-perusahaan yang memiliki rasio keuangan yang tidak sehat kemungkinan besar akan mengalami kebangkrutan.

Statistika juga sering dipakai oleh para auditor dalam melakukan audit. Mereka menggunakan prosedur sampling statistik ketika melakukan audit terhadap suatu institusi/lembaga. Misalnya, sebuah KAP (kantor akuntan publik) ingin mengetahui apakah jumlah piutang yang ditampilkan pada neraca klien cukup mewakili jumlah aktual piutang. Jika pemeriksaan dilakukan terhadap sejumlah besar piutang individu maka akan sangat menyita waktu dan biayanya mahal.

Pada kondisi yang seperti ini, biasanya staf auditor akan memilih sebagian dari rekening yang disebut sampel. Setelah memeriksa keakuratan rekening sampel, auditor menarik kesimpulan, apakah piutang jumlah yang ditampilkan pada neraca klien dapat diterima.

#### b. Ekonomi

Gejala-gejala ekonomi baik makro maupun mikro, berupa produktifitas ekonomi, laju inflasi, tingkat bunga, maupun pengangguran merupakan obyek pengamatan, kajian dan analisis ekonom (ahli ekonomi). Mereka menggunakan metode statistika untuk mendapatkan berbagai angka indeks, seperti indeks harga konsumen atau indeks harga produsen. Mereka juga menggunakan metode statistika untuk mengukur laju inflasi dari waktu ke waktu. Untuk melakukan peramalan, ekonomi menggunakan analisis runtun waktu (*time series analysis*), di samping analisis regresi untuk menganalisis dan menaksir kondisi perekonomian pada masa yang akan datang.

Pimpinan sebuah supermarket misalnya, sangat membutuhkan metode statistika sebagai alat bantu untuk mengambil keputusan mengenai penetapan jumlah *inventory* (persediaan) suatu barang agar tidak sampai berlebih sehingga menyebabkan barang lama mengendap di gudang yang akan berakibat pada kerusakan barang atau bahkan mungkin akan kehabisan barang tertentu yang menyebabkan kekecewaan pembeli. Termasuk juga perlu melakukan estimasi/penaksiran volume penjualan di masa yang akan datang. Untuk menyelesaikan masalah-masalah seperti ini diperlukan analisis statistik-kuantitatif tentang segala aspek kegiatan perusahaan agar dapat digunakan sebagai bahan dasar pengambilan keputusan mengenai kegiatan perusaha-

an di masa mendatang agar perusahaan tetap eksis dan *survive*. Penggunaan statistika lainnya bisa kita temui pada beberapa hal seperti: melakukan kendali mutu (*quality control*), pengujian terhadap metode dan produk baru, penyelidikan tentang preferensi konsumen, penaksiran potensi pasar bagi produk baru, penelitian mengenai potensi pasar di daerah baru, penetapan harga, penelitian terhadap efektifitas iklan, pengujian terhadap efektifitas metode penjualan yang berbeda.

#### c. Keuangan

Investor (pemodal) dan manajer umumnya menginvestasikan dananya dengan membeli surat berharga. Mereka ingin mendapatkan penerimaan keuntungan yang maksimal dengan resiko yang seminimal mungkin. Untuk itu, investor tidak membeli surat berharga sejenis saja. Mereka tidak ingin "meletakkan telur pada satu macam keranjang saja". Prinsip statistika dapat digunakan untuk mengukur penerimaan yang diharapkan (*expected return*), dan resiko harus ditanggung dapat diukur melalui ukuran statistika mengenai variabilitas atau dispersi.

Seorang analis keuangan selalu menggunakan data statistik untuk membantu membuat keputusan investasinya. Sebagai contoh dalam bisnis saham, para analis pasti melihat data-data keuangan yang ada seperti harga, rasio keuntungan, dividen dan lain-lain. Informasi saham sebuah perusahaan dibandingkan dengan informasi tentang rata-rata saham, seorang analis keuangan dapat mulai menarik kesimpulan, apakah suatu saham tertentu akan menguntungkan atau sebaliknya. Sebagai contoh, *Barron* (18 Februari 2008) melaporkan bahwa rata-rata dividen 30 saham di Dow Jones sebesar 2,45%. Sedangkan Altria Group menunjukkan hasil dividen sebesar 3,05%. Informasi statistik ini menunjuk-

kan bahwa dividen Altria Group lebih tinggi dari rata-rata dividen saham di Dow Jones. Oleh karena itu, seorang analis keuangan mungkin menyimpulkan bahwa dengan membeli Saham Altria Group bisa lebih menguntungkan. Informasi lain tentang Altria Group akan membantu analis untuk membeli, menjual, atau memberikan rekomendasi untuk menahan saham.

#### d. Produksi

Sebagian besar konsumen cenderung lebih mementingkan kualitas daripada kuantitas. Adanya pandangan seperti ini menyebabkan pentingnya bagi produsen untuk menerapkan kendali mutu (*quality control*) terhadap barang yang diproduksi. Ada banyak macam skema kendali mutu dalam statistika. Misalnya, skema kendali mutu rata-rata digunakan untuk memonitor rata-rata output. Sebagai contoh, mesin pengisi minuman kemasan air mineral akan mengisi sebanyak 200 ml ke dalam gelas plastik. Untuk ini, bagian produksi harus mengecek secara periodik dengan cara mengambil sampel gelas tersebut untuk dihitung rata-rata volumenya. Nilai ini ditulis di plot rata-rata. Jika rata-rata volum tersebut yang berada di atas nilai batas kendali maka berarti "*overflowing*" dan sebaliknya. Angka-angka ini yang akan digunakan oleh manajer produksi untuk mengambil keputusan produksi. Proses produksi bisa dilanjutkan jika plot rata-ratanya berada di antara batas kontrol atas dan batas kontrol bawah. Artinya skema rata-rata dapat membantu menentukan kapan penyesuaian yang diperlukan untuk memperbaiki proses produksi.

#### e. Manajemen

Konsumen ingin mendapatkan barang dan jasa dengan kualitas setinggi mungkin, dengan harga yang serendah mungkin. Dalam era pasar global, tingkat persaingan kian

ketat. Untuk dapat bersaing secara efektif, manajer harus selalu melakukan inovasi untuk meningkatkan mutu dan produktifitas. Metode statistika dapat digunakan untuk melakukan pengendalian mutu. Metode pengendalian mutu terpadu (*total quality control*) secara statistik digunakan oleh perusahaan-perusahaan modern untuk mengelola dan secara terus menerus mengembangkan proses perusahaan.

#### f. Pemasaran

Pemasaran merupakan sesuatu yang penting dalam bisnis. Produksi yang berlimpah tanpa diikuti dengan pemasaran yang setimpal, akan mematikan bisnis. Untuk itu manajer perlu mengetahui beberapa hal yang berkaitan dengan pemasaran produknya, terlebih bila akan diluncurkan produk baru. Manajer harus mengetahui tentang hal-hal berikut: Berapa besar pasar untuk suatu produk baru?, Keuntungan apa yang didapat dari adanya produk baru?, apa karakter, sikap, minat, dan opini konsumen yang tertarik terhadap suatu produk?. Untuk menjawab pertanyaan-pertanyaan tersebut, perlu dilakukan penelitian pasar, dan untuk itu diperlukan statistika.

Beberapa produsen berani membeli video rekaman hasil dari *scanner elektronik* di kasir suatu ritel. Hal ini biasanya akan dimanfaatkan sebagai alat untuk mengumpulkan data dalam kaitannya dengan riset pemasaran. Produsen juga berani membeli data dan ringkasan statistik mengenai perilaku konsumen pada saat promosi seperti harga khusus dan penggunaan *display* pada sebuah toko. Manajer dapat menganalisis *scanner* statistik pada kegiatan promosi untuk mendapatkan pemahaman lebih baik tentang hubungan antara kegiatan promosi dan penjualan. Analisis seperti itu

sering terbukti membantu menemukan strategi pemasaran untuk berbagai produk yang akan datang.

#### 5. Dalam Bidang Pemerintahan

Dalam pemerintahan statistika digunakan untuk berbagai macam tujuan seperti "sensus penduduk", yang merupakan salah satu prosedur yang paling dikenal. Indeks Penerapan statistika lainnya yang sekarang sangat populer adalah prosedur jajak pendapat atau *polling* (misalnya dilakukan sebelum pemilihan umum), serta jajak hitung cepat (perhitungan cepat hasil pemilu) atau *quick count*.

#### 6. Bagi Pembimbing Penelitian, Penguji Skripsi/Tesis dan Disertasi

Bagi seorang pembimbing penelitian, bila penelitiannya menggunakan metode kuantitatif, pemahaman dan penguasaan statistika merupakan suatu tuntutan, agar ia dapat memberikan bimbingan secara baik dan benar. Baik artinya ia bisa memberi arahan secara tepat, dan benar maksudnya arahan yang diberikan sesuai dengan kaidah dan dapat dipertanggungjawabkan secara ilmiah.

Tuntutan untuk memahami dan menguasai statistika juga berlaku bagi penguji skripsi/tesis dan disertasi. Hal ini dimaksudkan agar dapat melakukan penilaian secara benar dan obyektif. Dengan demikian akan dihasilkan alumnus yang berkualitas.

#### 7. Bagi Intelektual/Pembaca/Pengguna Informasi

Statistika berguna untuk membantu memahami diagram, grafik atau tabel yang sering disajikan dalam buku, artikel atau sajian informasi di surat kabar, majalah dan televisi. Dengan adanya pemahaman dan penguasaan statis-

tika —minimal dasar-dasarnya— para intelektual, pembaca, pengguna informasi dapat mengerti maksud dari grafik, diagram atau tabel-tabel yang dijumpai. Dengan demikian pemahaman tentang informasi yang dibaca dapat utuh.

*Soal-soal Bab 1*

1. Siapa saja tokoh yang berperan dalam perkembangan statistika? Jelaskan peran mereka masing-masing!
2. Jelaskan secara singkat perkembangan statistika di dunia!
3. Jelaskan pengertian dari istilah berikut!
  - a. Statistik
  - b. Statistika
4. Berikan 2 contoh statistik sebagai informasi dan sebagai ukuran sampel!
5. Jelaskan perbedaan antara statistika deskriptif dan statistika Inferensial!
6. Uraikan berbagai kegiatan yang ada dalam statistika deskriptif dan inferensial!
7. Jelaskan apa yang dimaksud dengan statistika parametrik dan nonparametrik!
8. Sebut dan jelaskan ciri khas statistika!
9. Berikan penjelasan singkat tentang kegunaan statistika dalam bidang penelitian!

*Bab 2*

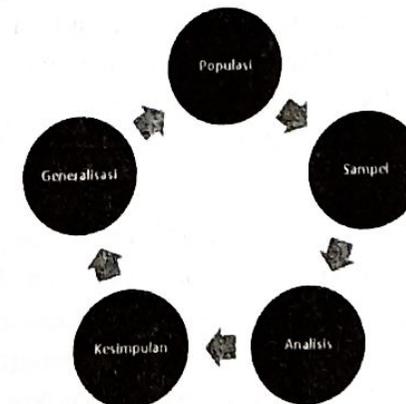
## POPULASI DAN SAMPEL

## A. Pendahuluan

Sebagaimana telah dijelaskan sebelumnya bahwa dalam statistika inferensial selalu berkaitan dengan penarikan kesimpulan berdasarkan data yang dianalisis dari sampel untuk menggambarkan karakteristik atau ciri dari populasi. Penarikan kesimpulan seperti ini biasanya dilakukan pada penelitian atau studi dengan menggunakan metode survei yang memakai data dari sampel tetapi kemudian hasil perhitungannya diperluas untuk menggambarkan atau menyimpulkan karakteristik dari populasi. Jadi, antara sampel dengan populasi terdapat hubungan yang sangat erat. Meskipun demikian, tidak semua penelitian selalu menggunakan sampel, melainkan langsung memakai populasi sehingga kesimpulan yang diperoleh tidak perlu digeneralisasi, melainkan langsung menyimpulkan karakteristik populasi tersebut.

Gambar 2.1

Diagram Alir Konsep Populasi dan Sampel



## B. Populasi

Populasi didefinisikan sebagai keseluruhan objek yang diteliti atau diamati. Dalam arti luas, populasi merupakan wilayah generalisasi yang terdiri atas obyek yang mempunyai karakteristik tertentu yang ditetapkan oleh peneliti untuk dipelajari kemudian ditarik kesimpulan (Boediono dkk, 2008). Populasi bukan hanya berupa orang, hewan, dan benda hidup lainnya tetapi juga bisa berupa benda mati. Populasi juga bukan sekedar jumlah yang diamati, tetapi juga seluruh karakteristik/sifat yang dimiliki oleh obyek tersebut.

Misalnya akan dilakukan penelitian di SMA X, maka SMA X merupakan populasi. Sekolah tersebut mempunyai sejumlah orang, hal ini berarti SMA X merupakan populasi dalam arti jumlah. Tetapi SMA X juga mempunyai karakteristik tertentu pada orang-orangnya. Misalnya efektivitas mengajar guru-gurunya, kedisiplinan siswa maupun gurunya /obyek, gaya kepemimpinan kepala sekolahnya, dan lain sebagainya. Yang demikian merupakan populasi dalam arti karakteristik.

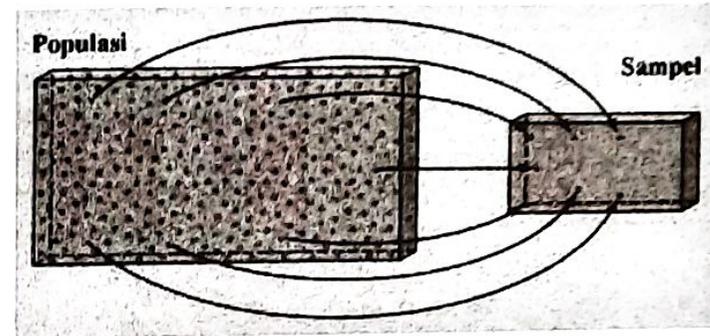
Satu orang pun dapat digunakan sebagai populasi, karena satu orang itu mempunyai berbagai karakteristik. Misalnya, gaya bicara, hobi, cara bergaul, model kepemimpinan dan lain-lain. Atau seorang siswa mempunyai catatan prestasi belajar dalam suatu mata pelajaran tertentu, dari semester I sampai VI. Catatan prestasi tersebut merupakan populasi.

Populasi menggambarkan sesuatu yang sifatnya *ideal* atau *teoritis*. Sedangkan sampel menggambarkan sesuatu yang sifatnya *nyata* atau *empiris*. Populasi dan sampel masing-masing mempunyai ciri/karakteristik yang dapat diukur. Karakteristik yang dihitung (diukur) dari populasi

disebut *parameter*, misalnya mean dilambangkan  $\mu$ , standar deviasi dilambangkan  $s$ , proporsi dilambangkan dengan  $P$ , koefisien korelasi dilambangkan  $\rho$ . Sedangkan karakteristik yang diukur dari sampel disebut *statistik*.

### Contoh 1.4

- Populasi mahasiswa D3 Perbankan Syariah di STAIN Pekalongan
- Populasi nasabah Bank Muamalat Kantor Cabang Purwokerto



Sumber: <http://e-learning.um.ac.id>

Gambar 1.2 Hubungan antara populasi dan sampel

Penelitian yang obyeknya seluruh anggota populasi disebut *sensus*. Sebagai contoh, kita mengenal sensus penduduk, ini berarti semua penduduk diambil datanya tanpa terkecuali baik laki-laki, perempuan, apapun pekerjaannya diteliti sehingga mengetahui jumlah penduduk Indonesia yang sebenarnya. Demikian pula kita mengenal sensus ekonomi, dimana seluruh karakteristik yang mencakup kegiatan ekonomi penduduk dijadikan obyek penelitian. Sedangkan penelitian survei yang obyeknya hanya sebagian dari anggota populasi disebut *sampling*.

Meskipun populasi merupakan gambaran yang ideal dari suatu obyek, tetapi penelitian dengan menggunakan populasi sangat jarang dilakukan. Hal ini umumnya disebabkan oleh faktor biaya yang banyak, tenaga yang dikeluarkan lebih banyak dan waktu untuk mengumpulkan data cukup lama. Penelitian yang sering dilakukan adalah dengan menggunakan sampel. Di samping alasan efektifitas dan efisiensi waktu tenaga dan biaya, tingkat akurasi dari generalisasi yang dilakukan juga tidak kalah akurat.

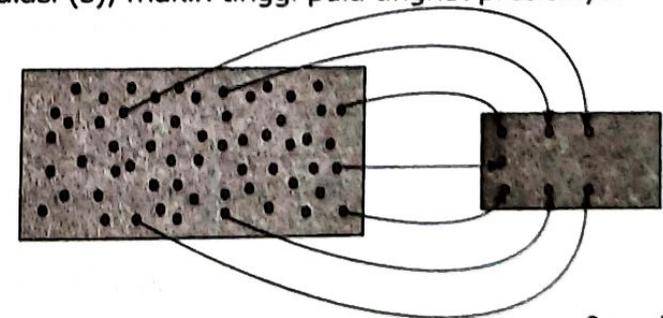
### C. Sampel

Sampel merupakan *bagian/perwakilan dari populasi yang benar-benar diamati*. Sampel yang baik adalah yang dapat mewakili seluruh populasi. Sampel juga disebut sebagai miniatur (mikrokosmos) populasi. Pada saat kita mengambil sampel dalam sebuah penelitian maka semua proses pengambilan sampel harus menggunakan aturan yang benar. Aturan yang benar dimaksud adalah pemilihan teknik samplingnya harus cocok dan jumlah sampel minimal yang harus diambil juga terpenuhi. Salah satu syarat mutlak dari pengambilan sampel adalah sampel yang diperoleh harus *representatif*. Sampel representatif yang dimaksud adalah sampel yang memiliki ciri/karakteristik yang sama atau relatif sama dengan ciri/karakteristik populasinya. Mungkin sekali lagi diantara kita akan bertanya, mengapa harus meneliti sampel? Mengapa tidak populasinya sekalian agar kesimpulannya tidak perlu diuji lagi. Banyak alasan yang bisa dikemukakan untuk menjawab pertanyaan tersebut. Kita tahu bahwa secara umum jumlah anggota populasi sangat banyak, sehingga tidak mungkin seluruh elemen diteliti, di samping itu juga karena keterbatasan waktu penelitian, biaya, dan

sumber daya manusia yang melakukan penelitian. Di sisi lain penelitian terhadap sampel bisa lebih reliabel daripada terhadap populasi, misalnya, karena elemen sedemikian banyaknya maka akan memunculkan kelelahan fisik dan mental para pencacahnya sehingga banyak terjadi kekeliruan. Apalagi jika populasinya homogen, maka penelitian terhadap seluruh elemen dalam populasi menjadi tidak masuk akal.

Adapun syarat sampel yang baik adalah:

- *Memiliki akurasi atau ketepatan yang tinggi*, yaitu tingkat ketidakadaan "bias" (kekeliruan) dalam sampel. Dengan kata lain makin sedikit tingkat kekeliruan yang ada dalam sampel, makin akurat sampel tersebut. Tolok ukur adanya "bias" atau kekeliruan adalah populasi. Agar sampel dapat memprediksi dengan baik populasi, sampel harus mempunyai selengkap mungkin karakteristik populasi
- *Memiliki tingkat presisi yang baik*, yaitu memiliki tingkat presisi estimasi. Presisi mengacu pada persoalan *sedekat mana estimasi kita dengan karakteristik populasi*. Presisi diukur oleh simpangan baku (*standard deviation*). Makin kecil perbedaan di antara simpangan baku yang diperoleh dari sampel ( $S$ ) dengan simpangan baku dari populasi ( $s$ ), makin tinggi pula tingkat presisinya.



Populasi

Sampel

Gambar 2.2 Ilustrasi Pengambilan Sampel

#### D. Mengapa Sampel?

Bayangkan bahwa Anda akan melakukan penelitian untuk mendapatkan informasi tentang persentase penduduk Indonesia yang memiliki akses internet. Jika tidak menggunakan sampel, Anda harus menanyai lebih dari 230 juta orang. Jelas bahwa melakukan wawancara ke lebih dari 230 juta orang memerlukan upaya besar di berbagai hal. Pertama, ada kebutuhan besar dalam hal waktu, kedua adalah biaya, karena perlu mempekerjakan banyak orang untuk membuat wawancara, membayar perjalanan mereka ke setiap desa dan lain-lain. Selain itu, perlu diingat pula bahwa sangatlah tidak mudah untuk sampai ke setiap daerah di Indonesia. Apalagi jika ketika hendak melakukan wawancara, ada yang sedang di rumah sakit, ada yang sedang dalam perjalanan ke luar kota atau luar negeri, dan lain-lain. Dalam keadaan seperti ini dan juga karena alasan ekonomi, maka akan lebih mudah untuk mewawancarai "sebagian tertentu dari populasi" yang disebut sampel. Sampel juga harus dipilih dengan cara yang tepat agar dapat diperoleh kesimpulan yang tepat untuk seluruh penduduk.

Berikutnya, misalkan akan dilakukan penelitian tentang rata-rata daya tahan hidup sebuah mesin motor. Untuk mengetahui daya tahan mesin motor, maka harus menggunakannya sampai mati/rusak. Jika harus mencoba mesin motor setiap orang, akan sangat menyita waktu, tenaga dan biaya. Jadi, yang harus dilakukan adalah memilih sampel yang sesuai dan kemudian kita bisa mengambil kesimpulan umum yang tepat.

Karena alasan yang baru saja disebutkan di atas, dalam banyak contoh akan lebih mudah untuk menggunakan sampel. Namun yang patut diperhatikan adalah jika ingin

mendapatkan kesimpulan yang sangat baik dari populasi, maka harus memastikan bahwa sampel yang terpilih merupakan sampel yang tepat. Sebagai contoh, untuk kasus akses internet di Indonesia, jika hanya memilih 25 orang dari 230 juta penduduk jelas tidak cukup, itu bukan sampel yang representatif. Demikian juga tidak akan mewakili jika hanya memilih 100 orang dari Kota Semarang, atau memilih semua teman Anda dan keluarga Anda saja.

Ada beberapa hal yang harus diperhatikan dalam pemilihan sampel:

- a. Metode pemilihan sampel (teknik sampling untuk digunakan)
- b. Ukuran sampel
- c. Derajat reliabilitas kesimpulan yang diperoleh, yaitu perkiraan kesalahan yang akan diperoleh (dalam hal probabilitas).

Pemilihan sampel yang tidak tepat dapat menyebabkan kesalahan lebih lanjut pada saat akan melakukan estimasi parameter yang sesuai dalam populasi. Ada beberapa penyebab terjadinya kesalahan, diantaranya adalah pewawancara yang bersifat parsial (tidak lengkap), bisa juga disebabkan karena orang yang diwawancarai tidak ingin menjawab pertanyaan-pertanyaan tertentu (atau tidak bisa menjawab). Secara ringkas, semua jenis kesalahan yang mungkin bisa dikelompokkan sebagai berikut:

- a. Kesalahan karena pemilihan anggota sampel: ini terjadi jika salah satu anggota populasi memiliki probabilitas tinggi untuk terpilih daripada yang lain tetapi ternyata tidak terpilih. Misalnya, kita ingin mengukur seberapa puas pengunjung dari sebuah pameran. Untuk itu,

dilakukan wawancara kepada beberapa dari pengunjung 10 sampai 12 di pagi hari saja. Ini berarti bahwa orang yang pergi ke pameran di sore hari tidak akan diwakili, ini artinya sampel tidak akan mewakili semua pengunjung. Salah satu cara untuk menghindari kesalahan jenis ini adalah memilih sampel sehingga semua pengunjung memiliki probabilitas yang sama untuk terpilih.

- b. Kesalahan karena jawaban: tidak tertutup kemungkinan bahwa beberapa anggota populasi tidak ingin atau tidak bisa menjawab pertanyaan-pertanyaan tertentu dalam wawancara. Atau bisa juga terjadi, ketika kita memiliki kuesioner (termasuk pertanyaan pribadi), beberapa anggota populasi tidak menjawab jujur. Ini merupakan kesalahan yang umumnya sangat sulit untuk dihindari, tetapi kita bisa memeriksa kejujuran dalam jawaban dengan cara memasukkan beberapa pertanyaan lain (pertanyaan filter) untuk mendeteksi apakah jawaban tersebut jujur atau tidak.

### E. Cara Menentukan Jumlah Sampel

Terdapat berbagai macam cara dalam menentukan jumlah sampel minimal yang harus diambil dari populasi, baik ukuran populasinya diketahui maupun tidak. Jumlah sampel yang dipilih harus mewakili populasi karena jumlah sampel mempengaruhi tingkat kesalahan yang terjadi. Semakin banyak jumlah sampel maka semakin kecil tingkat kesalahan dari generalisasi yang dilakukan, begitu pula sebaliknya (*vice versa*).

Ada beberapa faktor yang mempengaruhi jumlah sampel, diantaranya:

- a. Tingkat presisi yang diinginkan (*level of precisions*)

Semakin tinggi tingkat homogenitas populasi semakin kecil ukuran sampel yang boleh diambil; semakin rendah tingkat homogenitas populasi semakin besar ukuran sampel yang harus diambil.

- b. Derajat keseragaman (*degree of homogeneity*)

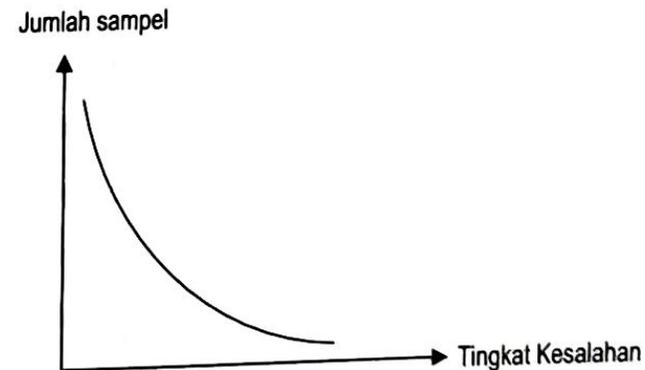
Semakin tinggi tingkat presisi yang diinginkan peneliti, semakin besar sampel yang harus diambil

- c. Banyaknya variabel yang diteliti dan rancangan analisis

Semakin banyak variabel yang akan dianalisis, misalnya dengan menggunakan rancangan analisis tabulasi silang atau uji *chi-square of independen* (uji chi kuadrat), mengingat adanya persyaratan pengujian hubungan antarvariabel yang tidak membolehkan adanya nilai frekuensi hasil penelitian  $< 1$ , maka ukuran sampelnya harus besar

- d. Biaya, waktu, dan tenaga yang tersedia.

Hubungan antara jumlah sampel dengan tingkat kesalahan digambarkan pada grafik berikut.



Grafik 2.1 Hubungan antara jumlah sampel dengan tingkat kesalahan

Secara lebih rinci tentang ukuran sampel dan teknik pengambilan sampel akan dibahas pada bab Teknik Sampling.

### Bab 3

## BERBAGAI HAL TENTANG DATA

### A. Pendahuluan

Dalam statistika, data merupakan bagian yang sangat penting. Tanpa data, statistika bagaikan badan tanpa nyawa. Sebab, apa yang akan dikumpulkan, apa yang akan disajikan, apa yang akan diolah, disimpulkan, diuji dan diinterpretasikan adalah kumpulan data. Sehingga keberadaan data menjadi mutlak.

Data yang dimaksud dalam statistika adalah *kumpulan informasi dari suatu obyek*. Data bisa berbentuk bilangan maupun nonbilangan. Misalnya data tentang jumlah pabrik tekstil di Provinsi Jawa Tengah menurut Kota/Kabupaten tahun 2010, maka data ini memuat bilangan yang menyatakan jumlah pabrik tekstil. Dalam contoh ini, data yang ada jelas berupa bilangan. Tetapi jika data tentang jenis kelamin, warna kesukaan, hobi, bahasa yang dikuasai, asal kenegaraan dan sebagainya, maka data yang terkumpul pasti bukan berupa bilangan (nonbilangan).

Apabila kumpulan informasi melibatkan semua anggota populasi, maka data itu disebut *data populasi*. Akan tetapi jika kumpulan informasi/nila tersebut hanya melibatkan sebagian saja dari anggota populasi, maka disebut *data sampel*.

### B. Sumber Data Statistik

Sumber data dalam penelitian adalah subjek dari mana data dapat diperoleh. Sumber data statistik sebenarnya bisa

berasal dari mana saja. Baik di sekitar kita, maupun di tempat yang jauh sekalipun dari tempat kita. Anggapan yang menyatakan bahwa data statistik hanya terdapat pada lembaga-lembaga pengumpul data seperti Badan Pusat Statistik, lembaga penelitian atau Departemen, maka anggapan seperti itu adalah salah. Data statistik terdapat di mana-mana asal kita bersikap kritis memperhatikan sekeliling kita, melakukan observasi dan pencatatan yang cermat. Dalam mengumpulkan data, sangat bergantung dari tujuan penelitian. Data bisa kita dapatkan dari supermarket, bank, toko, pasar, tempat olahraga, perpustakaan, perusahaan, tempat rekreasi, perguruan tinggi, sekolah-sekolah dan masih banyak tempat lain. Secara umum, kita bisa mengidentifikasi sumber data secara mudah dengan menggunakan prinsip 3P yaitu *Person* (sumber data berupa orang), *Place* (sumber data berupa tempat) dan *Paper* (sumber data berupa simbol).

### C. Penggolongan Data

Dalam statistika, kita mengenal beberapa jenis data tergantung konteksnya. Telah dijelaskan sebelumnya bahwa suatu data bisa berupa angka bisa juga berupa nonangka. Beberapa penggolongan berkaitan dengan data akan dijelaskan sebagai berikut:

#### a) Data Intern dan Ekstern

Yang dimaksud dengan data intern adalah data yang dikumpulkan oleh badan-badan tentang aktifitas dirinya dan hasilnya dipakai untuknya juga. Sedangkan data ekstern adalah data yang diambil oleh badan-badan dari luar dirinya.

#### b) Data Primer dan Sekunder

Data primer merupakan data yang diambil secara langsung dari sumbernya/objek yang diamati. Sedangkan data sekunder adalah data yang diambil tidak langsung dari sumbernya, bisa diperoleh dari data yang sudah ada maupun mengutip dari literatur.

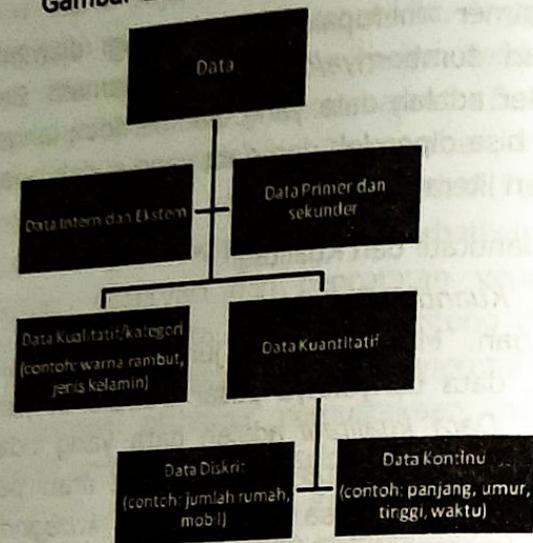
#### c) Data Kuantitatif dan Kualitatif

*Data Kuantitatif* adalah data yang berupa angka/bilangan. Misalnya data jumlah pengunjung suatu minimarket, data banyaknya pasar tradisional yang ada di Pekalongan. *Data kualitatif* adalah data yang tidak berupa bilangan. Biasanya berupa kata, kalimat atau pernyataan. Data kualitatif juga biasa disebut *data kategori*. Sebagai contoh data kualitatif adalah data jenis kelamin mahasiswa, data jenis-jenis rasa, data jenis warna kesukaan. Data kuantitatif terbagi menjadi dua yakni data diskrit dan kontinu.

#### d) Data diskrit dan Kontinu

*Data diskrit* merupakan data yang diperoleh dengan jalan menghitung/mencacah data yang ada. Misalnya data jumlah mahasiswa jurusan Syariah, jumlah barang yang ada di toko A dan sebagainya. Data ini kita peroleh dengan cara menghitung/mencacah satu persatu. Sedangkan *data kontinu* merupakan data yang diperoleh dengan cara mengukur. Data kontinu akan mempunyai nilai hanya jika berada dalam interval. Contohnya adalah data tinggi badan anak, berat badan, suhu, kelembaban, curah hujan dan lain-lain. Data-data ini kita dapatkan dengan cara mengukur dengan menggunakan alat ukur.

Gambar 3.1. Bagan Penggolongan Data



#### D. Validitas Data Statistik

Pada era demokrasi seperti ini, berbagai data statistik banyak bertebaran dimana-mana. Berbagai lembaga seperti-nya berlomba-lomba untuk menghasilkan data statistik. Hampir semua bidang kehidupan seperti-nya juga sudah terjangkau. Bidang-bidang seperti ekonomi, pendidikan, kependudukan, biologi, sosial, dll, telah banyak dilengkapi dengan data statistik. Semakin dipertegas dengan munculnya ilmu-ilmu baru hasil penggabungan antara suatu ilmu dengan statistik. Contohnya sosiometric yang merupakan hasil peleburan ilmu sosial dengan statistik, ada lagi econometric hasil penggabungan ekonomi dengan statistik, tidak ketinggalan biometric (biologi dan statistik), dan masih banyak lagi.

Kenapa data statistik benar-benar sudah menjamur pada berbagai bidang? Itu disebabkan karena manfaat data statistik yang luar biasa. Data statistik merupakan dasar

pengambilan keputusan yang sangat baik. Misalkan, penggunaan data statistik jumlah penduduk Indonesia untuk mendasari penentuan target berapa ton beras yang harus dihasilkan dalam satu tahun ini. Selain itu, ada data statistik jumlah penduduk miskin yang sangat berguna untuk mendasari pelaksanaan program pengentasan kemiskinan. Selain itu, data statistik juga mampu berperan sebagai alat pengevaluasi yang ampuh. Misal, data statistik mengenai pertumbuhan jumlah penduduk umur 0-4 dan 5-9 tahun, mampu mengevaluasi keberhasilan program Keluarga Berencana (KB). Hanya karena data statistik, program KB dipaksa mereformasi diri untuk menjadi lebih baik. Peran data statistik sebagai dasar dan alat pengevaluasi memang sungguh super sekali.

Melihat bagaimana berguna dan pentingnya data statistik, banyak pihak berharap data statistik memang benar-benar sesuai dengan keadaan nyata di lapangan. Data statistik diharapkan menjadi pemotret keadaan real, bukan pelukis yang cenderung subyektif. Data statistik diharuskan apa adanya tanpa rekayasa. Ekspektasi tinggi masyarakat yang mengharapkan data statistik memang benar apa adanya atau 100% benar. Hal itulah yang menyebabkan data statistik sering dijadikan kambing hitam apabila suatu hal (misal program pemerintah) tidak tepat sasaran. Apabila kasusnya seperti itu, data statistik sering dianggap kurang mampu mencerminkan kondisi real.

Data statistik harus diakui tidak akan mampu memenuhi ekspektasi masyarakat 100% benar atau dengan sempurna menggambarkan kondisi real. Kenapa bisa begitu? Mari kita telaah apa sih pengertian dari data statistik, yaitu data yang diperoleh lewat survey dan hanya sebagian obyek saja yang

diteliti. Misal survey untuk mengetahui jumlah penduduk miskin, tidak semua penduduk Indonesia dicacah, tetapi hanya sebagian saja lalu dicari berapa yang tergolong miskin, dan hasil itu digunakan untuk memperkirakan proporsi jumlah penduduk miskin Indonesia. Selalu ada kesalahan bila memakai cara itu dan tidak dapat dielakkan lagi.

Statistik "hanya" berusaha mendekatkan perkiraannya dengan nilai sesungguhnya. Selalu berjuang agar error berada pada posisi minimum. Bagaimana itu dapat dilakukan? Hal itu dilakukan salah satu terutamanya pada cara pemilihan sampel yang mampu mewakili kenyataan. Ilmu peluang juga digunakan sebagai senjata ampuh untuk memperkirakan data keseluruhan. Kesalahan yang masih ditolerir dalam data statistik antara lain sebesar 1% atau 5% untuk ilmu eksak dan pengetahuan alam, 10% atau 15% untuk ilmu sosial.

Sebenarnya bisa saja 100% benar, apabila tiap unit obyek dicacah (asumsi tidak ada kesalahan teknis), sehingga kesalahan dalam yang berasal dari pemilihan sampel tidak ada. Tetapi cara mencari data seperti itu memerlukan waktu lama, serta tenaga, dan biaya yang sangat besar. Contohnya sensus penduduk, pertanian, dan ekonomi di Indonesia yang hanya dilaksanakan setiap 10 tahun sekali. Sedangkan kebutuhan data yang *up to date* selalu ada dalam selang waktu tidak lama.

Solusi dari hal itu adalah data statistik yang untuk memperolehnya hanya membutuhkan waktu, tenaga, dan biaya yang amat lebih sedikit daripada data yang diperoleh lewat pencacahan semua unit obyek (sensus). Memang kebenaran data statistik yang pasti memiliki error, tetapi data statistik (perkiraan) itu selalu diusahakan sedekat-dekatnya dengan kondisi real, sehingga bisa digunakan untuk menjadi dasar dan alat pengevaluasi.

### E. Variabel

Fokus dari desain penelitian dan analisis statistik adalah studi tentang variabel. Pada saat Anda ingin mempelajari suatu fenomena, langkah pertama adalah mendefinisikan fenomena yang diteliti tersebut, dalam hal ini adalah menentukan variabel-variabel yang kita amati dan selanjutnya menentukan bagaimana cara Anda mengukur variabel tersebut. Proses tersebut dikenal dengan istilah *definisi operasional*. Informasi ilmiah yang dijelaskan dalam definisi operasional sangat membantu peneliti lain yang ingin melakukan penelitian dengan menggunakan variabel yang sama, karena berdasarkan informasi itu, ia akan mengetahui bagaimana caranya melakukan pengukuran terhadap variabel yang dibangun berdasarkan konsep yang sama. Dengan demikian, ia dapat menentukan apakah tetap menggunakan prosedur pengukuran yang sama atau diperlukan pengukuran yang baru. Konsep-konsep yang sudah diterjemahkan menjadi satuan yang sudah kita anggap lebih operasional (variabel dan konstruk), biasanya belum sepenuhnya siap untuk diukur. Karena variabel dan konstruk tersebut memiliki alternatif dimensi yang bisa diukur dengan cara berlainan. Contoh tentang variabel usia/umur. Cara pengukuran variabel tersebut bisa saja berbeda, pertama mungkin Anda mengukur usianya langsung secara numerik, misalnya 4, 12.5, 18, 31 tahun dst, atau bisa saja Anda mengukur berdasarkan kategori, misalnya Balita (0-5 th), Anak-anak (5 - 14), Remaja (14 - 24), Dewasa (25 - 54), Tua (55-64), dan Lansia (>65) tahun.

Jelas di sini bahwa untuk memahami suatu fenomena, kita harus memahami dulu istilah variabel dan skala pengukuran. Apabila Anda tidak menentukan secara jelas cara

pengukuran variabel yang ingin Anda pelajari, pada akhirnya Anda akan mengalami kebingungan dalam menentukan desain penelitian yang tepat serta dalam menentukan prosedur analisis statistik yang sesuai.

Variabel berasal dari kata *vary* (berubah) dan *able* (dapat). Jadi, secara harfiah variabel berarti dapat berubah, sehingga setiap variabel dapat diberi nilai dan nilai itu bisa berubah. Nilai tersebut bisa kuantitatif (terukur dan atau terhitung, dapat dinyatakan dengan angka) juga bisa kualitatif (jumlah dan derajat atributnya yang dinyatakan dengan nilai mutu). Variabel merupakan elemen penting dalam masalah penelitian. Dalam statistik, *variabel* didefinisikan sebagai konsep, kualitas, karakteristik, atribut, atau sifat-sifat dari suatu objek (orang, benda, tempat, dll) yang nilainya berbeda-beda antara satu objek dengan objek lainnya dan sudah ditetapkan oleh peneliti untuk dipelajari dan ditarik kesimpulannya. *Karakteristik* adalah ciri tertentu pada obyek yang kita teliti, yang dapat membedakan objek tersebut dari objek lainnya, sedangkan objek yang karakteristiknya sedang kita amati dinamakan *satuan pengamatan* dan angka atau kategori (nilai mutu) tertentu dari suatu objek yang kita amati dinamakan *variate* (nilai). Kumpulan nilai yang diperoleh dari hasil pengukuran atau penghitungan suatu variabel dinamakan dengan *data*.

Karakteristik yang dimiliki suatu pengamatan keadaannya berbeda-beda (berubah-ubah) atau memiliki gejala yang bervariasi dari satu satuan pengamatan ke satu satuan pengamatan lainnya, atau, untuk satuan pengamatan yang sama, karakteristiknya berubah menurut waktu atau tempat. Apabila karakteristik setiap satuan pengamatan semuanya sama, tidak beragam, maka bukan lagi merupakan variabel, melainkan konstanta.

Variabel adalah suatu sifat atau fenomena yang menunjukkan sesuatu yang dapat diamati dan nilainya berbeda-beda. Variabel merupakan konsep yang mempunyai variasi nilai dengan minimal ada dua variasi. Dalam statistika, nilai-nilai dari variabel tersebut merupakan data yang akan diolah dengan prosedur statistika. Bagian dari variabel disebut atribut. Sesuatu dikatakan variabel jika mempunyai nama, dapat diamati/diukur, nilainya berbeda-beda, memiliki definisi verbal, ada kelompok penggolongan/satuan.

Misalnya, *tinggi badan* merupakan suatu variabel. Nama variabelnya *tinggi badan*, dapat diamati/diukur (dengan alat ukur, bisa berupa meteran), definisi verbalnya adalah *jarak dari ujung kaki sampai ujung kepala*, nilainya bisa berbeda-beda, satuannya adalah *centimeter (cm)* atau bisa juga menggunakan satuan *meter (m)*.

## F. Skala Pengukuran

Pengukuran adalah dasar dari penyelidikan ilmiah. Segala sesuatu yang kita lakukan dimulai dengan pengukuran objek yang akan kita pelajari. Pengukuran adalah pemberian angka atau kode pada suatu obyek. Terdapat empat jenis skala pengukuran yaitu Nominal, Ordinal, Interval, Ratio (sering disingkat NOIR). Skala yang paling rendah adalah nominal dan yang tertinggi adalah skala rasio. Skala pengukuran yang lebih tinggi akan memiliki karakteristik skala pengukuran di bawahnya. Misalnya, skala rasio akan memiliki karakteristik nominal, interval, dan ordinal. Perbedaan skala secara rinci akan dijelaskan sebagai berikut:

### a. Skala Nominal

Skala nominal merupakan skala pengukuran yang paling rendah tingkatannya dan hanya bisa digunakan untuk klasi-

fikasi kualitatif atau kategorisasi. Skala nominal mempunyai ciri hanya untuk *membedakan* skala ukur yang satu dengan yang lain. Skala ini hanya mendasarkan diri pada konsep penggolongan kategori agar tidak tumpang tindih (*mutually exclusive*) dan tuntas (*exhaustive*). Angka yang diberikan tidak merefleksikan kedudukan dan tingkatan satu kategori terhadap kategori yang lain. Pada skala nominal, data hanya bisa diklasifikasikan ke dalam kategori-kategori.

Skala nominal disebut juga jenis pengukuran yang paling terbatas. Sebagai contoh, kita dapat mengatakan bahwa jenis kelamin ke 2 orang tersebut berbeda, satu perempuan dan satunya lagi laki-laki. Di sini kita bisa membedakan karakteristik keduanya, tetapi kita tidak bisa mengukur dan mengatakan mana yang "lebih" atau mana yang "kurang" dari kualitas yang diwakili oleh variabel tersebut. Kita hanya bisa memberikan kode/label pada kedua karakteristik tersebut, misalnya angka 0 untuk perempuan dan angka 1 untuk laki-laki. Kode/label angka tersebut bisa saja di tukar. Kode di sana hanya berfungsi sebagai pembeda antara kedua objek dan tidak menunjukkan urutan atau kesinambungan. Angka 1 tidak menunjukkan lebih tinggi atau lebih baik di banding 0.

Operator aritmetika yang bisa digunakan pada skala nominal hanya tanda "=" atau "≠".

Contoh-contoh variabel nominal lainnya adalah:

- jenis tanah,
- varietas,
- ras,
- warna,
- bentuk,

- kota,
- Golongan darah
- Jenis penyakit
- Agama
- Suku
- Nomor
  - KTP/SIM/Kartu
  - Pelajar

#### b. Skala Ordinal

Skala ordinal mendasarkan diri pada pengurutan (*order*) yang merupakan tingkatan-tingkatan dimulai dari yang terendah sampai yang tertinggi atau sebaliknya. Skor yang dibuat memiliki jenjang/urutan. Dalam hal ini kita bisa mengatakan A "lebih" baik dibanding B atau B "kurang" baik dibanding A, namun kita tidak bisa mengatakan seberapa banyak lebihnya A dibanding B. Dengan demikian, batas satu variasi nilai ke variasi nilai yang lain tidak jelas, sehingga yang dapat dibandingkan hanyalah apakah nilai tersebut lebih tinggi, sama, atau lebih rendah daripada nilai yang lain, namun kita tidak bisa mengatakan berapa perbedaan jarak (*interval*) diantara nilai-nilai tersebut.

Contoh umum skala ordinal adalah status sosial ekonomi keluarga. Sebagai contoh, kita tahu bahwa kelas menengah ke atas lebih tinggi status sosial ekonominya dibanding kelas menengah ke bawah, tapi kita tidak bisa mengatakan berapa lebihnya atau mengatakan bahwa kelas menengah ke atas 18 % lebih tinggi. Pemberian simbol/kode angka pada skala ordinal, selain berfungsi untuk membedakan karakteristik antar objek juga sudah menentukan urutan peringkat dari objek tersebut. Skala ordinal masih bersifat kualitatif.

Operator aritmetika yang bisa digunakan pada skala ordinal adalah tanda "=", "≠", "<" dan ">". Misal kode angka untuk kelas bawah = 0, menengah = 1, dan atas = 2. Angka 0 berbeda dengan 1 ataupun 2 (operator aritmetika: = dan ≠), 0 lebih rendah dibanding 1 (operator aritmetika: < dan >),

Contoh

- Variabel kelas ekonomi (1 untuk miskin, 2 untuk menengah dan 3 untuk kaya)
- Tingkat penghasilan penduduk (1 rendah, 2 sedang dan 3 tinggi)
- Tingkat pendidikan
- Tingkat keparahan penyakit
- Tingkat kesembuhan
- Derajat keganasan kanker

### c. Skala Interval

Skala interval merupakan ukuran yang tidak hanya mengurutkan obyek berdasarkan suatu atribut saja tetapi juga memberi informasi tentang interval/jarak antara satu obyek dengan obyek yang lain. Sebagai contoh, suhu, yang diukur dalam derajat Fahrenheit atau Celcius, merupakan skala interval. Kita dapat mengatakan bahwa suhu 40 derajat lebih tinggi daripada suhu 35 derajat, demikian juga suhu 20 derajat lebih tinggi dibanding dengan suhu 15 derajat. Perbedaan selisih suhu antara 40 dan 35 derajat nilainya sama dengan perbedaan suhu antara 20 dan 15 derajat, yaitu 5 derajat.

Jelas di sini bahwa pada skala interval, selain kita bisa membedakan (mengkategorikan), mengurutkan nilainya, juga bisa di hitung berapa perbedaannya/selisihnya dan jarak atau

intervalnya juga dapat dibandingkan. Perbedaan antara kedua nilai pada skala interval sudah punya makna yang berarti, berbeda dengan perbedaan pada skala ordinal yang maknanya tidak berarti. Misalnya, perbedaan antara suhu 40 dan 50 derajat dua kali lebih besar dibandingkan dengan perbedaan antara suhu 30 dan 35. Dengan demikian, selain sudah mencakup sekala nominal, juga sudah termasuk skala ordinal, tetapi nilai mutlaknya tidak dapat dibandingkan secara matematik, oleh karena batas-batas variasi nilai pada interval adalah arbiter (angka nolnya tidak absolut). Operator aritmetika yang bisa digunakan pada skala interval adalah tanda "=", "≠", "<", ">", "+", "-". Misal suhu:  $30 + 10 = 40$  derajat.

Contoh skala interval yang lain:

- Suhu
- IQ
- Nilai matematika
- Skor Tes Toefl
- IPK
- Beberapa indeks pengukuran tertentu

### d. Skala Rasio

Skala rasio diperoleh apabila selain informasi tentang urutan dan interval antara obyek, dipunyai juga informasi tambahan tentang jumlah absolut atribut yang dimiliki oleh salah satu dari obyek tersebut. Jadi ukuran rasio mirip dengan ukuran interval hanya jaraknya diukur tidak dari angka rata-rata atau angka minimal tetapi dari titik nol yang sesungguhnya (nol mutlak), sehingga memungkinkan menyatakan rasio atau perbandingan di antara kedua nilai, misalnya

x adalah dua kali lebih y. Contohnya adalah berat, tinggi, panjang, usia. Sebagai contoh, berat A = 70 kg, berat B = 35 kg, Berat C = 0 kg. Di sini kita bisa membandingkan rasio, misalnya kita bisa mengatakan bahwa berat A dua kali berat B. Berat C = 0 kg, artinya C tidak mempunyai bobot. Angka 0 di sini jelas dan berarti dan angka 0 menunjukkan nilai 0 mutlak. Memang agak sedikit susah dalam membedakan antara skala interval dengan rasio. Kuncinya adalah di angka 0, apakah nilai nol tersebut mutlak (berarti) atau tidak? Kita tidak saja dapat mengatakan bahwa suhu 200 derajat lebih tinggi daripada suhu 100 derajat, tetapi kita juga sudah dapat menyatakan dengan pasti bahwa rasionya benar dua kali lebih tinggi. Operator aritmetika yang bisa digunakan pada skala rasio adalah tanda "=", "≠", "<", ">", "+", "-", "x" dan "÷".

Contoh skala rasio lainnya:

- Hasil panen (diukur dari nol sejati)
- Banyaknya hasil produksi (diukur dari nol sejati juga)
- Waktu, panjang, tinggi, berat, usia
- Kadar zat dan jumlah sel tertentu
- Dosis obat, dll

**Catatan:** Skala interval tidak memiliki karakteristik rasio. Kebanyakan prosedur analisis data statistik tidak membedakan antara data yang diukur dalam skala interval dan rasio.

**Tabel 3.1**  
Ringkasan skala pengukuran

Nominal	Data Kategori	Mutually exclusive	=, ≠	Jenis Kelamin Wana Kulit
---------	---------------	--------------------	------	-----------------------------

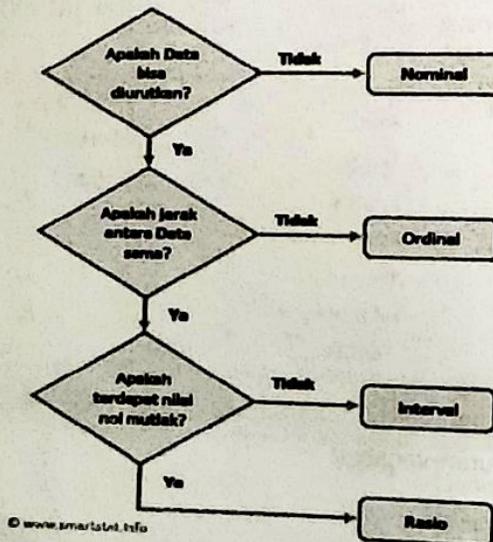
Ordinal	Data yang hanya bisa diurutkan dari kecil ke besar atau sebaliknya	Mutually exclusive Urutannya Pasti/Jelas	=, ≠ <, >	Status sosial ekonomi keluarga Peringkat Kelas Pangkat/Jabatan/Golongan
Interval	Selain mencakup karakteristik Nomina dan Ordinal, juga sudah bisa dilakukan operasi penjumlahan karena jarak antara datanya sudah jelas. Tidak mempunyai nilai nol mutlak	Mutually exclusive Urutannya Pasti Jarak antara kode sama	=, ≠, <, >, +, -	Suhu (Celsius & Fahrenheit) IQ (tingkat kecerdasan)
Rasio	Mencakup karakteristik Interval dan mempunyai nilai nol mutlak	Mutually exclusive Urutannya Pasti Jarak antara kode sama Terdapat nilai nol mutlak	=, ≠, <, >, +, -, x, ÷	Suhu (Kelvin) Waktu Panjang Berat Tinggi

Sumber: <http://smartstat.wordpress.com/2010/02/20/skala-pengukuran-variabel/>

Beberapa contoh penerapan bisa dilihat pada tabel berikut:

NO	ASAL MAHASISWA	JARAK (KM)	ASAL SEKOLAH	RATA NILAI UAN	AKREDITASI SEKOLAH
1	Batang	52	SMA	8,5	A
2	Pekalongan	56	SMK	8,28	C
3	Pemalang	60	MA	6,5	B
4	Kota Pekalongan	55	SMK	84,4	C
5	Tegal	76	MA	7,8	C
6	Semarang	35	MA	79,5	A
7	Kendal	45	MA	7,6	B
8	Ungaran	30	SMK	8,71	B
9	Cilacap	50	SMA	7,7	A
Skala	Nominal	Rasio	Nominal	Interval	Ordinal

Gambar 3.2 Flowchart untuk menentukan skala pengukuran variabel



### G. Angka Eksak dan Angka Pendekatan (*Aproksimatif*)

Sebagaimana yang disebutkan di awal, data merupakan kumpulan angka atau bilangan. Meskipun terdapat data kualitatif (bukan angka), tetapi jika akan diproses secara statistik maka data ini harus dikuantitatifkan dengan cara diberikan kode (*koding*) atau diberikan skor (*skoring*). Suatu angka atau bilangan jika bisa dinyatakan sampai unit terkecil, maka data seperti ini terdiri dari *angka-angka eksak*.

Contoh (Angka Eksak)

- Jumlah mahasiswa reguler D3 Perbankan Syariah di STAIN Pekalongan
- Jumlah penduduk Indonesia dalam *statistical pocketbook of Indonesia*
- Jumlah transmigran yang diterbitkan oleh Kementerian Tenaga Kerja dan Transmigrasi

Angka pendekatan (*aproksimatif*) merupakan angka yang diperoleh dari hasil pengukuran atau angka hasil pembulatan. Sebab, walaupun angka tersebut berbentuk bilangan bulat atau hanya ditulis beberapa digit, tetapi sesungguhnya angka tersebut merupakan angka pendekatan.

Contoh (Angka pendekatan/*aproksimatif*)

- Tinggi badan seseorang adalah 171 cm. Sesungguhnya angka ini adalah angka pendekatan. Bisa jadi tinggi badan "sesungguhnya" adalah 171,223...
- Jumlah penduduk kota Pekalongan 2,1 juta. Ini merupakan angka pendekatan. Angka sesungguhnya adalah mendekati 2,1 juta, misalnya 2.099.978.

Dalam perhitungan matematika, kita mengenal  $1/3 = 0,333333...$  ini merupakan angka pendekatan. Sedangkan  $1/5 = 0,2$  merupakan angka eksak.

### H. Angka Pendekatan sebagai Interval

Setiap angka pendekatan sebenarnya memiliki "angka yang benar" (*true value*). Nilai tersebut terletak pada sebuah interval yang memiliki batas bawah (*lowerbound*) dan batas atas (*superbound*). Secara logis, kita akan menerima bahwa hasil pengukuran panjang 16 m memiliki nilai sesungguhnya yang mendekati 16 m. Dan yang harus dicatat, nilai tersebut tidak boleh melebihi 0,5 m di bawah 16 atau 0,5 m di atas 16. Sebab, jika nilai tersebut ternyata 0,8 m di bawah 16 m, maka seharusnya dia lebih mendekati 15 m. Demikian juga, jika nilai tersebut terletak di atas 16 m melebihi 0,8 m maka sesungguhnya dia lebih dekat ke 17 m. Jadi jelas bahwa interval 16 m adalah 15,5 m – 16,5 m. 15,5 m diperoleh dari  $16 - 0,5$  dan 16,5 diperoleh dari  $16 + 0,5$ .

Jika kita menimbang berat badan diperoleh 69,998 kg, maka unit terakhir berat tersebut adalah 0,001. Sehingga untuk menentukan kedua batasnya adalah dengan mengurangi 69,998 kg dengan 0,0005 dan menambahkan 69,998 kg dengan 0,0005. Jadi kedua batasnya adalah 69,9975 dan 69,9985.

### I. Pembulatan Bilangan

Karena statistika memiliki ciri kuantitatif sehingga analisis data dalam statistika seringkali melibatkan data yang berbentuk bilangan. Bilangan yang diperoleh dalam pengukuran terkadang masih berupa bilangan desimal yang rumit. Untuk itu diperlukan pembulatan. Pembulatan bilangan dalam statistika selalu mengikuti kaidah umum dalam matematika. Berikut ini adalah beberapa aturan yang digunakan dalam pembulatan bilangan.

1. Jika angka yang harus dihilangkan adalah 4 atau kurang, maka angka terkanan yang mendahuluinya tetap.

Contoh: Rp. 75.234,- dibulatkan menjadi puluhan ribu rupiah maka diperoleh Rp. 75 ribu.

Angka yang harus dihilangkan adalah mulai dari 2 ke kanan. Karena angka ini kurang dari 4 maka angka yang mendahuluinya (yaitu 5) tetap.

2. Jika angka yang harus dihilangkan adalah lebih dari 5 atau angka 5 diikuti angka bukan nol maka angka yang mendahuluinya ditambah dengan 1.

Contoh: 176,51 kg dibulatkan menjadi 177 kg.

3. Jika angka yang harus dihilangkan hanya angka 5 atau angka 5 diikuti nol, maka angka yang mendahuluinya tetap jika genap dan ditambah 1 jika ganjil.

Contoh: 8,500 dibulatkan menjadi 8

19,5 dibulatkan menjadi 20

### Soal-soal Bab 3

1. Apa yang dimaksud dengan data? Jelaskan perbedaan antara data populasi dengan data sampel.
2. Sebutkan sumber-sumber data statistik yang anda ketahui di sekitar kita
3. Berikan masing-masing 5 contoh dari data intern, ekstern, primer, sekunder, kuantitatif, kualitatif, diskrit dan kontinu.
4. Sebutkan 2 contoh variabel dan berikan penjelasan mengapa ia bisa disebut variabel.

5. Apa tujuan pembagian skala ukuran dari suatu variabel? Carilah 3 contoh data yang masuk dalam kategori skala nominal, ordinal, interval dan rasio.
6. Apa yang dimaksud dengan angka eksak dan angka pendekatan? Berikan masing-masing 3 contoh
7. Carilah batas bawah dan batas atas (angka pendekatan sebagai interval) dari bilangan-bilangan berikut:
  - a. 2,345
  - b. 47,3223
  - c. 9186, 2354879
  - d. 1,00005
8. Bulatkanlah menjadi dua digit di belakang koma dari bilangan-bilangan berikut:
  - a. 12,345
  - b. 273,1355
  - c. 30,07519

*Bab 4*

## PENGUMPULAN DATA

## A. Pendahuluan

Pengumpulan data memegang peranan yang sangat penting dalam mendapatkan informasi yang akurat. Proses pengumpulan data harus dilakukan sebaik mungkin agar data yang terkumpul merupakan data yang bersifat *representatif*, valid, dan dapat dipertanggungjawabkan. Data yang diperoleh dengan cara yang benar akan menghasilkan kesimpulan yang valid. Sebaliknya data yang terkumpul dari proses yang salah akan menghasilkan kesimpulan bias yang dapat menyebabkan salah taksir dan akibat lainnya. Oleh karena itu dalam bab ini akan diberikan pembahasan penting bagaimana cara mendapatkan data yang valid sesuai dengan prosedur yang benar dalam pengumpulan data.

## B. Tahapan Pengumpulan Data

Pada tahapan pengumpulan data, bisa dibagi menjadi dua tahap yaitu tahap persiapan dan tahap pelaksanaan.

## Tahap Persiapan

- Menentukan dan merumuskan tujuan penelitian secara baik.
- Menentukan metode yang akan digunakan.
- Menentukan teknik pengumpulan data

## Soal-soal Bab 8

1. Diberikan data sebagai berikut: 2,3,4,5,6,5,4,3,4,5,6,-7,6,5,4,5,6,7,7,8,8,9,9,9,9. carilah:
  - a. Kuartil ke-1
  - b. Kuartil ke-2
  - c. Kuartil ke-3
  - d. Desil ke-3
  - e. Desil ke-5
  - f. Persentil ke-25
  - g. Persentil ke-50
  - h. Bandingkan hasil dari b, e, dan g. Berikan penjelasan!
2. Diberikan data umur pengunjung suatu pusat perbelanjaan sebagai berikut:

Umur	Frekuensi
5-9	5
10-14	10
15-19	15
20-24	10
25-29	7
30-34	3
Jumlah	50

Carilah:

- a. Kuartil ke-1
- b. Kuartil ke-2
- c. Kuartil ke-3
- d. Desil ke-3
- e. Desil ke-5
- f. Persentil ke-25
- g. Persentil ke-50
- h. Bandingkan hasil dari b, e, dan g. Berikan penjelasan!

## Bab 9

## UKURAN PENYEBARAN DATA

## A. Pendahuluan

Ukuran pemusatan data (tendensi sentral) hanya menunjukkan nilai pusat dari sebuah distribusi data dan tidak memberikan informasi apapun tentang sebaran data (*spread of the data*). Padahal jika kita hanya mengetahui nilai-nilai pusat data saja tanpa mengetahui pola sebaran data tersebut akan kehilangan informasi penting yang bisa digunakan untuk membantu mengambil keputusan.

Mengapa ukuran penyebaran data penting diketahui?

Misalkan anda akan menjadi *tour guide* bagi sekelompok orang akan melakukan rekreasi ke suatu tempat menggunakan dua buah mobil (masing-masing mobil berisi 6 orang). Informasi yang ada hanyalah rata-rata umur penumpang, dimana rata-rata umur penumpang mobil A 45 tahun dan rata-rata umur penumpang mobil B 25 tahun. Mungkin tanpa pikir panjang anda akan memilih mobil B dengan asumsi umur penumpangnya relatif lebih muda-muda dari penumpang mobil A. Setelah anda masuk mobil B ternyata penumpangnya berumur 1th, 49th, 2th, 48th, 1th, 49th dan 25th. Sedangkan penumpang mobil A berturut-turut adalah 26th, 34th, 35th, 33th, 25th, 27th. Sejenak anda mungkin tertegun menyadari telah salah mengambil keputusan yang diambil hanya berdasarkan informasi rata-rata saja.

Artinya, keputusan yang diambil berdasarkan informasi "rata-rata" (ukuran pemusatan data) bisa menyesatkan tanpa ada informasi tambahan yang berupa ukuran penyebaran data.

Ukuran penyebaran data (dispersi atau variabilitas) akan membantu mengetahui sejauh mana nilai-nilai yang diamati "menyebar" di sekitar pusat itu. Dengan kata lain, ukuran penyebaran data akan mengukur seberapa jauh jarak data yang satu dengan data lainnya (umumnya diukur dari nilai-nilai pusatnya, misalnya rata-rata).

Sifat-sifat ukuran penyebaran:

- jika semua data memiliki nilai yang sama maka nilai-nilai pusatnya (rata-rata) juga sama, hal ini mengakibatkan nilai dispersinya nol;
- jika data-data yang diamati cukup "berdekatan" maka data-datanya juga cukup "dekat" dengan nilai rata-rata, hal ini mengakibatkan nilai dispersinya rendah/kecil (tapi lebih besar dari nol), dan
- jika data-data yang diamati jaraknya cukup "berjauhan" maka data-datanya juga "jauh" dari rata-rata, hal ini akan mengakibatkan nilai dispersi tinggi/besar.

Ukuran penyebaran data disebut juga ukuran dispersi atau ukuran penyimpangan yaitu ukuran yang menggambarkan seberapa jauh suatu data menyimpang dari rata-ratanya (*an average*).

Kepentingan mengukur dispersi di sekitar rata-rata adalah untuk mengetahui pola penyebaran data terhadap rata-ratanya. Misalkan pada saat mengukur penyebaran nilai ujian siswa di sekitar rata-rata. Hal ini penting diketahui untuk mendapatkan informasi seberapa merata kemampuan

setiap siswa. Jika dispersinya kecil memberi makna bahwa kemampuan siswa cenderung merata dan sebaliknya.

Seringkali kita mendapati beberapa kelompok/populasi yang memiliki rata-rata sama namun tingkat dispersinya berbeda. Misalkan rata-rata pendapatan perkapita di daerah X dan Y sama, namun tingkat dispersinya berbeda. Jika tingkat dispersi pendapatan di daerah X lebih besar tingkat dispersinya, maka jurang perbedaan pendapatan antara si kaya dan si miskin di daerah X sangat timpang dibanding daerah Y.

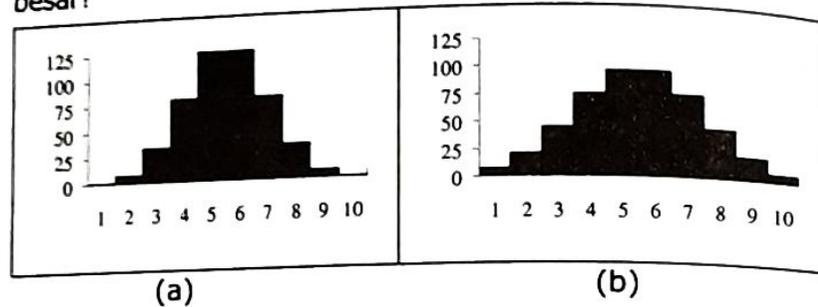


Setidaknya terdapat empat syarat ukuran dispersi dapat dikatakan baik, yaitu:

- Dihitung berdasarkan semua data dalam distribusi
- Bisa menjelaskan deviasi rata-rata
- Nilainya semakin meningkat jika distribusi data lebih heterogen (lebih beragam)
- Mudah dalam menghitung dan menafsirkan

Sebelum membahas lebih jauh tentang ukuran dispersi, ada baiknya kita memahami secara sekilas arti dari dispersi dengan menjawab pertanyaan visual berikut.

Manakah dari distribusi skor memiliki dispersi yang lebih besar?



Dari dua bentuk distribusi di atas, yang memiliki dispersi lebih besar adalah gambar (b) karena datanya lebih menyebar. Artinya, data-datanya kurang mirip (*similar*) satu sama lain.

### B. Jangkauan (*Range*)

Jangkauan merupakan ukuran penyebaran data yang paling sederhana.

Jangkauan dihitung dengan cara mencari jarak antara data/skor yang terendah (*lowest score*) dan data/skor tertinggi (*highest score*). Jangkauan digunakan sebagai ukuran, apabila dalam waktu yang sangat singkat kita ingin memperoleh gambaran tentang penyebaran data yang sedang kita selidiki, dengan mengabaikan faktor ketelitian atau kecermatan. Untuk mendapatkan jangkauan kita hanya memerlukan dua data yaitu data terbesar dan data terkecil. Rumus diperoleh dengan mengurangkan data terbesar dengan data terkecil atau:

$$J = \text{data terbesar} - \text{data terkecil}$$

Jika data terbesar dilambangkan dengan  $X_{\max}$  dan data terkecil dilambangkan dengan  $X_{\min}$  maka jangkauan menjadi:

$$J = X_{\max} - X_{\min}$$

Contoh 8.1:

Berat Badan (dalam kg): 50, 40, 60  $J$  = 20 kg

Berat Badan (dalam kg): 55, 30, 65, 40  $J$  = 35 kg

Berat Badan (dalam kg): 25, 40, 60, 40, 85  $J$  = 60 kg

Jika datanya berbentuk kelompok maka rumus jangkauan adalah:

Jangkauan = titik tengah kelas tertinggi - titik tengah kelas terendah

Contoh 8.2:

Data	Frekuensi	Titik Tengah
11-15	3	13
16-20	6	18
21-25	11	23
26-30	7	28
Jumlah	27	

Maka jangkauan dari data di atas adalah  $28 - 13 = 15$

Kebaikan dari jangkauan adalah bahwa dengan menggunakan jangkauan, dalam waktu singkat dapat diperoleh gambaran umum mengenai luas penyebaran data yang sedang kita hadapi. Namun demikian, jangkauan juga memiliki beberapa kelemahan yaitu:

- Akan sangat tergantung kepada nilai-nilai ekstrimnya. Dengan kata lain, besar-kecilnya jangkauan akan sangat ditentukan oleh nilai terendah dan nilai tertinggi yang terdapat dalam distribusi data, sehingga dengan demikian jangkauan sifatnya sangat labil dan kurang teliti.
- Jangkauan sebagai ukuran penyebaran data tidak memperhatikan distribusi yang terdapat didalam jangkauan itu sendiri.

Contoh 8.3 (kelemahan Jangkauan)

Data X:  $H=80, L=30 \quad J=80-30=50$

Data Y:  $H=95, L=45 \quad J=95-45=50$

Data Z:  $H=88, L=38 \quad J=88-38=50$

Contoh 8.4 (kelemahan jangkauan)

Apabila nilai tertinggi & terendah yang diperoleh oleh 8 mahasiswa adalah 80 dan 40, sehingga jangkauannya  $80-40=40$ . Dengan range 40 itu ada kemungkinan distribusi nilai itu adalah:

40,47,52,59,64,67,70,80 mungkin

40,40,40,40,40,40,40,80 mungkin juga

40,40,50,50,60,60,80,80 atau bentuk yang lainnya, yang jelas dengan mengetahui jangkauannya saja kita belum tahu secara pasti bagaimana sebenarnya distribusi data tersebut.

### C. Simpangan Kuartil (*Interquartile Range*)

Simpangan kuartil disebut juga jangkauan antar kuartil. Simpangan kuartil hanya mempertimbangkan bagian tengah (50%) dari distribusi data. Kuartil kedua membagi distribusi menjadi dua bagian (sama dengan median). Jadi, simpangan kuartil adalah kuartil ketiga minus kuartil pertama, yang

memberikan sebuah nilai diantara 50% data bagian tengah. Keterbatasan simpangan kuartil karena hanya didasarkan pada dua nilai yaitu Q3 dan Q1 saja. Adapun rumus dari simpangan kuartil adalah:

$$SK = \frac{Q_3 - Q_1}{2}$$

Contoh 9.1

Diketahui data 40,47,52,59,64,67,70, maka  $Q_3 = 67$  dan  $Q_1 = 47$  sehingga  $SK = (67 - 47)/2 = 10$

### D. Simpangan Persentil

Simpangan persentil disebut juga jangkauan antar persentil. Simpangan persentil dirumuskan sebagai persentil ke-90 dikurangi persentil ke-10 dibagi dua.

$$SP = \frac{P_{90} - P_{10}}{2}$$

Contoh 9.2

Jika diketahui data 40,43,45,47,50,52,55,57,59,60 maka  $P_{90} = 59$  dan  $P_{10} = 43$  sehingga  $SP = (59-43)/2 = 8$

### E. Simpangan rata-rata

Simpangan rata-rata menunjukkan seberapa menyimpang suatu data dari rata-ratanya. Kelemahan dari simpangan rata-rata adalah dalam menjumlahkan deviasi masing-masing skor atau deviasi masing-masing interval tersebut, tanda-tanda dalam deviasi diabaikan, dengan kata lain semua deviasi kita tandakan "plus", karena yang dijumlahkan hanya harga mutlak.

Untuk data tunggal, simpangan rata-rata dirumuskan dengan:

$$SR = \frac{\sum_{i=1}^n |x_i - \bar{x}|}{n}$$

Keterangan:

$x_i$  = data ke-I  
 $\bar{x}$  = rata-rata  
 $n$  = banyaknya data

Contoh 9.3

Diketahui data Berat badan (BB) dalam kilogram sebagai berikut:

50, 40, 60. Carilah simpangan rata-ratanya.

Penyelesaian:

Rata-rata = 50 kg, sehingga:

$$SR = \frac{\sum_{i=1}^n |x_i - \bar{x}|}{n} = \frac{|50 - 50| + |40 - 50| + |60 - 50|}{3} = \frac{0 + 10 + 10}{3} = \frac{20}{3} = 6,667$$

Untuk data kelompok menggunakan rumus:

$$SR = \frac{\sum_{i=1}^n f_i |x_i - \bar{x}|}{\sum_{i=1}^n f_i}$$

Keterangan:  $x_i$  adalah titik tengah kelas

Contoh 9.4

Carilah simpangan rata-rata dari data berikut:

Nilai	$f_i$	$x_i$	$ x_i - \bar{x} $	$f_i  x_i - \bar{x} $
35 - 39	5	37	11,78	58,9
40 - 44	15	42	6,78	101,7
45 - 49	20	47	1,78	35,6
50 - 54	15	52	3,22	48,3
55 - 59	10	57	8,22	82,2
60 - 64	5	62	13,22	66,1
Jumlah	70			392,8

Penyelesaian:

Sudah diketahui rata-ratanya adalah 48,78, maka simpangan rata-ratanya adalah:

$$SR = \frac{\sum_{i=1}^n f_i |x_i - \bar{x}|}{\sum_{i=1}^n f_i} = \frac{392,8}{70} = 5,6$$

## F. Variansi (Variance)

Variansi merupakan jumlah kuadrat semua deviasi nilai-nilai individual terhadap rata-rata kelompok. Dalam rangka mengatasi kelemahan simpangan rata-rata, Karl Pearson (ahli statistika) memberi jalan keluar sebagai berikut:

- Semua simpangan (baik *plus* atau *minus*) dikuadratkan lebih dulu, dengan demikian yang bertanda plus akan bertanda plus, dan minus akan bertanda plus karena dikuadratkan.
- Setelah simpangan dikuadratkan dan bertanda plus lalu dijumlahkan. Dicari rata-rata dan akhirnya

Secara aljabar, cara ini lebih baik (dalam usaha untuk menghilangkan tanda minus) dibandingkan dengan menggunakan harga mutlak (*absolut value*), seperti pada

kasus simpangan rata-rata. Dengan mengkuadratkan semua penyimpangan, jumlahkan, dan membagi dengan N, maka akan diperoleh varians. Simbol untuk varians sampel adalah  $s^2$ , sedangkan untuk varians dalam populasi menggunakan huruf Yunani  $\sigma^2$ .

Variansi untuk data tunggal

Dirumuskan dengan :

$$\text{var} = s^2 = \frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n-1}$$

Contoh 9.5

Contoh:

Diketahui data berat badan (kg): 50, 40, 60 dengan rata-rata = 50 kg, maka variansi dari data tersebut dapat dicari dengan cara:

$$\begin{aligned} \text{var} &= \frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n-1} = \frac{(50-50)^2 + (40-50)^2 + (60-50)^2}{3-1} \\ &= \frac{0+100+100}{3-1} = \frac{200}{2} = 100 \text{ kg} \end{aligned}$$

Contoh 9.6

Apabila ada 10 data yaitu: 5, 5, 6, 7, 8, 8, 9, 10, 11, 11 dengan rata-rata 8, maka variansinya adalah :

$$s^2 = \frac{[(5-8)^2 + (5-8)^2 + (6-8)^2 + (7-8)^2 + (8-8)^2 + (8-8)^2 + (9-8)^2 + (10-8)^2 + (11-8)^2 + (11-8)^2]}{10-1}$$

$$s^2 = \frac{9+9+4+1+1+4+9+9}{9}$$

$$s^2 = \frac{46}{9}$$

$$s^2 = 5\frac{1}{9}$$

Variansi untuk data kelompok

Dirumuskan dengan :

$$\text{var} = s^2 = \frac{\sum_{i=1}^n f_i (x_i - \bar{x})^2}{\sum_{i=1}^n f_i}$$

Contoh 9.7

Nilai	f <sub>i</sub>	X <sub>i</sub>	(x <sub>i</sub> - $\bar{x}$ ) <sup>2</sup>	f <sub>i</sub> (x <sub>i</sub> - $\bar{x}$ ) <sup>2</sup>
35 - 39	5	37	138,77	693,84
40 - 44	15	42	45,97	689,53
45 - 49	20	47	3,17	63,37
50 - 54	15	52	10,37	155,53
55 - 59	10	57	67,57	675,68
60 - 64	5	62	174,77	873,84
Jumlah	70		440,61	3151,79

Jadi variansinya adalah:

$$\text{var} = s^2 = \frac{\sum_{i=1}^n f_i (x_i - \bar{x})^2}{\sum_{i=1}^n f_i} = \frac{3151,79}{70} = 45,03$$

G. Simpangan Baku (Standar Deviasi)

Disebut standar deviasi karena deviasi rata-rata yang tadinya memiliki kelemahan, telah dibakukan atau distandarisasikan, sehingga lebih mantap, oleh karena itu dalam dunia analisis statistik, standar deviasi memiliki kedudukan yang sangat penting.

Merupakan akar dari variansi dan dirumuskan dengan:

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

### Mengapa harus dibagi n-1?

Pembagian dengan n-1 hanya berlaku pada sampel. Sebab membagi dengan n-1 akan meningkatkan nilai standar deviasi. Ini dilakukan karena sampel tidak memiliki varians (keragaman) yang sama persis dengan populasi. Sehingga dalam sampel pembagiannya menggunakan derajat bebas (N-1). Sebagai ilustrasi, perhatikan barisan bilangan berikut: 1/100, 1/10, 1/5, 1/4, 1/3, 1/2. Semakin kecil penyebutnya, maka bilangan tersebut akan semakin bertambah besar.

Derajat bebas adalah jumlah data yang bebas bervariasi sebelum yang terakhir ditentukan. Misalkan kita mempunyai tiga data, sebutlah X1, X2, dan X3. Kemudian kita telah mengetahui rata-ratanya, katakanlah 100. Maka nilai X1, X2, dan X3 yang bebas bervariasi hanya dua. Artinya, jika kita misalkan secara bebas memilih X1 = 150, X2 = 50, maka X3 sudah tidak bebas lagi. Itulah sebabnya pada beberapa analisis yang hanya bertujuan menduga satu parameter, derajat bebasnya adalah (n-1).

#### Contoh 9.8

Untuk data berat badan (kg): 50, 40, 60 dengan rata-rata = 50 kg diperoleh simpangan baku:

$$SD = \sqrt{\frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n-1}} = \sqrt{\frac{(50-50)^2 + (40-50)^2 + (60-50)^2}{3}} \\ = \sqrt{\frac{0+100+100}{3-1}} = \sqrt{\frac{200}{2}} = 10 \text{ kg}$$

#### Contoh 9.9

Sedangkan dari contoh 7.6, maka simpangan bakunya dapat ditentukan dengan mudah yaitu hanya dengan mencari akar positif dari variansi. Sehingga diperoleh:

$$s = \sqrt{\frac{46}{9}} = \frac{\sqrt{46}}{3}$$

### Kegunaan Standar Deviasi

Standar deviasi dan Mean banyak digunakan dalam dunia pendidikan dan penelitian. Dalam dunia pendidikan banyak digunakan pada evaluasi pendidikan, sedang dalam penelitian biasanya dalam perubahan suatu skor ke skor lainnya.

Secara rinci kegunaan itu adalah pada hal-hal berikut :

1. Untuk menetapkan Nilai Batas Lulus Aktual (NBLA), yaitu nilai batas kelulusan .  
Dengan rumus NBLA = Mean + 0,25 SD.....(7.17)
2. Untuk Mengubah *Raw Score* (Skor Mentah) ke dalam nilai standar Sekala 5 atau nilai huruf A, B, C, D, E. Dengan patokan :  
Nilai A: Mean + 1,5 SD ke atas  
Nilai B: Mean + 0,5 SD sampai Mean + 1,5 SD  
Nilai C: Mean -0,5 SD sampai Mean + 0,5 SD  
Nilai D: Mean -1, 5 SD sampai Mean - 0,5 SD  
Nilai E: Mean - 1,5 SD kebawah
3. Untuk Mengubah raw score menjadi nilai sekala sebalas (Standar eleven/STANEL)
4. Untuk mengubah data menjadi tiga kelompok secara ordinal. Misalnya mengelompokkan anak didik ke dalam 3 kelompok, yaitu kelompok pandai (peringkat atas), kelompok sedang ( peringkat bawah, dan kelompok kurang (peringkat bawah).  
Patokannya adalah :

Rangking atas: untuk yang bernilai di atas  $M + 1,5 SD$

Rangking Tengah untuk yang bernilai antara  $M-1,5 SD$  sampai  $M+1,5SD$  Rangking bawah untuk yang bernilai dibawah  $M-1,5 SD$ .

Selain itu standar deviasi sangat berguna untuk:

1. hal-hal yang berkaitan dengan kurva normal,
2. sebagai indeks variabilitas, artinya semakin besar standar deviasi maka distribusi data menjadi lebih bervariasi atau heterogen. Sebuah distribusi yang tidak memiliki dispersi akan memiliki standar deviasi 0. Contohnya adalah ketika setiap anggota tentara dalam satu peleton memiliki tinggi badan yang sama. Jadi, nilai terendah dari standar deviasi adalah 0, akan tetapi untuk nilai teringginya tidak terbatas.
3. untuk membandingkan satu distribusi dengan yang lain. Dapat digunakan untuk mengetahui bahwa satu daerah yang lebih homogen daripada daerah yang lain (pada variabel tertentu, misal pendapatan perkapita). Atau juga bisa digunakan untuk mengetahui bahwa siswa SMA X lebih homogen dalam hal nilai UAN daripada siswa SMA Y atau SMA Z.

#### H. Koefisien Variasi

Koefisien variasi termasuk ukuran dispersi relatif. Koefisien variasi digunakan untuk mengetahui variasi terkecil dari suatu distribusi data. Bisa juga digunakan untuk membandingkan dua atau lebih distribusi frekuensi. Terdapat beberapa cara dalam menentukan koefisien variasi dan semuanya dapat digunakan. Diantara beberapa cara tersebut akan diuraikan berikut.

$$\text{Rumus 1. } KV = \frac{SB}{X} \cdot 100\%$$

$$\text{Rumus 2. } KV = \frac{\text{Jangkauan}}{X} \cdot 100\%$$

$$\text{Rumus 3. } KV = \frac{SR}{X} \cdot 100\%$$

$$\text{Rumus 4. } KV = \frac{SK}{X} \cdot 100\%$$

$$\text{Rumus 5. } KV = \frac{Q3 - Q1}{Q3 + Q1} \cdot 100\%$$

#### Soal-soal Bab 9

1. Apabila diketahui data nilai ujian STATISTIKA I dari 15 orang mahasiswa yaitu: 45,47,53,54,55,58,58,59,62,-64,64, 64,65,70 dan 72.

Carilah

- a. Jangkauan
- b. simpangan rata-rata
- c. variansi
- d. simpangan baku

2. Jika nilai ujian mata kuliah "Matematika Keuangan" dikelompokkan dari seluruh mahasiswa tingkat I yang berjumlah 150 orang dalam 6 kelas interval, hasilnya adalah sebagai berikut :

Kelas Interval	Frekuensi
31 - 40	5
41 - 50	10
51 - 60	35
61 - 70	50
71 - 80	30
81 - 90	20

Carilah:

- Jangkauan
  - simpangan rata-rata
  - variansi
  - simpangan baku
3. Harga 5 mobil bekas masing-masing adalah Rp.5 Juta, Rp.5,5 Juta, Rp. 7 Juta, Rp.5 Juta serta Rp 4,25 Juta dan harga 5 ayam masing-masing Rp. 65.000, Rp.80.000, Rp.90.000, Rp45.000 dan Rp.15.000. Hitunglah simpangan baku harga mobil dan harga ayam. Manakah yang lebih bervariasi, harga mobil atau harga ayam?
4. Berat badan dua kelompok pasien di sebuah rumah sakit telah ditimbang. Kelompok pertama terdiri dari 10 anak berumur antara 1-15 th, sedangkan kelompok kedua terdiri dari 10 orang dewasa berumur antara 18-40 th dan menghasilkan data berikut:

Kelompok 1	Kelompok 2
11	31
13	43
15	45
15	51
17	54
21	61
22	66
28	73
32	75
34	83

Carilah koefisien variasinya menggunakan 5 rumus di atas. Dari hasil tsb apakah hasilnya konsisten. Berat badan kelompok manakah yang lebih bervariasi?

*Bab 10*

## SKEWNESS DAN KURTOSIS

## A. Pendahuluan

Dalam bab sebelumnya, Anda telah belajar kapan dan bagaimana menghitung beberapa ukuran tendensi sentral, ukuran dispersi dan ukuran letak. Bab ini membahas dua sifat penting dari suatu data yaitu kemiringan dan kurtosis. Ukuran kemiringan mewakili tingkat kesimetrisan data sedangkan kurtosis mewakili tingkat keruncingan data. Secara konseptual, kemiringan menggambarkan sisi mana dari suatu distribusi memiliki ekor panjang. Jika ekor panjang di sebelah kanan, maka kemiringan tersebut ke kanan atau positif, jika ekor panjang di sebelah kiri, maka kemiringan tersebut ke kiri atau negatif.

Kemiringan ke kanan adalah sesuatu yang umum terjadi ketika variabel yang dikaji memiliki batas di sebelah kiri tapi tak terbatas di sebelah kanan. Misalnya, durasi waktu respon, biasanya kurva yang terbentuk akan "miring ke kanan", karena tidak dapat mengambil nilai yang kurang dari nol; variabel keuangan (pendapatan, kekayaan, harga) biasanya memiliki kemiringan positif karena jarang mengambil nilai yang kurang dari nol, dan berat badan tubuh orang dewasa memiliki kemiringan positif karena sebagian besar orang lebih dekat dengan batas bawah daripada batas atas berat badan yang pada umumnya. Kemiringan negatif (kiri) adalah kurang umum dalam praktek, tetapi bisa terjadi bila variabel cenderung lebih dekat dengan maksimum daripada nilai

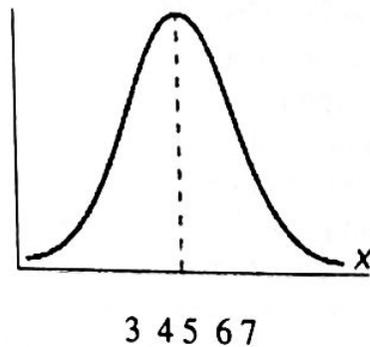
minimum. Sebagai contoh, nilai pada ujian yang soalnya mudah cenderung memiliki kemiringan kiri, dengan skor yang paling mendekati 100% dan nilai yang lebih rendah akan mengekor ke kiri.

**B. Skewness**

Kemiringan (*skewness*) berarti ketidaksimetrisan. Di beberapa literatur, *skewness* disebut dengan "kemencengan" atau "kecondongan". Sebuah distribusi dikatakan simetris apabila nilai-nilainya tersebar merata disekitar nilai rata-ratanya. Sebagai contoh, distribusi data berikut simetris terhadap nilai rata-ratanya, yaitu 5.

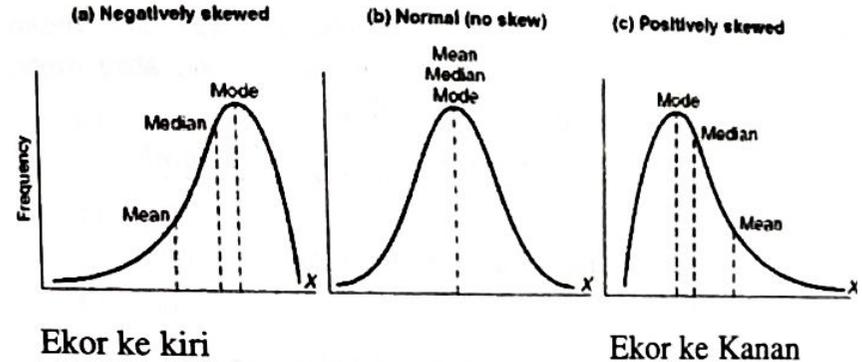
Data	Frekuensi
3	3
4	5
5	7
6	5
7	3

Jika digambarkan dalam grafik (histogram) diperoleh:



Data yang terdistribusi secara simetris maka kemiringannya sama dengan nol (not skewed,  $Sk=0$ ).

Kemiringan juga dapat dikatakan sebagai tingkat pergeseran suatu distribusi dari titik simetrinya. Terdapat tiga model kemiringan dari suatu distribusi data yaitu: miring ke kiri (*negatively skewed*), miring ke kanan (*positively skewed*) dan normal (*symmetric/no skew*). Sebuah distribusi dikatakan miring ke kanan jika memiliki "ekor" yang ditarik ke arah kanan (positif). Sedangkan sebuah distribusi disebut miring ke kiri jika memiliki "ekor" yang ditarik ke arah kiri (negatif).



Gambar 10.1 Tiga model kemiringan

Sumber:

[http://www.southalabama.edu/coe/bset/johnson/lectures/lec15\\_files/image014.jpg](http://www.southalabama.edu/coe/bset/johnson/lectures/lec15_files/image014.jpg)

**C. Menghitung Indeks Kemiringan (Skewness Index)**

Ada banyak metode untuk menghitung indeks kemiringan. Tidak semua program komputer menghitung kemiringan dengan cara yang sama. Jika Anda menggunakan

program komputer untuk memperoleh indeks kemiringan dan keruncingan, pastikan Anda tahu bagaimana menghitungnya.

Beberapa cara menghitung kemiringan yang dibahas disini adalah metode Pearson, metode Bowley, metode persentil, dan metode momen. Selengkapnya akan dibahas berikut ini.

### 1. Metode Pearson

Kesimetrisan suatu distribusi dengan cepat dapat diduga, apabila mean, median atau modusnya telah diketahui, yakni dengan menggunakan/melihat koefisien kemiringan (*skewness coefficient*). Menurut Pearson, menghitung indeks kemiringan cukup dengan menghitung tiga kali mean dikurangi median dibagi dengan standar deviasi, atau mean dikurangi modus dibagi standar deviasi.

$$\alpha = \frac{\bar{X} - \text{Mod}}{S} \quad \text{atau} \quad \alpha = \frac{3(\bar{X} - \text{Med})}{S}$$

$\alpha$  = derajat kemiringan Pearson

Kriteria :

1.  $\alpha = 0$ , maka distribusi datanya simetri
2.  $\alpha < 0$ , maka distribusi datanya miring ke kiri
3.  $\alpha > 0$ , maka distribusi datanya miring ke kanan

#### Contoh 10.1

Dari data 3, 3, 3, 4, 4, 5, 5 tentukanlah kemiringannya.

Dari data tersebut diperoleh:

- rata-rata ( $\bar{x}$ ) = 4,142857;
- median = 4;
- modus = 3; dan simpangan baku  $S = 1,46385$ .

Sehingga menurut Pearson diperoleh kemiringan:

$$\alpha = \frac{\bar{x} - \text{Mod}}{S} = \frac{4,142857-3}{1,46385} = 0,78072$$

atau

$$\alpha = \frac{3(\bar{x} - \text{Med})}{S} = \frac{3(4,142857-4)}{1,46385} = 0,29277$$

Kesimpulannya: distribusi data sedikit miring ke kanan

### 2. Metode Bowley (kuartil)

$$\alpha = \frac{Q_3 + Q_1 - Q_2}{Q_3 - Q_1}$$

- a) Jika  $Q_3 - Q_2 = Q_2 - Q_1$  atau  $Q_3 + Q_1 - 2Q_2 = 0$  maka  $\alpha = 0$  dan distribusi datanya simetri
- b) Jika  $Q_1 = Q_2$  maka  $\alpha = 1$  dan distribusi datanya miring ke kanan
- c) Jika  $Q_2 = Q_3$  maka  $\alpha = -1$  dan distribusi datanya miring ke kiri

#### Contoh 10.2.

Jika data pada contoh soal 10.1 dicari koefisien kemiringan menggunakan metode Bowley akan didapatkan:

$Q_1 = 3$ ,  $Q_2 = 4$  dan  $Q_3 = 5$ . Sehingga

$$\alpha = \frac{Q_3 + Q_1 - Q_2}{Q_3 - Q_1} = \frac{5+3-4}{5-3} = \frac{4}{2} = 2 \quad \text{berarti distribusi}$$

datanya miring ke kanan

### 3. Metode Persentil

$$\alpha = \frac{(P_{90} - P_{80}) - (P_{80} - P_{10})}{P_{90} - P_{10}}$$

## Contoh 10.3

Data jumlah cerita buku yang dimiliki oleh 10 siswa SD kelas lima adalah sebagai berikut: 7,8,8,9,9,11,12,15,17,17. Bagaimana kemiringan datanya jika menggunakan metode persentil?

$$\alpha = \frac{(P_{90} - P_{10}) - (P_{80} - P_{20})}{P_{90} - P_{10}} = \frac{(17-15) - (15-7)}{15-7} = \frac{-10}{8} = -1,25 \rightarrow \text{miring ke kiri}$$

## 4. Metode Momen

Untuk data tunggal (belum dikelompokkan) maka rumus yang digunakan untuk mencari kemiringan adalah

$$\alpha_3 = \frac{\sum (x_i - \bar{x})^3}{nS^3}$$

Sedangkan untuk data kelompok digunakan formula berikut

$$\alpha_3 = \frac{\sum f(x_i - \bar{x})^3}{\sum fS^3}$$

1. Jika  $\alpha_3 = 0$ , maka distribusi datanya simetri
2. Jika  $\alpha_3 < 0$ , maka distribusi datanya miring kiri
3. Jika  $\alpha_3 > 0$ , maka distribusi datanya miring kanan

## Contoh 10.4

Jika diketahui data 11 12 14 15 16 17 20, carilah kemiringannya menggunakan metode momen!

Mean = 15 dan simpangan baku = 3,05505

$$\begin{aligned} \alpha_3 &= \frac{\sum (x_i - \bar{x})^3}{nS^3} \\ &= \frac{(11-15)^3 + (12-15)^3 + (14-15)^3 + (15-15)^3 + (16-15)^3 + (17-15)^3 + (20-15)^3}{7 \times 3,05505^3} \\ &= \frac{42}{199,5966} \\ &= 0,210424 \end{aligned}$$

## Contoh 10.5

Berapa koefisien kemiringan dari data frekuensi dibawah ini:

Usia Wanita	Frekuensi
10-19	4
20-29	7
30-39	13
40-49	5
50-60	3
Jumlah	32

Buat tabel perhitungan sebagai berikut:

Usia Wanita	Frek	Xi	fixi	(xi-mean) <sup>2</sup>	f <sub>i</sub> (xi-mean) <sup>2</sup>	f <sub>i</sub> (xi-mean) <sup>3</sup>
10-19	4	14,5	58	351,563	1406,25	-6591,8
20-29	7	24,5	171,5	76,5625	535,938	-669,92
30-39	13	34,5	448,5	1,5625	20,3125	1,95313
40-49	5	44,5	222,5	126,563	632,813	1423,83
50-60	3	54,5	163,5	451,563	1354,69	9595,7
Jumlah	32		1064		3950	3759,77

Dari hasil perhitungan diperoleh:

Mean = 33,25, simpangan baku = 11,11024 dan diperoleh kemiringan sebesar:

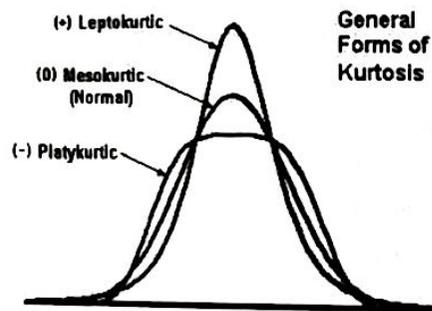
$$\alpha_3 = \frac{\sum f(x_i - \bar{x})^3}{\sum fS^3} = \frac{3759,77}{32 \times 11,11024^3} = 0,085672 \rightarrow \text{sedikit miring ke kanan}$$

#### D. Kurtosis (keruncingan)

Kurtosis adalah tingkat keruncingan dari suatu distribusi. Sebuah distribusi normal adalah distribusi mesokurtik. Sebuah distribusi leptokurtik mumi memiliki puncak yang lebih tinggi dibandingkan distribusi normal dan memiliki ekor yang lebih berat. Sebuah distribusi platikurtik mumi memiliki puncak lebih rendah dari distribusi normal dan ekor lebih ringan.

Ada tiga jenis derajat keruncingan:

1. Leptokurtis (Puncak Runcing)
2. Mesokurtis (Puncak Normal)
3. Platikurtis (Puncak Rendah)



Gambar 10.2 Bentuk-bentuk keruncingan

#### E. Perhitungan Kurtosis

Dalam menghitung kurtosis atau keruncingan didasarkan pada beberapa formula. Dalam buku ini akan dibahas dua cara yang cukup populer yaitu:

#### 1. Metode persentil

Kurtosis pada metode persentil dilambangkan dengan  $\kappa$  (baca: kappa), dengan rumus:

$$\kappa = \frac{SK}{P90 - P10} = \frac{\frac{1}{2}(Q3 - Q1)}{P90 - P10}$$

#### Contoh 10.6

Data diambil dari contoh 10.3. (7,8,8,9,9,11,12,-15,17,17) Keruncingan-nya dihitung sebagai berikut:

$$\kappa = \frac{\frac{1}{2}(Q3 - Q1)}{P90 - P10} = \frac{\frac{1}{2}(15 - 8)}{17 - 7} = \frac{3,5}{10} = 0,35$$

#### 2. Metode Momen

Untuk mendapatkan koefisien keruncingan dengan menggunakan metode momen digunakan rumus berikut:

$$\alpha_4 = \frac{\sum (x_i - \bar{x})^4}{nS^4}$$

Kriteria keruncingan untuk model ini adalah:

- Jika  $\alpha_4 = 3 \rightarrow$  mesokurtik
- Jika  $\alpha_4 > 3 \rightarrow$  leptokurtik
- Jika  $\alpha_4 < 3 \rightarrow$  platikurtik

#### Contoh 10.7

Dari soal pada contoh 10.4 diperoleh koefisien kurtosis sebesar

$$\alpha_4 = \frac{\sum (x_i - \bar{x})^4}{nS^4} = \frac{349387}{487578,1} = 1,395524 \rightarrow \text{platikurtik}$$

## Bab 11

## TEKNIK SAMPLING

## A. Pendahuluan

Teknik sampling adalah strategi yang diterapkan oleh peneliti selama proses pengambilan sampel. Proses ini dilakukan ketika para peneliti bertujuan untuk menarik kesimpulan untuk seluruh populasi setelah melakukan penelitian pada sampel yang diambil dari populasi yang sama. Jika sampling dilakukan dengan metode yang tepat, analisis statistik dari suatu sampel dapat digunakan untuk menggeneralisasikan keseluruhan populasi. Agar hasil penelitian yang dilakukan terhadap sampel masih tetap bisa dipercaya, dalam artian masih bisa mewakili karakteristik populasi, maka cara penarikan sampelnya harus dilakukan secara seksama. Cara pemilihan sampel inilah yang kemudian dikenal dengan nama teknik sampling atau teknik pengambilan sampel. Secara konseptual, teknik sampling adalah kaidah-kaidah dalam menentukan besarnya sampel dan obyek yang menjadi sampel. Jika terjadi kesalahan dalam pengambilan sampel, maka penarikan kesimpulan untuk populasi menjadi tidak akurat. Sebagai contoh, Pada tahun 2002 kita dikejutkan oleh penelitian yang menyatakan bahwa 90% mahasiswa Yogyakarta telah melakukan hubungan seksual. Setelah diteliti, ternyata peneliti dalam mengambil sampel tidak menggunakan teknik pengambilan sampel yang benar. Peneliti hanya mengambil sampel kepada obyek yang mendukung hipotesisnya.

## Syarat Sampel yang Baik

Secara umum, sampel yang baik adalah yang dapat mewakili sebanyak mungkin karakteristik populasi. Dalam

## Soal-soal Bab 10

1. Jelaskan apa yang dimaksud dengan kurva:
  - a. Model positif
  - b. Model negatif
  - c. Simetris
  - d. Mesokurtik
  - e. Leptokurtik
  - f. Platikurtik
2. Berikan contoh peristiwa yang memungkinkan bisa menghasilkan kurva model positif maupun negatif.
3. Diberikan data 2,3,4,5,4,4,5,6,6,7,5,4,6,7. Carilah kemiringan dan keruncingannya (dengan semua cara/metode)
4. Pada pengukuran kadar haemoglobin (Hb) terhadap 50 wanita hamil diperkirakan diperoleh data berikut:

Hb (gr%)	Frekuensi
07-08	3
09-10	7
11-12	17
13-14	10
15-16	5

Tentukan koefisien kemiringan pearson dan keruncingan momen-nya!

bahasa pengukuran, artinya sampel harus valid, yaitu bisa mengukur sesuatu yang seharusnya diukur. Kalau yang ingin diukur adalah masyarakat Sunda sedangkan yang dijadikan sampel adalah hanya orang Banten saja, maka sampel tersebut tidak valid, karena tidak mengukur sesuatu yang seharusnya diukur (orang Sunda).

Sampel yang valid ditentukan oleh dua pertimbangan.

**Pertama: Akurasi atau ketepatan**, yaitu tingkat ketidak-adaan "bias" (kekeliruan) dalam sampel. Dengan kata lain makin sedikit tingkat kekeliruan yang ada dalam sampel, makin akurat sampel tersebut. Tolok ukur adanya "bias" atau kekeliruan adalah populasi.

Cooper dan Emory (1995) menyebutkan bahwa "*there is no systematic variance*" yang maksudnya adalah tidak ada keragaman pengukuran yang disebabkan karena pengaruh yang diketahui atau tidak diketahui, yang menyebabkan skor cenderung mengarah pada satu titik tertentu. Sebagai contoh, jika ingin mengetahui rata-rata luas tanah suatu perumahan, lalu yang dijadikan sampel adalah rumah yang terletak di setiap sudut jalan, maka hasil atau skor yang diperoleh akan bias. Kekeliruan semacam ini bisa terjadi pada sampel yang diambil secara sistematis.

Contoh *systematic variance* yang banyak ditulis dalam buku-buku metode penelitian adalah jajak-pendapat (*polling*) yang dilakukan oleh *Literary Digest* (sebuah majalah yang terbit di Amerika tahun 1920-an) pada tahun 1936. Mulai tahun 1920, 1924, 1928, dan tahun 1932 majalah ini berhasil memprediksi siapa yang akan jadi presiden dari calon-calon presiden yang ada. Sampel diambil berdasarkan petunjuk dalam buku telepon dan dari daftar pemilik mobil. Namun pada tahun 1936 prediksinya salah. Berdasarkan jajak

pendapat, di antara dua calon presiden (Alfred M. Landon dan Franklin D. Roosevelt), yang akan menang adalah Landon, namun meleset karena ternyata Roosevelt yang terpilih menjadi presiden Amerika.

Setelah diperiksa secara seksama, ternyata *Literary Digest* membuat kesalahan dalam menentukan sampel penelitiannya. Karena semua sampel yang diambil adalah mereka yang memiliki telepon dan mobil, akibatnya pemilih yang sebagian besar tidak memiliki telepon dan mobil (kelas rendah) tidak terwakili, padahal Rosevelt lebih banyak dipilih oleh masyarakat kelas rendah tersebut. Dari kejadian tersebut ada dua pelajaran yang diperoleh: (1) keakuratan prediktibilitas dari suatu sampel tidak selalu bisa dijamin dengan banyaknya jumlah sampel, (2) agar sampel dapat memprediksi dengan baik populasi, sampel harus mempunyai selengkap mungkin karakteristik populasi.

**Kedua: Presisi.** Kriteria kedua, sampel yang baik adalah memiliki tingkat presisi estimasi. Presisi mengacu pada persoalan **sedekat mana estimasi kita dengan karakteristik populasi**. Contoh: Dari 300 pegawai konveksi, diambil sampel 50 orang. Setelah diukur ternyata rata-rata perhari, setiap orang menghasilkan 10 potong celana panjang. Namun berdasarkan laporan harian, pegawai bisa menghasilkan per harinya rata-rata 14 unit. Artinya di antara laporan harian yang dihitung berdasarkan populasi dengan hasil penelitian yang dihasilkan dari sampel, terdapat perbedaan 4 unit. Makin kecil tingkat perbedaan di antara rata-rata populasi dengan rata-rata sampel, maka makin tinggi tingkat presisi sampel tersebut.

Namun yang harus kita pahami bahwa "belum pernah ada sampel yang bisa mewakili karakteristik populasi

sepenuhnya". Oleh karena itu dalam setiap penarikan sampel senantiasa melekat keasalahan-kesalahan, yang dikenal dengan nama *sampling error*. Presisi diukur oleh simpangan baku (*standard error*). Makin kecil perbedaan di antara simpangan baku yang diperoleh dari sampel (S) dengan simpangan baku dari populasi ( $\sigma$ ), makin tinggi pula tingkat presisinya. Meskipun tidak selalu, tingkat presisi mungkin bisa meningkat dengan cara menambahkan jumlah sampel, karena kesalahan mungkin bisa berkurang kalau jumlah sampelnya ditambah. Dengan contoh di atas tadi, mungkin saja perbedaan rata-rata di antara populasi dengan sampel bisa lebih sedikit, jika sampel yang ditariknya ditambah. Katakanlah dari 50 sampel menjadi 75 sampel.

## B. Teknik Penentuan Ukuran Sampel

Ukuran sampel adalah jumlah sampel minimal yang harus diambil dari populasi agar sampel dikatakan *representatif* (mewakili). Penentuan ukuran sampel tergantung dari populasinya yakni dilihat dari homogenitas populasi dan besar populasi. Jika populasi sangat homogen dan ukuran sangat kecil, maka bisa diambil sampel dengan ukuran kecil tetapi jika populasi heterogen dan ukurannya besar, maka diperlukan sampel dalam jumlah yang besar. Berikut berbagai cara dalam menentukan ukuran sampel.

### 1. Pendapat Slovin

Ketika suatu penelitian tidak mungkin melibatkan seluruh anggota populasi (misalnya populasi penduduk DKI Jakarta), maka pastilah seseorang akan mengambil sampel dengan menggunakan teknik *random sampling*. Untuk menggunakan rumus Slovin memungkinkan peneliti untuk mengambil

sampel dengan tingkat akurasi yang diinginkan. Artinya, peneliti memiliki keleluasaan tentang berapa besar ukuran sampel yang perlu diambil dengan cara menentukan besarnya tingkat kesalahan (*error*) yang diinginkan.

### Kapan Rumus Slovin digunakan?

Pada saat populasi diambil sampelnya maka harus memperhitungkan tingkat kepercayaan dan margin kesalahan. Terkadang dari sebuah populasi kita banyak mendapatkan informasi, kadang-kadang hanya sedikit informasi dan bahkan kita tidak mendapatkan informasi sama sekali. Sebagai contoh, kita mungkin tahu bahwa populasi berdistribusi normal (misalnya tinggi badan, berat badan atau IQ), kita mungkin tahu bahwa ada distribusi bimodal (seperti yang sering terjadi dengan nilai kelas di kelas matematika) atau kita mungkin tidak tahu tentang bagaimana populasi akan berperilaku (seperti pemungutan suara mahasiswa untuk mendapatkan pendapat mereka tentang budaya akademik kampus). Rumus Slovin adalah digunakan ketika informasi tentang perilaku populasi tidak diketahui sama sekali.

Slovin mengemukakan bahwa jumlah sampel minimal agar representatif adalah:

$$n = \frac{N}{1 + Np^2}$$

keterangan:

- n = jumlah sampel
- N = jumlah Populasi
- p = prosentase kelonggaran kesalahan (misal 2%)

Untuk menggunakan rumus tersebut, angka pertama yang pertama ditentukan adalah toleransi kesalahan yang

diinginkan. Sebagai contoh, pada tingkat kepercayaan 95 persen (memberikan margin kesalahan 0,05), atau mungkin memerlukan akurasi lebih ketat dari tingkat kepercayaan 98 persen (dengan *margin of error* 0,02).

#### Contoh 11.1

Misalkan terdapat 1.000 karyawan pemerintah kota dan Anda ingin melakukan survei untuk mengetahui alat-alat yang paling cocok untuk pekerjaan mereka. Misalkan diambil *margin of error* sebesar 0,05. Dengan menggunakan rumus Slovin, maka jumlah sampel yang akan disurvei adalah  $n = N/(1 + Ne^2) = 1.000/(1 + 1000 * 0,05 * 0,05) = 286$

### 2. Pendapat Gay

Menurut Gay, ukuran minimum sampel yang dapat diterima berdasarkan desain penelitian yang digunakan adalah:

- Untuk penelitian deskriptif, minimal 10% populasi. Untuk populasi relatif kecil minimal 20% populasi.
- Untuk penelitian menggunakan teknik analisis korelasional minimal 30 obyek.

#### Contoh 11.2:

Seorang peneliti ingin mengetahui tingkat kesukaan masyarakat di kota A terhadap suatu produk makanan. Untuk tujuan tersebut ia melakukan pengamatan terhadap penduduk kota tersebut. Bila jumlah penduduk kota 1000, maka sampel yang harus diambil sebanyak:

$$1000 \quad 0\% = 100 \text{ penduduk.}$$

### 3. Pendapat Kracjie

Hampir sama dengan Slovin, hanya Kracjie dalam melakukan perhitungan ukuran sampel didasarkan atas

kesalahan 5%. Jadi sampel yang diperoleh itu mempunyai kepercayaan 95% terhadap populasi. Kracjie selanjutnya membuat tabel (tabel 11.1), sehingga kita tidak perlu melakukan perhitungan yang rumit.

Tabel 11.1

Tabel Kracjie untuk Penentuan Jumlah Sampel

Populasi (N)	Sampel (n)	Populasi (N)	Sampel (n)	Populasi (N)	Sampel (n)
10	10	220	140	1200	291
15	14	230	144	1300	297
20	19	240	148	1400	302
25	24	250	152	1500	306
30	28	260	155	1600	310
35	32	270	159	1700	313
40	36	280	162	1800	317
45	40	290	165	1900	320
50	44	300	169	2000	322
55	48	320	175	2200	327
60	52	340	181	2400	331
65	56	360	186	2600	335
70	59	380	191	2800	338
75	63	400	196	3000	341
80	66	420	201	3500	346
85	70	440	205	4000	351
90	73	460	210	4500	354
95	76	480	214	5000	357

100	80	500	217	6000	361
110	86	550	226	7000	364
120	92	600	234	8000	367
130	97	650	242	9000	368
140	103	700	248	10000	370
150	108	750	254	15000	375
160	113	800	260	20000	377
170	118	850	265	30000	379
180	123	900	269	40000	380
190	127	950	274	50000	381
200	132	1000	278	75000	382
210	136	1100	285	1000000	384

### C. Teknik Pengambilan Sampel

Secara umum, terdapat dua teknik dalam mengambil sampel. Yaitu dengan cara acak (*random sampling*) yang biasa disebut *probability sampling* dan tidak acak atau *non-probability sampling*. Pengambilan sampel secara acak merupakan pengambilan sampel yang memberikan peluang sama bagi setiap anggota populasi untuk menjadi anggota sampel. Metode pengambilan sampel secara acak dianggap terbaik, karena peneliti terbebas dari subyektifitas. Generalisasi hasil penelitian sampel terhadap populasi bisa lebih dipertanggungjawabkan.

#### Contoh 11.3 (Probability Sampling)

Jika elemen populasinya ada 200 dan yang akan dijadikan sampel adalah 50, maka setiap elemen tersebut mempunyai kemungkinan sama untuk terpilih, dengan besar kemungkinan untuk bisa terpilih menjadi sampel adalah  $50/200$ .

Sedangkan yang dimaksud dengan *nonrandom sampling* atau *nonprobability sampling*, setiap elemen populasi tidak mempunyai kemungkinan yang sama untuk dijadikan sampel.

#### Contoh 11.4. (Nonprobability Sampling)

Misal kita ingin mengambil sampel pengguna kompor berbahan bakar elpiji. Dari 25 sampel yang terpilih, ternyata rumahnya berdekatan dengan rumah peneliti, sedangkan yang lainnya, karena jauh, tidak dipilih; artinya kemungkinannya 0 (nol).

Dua jenis teknik pengambilan sampel di atas mempunyai tujuan yang berbeda. Jika peneliti ingin hasil penelitiannya bisa dijadikan ukuran untuk mengestimasi populasi, atau istilahnya adalah melakukan generalisasi maka seharusnya sampel *representatif* dan diambil secara acak. Namun jika peneliti tidak mempunyai kemauan melakukan generalisasi hasil penelitian maka sampel bisa diambil secara tidak acak. Sampel tidak acak biasanya juga diambil jika peneliti tidak mempunyai data pasti tentang ukuran populasi dan informasi lengkap tentang setiap elemen populasi. Contohnya, jika yang diteliti populasinya adalah konsumen teh botol, kemungkinan besar peneliti tidak mengetahui dengan pasti berapa jumlah konsumennya, dan juga karakteristik konsumen. Karena dia tidak mengetahui ukuran populasi yang tepat, bisakah dia mengatakan bahwa 200 konsumen sebagai sampel dikatakan *representatif*?. Kemudian, bisakah peneliti memilih sampel secara acak, jika tidak ada informasi yang cukup lengkap tentang diri konsumen?. Dalam situasi yang demikian, pengambilan sampel dengan cara acak tidak dimungkinkan, maka tidak ada pilihan lain kecuali sampel diambil dengan cara tidak acak atau *nonprobability sampling*, namun dengan konsekuensi hasil penelitiannya tersebut tidak

bisa digeneralisasikan. Jika ternyata dari 200 konsumen teh botol tadi merasa kurang puas, maka peneliti tidak bisa mengatakan bahwa sebagian besar konsumen teh botol merasa kurang puas terhadap teh botol.

Pada setiap jenis teknik pemilihan tersebut, terdapat beberapa teknik yang lebih spesifik lagi. Pada sampel acak (*random sampling*) dikenal dengan istilah *simple random sampling*, *stratified random sampling*, *cluster sampling*, *systematic sampling*, dan *area sampling*. Pada *nonprobability sampling* dikenal beberapa teknik, antara lain adalah *convenience sampling*, *purposive sampling*, *quota sampling*, *snowball sampling*

#### a. Probability/Random Sampling

Syarat pertama yang harus dilakukan untuk mengambil sampel secara acak adalah memperoleh atau membuat kerangka sampel atau dikenal dengan nama *sampling frame*. Yang dimaksud dengan kerangka sampel adalah daftar yang berisikan setiap elemen populasi yang bisa diambil sebagai sampel. Elemen populasi bisa berupa data tentang orang/-binatang, tentang kejadian, tentang tempat, atau juga tentang benda.

##### Contoh 11.5 (*sampling frame*)

Misalkan populasi penelitian adalah mahasiswa perguruan tinggi A

Maka *sampling frame*-nya adalah *daftar semua mahasiswa* yang terdaftar di perguruan tinggi tersebut selengkap mungkin yang meliputi nama, NIM, jenis kelamin, alamat, usia, dan informasi lain yang berguna bagi penelitiannya. Dari daftar ini, peneliti akan bisa secara pasti mengetahui jumlah populasinya (N).

##### Contoh 11.6 (*sampling frame*)

Jika populasinya adalah rumah tangga dalam sebuah kota.

Maka *sampling frame*-nya adalah daftar seluruh rumah tangga kota tersebut.

Jika populasinya adalah wilayah Jawa Barat, maka peneliti harus mempunyai peta wilayah Jawa Barat secara lengkap. Kabupaten, Kecamatan, Desa, Kampung. Lalu setiap tempat tersebut diberi kode (angka atau simbol) yang berbeda satu sama lainnya.

Di samping *sampling frame*, peneliti juga harus mempunyai alat yang bisa dijadikan penentu sampel. Dari sekian elemen populasi, elemen mana saja yang bisa dipilih menjadi sampel?. Alat yang umumnya digunakan adalah Tabel Angka Random, kalkulator, atau undian. Pemilihan sampel secara acak bisa dilakukan melalui sistem undian jika elemen populasinya tidak begitu banyak. Tetapi jika sudah ratusan, cara undian bisa mengganggu konsep "acak" atau "random" itu sendiri.

#### 1. Sampel Acak Sederhana (*Simple Random Sampling*)

Cara atau teknik ini dapat dilakukan jika analisis penelitiannya cenderung deskriptif dan bersifat umum. Perbedaan karakter yang mungkin ada pada setiap unsur atau elemen populasi tidak merupakan hal yang penting bagi rencana analisisnya. Misalnya, dalam populasi ada wanita dan pria, atau ada yang kaya dan yang miskin, ada manajer dan bukan manajer, dan perbedaan-perbedaan lainnya. Selama perbedaan gender, status kemakmuran, dan kedudukan dalam organisasi, serta perbedaan-perbedaan lain tersebut bukan merupakan sesuatu hal yang penting dan mempunyai pengaruh yang signifikan terhadap hasil penelitian, maka

peneliti dapat mengambil sampel secara acak sederhana. Dengan demikian setiap unsur populasi harus mempunyai kesempatan sama untuk bisa dipilih menjadi sampel. Prosedurnya adalah:

- Susun sampling frame
- Tetapkan jumlah sampel yang akan diambil
- Tentukan alat pemilihan sampel
- Pilih sampel sampai dengan jumlah terpenuhi

## 2. Sampel Acak Berstrata (Stratified Random Sampling)

Karena unsur populasi berkarakteristik heterogen, dan heterogenitas tersebut mempunyai arti yang signifikan pada pencapaian tujuan penelitian, maka peneliti dapat mengambil sampel dengan cara ini. Misalnya, seorang peneliti ingin mengetahui sikap manajer terhadap satu kebijakan perusahaan. Dia menduga bahwa manajer tingkat atas cenderung positif sikapnya terhadap kebijakan perusahaan tadi. Agar dapat menguji dugaannya tersebut maka sampelnya harus terdiri atas paling tidak para manajer tingkat atas, menengah, dan bawah. Dengan teknik pemilihan sampel secara acak berstrata, maka dia akan memperoleh manajer di ketiga strata tersebut, yaitu strata manajer atas, manajer menengah dan manajer bawah. Dari setiap strata tersebut dipilih sampel secara acak. Prosedurnya :

- Siapkan sampling frame
- Bagi sampling frame tersebut berdasarkan strata yang dikehendaki
- Tentukan jumlah sampel dalam setiap strata
- Pilih sampel dari setiap strata secara acak.

Pada saat menentukan jumlah sampel dalam setiap stratum, peneliti dapat menentukan secara (a) proposional,

(b) tidak proposional. Yang dimaksud dengan proposional adalah jumlah sampel dalam setiap strata sebanding dengan jumlah unsur populasi dalam strata tersebut. Sedangkan yang dimaksud adalah jumlah dalam setiap strata tidak sebanding. Hal ini terjadi jika jumlah unsur atau elemen di salah satu atau beberapa strata sangat sedikit.

Contoh 11.7 (pengambilan sampel berstrata secara proporsional)

Misalkan terdapat 160 manajer dengan perincian:

manajer tingkat atas (I)	: 15 manajer
manajer tingkat menengah (II)	: 45 manajer
manajer tingkat bawah (III)	: 100 manajer.

Jika jumlah sampel yang akan diambil 100 manajer, maka:

strata I diambil	: $(15:160) \times 100 = 9$ manajer
strata II diambil	: $(45:160) \times 100 = 28$ manajer
stratum III diambil	: $(100:160) \times 100 = 63$ manajer.

Contoh 11.8 (pengambilan sampel berstrata secara tidak proporsional)

Misalnya saja, kalau dalam stratum manajer kelas atas (I) hanya ada 4 manajer, maka peneliti bisa mengambil semua manajer dalam stratum tersebut, dan untuk manajer tingkat menengah (II) ditambah 5, sedangkan manajer tingkat bawah (III), tetap 63 orang.

## 3. Sampel Pengelompokan/Gugus (*Cluster Sampling*)

Teknik ini biasa juga diterjemahkan dengan cara pengambilan sampel berdasarkan gugus. Berbeda dengan teknik pengambilan sampel acak yang distratifikasikan, di

mana setiap unsur dalam satu stratum memiliki karakteristik yang homogen (stratum A: laki-laki semua, stratum B: perempuan semua), maka dalam sampel gugus, setiap gugus boleh mengandung unsur yang karakteristiknya berbeda-beda atau heterogen. Misalnya, dalam satu organisasi terdapat 100 departemen. Dalam setiap departemen terdapat banyak pegawai dengan karakteristik berbeda pula. Beda jenis kelaminnya, beda tingkat pendidikannya, beda tingkat pendapatnya, beda tingkat manajerialnya, dan perbedaan-perbedaan lainnya. Jika peneliti bermaksud mengetahui tingkat penerimaan para pegawai terhadap suatu strategi yang segera diterapkan perusahaan, maka peneliti dapat menggunakan cluster sampling untuk mencegah terpilihnya sampel hanya dari satu atau dua departemen saja. Prosedur:

- a. Susun sampling frame berdasarkan gugus (pada kasus di atas, elemennya ada 100 departemen)
- b. Tentukan berapa gugus yang akan diambil sebagai sampel
- c. Pilih gugus sebagai sampel dengan cara acak
- d. Teliti setiap pegawai yang ada dalam gugus sampel

#### 4. Sampel Sistematis (*Systematic Sampling*)

Jika peneliti dihadapkan pada ukuran populasi yang banyak dan tidak memiliki alat pengambil data secara random, cara pengambilan sampel sistematis dapat digunakan. Cara ini menuntut kepada peneliti untuk memilih unsur populasi secara sistematis, yaitu unsur populasi yang bisa dijadikan sampel adalah yang "keberapa". Misalnya, setiap unsur populasi yang keenam, yang bisa dijadikan sampel. Soal "keberapa"-nya satu unsur populasi bisa dijadikan sampel tergantung pada ukuran populasi dan ukuran sampel. Misalnya, dalam satu populasi terdapat 5000 rumah. Sampel

yang akan diambil adalah 250 rumah dengan demikian interval di antara sampel kesatu, kedua, dan seterusnya adalah 25. Prosedurnya :

1. Susun sampling frame
2. Tetapkan jumlah sampel yang ingin diambil
3. Tentukan K (kelas interval)
4. Tentukan angka atau nomor awal di antara kelas interval tersebut secara acak atau random - biasanya melalui cara undian saja.
5. Mulailah mengambil sampel dimulai dari angka atau nomor awal yang terpilih.
6. Pilihlah sebagai sampel angka atau nomor interval berikutnya

#### 5. Sampel Wilayah (*Area Sampling*)

Teknik ini dipakai ketika peneliti dihadapkan pada situasi bahwa populasi penelitiannya tersebar di berbagai wilayah. Misalnya, seorang marketing manager sebuah stasiun TV ingin mengetahui tingkat penerimaan masyarakat Bali atas sebuah mata tayangan, teknik pengambilan sampel dengan *area sampling* sangat tepat. Prosedurnya:

- a. Susun *sampling frame* yang menggambarkan peta wilayah (Provinsi), Kabupaten, Kotamadya, Kecamatan, Desa.
- b. Tentukan wilayah yang akan dijadikan sampel (Kabupaten?, Kotamadya?, Kecamatan?, Desa?)
- c. Tentukan berapa wilayah yang akan dijadikan sampel penelitiannya.
- d. Pilih beberapa wilayah untuk dijadikan sampel dengan cara acak atau random. Kalau ternyata masih terlampau banyak responden yang harus diambil datanya, bagi lagi wilayah yang terpilih ke dalam sub wilayah.

### b. Sampel Tidak Acak (*Nonprobability/Nonrandom Sampling*)

Seperti telah diuraikan sebelumnya, jenis sampel ini tidak dipilih secara acak. Tidak semua unsur atau elemen populasi mempunyai kesempatan sama untuk bisa dipilih menjadi sampel. Unsur populasi yang terpilih menjadi sampel bisa disebabkan karena kebetulan atau karena faktor lain yang sebelumnya sudah direncanakan oleh peneliti.

#### 1. Sampel yang dipilih dengan pertimbangan kemudahan (*Convenience Sampling/incidental sampling*)

Dalam memilih sampel, peneliti tidak mempunyai pertimbangan lain kecuali berdasarkan kemudahan saja. Seseorang diambil sebagai sampel karena kebetulan orang tadi ada di situ atau kebetulan dia mengenal orang tersebut. Oleh karena itu ada beberapa penulis menggunakan istilah *accidental sampling* —tidak disengaja— atau juga *captive sample* (man-on-the-street) Jenis sampel ini sangat baik jika dimanfaatkan untuk penelitian penjajagan, yang kemudian diikuti oleh penelitian lanjutan yang sampelnya diambil secara acak (*random*). Beberapa kasus penelitian yang menggunakan jenis sampel ini, hasilnya ternyata kurang obyektif.

#### 2. Sampel yang dipilih dengan tujuan tertentu (*Purposive Sampling*)

Sesuai dengan namanya, sampel diambil dengan maksud atau tujuan tertentu. Seseorang atau sesuatu diambil sebagai sampel karena peneliti menganggap bahwa seseorang atau sesuatu tersebut memiliki informasi yang diperlukan bagi penelitiannya. Dua jenis sampel ini dikenal dengan *nama judgement dan quota sampling*.

#### a. *Judgment Sampling*

Sampel dipilih berdasarkan penilaian peneliti bahwa dia adalah pihak yang paling baik untuk dijadikan sampel penelitiannya.. Misalnya untuk memperoleh data tentang bagaimana satu proses produksi direncanakan oleh suatu perusahaan, maka manajer produksi merupakan orang yang terbaik untuk bisa memberikan informasi. Jadi, *judgment sampling* umumnya memilih sesuatu atau seseorang menjadi sampel karena mereka mempunyai "*information rich*". Dalam program pengembangan produk (*product development*), biasanya yang dijadikan sampel adalah karyawannya sendiri, dengan pertimbangan bahwa kalau karyawan sendiri tidak puas terhadap produk baru yang akan dipasarkan, maka jangan terlalu berharap pasar akan menerima produk itu dengan baik.

#### b. *Quota Sampling*

Teknik sampel ini adalah bentuk dari sampel distratifikasi secara proposional, namun tidak dipilih secara acak melainkan secara kebetulan saja. Misalnya, di sebuah kantor terdapat pegawai laki-laki 60% dan perempuan 40% . Jika seorang peneliti ingin mewawancarai 30 orang pegawai dari kedua jenis kelamin tadi maka dia harus mengambil sampel pegawai laki-laki sebanyak 18 orang sedangkan pegawai perempuan 12 orang. Sekali lagi, teknik pengambilan ketiga puluh sampel tadi tidak dilakukan secara acak, melainkan secara kebetulan saja.

#### 3. Sampel Bola Salju (*Snowball Sampling*)

Cara ini banyak dipakai ketika peneliti tidak banyak tahu tentang populasi penelitiannya. Dia hanya tahu satu atau dua orang yang berdasarkan penilaiannya bisa dijadikan sampel.

Karena peneliti menginginkan lebih banyak lagi, lalu dia minta kepada sampel pertama untuk menunjukan orang lain yang kira-kira bisa dijadikan sampel. Misalnya, seorang peneliti ingin mengetahui pandangan kaum lesbian terhadap lembaga perkawinan. Peneliti cukup mencari satu orang wanita lesbian dan kemudian melakukan wawancara. Setelah selesai, peneliti tadi minta kepada wanita lesbian tersebut untuk bisa mewawancarai teman lesbian lainnya. Setelah jumlah wanita lesbian yang berhasil diwawancarainya dirasa cukup, peneliti bisa mengentikan pencarian wanita lesbian lainnya. Hal ini bisa juga dilakukan pada pencandu narkotik, para gay, atau kelompok-kelompok sosial lain yang eksklusif (tertutup).

## Bab 12

### ANGKA INDEKS

#### A. Pendahuluan

Dalam dunia bisnis, seorang manajer sering berurusan dengan angka-angka yang berubah dari waktu ke waktu, misalnya harga yang dibayarkan untuk pembelian bahan baku, jumlah karyawan dan pelanggan, pendapatan dan keuntungan tahunan, dan lain-lain. Perubahan angka seperti ini bisa dijelaskan dengan menggunakan angka indeks. Jadi, angka indeks adalah suatu angka yang digunakan untuk menggambarkan perubahan yang terjadi terhadap variabel ekonomi ataupun sosial dari kegiatan yang sama dalam dua waktu yang berbeda. Angka indeks biasanya berkaitan dengan uang atau tenaga kerja. Angka indeks sering digunakan untuk menunjukkan perubahan dalam bidang ekonomi, industri dan pasar.

Angka indeks sangat dibutuhkan dalam kegiatan bisnis agar dapat mengetahui maju mundurnya atau naik turunnya suatu usaha atau kegiatan. Jadi tujuan pembuatan angka indeks sebetulnya adalah untuk mengukur secara kuantitatif terjadinya perubahan dalam dua waktu yang berlainan misalnya indeks harga untuk mengukur perubahan harga (berapa kenaikannya atau penurunannya), indeks produksi untuk mengetahui perubahan yang terjadi dalam kegiatan produksi, indeks biaya hidup untuk mengukur tingkat inflasi, dan lain-lain.

### B. Komponen Penyusunan Angka Indeks

Di dalam membuat angka indeks diperlukan dua macam waktu, yaitu waktu dasar (*base period*) dan waktu yang bersangkutan atau sedang berjalan (*current period*). Waktu dasar adalah waktu dimana suatu kegiatan (kejadian) dipergunakan sebagai dasar perbandingan. Waktu yang sedang berjalan ialah waktu dimana suatu kegiatan (kejadian) dipergunakan sebagai dasar perbandingan terhadap kegiatan (kejadian) pada waktu dasar.

#### Contoh 12.1

Jumlah produksi barang A yang dihasilkan oleh PT. Azkiya selama tahun 2005 dan 2006 masing-masing adalah 150 ton dan 225 ton. Hitunglah indeks produksi masing-masing tahun.

#### Penyelesaian:

Jika dibuat indeks produksi tahun 2006 dengan waktu dasar 2005, maka produksi pada tahun 2005 dipergunakan untuk dasar perbandingan, sedangkan produksi tahun 2006 (waktu bersangkutan) akan diperbandingkan terhadap produksi tahun 2005 tadi.

Indeks produksi 2006 adalah:

$$225/150 \times 100\% = 150\% \text{ (ada kenaikan produksi 50\%)}$$

### C. Indeks Harga Relatif Sederhana dan Agregatif

Indeks harga relatif sederhana (*simple relative price index*) ialah indeks yang terdiri dari satu macam barang saja, baik untuk indeks produksi maupun indeks harga. Misalnya indeks produksi beras, indeks produksi karet, indeks produksi

ikan, indeks harga beras, indeks harga karet, indeks harga ikan, dan sebagainya.

Rumus Indeks harga relatif sederhana :

$$I_{t_0} = \frac{P_t}{P_0} \times 100\%$$

Keterangan:

$I_{t_0}$  = indeks harga pada waktu  $t$  dengan waktu dasar 0

$P_t$  = harga pada waktu  $t$

$P_0$  = harga pada waktu 0

#### Contoh 12.2

Tabel dibawah ini menyajikan data rata-rata perdagangan beberapa hasil pertanian di Jakarta dari tahun 1992 - 1997. Hitunglah indeks harga beras pada tahun 1995, 1996, dan 1997 dengan waktu dasar tahun 1992

Jenis Pertanian	1992	1993	1994	1995	1996	1997
Beras	66.368	67.337	81.522	100.209	101.382	111.183
Jagung Kuning	34.877	39.829	45.850	50.000	62.740	66.208
Kacang Kedelai	110.505	116.458	121.542	115.052	114.800	125.733
Kacang Hijau	111.528	111.063	127.108	128.750	163.042	192.771
Kacang Tanah	161.243	198.271	209.542	200.000	228.792	223.250
Ketela Pohon	15.433	13.853	20.538	26.944	26.079	24.311
Ketela Rambat	22.033	22.273	29.831	36.698	35.688	35.131
Kentang	46.984	55.110	85.183	82.404	93.713	121.920

Penyelesaian:

Untuk tahun 1995:

$$I_{95/92} = \frac{P_{95}}{P_{92}} \times 100\% = \frac{100.209}{66.368} \times 100\% = 150,99\%$$

Untuk tahun 1996:

$$I_{96/92} = \frac{P_{96}}{P_{92}} \times 100\% = \frac{101.382}{66.368} \times 100\% = 152,76\%$$

Untuk tahun 1997 :

$$\begin{aligned} I_{97/92} &= P_{97}/P_{92} \times 100\% \\ &= 111.183/66.368 \times 100\% \\ &= 167,52\% \end{aligned}$$

*Indeks agregatif relatif sederhana* merupakan indeks yang terdiri dari beberapa barang (kelompok barang). Indeks ini digunakan untuk melihat perkembangan kuantitas barang dan jasa dengan dibandingkan dengan tahun dasar, misalnya indeks harga 9 macam bahan pokok, indeks impor Indonesia, indeks ekspor Indonesia, indeks harga bahan makanan, indeks biaya hidup, indeks hasil penjualan suatu perusahaan (lebih dari satu barang yang dijual), dan lain-lain.

$$I_t = \frac{\sum H_t}{\sum H_0} \times 100\%$$

Keterangan:

$H_t$ : Jumlah harga pada tahun t

$H_0$ : Jumlah harga pada tahun dasar

Contoh 12.3

Carilah indeks agregatif relatif sederhana pada tahun 2002

Bulan	2000	2001	2002	2003	2004
Januari	3500	3800	4100	4200	3850
Februari	3800	3450	4120	4250	3800

Maret	3400	3600	3950	4150	3900
April	4000	3900	3890	4050	3950
Mei	4200	4150	3950	3900	4000
Juni	3900	3950	4000	4100	3990
Total	22800	22850	24010	24650	23490
Indeks Harga	100	100	105	108	103

$$I_{2002} = \frac{\sum H_{2002}}{\sum H_{2000}} \times 100\% = \frac{24.010}{22.800} \times 100\% = 105\%$$

Jadi indeks agregatif relatif sederhana pada tahun 2002 adalah 105%. Artinya pada tahun 2002 terjadi kenaikan sebesar 105% - 100% = 5%

*Soal-soal Bab 12*

1. Tabel dibawah ini menyajikan data produksi Tanaman Bahan Makanan menurut jenis, dari tahun 1993 - 1998. Hitunglah indeks produksi kacang tanah tahun 1996, 1997, dan 1998 dengan waktu dasar 1993

Jenis Pertanian	1992	1993	1994	1995	1996	1997
Padi Sawah	45.559	43.959	46.806	48.188	46.592	45.711
Padi Ladang	2.622	2.682	2.938	2.913	2.785	2.761
Jagung	6.460	6.869	8.246	9.307	8.711	10.059
Ubi Kayu	17.285	15.729	15.441	17.002	15.134	14.728
Ubi Jalar	2.088	1.845	2.171	2.017	1.847	1.928
Kacang Tanah	639	632	760	738	688	691
Kedelai	1.709	1.565	1.680	1.517	1.357	1.306
Jumlah	76.362	73.281	78.042	81.682	77.114	77.184

2. Dari data pada no 1, carilah indeks agregat relatif sederhana pada tahun 1996.

## DAFTAR PUSTAKA

- Alghifari (1997), *Statistika Ekonomi 1*, Edisi ke-2, Bag. Penerbitan STIE YKPN, Yogyakarta.
- Anto Dajan (1986), *Pengantar Metode Statistik Jilid I*. Jakarta: LP3ES.
- Boediono dan I Wayan Kosler (2004), *Statistika dan Probabilitas*, Remaja Rosydakarya, Bandung.
- Hasan Mustafa (2000), *Teknik Sampling*, diambil dari [www.unpar.ac.id](http://www.unpar.ac.id).
- Jonathan Sarwono (2005), *Teori dan Riset Pemasaran*, Andi Offset Yogyakarta
- Salafudin (2006), *Statistika Terapan untuk Penelitian Sosial*. Pekalongan: STAIN Press
- Sugiyono (2009), *Statistika untuk Penelitian*, Alfabeta, Bandung.
- Uma Sekaran (1999), "*Research Methods for Business: A Skill Building Approach*", second edition, John Willey dan Sons, Inc., New York.
- Uma Sekaran, "Research Methods for Business", 2003, p. 238

## Internet:

<http://administrasipendidikan.org>

<http://www.antara-sumbar.com>

<http://sahambbri.files.wordpress.com>. Di akses tanggal 13 Pebruari 2010

- <http://www.antara-sumbar.com>. Diakses tanggal 13 Pebruari 2010
- <http://www.ak-online.de>. Diakses 13 Pebruari 2010
- <http://administrasipendidikan.org>. Diakses tanggal 13 Pebruari 2010
- [www.slideshare.net/.../statistika-deskriptif-bab-07-angka-indeks](http://www.slideshare.net/.../statistika-deskriptif-bab-07-angka-indeks). Diakses tanggal 12 Maret 2010
- <http://www.uib.no/People/ngbnk/kurs/notes/node5.html>
- <http://www.experiment-resources.com/statistical-sampling-techniques.html>
- <http://userpages.umbc.edu/~nmiller/POLI300/MEASURES%20OF%20DISPERSION.ppt>
- <http://www.academic.marist.edu/~jzbv/Statistics/CentralTendency.ppt>
- <http://people.ucsc.edu/~flannery/econ113/powerpoints/descriptivestatistics.ppt>
- <http://smartstat.wordpress.com/2010/02/20/skala-pengukuran-variabel/>
- <http://birokrasi.kompasiana.com/2011/12/26/data-statistik-tidak-mungkin-100-benar/>
- <http://globviews.files.wordpress.com/2009/03/10-metode-pengumpulan-data-instrumen-compatibility-mode.pdf>
- <http://mvpprograms.com/help/mvpstats/distributions/SkewnessKurtosis>
- <http://www.utexas.edu/lbj/sites/default/files/file/news/Skew.pdf>

## BIOGRAFI SINGKAT PENULIS 1



Yusuf Nalim, lahir pada tanggal 05 Januari 1978 di Cilacap, Jawa Tengah. Setelah menamatkan pendidikan Sekolah Menengah Atas (SMA) di SMAN 1 Kroya, penulis melanjutkan studi Sarjana di Jurusan Matematika Universitas Negeri Surabaya, selesai tahun 2002. Pada tahun 2003 penulis melanjutkan ke jenjang Magister di Jurusan Statistika Institut Teknologi Sepuluh Nopember (ITS) Surabaya. Semasa SMA penulis cukup aktif terlibat dalam kegiatan organisasi, antara lain sebagai pengurus OSIS dan aktif di Organisasi Kepramukaan. Sejak 2008, penulis merupakan Dosen Tetap pada STAIN Pekalongan dengan bidang mata ajar Statistika Deskriptif, Statistika Inferensial, Matematika Ekonomi dan Bisnis serta pernah mengajar Riset Operasi. Sebelumnya penulis pernah mengajar di STIKES Al-Irsyad Al-Islamiyyah Cilacap dengan mengampu mata kuliah Biostatistika (Prodi D3 Kebidanan dan S1 Keperawatan) dan Statistika Farmasi (Prodi D3 farmasi). Saat ini penulis aktif sebagai pengajar dan peneliti. Jabatan profesi saat ini adalah sebagai Sekretaris sekaligus merangkap sebagai Koordinator Divisi Informasi dan Analisis Data pada Lembaga Survei Validita.

## BIOGRAFI SINGKAT PENULIS 2



Salafudin, lahir pada tanggal 25 Agustus 1965, di Desa Debong Wetan, Kecamatan Dukuhturi, Kabupaten Tegal, Jawa Tengah, dari pasangan suami istri Tarmudi - Fatimah. Menyelesaikan pendidikan SD di MI Muhammadiyah, SMP I Tegal dan SMA 1 Tegal jurusan IPA. Sejak awal menyukai Matematika. Kesukaannya pada matematika menghantarkannya masuk Jurusan Matematika, Fakultas MIPA Universitas Diponegoro (UNDIP) Semarang melalui jalur PMDK. Studi S1 diselesaikan pada tahun 1994. Studi S2 di Fakultas Pascasarjana Universitas Gadjah Mada (UGM) Yogyakarta pada bidang keilmuan yang sama diselesaikan pada tahun 2004. Salah satu obsesinya adalah mengimplementasikan interkoneksi dan integrasi ilmu pengetahuan dengan agama. Untuk itu penulis menempuh pendidikan S3 mengambil bidang Pendidikan Islam di Universitas Ibn Khaldun Bogor.

Penulis mengajar di STAIN Pekalongan sejak tahun 1999, mengampu mata kuliah Statistika, Ilmu Alamiah Dasar (IAD), Matematika Dasar, Statistika Deskriptif dan Metode Kuantitatif. Sebelumnya, penulis pernah mengajar di MAN 2 Pekalongan dan SMA Muhammadiyah 1 Pekalongan, juga sebagai tutor di Lembaga Bimbingan Belajar Primagama Cabang Tegal.

Selain sebagai dosen, penulis juga aktif dalam kegiatan sosial dan penyelenggaraan pendidikan. Pada tahun 1995,

mendirikan Yayasan Pendidikan dan Pengembangan Sumber Daya Insani (YP2SI) Al Ummah. Pada lembaga yang bergerak dalam bidang pendidikan dan dakwah ini, yang antara lain menaungi PAUD, RAIT, TKIT, SDIT Ulul Albab, SMPIT As Salam Boarding School, serta Al Kindi Club, penulis sebagai ketua hingga tahun Januari 2008, dan kini sebagai ketua Dewan Pembina.

Penulis mempunyai istri Fathia Rahmawati, seorang pendidik yang mengajar di MTs IN Banyuurip Ageng Kota Pekalongan, dan anak 3 : Shofwa Fathina (FE UNDIP), M. Ilzam Falah (Kelas II SMAIT Nur Hidayah Sukoharjo) dan M. Izzuddin (Kelas I MA Husnul Khotimah PP Husnul Khotimah Kuningan).

# STATISTIKA

## Deskriptif

Saat ini penggunaan statistika telah menyentuh semua bidang ilmu pengetahuan, mulai dari bidang eksak (astronomi, geologi, teknik) hingga bidang sosial (psikologi, komunikasi, linguistik). Metodologi statistika banyak mempengaruhi perkembangan bidang-bidang ekonomi, biologi dan cabang-cabang terapannya. Akibatnya lahirlah ilmu-ilmu gabungan seperti ekonometrika (ekonomi dan statistika), biometrika (biologi dan statistika), dan psikometrika (psikologi dan statistika).

Kandungan materi dalam buku ini sebenarnya merupakan materi dasar dari ilmu statistika. Kajian dalam statistika deskriptif hanya bertujuan memberikan gambaran ringkas (ringkasan) tentang data (populasi/sampel) dan pengamatan yang telah dilakukan terhadap suatu peristiwa. Lingkup kajian utama dari statistika deskriptif adalah bagaimana melakukan proses pengumpulan data dalam penelitian, pengolahan data, penyajian data dalam bentuk tabel, grafik serta, mengetahui karakteristik khas dari masing-masing data dalam bentuk rata-rata, median, modus, simpangan, kuartil dan lain-lain. Untuk mengetahui perkembangan suatu peristiwa kuantitatif maka diperlukan pengetahuan mengenai angka indeks.

Pembahasan buku ini meliputi 12 Bab, dimulai dari Konsep Dasar Statistika (Bab 1) dan diakhiri dengan ulasan mengenai Angka Indeks (Bab 12). Penulis sengaja membagi beberapa materi bahasan dalam bab tersendiri dengan harapan agar penjelasannya lebih rinci dan mudah dipahami. Seperti materi tentang Populasi dan Sampel (Bab 2) serta materi tentang Data (Bab 3). Dalam banyak literatur, materi tentang populasi, sampel dan data seringkali hanya dijadikan subbab.

Semoga dengan kehadiran buku ini bisa menjadikan nilai tambah bagi khazanah keilmuan statistika dan sekaligus bisa menjadi rujukan bagi pembaca dari semua kalangan (mahasiswa, pengajar, praktisi, peneliti dan lain-lain) yang ingin mengetahui lebih dalam tentang dasar-dasar statistika.

Diterbitkan oleh

**STAIN Pekalongan Press**

Jl. Kusumabangsa No 9 Pekalongan 51114

Telp (0285) 412575 Fax (0285) 423418

e-mail: p3mstainpkl@yahoo.co.id

ISBN 978-979-3968-60-5



9 789793 968605 >