

1.1 Tujuan dari Statistika

Definisi Statistika

Statistika adalah cabang ilmu yang mempelajari tentang bagaimana mengumpulkan, menganalisis dan menginterpretasikan data. Atau dengan kata lain, statistika menjadi semacam alat dalam melakukan suatu riset empiris.

Dalam menganalisis data, para ilmuwan menggambarkan persepsinya tentang suatu fenomena. Deskripsi yang sudah stabil tentang suatu fenomena seringkali mampu menjelaskan suatu teori. (Walaupun demikian, orang dapat saja berargumentasi bahwa ilmu biasanya menggambarkan *bagaimana* sesuatu itu terjadi, bukannya *mengapa*). Penemuan teori baru merupakan suatu proses kreatif yang didapat dengan cara mereka ulang informasi pada teori yang telah ada atau mengesktrak informasi yang diperoleh dari dunia nyata. Pendekatan awal yang umumnya digunakan untuk menjelaskan suatu fenomena adalah *statistika deskriptif*.

Statistika Deskriptif

Statistika deskriptif adalah tehnik yang digunakan untuk mensarikan data dan menampilkannya dalam bentuk yang dapat dimengerti oleh setiap orang. Hal ini melibatkan proses kuantifikasi dari penemuan suatu fenomena. Berbagai statistik sederhana, seperti rata-rata, dihitung dan ditampilkan dalam bentuk tabel dan grafik. Statistika deskriptif dapat memberikan pengetahuan yang signifikan pada kejadian fenomena yang belum dikenal dan mendeteksi keterkaitan yang ada di dalamnya. Tetapi dapatkah statistika deskriptif memberikan hasil yang bisa diterima secara ilmiah? Statistik merupakan suatu alat pengukuran yang berhubungan dengan keragaman pada karakteristik objek-objek yang berbeda .

Objek yang belum dikenal tidaklah mewakili populasi objek yang memiliki "quantifiabel feature" melalui penyelidikan. Namun demikian, keragaman bisa menjadi hasil dari keberagaman yang lainnya (karena acak atau terkontrol). Pada ilmu fisika, yang sangat berkaitan dengan ekstraksi dan formulasi persamaan matematik tidak menyisakan banyak tempat untuk fluktuasi acak. Pada ilmu statistika, fluktuasi seperti itu dapat dijadikan model. Hubungan relasi statistik selanjutnya merupakan hubungan relasi yang menerangkan suatu proporsi perubahan stokastik yang pasti.

Statistika Induktif

Berbeda dengan fisika, hubungan atau relasi empiris yang diobservasi pada ilmu alam, sosiologi dan psikologi (dan bidang pilhan lain misalnya ekonomi) bersifat statis. Pada bidang-bidang ini, pekerjaan empiris dilaksanakan berdasarkan percobaan-percobaan atau survey sampel. Pada kasus lainnya, seluruh populasi tidak dapat diobservasi-karena berbagai alasan ekonomis ataupun praktis. Mengambil kesimpulan tentang suatu populasi berdasarkan data dari sampel yang terbatas merupakan tujuan dari suatu proses pengambilan keputusan *inferensial* atau *statistik induktif*.

Perubahan di sini merupakan suatu refleksi variasi pada sampel dan proses pengambilan sampel.

Statistika dan Proses Ilmiah

Dengan ketergantungan pada tahap penelitian ilmiah, data diuji dengan tingkatan informasi yang berbeda. Data dapat dikumpulkan untuk meng-eksplere sebuah fenomena pada pendekatan pertama, dan dapat juga

digunakan untuk memberikan uji hipotesis (pembuktian) secara statistik tentang struktur karakteristik melalui penyelidikan. Kemudian, statistik digunakan pada seluruh tahap proses ilmiah dimana fenomena yang terukur ikut terlibat. Disini, konsep kita tentang kemampuan mengukur menjangkau luas berbagai proposisi menarik dan ilmiah. Sebagai contoh, sebuah proposisi misalnya "seekor lebah bumble sedang terbang". Dengan menghitung jumlah kejadian tersebut dalam berbagai setting yang berbeda, kita sedang mengukur kejadian fenomena alam. Pada tahap dasar ini kita mencoba untuk menduga kemungkinan datangnya seekor lebah bumble pada keadaan yang khusus (misalnya pada hari-hari turun hujan di musim panas di Berlin).

1.2 Penelitian secara Statistik

Merancang Penelitian Menurut Kaidah Statistik

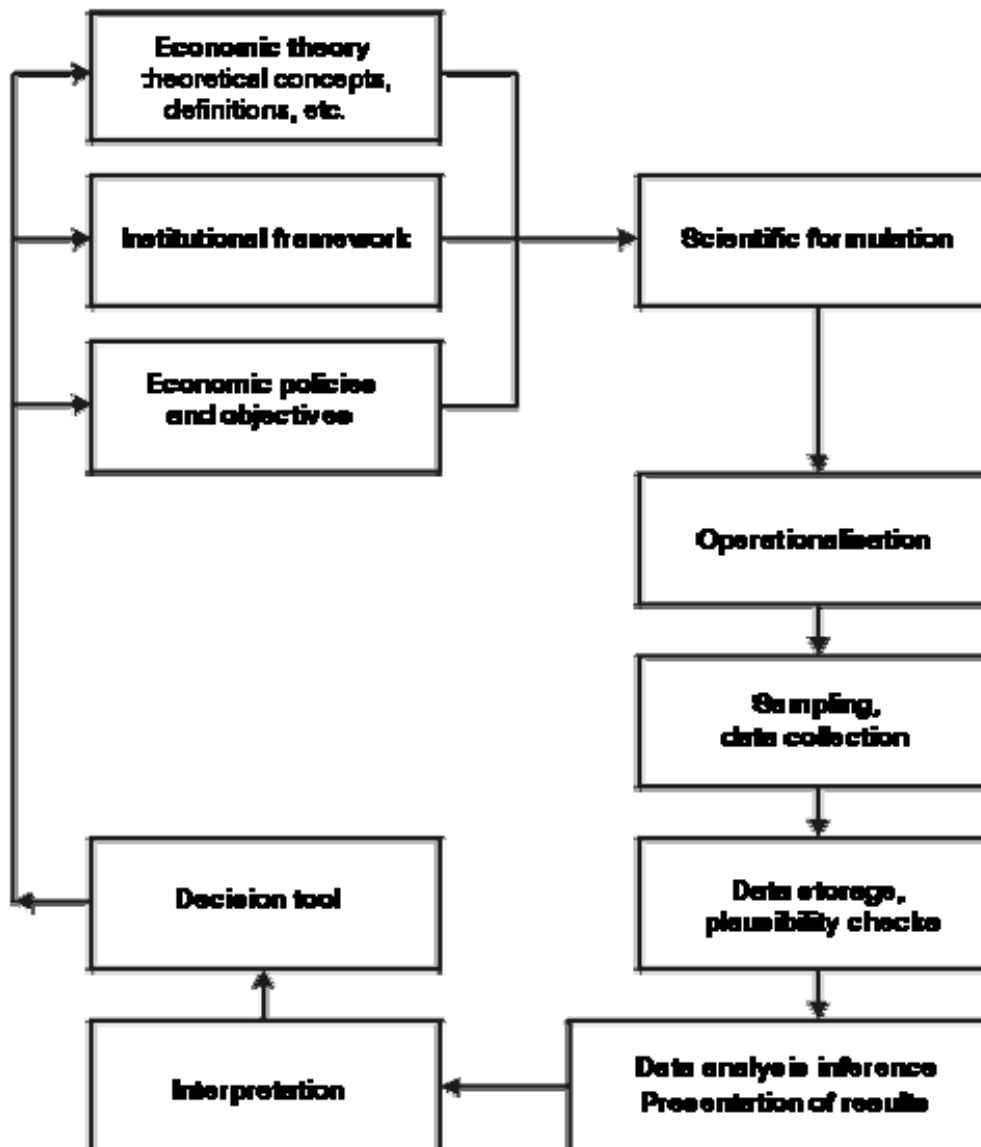
Penggunaan statistika dalam suatu penelitian sering melibatkan langkah-langkah berikut:

1. Merancang penelitian: perkembangan sasaran, terjemahan konsep teoritis menjadi fenomena yang dapat diobservasi (sebagai contoh, variabel-variabel), setting lingkungan (misalnya, dalam menentukan parameter mana yang konstan dilaksanakan) proyeksi biaya dan sebagainya.
2. Pengambilan data
 - Data primer: data yang dikumpulkan oleh institusi yang melaksanakan penelitian
 - Survey:
 - menyimpan data tanpa menguji kontrol melalui kondisi lingkungan yang bisa jadi mempengaruhi observasi-observasi.
 - Mengobservasi seluruh anggota populasi (*sensus*) atau mengambil sampel (*sample survey*)
 - Mengumpulkan data dengan wawancara atau dengan pengukuran
 - Dokumentasi data melalui daftar pertanyaan, protokol dan lain sebagainya.
 - Pribadi berhadapan dengan observasi tak langsung (misalnya wawancara pribadi, pemberian daftar pertanyaan melalui pos, telepon dan lain-lain)
 - Eksperimen atau percobaan: mengontrol variabel-variabel secara aktif untuk mendapatkan pengaruhnya pada variabel-variabel lainnya.
 - Penyimpanan otomatis: mengobservasi atau mengamati data ketika sedang dibangkitkan, misalnya dalam suatu proses produksi.
 - Data Sekunder: menggunakan data yang sudah tersedia, baik dari sumber internal ataupun eksternal.
3. Mengatur data
4. Analisa: mengaplikasikan alat-alat atau metode statistika.
5. Interpretasi: kesimpulan mana yang mendukung informasi kuantitatif yang dibangkitkan dengan prosedur statistik?

Sumber-sumber Data Statistik

- Statistik publik
- Statistik swasta
- Organisasi Internasional

Gambar berikut mengilustrasikan urutan langkah-langkah dalam penyelidikan statistik:



1.3 Elemen dan Populasi Statistik

Elemen-elemen Statistik

Objek-objek yang sifatnya/atributnya diobservasi atau diukur untuk keperluan statistik disebut elemen. Untuk mengenali semua elemen yang relevan dengan suatu penyelidikan tertentu, seseorang harus mengelompokkan karakteristik definisi mereka seperti dimensi, temporal dan tempat.

Contoh: Sensus populasi di Jerman.

- jenis karakteristik: penduduk Jerman.
- tempat: alamat tetap di Republik Federal Jerman.
- temporal: tanggal sensus

Populasi

Semesta elemen statistik yang ditutupi oleh sebuah himpunan spesifikasi tertentu disebut populasi.

Secara umum, dengan meningkatkan jumlah kriteria untuk disesuaikan dengan elemen-elemen tersebut akan memberikan hasil pendugaan yang lebih baik dan lebih homogen.

Pada sensus, seluruh elemen populasi diteliti. Menyimpan informasi dari porsi populasi tersebut menghasilkan survey sampel. Ukuran populasi bisa berhingga atau tak berhingga.

Jumlah elemen yang terdapat pada populasi bisa berubah kapan saja, dimana beberapa elemen keluar dari populasi dan elemen lain masuk. Sensitivitas populasi yang berjalan menurut waktu ini harus diperhatikan ketika melakukan penyelidikan statistik.

1.4 Peubah Statistik

Sebuah karakteristik yang bisa diamati dari suatu **elemen statistik** disebut variabel.

Nilai-nilai aktual yang diasumsikan oleh variabel statistik dinamakan observasi, amatan, pengukuran atau data. Himpunan nilai-nilai yang mungkin diambil oleh sebuah variabel dinamakan ruang sampel atau ruang contoh.

Variabel dilambangkan dengan huruf besar X, Y, \dots , sedangkan realisasinya ditulis menggunakan huruf kecil: x_1, x_2, \dots , indeks dalam hal ini merefleksikan elemen statistik yang dijadikan sampel.

Variabel	Observasi
X	x_1, x_2, x_3, \dots
Y	y_1, y_2, y_3, \dots

Adalah penting untuk membedakan antara variabel yang digunakan untuk identifikasi dan variabel target.

Variabel-variabel Identifikasi

Elemen-elemen populasi harus ditentukan spesifikasinya ketika menentukan suatu himpunan dengan nilai yang tetap. Sebagai contoh, dalam hal membatasi penelitian statistik terhadap orang dengan jenis kelamin perempuan akan melibatkan setting identifikasi variabel jenis "kelamin" dengan "perempuan".

Variabel-variabel Target

Variabel-variabel ini merupakan karakteristik minat/kepentingan, fenomena yang dieksplorasi oleh alat teknis statistik. Misalnya, usia orang yang termasuk dalam suatu populasi tertentu.

Contoh : Sasaran penelitian statistik adalah untuk mengeksplor atau menggali struktur sosio-ekonomi Berlin pada tanggal 21 Desember 1995.

Variabel-variabel identifikasi yang dipilih:

- resmi: penduduk
- tempat: alamat tetap di Berlin
- temporal: 31 Desember 1995

Eelemen statistik: Penduduk Berlin yang terdaftar pada tanggal 31 Desember 1995.

Populasi: semua penduduk Berlin pada tanggal 31 Desember 1995.

Variabel target yang mungkin:

Simbol	Variabel	Ruang Sampel
<i>X</i>	Usia (dihitung dalam tahun)	$\{0, 1, 2, \dots\}$
<i>S</i>	Jenis kelamin	{perempuan, laki-laki}
<i>T</i>	Status	{tidak kawin, kawin, cerai}
<i>Y</i>	Pendapatan bulanan	

1.5 Skala-skala Pengukuran

Variabel-variabel acak tersebut bisa mengambil nilai-nilai yang berbeda:

Simbol	Variabel	Ruang Sampel
X	Usia (dalam tahun)	$\{0, 1, 2, \dots\}$
S	Jenis kelamin	{perempuan, laki-laki}
T	Status	{tidak kawin, kawin, cerai}
Y	Pendapatan bulanan	$[0, \infty)$

Variabel-variabel tersebut bisa dikelompokkan menjadi variabel kuantitatif, sebagai contoh, variabel bernilai numerik (usia dan pendapatan) serta variabel kualitatif, variabel kategorik (jenis kelamin, status perkawinan). Karena nilai-nilai numerik biasa digunakan untuk melambangkan observasi variabel kualitatif, bisa jadi yang muncul adalah kuantitatif.

Namun demikian, penentuan sintetis semacam itu tidak berasal dari kualitas yang sama seperti pengukuran numerik yang muncul alami ketika mengobservasi sebuah fenomena.

Perbedaan tajam antara variabel kuantitatif dan variabel kualitatif terletak pada sifat-sifat **pengukuran skala** aktual, terutama pada kemampuannya untuk digunakan sebagai input pada suatu metode statistik. Dalam mengembangkan peralatan baru, ahli statistik membuat asumsi tentang skala-skala pengukuran yang diperbolehkan.

Sebuah pengukuran adalah suatu penentuan/penugasan pada sebuah observasi. Beberapa pengukuran muncul lebih alami dibandingkan yang lainnya. Sebagai contoh, dengan mengukur tinggi seseorang, kita menggunakan yardstick atau alat ukur yang meyakinkan kemampuan perbandingan antara observasi-observasi hingga ketelitian satuan-satuan yang digunakan (misalnya inci atau sentimeter). Sebaliknya, nilai sekolah menunjukkan klasifikasi yang relatif kasar yang mengindikasikan ranking tertentu tetapi dengan memasukkan para siswa kedalam kategori yang sama. Nilai yang ditentukan untuk pernyataan kualitatif semisal 'sangat baik', 'sedang' dan sebagainya adalah nilai sembarang namun singkat dan praktis dalam menilai prestasi seseorang. Dikarenakan tidak adanya alasan konseptual pada skala penilaian di sekolah, maka tidak diperlukan penafsiran 'jarak' antara nilai-nilai tersebut.

Jelasnya, pengukuran yang cermat memberikan informasi yang lebih banyak daripada hanya sekedar nilai sekolah, sebagaimana halnya jarak-jarak antara pengukuran bisa dibandingkan dengan konsisten. Pernyataan seperti "Tinggi Tom dua kali tinggi anaknya" atau "Manuela lebih kecil 35 cm dari pasangannya" atau pernyataan-pernyataan sejenisnya masih bisa diterima. Karena metode statistik dikembangkan dengan dasar matematika, maka skala yang bisa digunakan juga didefinisikan dengan menggunakan konsep matematika.

Ini semua adalah transformasi yang dapat digunakan tanpa kehilangan informasi. Semakin luas jangkauan transformasi yang digunakan, semakin sedikit informasi yang bisa disampaikan oleh skala. Tabel berikut ini memberikan daftar skala pengukuran umum untuk meningkatkan isi permintaan informasi. Skala-skala yang membawa banyak informasi selalu bisa ditransformasikan menjadi skala yang memberi sedikit informasi.

Variabel	Skala Pengukuran	Pernyataan	Transformasi yang digunakan
Kualitatif	Skala Nominal	Ekuivalen	any equivalence preserving mapping
Kategori	Skala Ordinal	ekuivalen, berurut	any order preserving mapping
Kuantitatif	Skala Interval	ekuivalen, berurut,	$y = \alpha x + \beta, \alpha > 0$
Metrik		jarak	
	Skala Ratio	ekuivalen, berurut,	
		jarak, rasio	
	Skala Absolute	ekuivalen, berurut,	fungsi identitas
		jarak, rasio,	
		tingkat absolute	

1.6 Variabel yang Bersifat Kualitatif

Skala Nominal

Skala yang paling primitif, yaitu yang hanya bisa menyatakan apakah dua buah nilai sama atau tidak, adalah skala nominal dan murni kualitatif. Jika sebuah ruang sampel eksperimen terdiri dari kategori tanpa urutan yang alami, maka variabel acak yang berkaitan terskala secara nominal.

Angka-angka/jumlah berbeda yang ditugaskan untuk memberi hasil biasanya mengindikasikan apakah sembarang dua keluaran (outcome) sama atau tidak.

Sebagai contoh, angka-angka yang ditentukan pada opini politik yang berbeda bisa menolong dalam mengkompilasi hasil dari daftar pertanyaan.

Namun untuk membandingkan dua opini, kita hanya bisa merelasikannya sebatas kesamaan jenis atau tidak. Angka/jumlah tersebut tidak menghasilkan ranking. Variabel binary atau Variabel dikotomis adalah variabel yang memiliki tepat dua outcome yang eksklusif satu sama lainnya.

Jika angka-angka indikator yang ditentukan tersebut menyampaikan informasi tentang ranking kategori, maka variabel binari dapat juga dianggap terskala secara ordinal. Jika kategori (peristiwa) yang merupakan ruang sampel tersebut bersifat eksklusif satu dengan lainnya, misalnya sebuah elemen statistik bisa berhubungan dengan lebih dari satu kategori, maka variabel tersebut dinamakan kumulatif. Sebagai contoh, seseorang bisa menerima kualifikasi kategori profesional yang berbeda. Tapi hanya bisa satu pekerjaan saja yang fulltime (menurut definisi).

Skala Ordinal

Jika jumlah yang ditentukan untuk pengukuran memperlihatkan suatu ranking alami, maka variabel tersebut diukur dengan skala ordinal. Pada skala ordinal, jarak antara nilai-nilai yang berbeda tidak dapat ditafsirkan -sebuah variabel yang diukur pada skala ordinal bagaimanapun bukan kuantitatif. Sebagai contoh, nilai sekolah merefleksikan tingkatan prestasi yang berbeda-beda. Walaupun demikian tidak ada alasan untuk menyatakan bahwa pekerjaan yang memperoleh nilai "4" dua kali lebih baik dari pekerjaan yang beroleh nilai "2". Karena jumlah yang ditentukan pada pengukuran tersebut merefleksikan ranking secara relatif antara satu dengan lainnya, maka nilai tersebut dinamakan nilai rank. Terdapat banyak contoh variabel yang terskala ordinal dalam bidang psikologi, sosiologi, studi bisnis dan lain sebagainya. Skala dapat ditentukan dengan mengusahakan mengukur semacam konsep seperti status sosial, inteligensia, tingkatan agresi atau tingkat kepuasan.

1.7 Variabel Kuantitatif

Pengukuran variabel kuantitatif dapat juga ditafsirkan dengan istilah jarak antara berbagai observasi tanpa adanya pengurutan yang alami.

Skala Interval

Jika jarak antara pengukuran ditafsirkan lebih mendalam, maka variabel tersebut diukur pada skala interval.

Berlawanan dengan skala rasio, rasio pengukuran tidak mempunyai pengertian yang mendasar, dan untuk skala ordinal tidak memiliki nilai nol. Sebagai contoh, suhu yang diukur dalam derajat celsius bisa ditafsirkan dalam urutan kelas-kelas yang lebih tinggi atau lebih rendah. Namun demikian, tinggi suhu bertemperatur 20 derajat celsius tidak bisa dikatakan dua kali tinggi suhu bertemperatur 10 derajat. Ingat kembali temperatur ekuivalensi dalam fahrenheit. Dengan mengkonversikan temperatur dari celsius ke fahrenheit atau sebaliknya akan mengikutsertakan pertukaran titik nol.

Skala Rasio (Nisbah)

Nilai-nilai variabel yang terukur pada skala rasio dapat ditafsirkan baik dengan istilah jarak *ataupun* rasio. Skala rasio atau nisbah ini membawa informasi yang lebih daripada skala interval, namun hanya interval-interval (jarak diantara observasi) yang memiliki arti kuantitatif.

Fenomena yang diukur pada skala rasio mempunyai elemen nol alami, yang menunjukkan kekurangan total dari atributnya. Namun adanya satuan pengukuran alami tidak begitu diperlukan. Contohnya adalah berat, tinggi, usia, dan lain sebagainya.

Skala Absolut

Skala absolut merupakan skala metrik yang memiliki satuan pengukuran alami. Pengukuran skala absolut sendiri adalah pengukuran, satu-satunya pengukuran tanpa alternatif. Contoh: Semua fenomena yang dapat dihitung seperti jumlah orang dalam sebuah ruangan atau jumlah bola dalam sebuah keranjang.

Variabel Diskret

Sebuah variabel diskret yang mengambil himpunan nilai-nilai berhingga atau tak hingga disebut diskret. Contoh: Produksi mobil perbulan atau jumlah bintang di ruang angkasa.

Variabel Kontinyu

Suatu variabel metrik dikatakan kontinyu jika mempunyai sejumlah tak hingga nilai-nilai pada sembarang interval. Contoh: Petrol yang terjual pada periode waktu tertentu.

Namun dalam prakteknya banyak variabel kontinyu diukur secara diskret dikarenakan terbatasnya tingkat ketelitian peralatan ukur fisik. Untuk mengetahui usia seseorang dapat dilakukan, tetapi tidak terlalu tepat.

1.8 Pengelompokan Data Kontinu

Misalkan ada data tentang tinggi badan dari 100 siswa laki-laki. Untuk memperoleh informasi sekilas dari distribusi tinggi badan, maka kita mulai membaca data mentah. Seorang yang ahli akan segera menemukan bahwa kira-kira lebih 10 observasi menjadi tidak berarti jika tanpa melakukan proses simplifikasi. Secara intuitif, individu-individu tersebut dikelompokkan berdasarkan kesamaan tinggi badan. Sekilas informasi tentang data dapat diperoleh dengan memperhatikan besarnya pengelompokan ini daripada hanya membaca data secara umum. Disamping telah mendapat informasi detail tentang tinggi badan yang sebenarnya, diperoleh juga gambaran umum yang lebih jelas tentang data.

Sampel data yang dikumpulkan dari variabel yang **kontinu** atau quasi-kontinu dapat disusun dengan membaginya ke dalam ruang sample yang saling mutually exclusive. Menghitung banyaknya kejadian yang muncul dalam tiap kelas atau kelompok berarti mendiskripsikan data secara ringkas. Mengelompokkan data ke dalam kelas-kelas tertentu akan sangat membantu kita dalam memahami struktur data atau penyebaran data dalam ruang sample. Kelas-kelas adalah interval-interval yang tidak saling beririsan (non-overlapping) yang dibatasi oleh batas-batas atas dan bawah yang sering disebut class boundaries.

Hilangnya informasi terjadi dari penukaran nilai aktual karena ukuran dan lokasi kelas-kelasnya. Jika seseorang menggunakan kelas yang terlalu sedikit, maka pola-pola yang berguna bisa jadi tidak terlihat. Terlalu banyak kelas juga bisa menghalangi nilai pengelompokan yang jelas.

Batas Kelas

Nilai atas dan bawah suatu kelas disebut batas kelas.

Sebuah kelas j dispesifikasikan oleh batas bawahnya yaitu a_j dan batas atas a_j^* ($j = 1, \dots, k$), dimana

$a_j^* = a_{j+1}$ ($j = 1, \dots, k - 1$), misalnya, batas atas dari kelas ke j dan batas bawah dari kelas ke $(j + 1)$ sama.

$a_j^* < a \leq a_j^*$ atau $a_j^* \leq a < a_j^*$ ($j = 1, \dots, k$), misalnya, batas kelas dapat dijadikan tanda untuk kelas-kelas lain yang terpisah.

Contoh

1st alternatif		2nd alternatif	
kurang dari 10	< 10	≤ 10	kurang dari atau sama dengan 10
10 hingga kurang dari 12	$\geq 10, < 12$	$> 10, \leq 12$	besar dari 10 hingga kecil dari atau sama dengan 12

12 hingga kecil dari 15	$\geq 12, < 15$	$> 12, \leq 15$	besar dari 12 hingga kecil dari atau sama dengan 15
15 atau besar dari	≥ 15	> 15	besar dari 15

Jika pengukuran variabel-variabel (secara teoritis) tidak memiliki batas diklasifikasikan, kelas yang terletak paling kiri- dan/atau paling kanan secara berturut-turut akan menuju $-\infty$, $+\infty$, misalnya, mereka membentuk suatu interval yang setengah terbuka.

Lebar Kelas

Dengan mengambil selisih diantara dua batas dari suatu kelas akan diperoleh lebar kelas (kadang-kadang merujuk ke ukuran kelas):

$$\Delta a_j = a_j^* - a_j^* \quad (j = 1, \dots, h)$$

Kelas-kelas tidak mesti memiliki lebar yang sama.

Titik tengah kelas

Suatu titik tengah kelas a_j dapat ditafsirkan sebagai nilai yang mewakili kelas tersebut, jika pengukuran yang dilakukan berdistribusi genap atau simetris.

$$a_j = \frac{a_j^* + a_j^*}{2} \quad (j = 1, \dots, h)$$

1.9 Barisan Statistik dan Frekuensi Statistik

Barisan Statistik

Ketika menyimpan data, kita membangkitkan sebuah barisan statistik. Bentuk aslinya yaitu yang belum diproses disebut *data mentah*. Diberikan sebuah level **skala** yang sesuai (misalnya, sebuah skala ordinal), kita bisa menyortir data kemudian membuat sebuah *barisan yang berurut*.

Data yang dikumpulkan pada titik-titik yang sama dalam suatu waktu atau pada periode waktu yang berbeda pada elemen yang berbeda disebut data cross-section.

Data yang dikumpulkan pada titik-titik yang berbeda dalam suatu waktu atau pada periode waktu yang berbeda pada elemen yang sama disebut data time-series. Barisan observasi tersebut diurutkan pada sumbu axis sebagai waktu.

Frekuensi

Sejumlah observasi yang termasuk ke dalam kelas yang ada disebut **frekuensi**.

Kelas-kelas dikonstruksi untuk menyingkatkan data kontinyu atau quasi-kontinyu dengan menggunakan frekuensi.

Pada data diskret kita kadang-kadang kebetulan menemukan apa disebut ikatan, misalnya dua observasi atau lebih yang mengambil nilai yang sama. Maka data diskret tidak memerlukan pengelompokan yang bertujuan untuk menghitung frekuensi.

Frekuensi Absolut

Menghitung jumlah observasi yang mengambil nilai yang spesifik menghasilkan frekuensi absolut:

$$h(X = a_j) = h(a_j) = h_j$$

Ketika data dikelompokkan, *frekuensi-frekuensi kelas absolute* dihitung sebagai berikut:

$$h(a_j) = h(a_j^l \leq X < a_j^u)$$

Sifat:

$$0 \leq h(a_j) \leq n$$

$$\sum_j h(a_j) = n$$

Frekuensi Relatif

Bagian observasi yang mengambil nilai spesifik atau masuk ke dalam kelas yang spesifik disebut frekuensi relatif, frekuensi absolut yang memiliki standar dari jumlah total observasi.

$$f(x_j) = \frac{h(x_j)}{n}$$

Sifat:

$$0 \leq f(x_j) \leq 1$$

$$\sum_j f(x_j) = 1$$

Distribusi Frekuensi

Dengan mengstandarisasi frekuensi-frekuensi kelas untuk data yang dikelompokkan oleh ukuran kelasnya masing-masing, frekuensi untuk kelas yang memiliki ukuran berbeda dibuat hingga bisa dibandingkan. Frekuensi yang muncul kemudian bisa dikumpulkam untuk membentuk sebuah distribusi frekuensi.

$$\hat{h}(x_j) = \frac{h(x_j)}{x_j^u - x_j^l}$$

$$\hat{f}(x_j) = \frac{f(x_j)}{x_j^u - x_j^l}$$

dimana x_j^l, x_j^u adalah batas kelas atas dan bawah dengan $x_j^l < x \leq x_j^u$.
