

**PENGANTAR**

# *Statistik Sosial*

Zakirman  
Zulfa



STKIP PGRI Sumbar Press  
Jalan Diponegoro No. 100  
Gugur II, Kota Pekanbaru 28112 Riau

0 1689 70037 6 003762



9 786237 003762



STKIP PGRI Sumbar Press

ISBN 978-623-7003-76-2

# **Pengantar Statistik Sosial**

**Zakirman, Zulfa**



**STKIP PGRI Sumbar Press**

## **Pengantar Statistik Sosial**

Pengarang : Zakirman, Zulfa

**ISBN: 978-623-7003-76-2**

Editor:  
Villia Anggraini

Foto:  
Zulfa

Desain Sampul :  
Marizal Hodhol

Ilustrasi Dalam:  
Zulfa

Tata Layout:  
Harizqi Azri

Penerbit:  
STKIP PGRI Sumbar Press

*Kantor Pusat:*

Gd. A Lt. 2 Kampus I STKIP PGRI Sumatera Barat Jl. Gajah Mada Gunung Pangilun Kota Padang, Phonecell/Telp: 085365372924/ (0751) 7053731. Email: [p3mstkipgrisumbar@gmail.com](mailto:p3mstkipgrisumbar@gmail.com)

Cetakan Pertama, 2020

Hak cipta dilindungi undang-undang

Dilarang memperbanyak karya tulis ini dalam bentuk dan dengan cara apapun tanpa ijin tertulis dari penerbit

## KATA PENGANTAR

Syukur Alhamdulillahirabbil'aalamin, segala puji syukur penulis ucapkan kepada Allah Yang Maha Penyayang. Tanpa karunia-Nya mustahil naskah buku ini diselesaikan dengan baik. Buku Ajar Pengantar Statistik Sosial ini ditulis oleh **Dr. Zakirman, M.Pd**, dan **Zulfa, M.Pd, M.Hum**, Mata kuliah ini merupakan mata kuliah di program studi pendidikan Sejarah STKIP PGRI Sumatera Barat. Mata kuliah merupakan mata kuliah yang membutuhkan pemahaman, penguasaan mahasiswa dalam konsep dan teori pengantar statistik sosial. Mata kuliah ini bertujuan memberikan pengetahuan, pemahaman dan keterampilan kepada mahasiswa tentang memahami konsep dan teori-teori Statistik sosial dan implementasinya dalam masalah sosial maupun pendidikan, sehingga dalam mengkaji masalah sosial dapat diterapkan pendekatan kuantitatif. terselesainya penulisan buku ini juga tidak terlepas dari bantuan banyak pihak. Karena itu, Tim penulis menyampaikan terimakasih yang tidak terhingga kepada Bapak Ibu Ketua STKIP PGRI Sumatera Barat Dr. Zuzmelia, M.Si, Wakil Ketua I Sri Imelwaty Ph.D, M.Pd, Wakil Ketua II Ibu Liza Husnita, M.Pd dan Wakil Ketua III Bapak Jaruddin Ph.D. Dan semua bentuk kemudahan yang telah diberikan oleh UP3M, Kampus STKIP PGRI Sumatera Barat, dan Yayasan PGRI Padang benar-benar bermanfaat bagi Tim Penulis Buku Ajar Pengantar Statistik Sosial. Selain itu, Tim penulis Buku Ajar juga menyampaikan rasa terimakasih segenap kepada Program Studi Pendidikan Sejarah untuk pimpinan prodi dan semua dosen, Teman-teman seperjuangan di Program Pasca S3 bunda Dr. Gustina, M.Pd, Dian Hasfera, M.I.Kom , , Khairul Amri, M.Pd, dan Siti Aisyah, S.Pdi, M.Pd Terima kasih Untuk semua bantuan motivasi, dan saran-sarannya. Meskipun telah berusaha untuk menghindari kesalahan, Tim penulis menyadari juga bahwa buku ini masih mempunyai kelemahan sebagai kekurangannya. Karena itu, tim penulis berharap agar pembaca berkenan menyampaikan kritik. Dengan segala pengharapan dan keterbukaan tim penulis menyampaikan rasa terimakasih dengan setulus-tulusnya. Kritik merupakan perhatian agar dapat menuju kesempurnaan buku ajar ini. Akhir kata tim penulis berharap agar buku ini dapat menuju kesempurnaan. Dan berharap agar buku ini dapat membawa manfaat kepada mahasiswa. Tim penulis berharap semoga buku ini dapat bermanfaat bagi mahasiswa yang fokus ke penelitian pendidikan kuantitatif. Akhir kata tim penulis berharap agar buku ajar ini dapat membawa manfaat kepada mahasiswa.

Padang, 19 September 2020

Tim Penulis

## DAFTAR ISI

KATA PENGANTAR.....	i
DAFTAR ISI.....	ii
CHAPTER 1_PENYAJIAN DATA STATISTIK.....	1
CHAPTER 2_DISTRIBUSI FREKUENSI.....	8
CHAPTER 3_UKURAN PEMUSATAN DAN LETAK DATA.....	12
CHAPTER 4_UKURAN PENYEBARAN DATA.....	16
CHAPTER 5 HIPOTESIS PENELITIAN PENDIDIKAN.....	20
CHAPTER 6 PENARIKAN SAMPEL DALAM PENELITIAN .....	25
CHAPTER 7_PENGANTAR STATISTIK PENDIDIKAN DAN SOSIAL.....	29
CHAPTER 8 UJI T DALAM STATISTIK PARAMETRIK.....	38
CHAPTER 9_ONE WAY ANOVA/ANOVA 1 ARAH .....	44
CHAPTER 10 ANALISIS REGRESI DAN KORELASI SEDERHANA.....	51
CHAPTER 11_ANALISIS FAKTOR .....	58
CHAPTER 12_UJI DALAM STATISTIK NON PARAMETRIK (MANN WHITNEY) .....	63
REFERENSI .....	68

## CHAPTER 1

### PENYAJIAN DATA STATISTIK

#### **Tinjauan Umum Statistik**

Statistik dapat didefinisikan sebagai ilmu yang berkaitan dengan mengembangkan dan mempelajari serta metode untuk mengumpulkan, menganalisis, menafsirkan dan menyajikan data empiris. Kegunaan statistik dapat terlihat hampir dalam keseluruhan bidang disiplin penelitian, seperti pendidikan, social, eksakta, ekonomi, manajemen, pertanian dan masih banyak lainnya. Perhitungan untuk statistik saat ini di dominasi oleh perhitungan dengan aplikasi pintar seperti *SPSS* dan beberapa peneliti juga menggunakan *Microsoft Excel* sebagai bantuan dalam penyelesaian perhitungan. Dua gagasan yang bersifat mendasar dalam bidang statistik adalah ketidakpastian dan variasi. Beberapa kondisi dan situasi yang kita temui dalam sains (atau lebih umum dalam kehidupan) memiliki hasil yang bersifat tidak pasti.

Probabilitas dapat dikatakan sebagai bahasa matematika yang digunakan untuk mendiskusikan kejadian yang tidak pasti dan probabilitas memainkan peran kunci dalam statistik. Setiap upaya pengukuran atau pengumpulan data tunduk pada sejumlah sumber variasi. Dengan ini dapat diartikan bahwa jika pengukuran yang sama diulang, maka jawabannya akan ada kemungkinan terjadinya perubahan.

Sederhananya, statistik adalah berbagai prosedur untuk mengumpulkan, mengatur, menganalisis dan menyajikan data kuantitatif. 'Data' adalah istilah untuk fakta-fakta yang telah diperoleh dan kemudian dicatat, dan untuk ahli statistik, 'data' biasanya mengacu pada data kuantitatif yang berupa angka. Statistik dapat membantu mengubah data menjadi informasi; yaitu, data yang telah ditafsirkan, dipahami dan berguna bagi penerima.

Dalam istilah yang sangat luas, statistik dapat dibagi menjadi dua cabang, yaitu statistik deskriptif dan inferensial. Statistik deskriptif berkaitan dengan data kuantitatif dan metode untuk mendeskripsikannya. Statistik inferensial (analitis) membuat kesimpulan tentang populasi (seluruh kelompok orang atau lebih) dengan menganalisis data yang dikumpulkan dari sampel (bagian dari seluruh kelompok) dan berhubungan dengan metode yang memungkinkan penarikan kesimpulan dari data ini.

### **A. Penyajian Data Statistik**

Data biasanya dikumpulkan dalam format mentah dan dengan demikian informasi yang terdapat didalamnya sulit dipahami. Oleh karena itu, data mentah perlu dirangkum, diproses, dan dianalisis. Informasi yang berasal dari data mentah harus disajikan dalam format yang efektif, jika tidak, akan menjadi kerugian bagi penulis itu sendiri maupun bagi pembaca. Teknik penyajian data dan informasi dapat berupa: tekstual, tabular, dan grafik. Teks adalah metode utama untuk menjelaskan temuan, menguraikan dan memberikan informasi kontekstual. Sebuah tabel paling cocok untuk merepresentasikan informasi individu dan merepresentasikan informasi kuantitatif dan kualitatif. Grafik adalah alat visual yang sangat efektif karena menampilkan data secara sekilas, memfasilitasi perbandingan, dan dapat mengungkap hubungan dalam data seperti perubahan dari waktu ke waktu, distribusi frekuensi, dan korelasi dari keseluruhan. Teks, tabel, dan grafik untuk presentasi data dan informasi adalah alat komunikasi yang sangat kuat, dapat membuat artikel mudah dimengerti, menarik dan mempertahankan minat pembaca, serta efisien menyajikan informasi kompleks dalam jumlah besar.

Bagian utama dari Statistik adalah tampilan data yang dirangkum. Data awalnya dikumpulkan dari sumber yang diberikan, apakah itu eksperimen, survei, observasi, dan disajikan dalam bentuk sebagai berikut:

## 1. Metode teks

Teks adalah metode utama untuk menyampaikan informasi dan digunakan untuk menjelaskan hasil serta memberikan informasi kontekstual. Data pada dasarnya disajikan dalam paragraf atau kalimat. Teks dapat digunakan untuk memberikan interpretasi atau penekanan data tertentu. Jika informasi kuantitatif yang ingin disampaikan terdiri dari satu atau dua angka, lebih tepat menggunakan bahasa tulisan dari tabel atau grafik. Misalnya, informasi tentang tingkat kejadian delirium setelah anestesi di 2016-2017 dapat disajikan dengan penggunaan beberapa nomor: "Insiden tingkat delirium setelah anestesi adalah 11% pada tahun 2016 dan 15% di 2017; tidak ada perbedaan signifikan dari tingkat insiden yang ditemukan antara dua tahun." Jika informasi ini disajikan dalam grafik atau tabel, tentu akan menghabiskan ruang yang cukup besar. Jika lebih banyak data yang akan disajikan, atau informasi lainnya seperti yang menyangkut data terkini yang harus disampaikan, tabel atau grafik akan lebih tepat. Secara alami, data yang disajikan dalam teks akan membutuhkan waktu lebih lama untuk pembacaan dan pemahamannya.

## 2. Penyajian dalam bentuk tabel

Tabel berperan dalam menyampaikan informasi yang telah dikonversi menjadi kata-kata atau angka dalam baris dan kolom, dan telah digunakan mulai 2.000 tahun lalu. Siapapun dengan tingkat keaksaraan yang rendah dapat dengan mudah memahami informasi yang disajikan dalam sebuah tabel. Tabel dapat dikatakan sebagai metode yang paling tepat untuk menyajikan informasi individu, dan dapat menyajikan informasi kuantitatif dan kualitatif. Kelebihan tabel adalah dapat menyajikan informasi secara akurat dan itu tidak dapat dalam penyajian secara grafik. Sebuah angka seperti "132.145852" dapat diekspresikan secara akurat dalam sebuah tabel. Kekuatan lain adalah informasi dengan unit yang berbeda dapat disajikan bersama.



Misalnya, tekanan darah, denyut jantung, jumlah obat yang diberikan, dan waktu anestesi bisa disajikan bersama dalam satu tabel. Dapat disimpulkan bahwa tabel berguna untuk meringkas dan membandingkan informasi kuantitatif yang berbeda variabel. Namun pada sisi lain dalam penggunaan tabel, interpretasi informasi membutuhkan waktu yang lebih lama jika dibandingkan dengan grafik.

**Table 1.** Modified Table in Lee and Kim's Research (Adapted from Korean J Anesthesiol 2017; 70: 39-45)

Variable	Group	Baseline	After drug	1 min	3 min	5 min
SBP	C	135.1 ± 13.4	139.2 ± 17.1	186.0 ± 26.6*	160.1 ± 23.2*	140.7 ± 18.3
	D	135.4 ± 23.8	131.9 ± 13.5	165.2 ± 16.2* <sup>†</sup>	127.9 ± 17.5 <sup>†</sup>	108.4 ± 12.6 <sup>†,†</sup>
DBP	C	79.7 ± 9.8	79.4 ± 15.8	104.8 ± 14.9*	87.9 ± 15.5*	78.9 ± 11.6
	D	76.7 ± 8.3	78.4 ± 6.3	97.0 ± 14.5*	74.1 ± 8.3 <sup>†</sup>	66.5 ± 7.2 <sup>†,†</sup>
MBP	C	100.3 ± 11.9	103.5 ± 16.8	137.2 ± 18.3*	116.9 ± 16.2*	103.9 ± 13.3
	D	97.7 ± 14.9	98.1 ± 8.7	123.4 ± 13.8* <sup>†</sup>	95.4 ± 11.7 <sup>†</sup>	83.4 ± 8.4 <sup>†,†</sup>

Gambar 1. Salah Satu Contoh Penyajian Data Dalam Bentuk Tabel

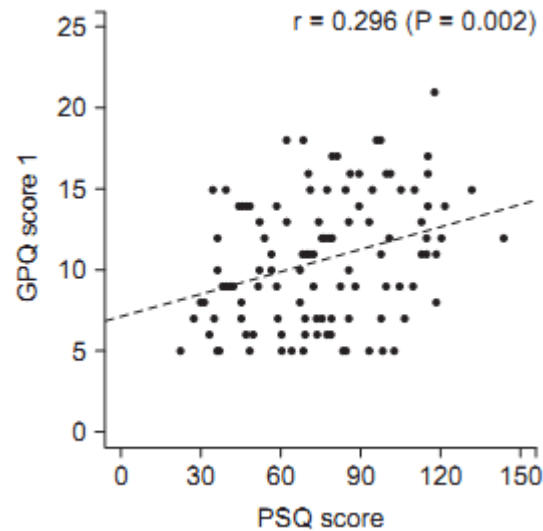
### 3. Penyajian dalam bentuk grafik

Pada pembahasan sebelumnya, tabel dapat digunakan untuk menyajikan berbagai informasi yang didapat dalam kegiatan penelitian. Grafik dapat digunakan untuk menyederhanakan informasi kompleks dengan menggunakan gambar dan menekankan pola data, dan berguna untuk meringkas, menjelaskan, atau menjelajahi data kuantitatif. Grafik efektif untuk menyajikan data dalam jumlah besar dan bisa digunakan sebagai pengganti tabel untuk menyajikan set data kecil. Format penyajian data dalam grafik harus dipertimbangkan agar pembaca dan pengulas dapat dengan mudah memahami informasi. Penyajian data menggunakan grafik dapat diklasifikasikan menjadi beberapa kelompok, yaitu:

#### a. Sebaran menggunakan plot (*Scatter Plot*)

*Scatter Plot* menyajikan data pada sumbu x dan y dan digunakan untuk menyelidiki hubungan antara dua variabel. Sebuah titik mewakili setiap individu atau objek, dan hubungan antara dua variabel dapat dipelajari dengan menganalisis pola-pola di beberapa poin. Garis

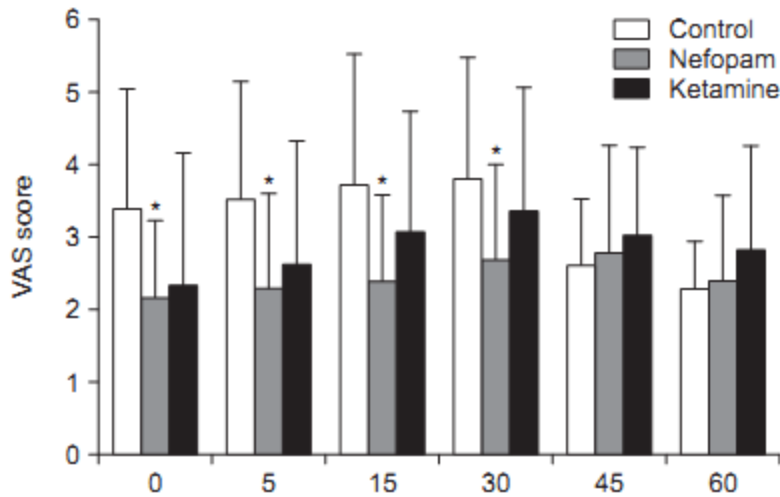
regresi ditambahkan ke grafik untuk menentukan apakah hubungan antara dua variabel dapat dijelaskan atau tidak.



Gambar 2. Sebaran Menggunakan *Scatter Plot*

b. Grafik batang dan Histogram

Grafik batang digunakan untuk menunjukkan dan membandingkan nilai dalam sebuah kategori atau kelompok diskrit, dan frekuensi atau ukuran parameter lainnya. Penyajian data dalam bentuk grafik batang dapat dibuat secara vertikal atau horizontal tergantung pada jumlah kategori, ukuran atau kompleksitas masing-masing kategori,. Tinggi (atau panjang) bar mewakili jumlah informasi dalam suatu kategori. Grafik batang fleksibel, dan dapat digunakan dalam format bar yang dikelompokkan atau dibagi lagi dalam dua atau lebih set data di setiap kategori. Salah satu bentuk grafik batang vertikal adalah bilah vertikal yang ditumpuk dalam sebuah penyajian data. Tumpukan grafik batang vertikal digunakan untuk membandingkan jumlah masing-masing kategori, dan menganalisis bagian-bagian dari suatu kategori. Grafik batang yang ditumpuk secara vertikal sangat baik dari aspek visualisasi, tidak memiliki garis referensi, membuat perbandingan bagian menjadi lebih jelas.



Gambar 3. Sampel Penyajian Data dalam Bentuk Grafik Batang

Penyajian data menggunakan grafik tidak hanya terfokus pada bentuk batang saja, namun dalam berbagai penyajian lainnya ada beberapa jenis bentuk grafik lainnya yaitu: bentuk lingkaran, bentuk garis, bentuk titik, berbentuk kotak dan lain-lain.

### Tugas 1

*Kerjakanlah soal-soal berikut dengan cermat dan teliti.*

Isilah titik-titik berikut dengan angka yang memiliki rentang dari 0 sampai 30

**“Data indeks pencemaran udara pada beberapa kota di Indonesia”** (soal 1 – 2)

Hari	PDG	JKT	PKU	BDG	MKSR	BALI	MDN	PONT
Senin	8	...	16	12	...	20	24	22
Selasa	4	...	20	...	6	...	...	...
Rabu	...	...	22	10	...	22	...	20
Kamis	10	...	...	...	8	...	10	...
Jumat	...	...	24	12	...	18	...	16
Sabtu	12	...	28	...	8	...	12	...
Minggu	10	...	...	16	...	16	...	14

1. Buatlah kesimpulan dalam bentuk teks mengenai ringkasan tabel diatas!
2. Gambarkanlah grafik sesuai dengan data indeks pencemaran diatas untuk kota:
  - a. Padang (batang)
  - b. Medan (lingkaran)
  - c. Bali dan pekanbaru (batang)
  - d. Keseluruhan kota (grafik batang bertumpuk)
3. Jelaskanlah keuntungan penggunaan grafik dalam penyajian data?
4. Uraikanlah konteks dan cakupan statistik dalam bidang penelitian?

## CHAPTER 2

### DISTRIBUSI FREKUENSI

Langkah berikutnya setelah selesainya pengumpulan data adalah mengatur data ke dalam bentuk yang bermakna sehingga pembacaan data menjadi lebih mudah. Salah satu metode umum untuk mengatur data adalah membangun distribusi frekuensi.

#### **A. Distibusi Frekuensi**

Distribusi frekuensi adalah tabulasi terorganisir/representasi grafis dari jumlah individu dalam setiap kategori pada skala pengukuran. Hal ini memungkinkan peneliti untuk melihat seluruh data dengan mudah. Ini menunjukkan apakah pengamatan tinggi atau rendah dan juga apakah data terkonsentrasi di satu area atau tersebar di seluruh skala. Dengan demikian, distribusi frekuensi menyajikan gambaran tentang bagaimana pengamatan individu didistribusikan dalam skala pengukuran.

*Short definition:*

*Distribusi Frekuensi: nilai dan frekuensinya (seberapa sering setiap nilai terjadi).*

Distribusi frekuensi merupakan ikhtisar dari semua nilai yang berbeda dalam beberapa variable. Distribusi frekuensi memberitahu bagaimana frekuensi didistribusikan di atas nilai. Distribusi frekuensi sebagian besar digunakan untuk meringkas variabel kategori. Itu karena variabel cenderung memiliki banyak nilai yang berbeda.

Distribusi frekuensi dapat disajikan sebagai tabel frekuensi, histogram atau diagram batang. Baik histogram dan diagram batang menyediakan tampilan visual menggunakan kolom,

dengan sumbu y mewakili jumlah frekuensi, dan sumbu x yang mewakili variabel yang akan diukur.

## **B. Tabel Frekuensi**

Tabel frekuensi (distribusi) menunjukkan kategori pengukuran yang berbeda dan jumlah pengamatan dalam setiap kategori. Sebelum merancang tabel frekuensi, seseorang harus memiliki gagasan tentang kisaran (nilai minimum dan maksimum). Rentang ini dibagi menjadi interval arbitrer yang disebut "interval kelas." Jika interval kelas terlalu banyak, maka tidak akan ada pengurangan dalam jumlah besar data. Di sisi lain, jika interval sangat sedikit, maka bentuk distribusi itu sendiri tidak dapat ditentukan. Umumnya, interval yang memadai berkisar antara 6 sampai 14 kelas. Lebar kelas dapat ditentukan dengan membagi rentang pengamatan dengan jumlah kelas. Berikut ini adalah beberapa panduan mengenai lebar kelas:

1. Dianjurkan untuk memiliki lebar kelas yang sama. Lebar kelas yang tidak merata hanya boleh digunakan jika ada kesenjangan besar dalam data.
2. Interval kelas harus saling eksklusif dan tidak tumpang tindih.
3. Kelas terbuka di sisi bawah dan atas (misalnya,  $<10, > 100$ ) dan harus dihindari.

## **C. Grafik Distribusi Frekuensi**

Grafik distribusi frekuensi adalah ilustrasi diagram dari informasi dalam tabel frekuensi. Histogram adalah representasi grafis dari variabel yang berada dalam sumbu X dan jumlah pengamatan (frekuensi) dalam sumbu Y. Persentase dapat digunakan jika tujuannya adalah untuk membandingkan dua histogram yang memiliki jumlah subyek yang berbeda. Histogram digunakan untuk menggambarkan frekuensi ketika data diukur pada interval atau skala rasio.

Ada empat karakteristik penting dari distribusi frekuensi, yaitu:

1. Ukuran tendensi sentral dan lokasi (mean, median, mode)

2. Ukuran dispersi (kisaran, varians, standar deviasi)
3. Luasnya simetri / asimetri (kemiringan)
4. Kerataan atau puncaknya (kurtosis).

#### **D. Distribusi Frekuensi Relative**

Distribusi frekuensi relatif adalah distribusi di mana frekuensi relatif dicatat terhadap setiap interval kelas. Frekuensi relatif dari suatu kelas adalah frekuensi yang diperoleh dengan membagi frekuensi dengan frekuensi total. Frekuensi relatif adalah proporsi frekuensi total yang ada dalam interval kelas tertentu dalam distribusi frekuensi.

##### **1. Tabel Distribusi Frekuensi Relatif**

Jika tabel distribusi frekuensi diubah menjadi frekuensi relatif maka tabel distribusi frekuensi disebut sebagai tabel distribusi frekuensi relatif. Untuk satu set data yang terdiri dari  $n$  nilai. Jika  $f$  adalah frekuensi nilai tertentu maka rasio ' $\frac{f}{n}$ ' disebut frekuensi relatifnya. *[Kesalahan Pemrosesan Matematika/ Math Processing Error]*

Contoh soal:

Find the relative frequency from the data given below:  
ctt:  $n = 30$ .

Class interval	Frequency
20-25	10
25-30	12
30-35	8
35-40	20
40-45	11
45-50	4
50-55	5

Relative frequency distribution table for the given data.

Here  $n = 70$

Class interval	Frequency (f)	Relative Cumulative Frequency ( <i>[Math Processing Error]</i> $f/n$ )
20-25	10	$10 / 70 = 0.143$
25-30	12	$12 / 70 = 0.171$
30-35	8	$8 / 70 = 0.114$
35-40	20	$20 / 70 = 0.286$
40-45	11	$11 / 70 = 0.157$
45-50	4	$4 / 70 = 0.057$
50-55	5	$5 / 70 = 0.071$
<b>Total</b>	$n = 70$	

## 2. Distribusi Frekuensi Berkelompok

Distribusi frekuensi berkelompok adalah daftar variabel yang tertera [Kesalahan Pemrosesan Matematika], dimasukkan ke dalam satu grup dalam satu kolom dengan daftar pada kolom kedua/kolom frekuensi. Distribusi frekuensi yang dikelompokkan adalah interval kelas dan frekuensi yang sesuai dalam sebuah bentuk tabel.

### Tugas 2

1. These are the numbers of newspapers sold at a local shop over the last 10 days:

22, 20, 18, 23, 20, 25, 22, 20, 18, 20

Buatkanlah tabel distribusi frekuensi untuk data penjualan Koran tersebut?

2. Diketahui data hasil pengukuran tinggi badan beberapa orang anak sebagai berikut (cm):

150	155	145	140	150	145
160	155	150	145	140	160
160	150	145	140	140	160

- a. Buatlah tabel distribusi frekuensi data tinggi badan anak tersebut!
- b. Tentukanlah modus (data tinggi badan yang sering muncul dari pengukuran tersebut)?



## CHAPTER 3

### UKURAN PEMUSATAN DAN LETAK DATA

Ukuran pemusatan data dapat dimaknai sebagai sembarang ukuran yang menunjukkan pusat segugus data, yang telah diurutkan dari yang terkecil sampai yang terbesar atau sebaliknya dari yang terbesar sampai yang terkecil. Salah satu manfaat dari ukuran pemusatan data adalah untuk membandingkan dua (populasi). Pada kondisi nyata akan sangat sulit untuk membandingkan masing-masing anggota dari masing-masing populasi atau masing-masing anggota data contoh. Nilai ukuran pemusatan ini dibuat sedemikian rupa sehingga cukup mewakili seluruh nilai dalam data yang akan dianalisis.

#### A. Mean (Nilai Rata-Rata)

Mean adalah dalam bahasa sederhana dapat diartikan sebagai rata-rata. Mean dapat dihitung dengan mengambil semua nilai dalam satu set-data dan membaginya dengan jumlah total nilai dalam set-data tersebut.

Contoh perhitungan rata-rata (mean):

##### 1. Data tunggal

Nilai ulangan matematika 15 siswa kelas XIIPA adalah 7, 8, 6, 4, 10, 5, 9, 7, 3, 8, 6, 5, 8, 9, dan

7. Tentukan nilai rata-ratanya?

Penyelesaian:

$$\text{mean} = \frac{7 + 8 + 6 + 4 + 10 + 5 + 9 + 7 + 3 + 8 + 6 + 5 + 8 + 9 + 7}{15} = 6.8$$

Jadi nilai rata-rata/mean dari data diatas adalah 6.8

## 2. Data berkelompok

$$\bar{x} = \frac{\sum_{i=1}^n f_i x_i}{\sum_{i=1}^n f_i} \text{ atau } \bar{x} = \bar{x}_s + \frac{\sum_{i=1}^n f_i d_i}{\sum_{i=1}^n f_i}$$

Keterangan:

$x_i$  = nilai tengah data ke-i

$f_i$  = frekuensi data ke -i

$x_s$  = rata-rata sementara (dipilih pada interval dengan frekuensi terbesar)

$d_i$  = simpangan ke-i (selisih nilai  $x_i$  dengan nilai  $x_s$ )

### **B. Median (Nilai Tengah)**

Berbeda dengan mean, median berkaitan dengan nilai tengah. Median dapat didefinisikan sebagai angka yang berada di tengah-tengah kumpulan data yang diatur dari terendah ke tertinggi atau dari tertinggi ke terendah. Median tidak mewakili rata-rata yang sebenarnya, dalam mendapatkan median terlebih dahulu dilakukan pengurutan data dari nilai terkecil hingga nilai terbesar. Misalnya 3, 3, 6, 13, 100 = 6, Mediannya adalah 6. Jika ada dua nilai tengah, median setengah di antara keduanya. Ini mungkin bukan bilangan bulat. Dalam kajian ekonomi, Median ini sangat berguna untuk menggambarkan hal-hal seperti gaji.

### **C. Modus/Mode (Nilai yang Sering Muncul)**

Mode dari satu set nilai data adalah nilai yang paling sering muncul. Seperti mean statistik dan median, mode adalah cara mengungkapkan, dalam (biasanya) nomor tunggal, informasi penting tentang variabel acak atau populasi. Untuk menentukan nilai mean, dapat dilakukan dengan mengurutkan data terlebih dahulu, dimulai dari data terkecil hingga terbesar. Kemudian

dapat dihitung frekuensi kemunculan angka tersebut. Untuk data tunggal, biasanya akan terlihat jelas angka yang menjadi modus dalam data penelitian.

#### **D. Kuartil**

Kuartil adalah nilai yang membagi suatu data terurut menjadi empat bagian yang sama. Kuartil dilambangkan dengan Q . Jenis kuartil ada 3, yaitu kuartil pertama (Q1) , kuartil kedua (Q2), dan kuartil ketiga (Q3).

#### **E. Desil Persentil**

Desil merupakan nilai yang membagi data menjadi sepuluh bagian sama besar. Desil sering dilambangkan dengan D. jenis ada 6, yaitu D1 , D2 , D3, .....,D9. Persentil merupakan nilai yang membagi data menjadi seratus bagian sama besar. Persentil sering dilambangkan dengan P. jenis persentil ada 99, yaitu P1, P2, P3 ... P99.

#### **F. Pencilan**

Dalam statistik, outlier atau “pencilan” adalah sebuah datum yang menyimpang cukup jauh dari datum lainnya di dalam satu sampel atau kumpulan datum (kumpulan datum disebut data). Pencilan di dalam satu kumpulan datum dapat menjadi peringatan bagi ahli statistik akan adanya ketidaknormalan atau kesalahan eksperimen pada pengukuran yang diambil, yang dapat membuat ahli statistik membuang pencilan tersebut dari kumpulan datum. Jika ahli statistik membuang pencilan dari kumpulan datum, kesimpulan yang diambil dari penelitian dapat menjadi sangat berbeda. Oleh karena itu, mengetahui cara menghitung dan menganalisis pencilan sangat penting untuk memastikan pengertian yang benar dari sebuah kumpulan datum statistik.

### Tugas 3

Tentukanlah mean, median dan modus dalam data berikut:

3. 13, 18, 13, 14, 13, 16, 14, 21, 13?
4. 10, 9, 8, 7, 6, 6, 8, 11, 12, 14, 10, 10, 8, 10, 7?
5. 1, 2, 4, 7?
6. 8, 9, 10, 10, 10, 11, 11, 11, 12, 13?

## CHAPTER 4

### UKURAN PENYEBARAN DATA

#### A. Jangkauan

Dalam statistika, jangkauan dari sebuah kumpulan data dikenal sebagai selisih antara nilai paling besar dengan nilai paling kecil. Dalam bahasa yang sederhana, jangkauan adalah perbedaan antara nilai terbesar dan terkecil. Namun, dalam statistik deskriptif, konsep jangkauan ini memiliki makna yang lebih kompleks. Kisarannya adalah ukuran interval terkecil yang berisi semua data dan memberikan indikasi dispersi statistik.

Langkah yang harus dilakukan dalam menentukan jangkauan data adalah mengurutkan data tersebut dari yang terkecil hingga ke terbesar. Tahapan dalam menentukan jangkauan sebagai berikut:

1. Susunlah sekumpulan angka dari yang terkecil hingga yang terbesar.
2. Identifikasi angka-angka terkecil dan terbesarnya di dalam kumpulan data.
3. Kurangkan angka yang terkecil dari yang terbesar.
4. Tuliskan hasil jangkauan.

#### Jenis-jenis jangkauan:

1. *Jangkauan interkuartil* adalah selisih antara kuartil ketiga dan kuartil pertama.

$$H = Q_3 - Q_1$$

#### Keterangan :

H = jangkauan interkuartil

$Q_3$  = kuartil ketiga

$Q_1$  = kuartil pertama

2. *Simpangan kuartil ( jangkauan semi interkuartil)*

Simpangan kuartil adalah setengah dari selisih kuartil ketiga dan kuartil pertama.

$$S_k = \frac{1}{2} Q_3 - Q_1$$

**Keterangan :**

$S_k$  = simpangan kuartil

$Q_3$  = kuartil ketiga

$Q_1$  = kuartil pertama

## **B. Ragam**

Ragam atau variasi adalah nilai yang menunjukkan besarnya penyebaran data pada kelompok data. Ragam atau variasi dilambangkan dengan  $s^2$ .

### **1. Variasi untuk data tunggal**

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

**Keterangan :**

$s^2$  = variasi

$x_i$  = data ke -i

$\bar{x}$  = rata-rata hitung

$n$  = banyak data

### **2. Variasi untuk data bergolong (berkelompok)**

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

**Keterangan :**

$s^2$  = variasi

$x_i$  = data ke -i

$\bar{x}$  = rata-rata hitung

$f_i$  = frekuensi data ke-i

**C. Simpangan Baku**

Simpangan baku atau disebut juga deviasi standar merupakan akar dari jumlah kuadrat deviasi dibagi banyaknya data. Simpangan baku sering dilambangkan dengan s.

**1. Simpangan baku untuk data tunggal**

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

**Keterangan :**

S = simpangan baku

$x_i$  = data ke -i

$\bar{x}$  = rata-rata hitung

n = banyak data

**2. Simpangan baku untuk data bergolong (berkelompok)**

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

**Keterangan :**

$s$  = simpangan baku

$x_i$  = data ke  $-i$

$\bar{x}$  = rata-rata hitung

$f_i$  = frekuensi data ke- $i$

**Tugas 4**

1. Diberikan sampel dengan data sebagai berikut: 8, 7, 10, 11, 14. Tentukanlah:
  - a. Rata-rata data
  - b. Simpangan baku dan varians data
2. Tentukanlah varians dan simpangan baku untuk nilai ujian 80 mahasiswa yang terangkum dalam tabel berikut ini:

Nilai Ujian	$f_i$	$X_i$
31-40	1	35.5
41-50	2	45.5
51-60	5	55.5
61-70	15	65.5
71-80	25	75.5
81-90	20	85.5
91-100	12	95.5
Jumlah	80	-



## CHAPTER 5

### HIPOTESIS PENELITIAN PENDIDIKAN

#### **A. Defenisi Hipotesis**

Hipotesis dalam statistik adalah asumsi tentang parameter populasi. Asumsi ini mungkin atau mungkin tidak benar. Hipotesis statistic sering juga disebut sebagai analisis data konfirmatori, dimana hipotesis yang dapat diuji berdasarkan pengamatan proses yang dimodelkan melalui serangkaian variabel acak. Uji hipotesis statistik adalah metode inferensi statistik. Umumnya, dua set data statistik dibandingkan, atau satu set data yang diperoleh dengan sampling dibandingkan terhadap kumpulan data sintetis dari model yang ideal. Sebuah hipotesis diusulkan untuk hubungan statistik antara dua set data, dan ini dibandingkan sebagai alternatif untuk hipotesis nol yang ideal yang tidak mengusulkan hubungan antara dua set data. Perbandingan ini dianggap signifikan secara statistik jika hubungan antara kumpulan data akan menjadi realisasi yang tidak mungkin dari hipotesis nol sesuai dengan probabilitas ambang — tingkat signifikansi. Tes hipotesis digunakan dalam menentukan apakah hasil penelitian akan mengarah pada penolakan hipotesis nol untuk tingkat signifikansi yang ditentukan sebelumnya.

#### **B. Merancang Hipotesis**

Dalam merancang sebuah hipotesis penelitian penting terlebih dahulu peneliti untuk mengkaji teori dasar yang berkaitan dengan permasalahan yang akan diteliti. Teori yang relevan dapat menjadi acuan dan dasar dalam penyusunan sebuah hipotesis. Hipotesis yang baik dan ideal dalam sebuah penelitian tidak disusun berdasarkan opini namun kebenaran teori yang sudah teruji dan dapat dipercaya. Contoh hipotesis:

Dalam sebuah penelitian seorang mahasiswa pendidikan, ia akan meneliti berkaitan dengan implementasi penggunaan model pembelajaran Snowball Throwing dalam meningkatkan hasil belajar sejarah siswa. Kemudian dalam penyusunan hipotesisnya, peneliti harus terlebih dahulu mengkaji seberapa besar pengaruh model pembelajaran Snowball Throwing dalam meningkatkan hasil belajar sejarah siswa. Setelah menemukan pijakan teori yang kuat, peneliti dapat merumuskan hipotesis sebagai berikut:

Ho : Tidak terdapat perbedaan yang signifikan antara hasil belajar sejarah siswa yang diajar menggunakan model pembelajaran Snowball Throwing dengan pembelajaran konvensional

Hi : Terdapat perbedaan yang signifikan antara hasil belajar sejarah siswa yang diajar menggunakan model pembelajaran Snowball Throwing dengan pembelajaran konvensional

Artinya dalam penelitian ini peneliti akan menguji hipotesis untuk mengetahui apakah terdapat perbedaan yang signifikan antara siswa yang diajar menggunakan model pembelajaran Snowball Throwing (eksperimen) dengan yang belajar secara konvensional (control). Hasil analisis dilanjutkan dengan penggunaan uji t pada pembahasan berikutnya.

Dalam literatur statistik, pengujian hipotesis statistik memainkan peran mendasar, langkah-langkah yang perlu diperhatikan dalam merancang hipotesis adalah sebagai berikut:

1. Ada hipotesis penelitian awal yang kebenarannya tidak diketahui. Langkah pertama adalah menyatakan hipotesis nol dan alternatif yang relevan. Ini penting, karena salah menyatakan hipotesis akan ber lumpur sisa proses.
2. Langkah kedua adalah mempertimbangkan asumsi statistik yang dibuat tentang sampel dalam melakukan tes; misalnya, asumsi tentang independensi statistik atau tentang bentuk

distribusi pengamatan. Ini sama pentingnya dengan asumsi yang tidak valid akan berarti bahwa hasil tes tidak valid.

3. Tentukan tes mana yang tepat, dan nyatakan statistik uji yang relevan.
4. Turunkan distribusi statistik uji di bawah hipotesis nol dari asumsi. Sebagai contoh, statistik tes mungkin mengikuti distribusi t Student atau distribusi normal.
5. Pilih tingkat signifikansi ( $\alpha$ ), ambang batas probabilitas di bawah ini yang mana hipotesis nol akan ditolak. Nilai-nilai umum adalah 5% dan 1%.
6. Lakukan analisis data
7. Putuskan untuk menolak hipotesis nol atau menerimanya. Aturan keputusan adalah menolak hipotesis nol  $H_0$  jika nilai yang diamati berada di wilayah kritis, dan untuk menerima atau "gagal menolak" hipotesis sebaliknya.

### **C. Uji Satu Pihak dan Uji Dua Pihak**

Dalam merancang sebuah hipotesis penelitian, peneliti sangat bergantung kepada teori yang berkaitan dengan permasalahan dalam penelitian. Jika dalam penelitian ditemukan berbagai variasi teori sehingga memungkinkan untuk melakukan pengujian hipotesis lanjut dalam bentuk modifikasi, peneliti dapat merumuskan hipotesis dalam bentuk tingkatan perbandingan. Seperti peneliti dapat merancang hipotesis sebagai berikut:

$H_0$  : Tidak terdapat perbedaan yang signifikan antara penggunaan model pembelajaran A dengan model pembelajaran B

$H_1$  : Model pembelajaran A lebih baik dalam meningkatkan hasil belajar siswa dibandingkan dengan penggunaan model pembelajaran B

Jika hipotesis dirancang dengan salah satu penguatan pada variable penelitian, uji ini disebut uji satu pihak. Uji satu pihak dapat dibagi menjadi dua kategori yaitu uji pihak kanan dan uji pihak

kiri. Untuk uji pihak kanan, penguatan rancangan hipotesis yang diinginkan bersifat positif dengan kalimat yang ditandai dengan kata “lebih baik, lebih bagus, dan lain-lain”. Begitu sebaliknya, jika uji pihak kiri yang ingin ditonjolkan dalam penelitian, peneliti dapat menggunakan kata “tidak lebih baik, kurang baik, dan lain-lain”. Hal lain yang perlu diingat saat menggunakan uji psatu pihak adalah nantinya saat dilakukan perbandingan antara nilai yang dihitung dengan alfa, peneliti harus melakukan perbandingan dengan nilai setengah alfa (karena uji satu pihak). Sebagai contoh, hasil perhitungan di dapat 0.020, sedangkan alfa yang telah disepakati adalah 0.05. Sehingga jika uji pihak kanan/kiri dilakukan, nilai hitung yang didapat (0.020) harus dibandingkan dengan setengah alfa, yaitu 0.025.

Untuk uji dua pihak berlaku ketentuan sebagai berikut:

1. Uji dua pihak untuk melihat perbedaan hasil dari variable yang diteliti (tidak melihat baik/buruknya)
2. Analisis perbandingan dapat dilakukan langsung dengan membandingkan hasil perhitungan dengan nilai alfa yang telah disepakati

Contoh hipotesis uji dua pihak:

Ho : Tidak terdapat perbedaan yang signifikan antara hasil belajar sejarah siswa yang diajar menggunakan model pembelajaran Jigsaw dengan pembelajaran konvensional

Hi : Terdapat perbedaan yang signifikan antara hasil belajar sejarah siswa yang diajar menggunakan model pembelajaran Jigsaw dengan pembelajaran konvensional

#### **Tugas 4**

3. Buatlah definisi hipotesis menurut pendapatmu sendiri?
4. Kemukakanlah perbedaan antara hipotesis nol dan hipotesis alternative?
5. Apa saja hal yang perlu diperhatikan sebelum merancang sebuah hipotesis?
6. Berikanlah masing-masing contoh hipotesis untuk uji satu pihak dan dua pihak?
7. Apa saja hal mendasar yang menjadi perbedaan hipotesis untuk uji satu pihak dan dua pihak?

## CHAPTER 6

### PENARIKAN SAMPEL DALAM PENELITIAN

Di bagian metodologi penelitian, peneliti akan diminta untuk memberikan rincian tentang populasi dan sampel studi. Bagian-bagian ini merupakan bagian tersulit yang umum bagi banyak peneliti, karena peneliti sering gagal untuk membedakan antara populasi mereka dengan sampel mereka. Jadi, apa sebenarnya perbedaan antara keduanya?

Pertama, sampel yang akan diteliti adalah kelompok individu yang benar-benar berpartisipasi dalam studi/penelitian. Mereka adalah orang-orang yang akan diwawancarai (mis., Dalam penelitian kualitatif) atau yang benar-benar menyelesaikan survei (mis., Dalam penelitian kuantitatif).

Di sisi lain, populasi adalah kelompok orang yang lebih luas yang ingin peneliti generalisasikan dalam hasil penelitian. Sampel akan selalu menjadi bagian dari populasi. Populasi yang tepat akan bergantung pada ruang lingkup studi dan penelitian. Misalnya, katakanlah pertanyaan penelitian ingin menanyakan apakah ada hubungan antara kecerdasan emosi dan kepuasan kerja pada perawat. Dalam hal ini, populasi mungkin perawat di Amerika Serikat. Namun, jika ruang lingkup penelitian lebih sempit (misalnya, jika penelitian berkaitan dengan masalah lokal atau spesialisasi / industri tertentu), maka populasi akan lebih spesifik, seperti "perawat di negara bagian Florida" atau "perawat praktik berlisensi di Amerika Serikat."

Untuk menjawab pertanyaan penelitian tertentu, peneliti perlu menyelidiki area atau kelompok tertentu, yang akan ditarik kesimpulan dari penelitian. Ketika populasi penelitian sangat kecil, kita perlu mempelajari sebagian dari populasi tersebut, yang perlu data yang didapatkan bersifat umum dan dapat menjadi perwakilan dalam keseluruhan data. Ini dikenal

sebagai sampel, dan pemilihan komponen sampel yang akan memberikan pandangan yang mewakili keseluruhan dikenal sebagai teknik sampling.

Untuk membuat sampel, peneliti harus terlebih dahulu mengidentifikasi jumlah total orang dalam populasi penelitian. Informasi ini mungkin tersedia di berbagai sumber, ini dikenal sebagai kerangka sampling. Pengambilan sampel dapat dilakukan dengan probabilitas atau non-probabilitas. Ini adalah keputusan desain penelitian yang penting, dan salah satu yang akan tergantung pada faktor-faktor seperti apakah teori di balik penelitian adalah positivis atau idealis, apakah metode kualitatif atau kuantitatif digunakan dll. Perhatikan bahwa dua metode tidak saling eksklusif, dan mungkin digunakan untuk tujuan yang berbeda pada titik yang berbeda dalam penelitian, katakan purposive sampling untuk mengetahui sikap kunci, diikuti oleh pendekatan yang lebih umum dan acak.

### **Probability sampling**

Dalam sampling probabilitas, setiap anggota populasi penelitian tertentu memiliki kesempatan yang sama untuk dipilih. Ini melibatkan, secara harfiah, pemilihan responden secara acak dari kerangka sampling, setelah memutuskan ukuran sampel. Jenis sampling ini lebih mungkin jika orientasi teoritis dari penelitian adalah positivis, dan metodologi yang digunakan cenderung bersifat kuantitatif. Pengambilan sampel kemungkinan bisa:

1. acak - pemilihannya benar-benar acak, dan sejumlah tertentu dari total populasi dipilih sepenuhnya secara acak.
2. sistematis - setiap elemen  $n$  dari populasi dipilih. Hal ini dapat menimbulkan masalah jika interval pemilihan berarti bahwa elemen-elemen tersebut memiliki karakteristik:
3. stratifikasi acak - populasi dibagi menjadi segmen, misalnya, di Universitas, kita dapat membagi populasi menjadi akademik, administrator, dan akademik terkait (staf

profesional terkait). Secara acak masing-masing kelompok kemudian dipilih. Ini memiliki keuntungan yang memungkinkan kita untuk mengkategorikan populasi sesuai dengan fitur-fitur tertentu.

4. klaster - subkelompok tertentu dipilih secara acak. Subkelompok ini mungkin didasarkan pada wilayah geografis tertentu, katakanlah kita dapat memutuskan untuk mengambil sampel wilayah tertentu di negara tersebut.

### **Non probability sampling**

Di sini, populasi tidak memiliki peluang yang sama untuk dipilih; sebagai gantinya, pemilihan terjadi menurut beberapa faktor seperti:

1. conidental/accidental - hadir pada waktu tertentu. Ini adalah cara mudah untuk mendapatkan sampel, tetapi mungkin tidak benar-benar akurat, karena faktor yang dipilih didasarkan pada keinginan, bukan pada pemahaman yang benar tentang karakteristik sampel.
2. purposive - orang dapat memilih secara sengaja karena pandangan mereka relevan dengan masalah yang bersangkutan. Namun kelemahan tekniknya adalah subjektivitasnya - pandangan tentang kriteria pemilihan dapat berubah selama durasi penelitian. Penggunaan dapat terbuat dari:
  - 'Teknik informan kunci' - yaitu orang-orang dengan pengetahuan khusus;
  - menggunakan orang pada titik-titik yang dipilih dalam heirarki organisasi;
  - bola salju, dengan satu orang didekati dan kemudian menyarankan orang lain.
3. kuota - asumsi dibuat bahwa ada subkelompok dalam populasi, dan kuota responden dipilih untuk mencerminkan keragaman ini. Subkelompok ini harus representatif secara keseluruhan, tetapi yang harus diperhatikan adalah dalam menarik kesimpulan untuk



seluruh populasi. Misalnya, sampel kuota yang diambil di Negara Bagian New York tidak akan mewakili seluruh Amerika Serikat.

Metode non probability sampling lebih mungkin digunakan dalam penelitian kualitatif, dengan tingkat kolaborasi yang lebih besar dengan responden yang memiliki peluang untuk pengumpulan data yang lebih detail.

## **TUGAS 6**

1. Jelaskanlah perbedaan antara populasi dan sampel?
2. Secara umum, terbagi menjadi berapakah teknik sampling dalam penelitian?
3. Kemukakanlah perbedaan antara probably sampling dengan non probably sampling?
4. Jelaskanlah teknik sampling yang dipilih jika kita ingin meneliti tentang guru pada suatu wilayah?

## CHAPTER 7

### PENGANTAR STATISTIK PENDIDIKAN DAN SOSIAL

#### A. Statistik Parametrik dan Non Parametrik

Sebagian besar pembelajaran statistik cenderung berfokus pada statistik parametrik; namun pada kondisi tertentu dalam sebuah penelitian, nantinya peneliti akan menemukan bahwa statistic yang digunakan bukan hanya parametric, tetapi ada kemungkinan penggunaan statistic non parametrik. Berikut ini adalah beberapa perbedaan antara statistik parametrik dan nonparametrik.

Statistik parametrik adalah uji statistik berdasarkan asumsi yang mendasari tentang distribusi data. Dengan kata lain, statistik parametrik didasarkan pada parameter kurva normal. Karena statistik parametrik didasarkan pada kurva normal, data harus memenuhi asumsi tertentu, atau statistik parametrik tidak dapat dihitung dan dilanjutkan. Sebelum menjalankan statistik parametrik terlebih dahulu kita harus memastikan untuk menguji asumsi dan prasyarat seperti uji normalitas dan homogenitas. Statistic parametric ditandai dengan adanya uji normalitas dan homogenitas sebagai prasyarat.

Adapun keuntungan dari penggunaan prosedur-prosedur dari statistik parametrik adalah sebagai berikut.

1. Syarat-syarat parameter dari suatu populasi yang menjadi sampel biasanya tidak diuji dan dianggap memenuhi syarat, Pengukuran terhadap data dilakukan dengan kuat.
2. Observasi bebas satu sama lain dan ditarik dari populasi yang berdistribusi normal serta Memiliki varian yang homogeny.

Sedangkan kekurangan dari penggunaan prosedur-prosedur model statistik parametrik adalah :

1. Populasi harus memiliki varian yang sama.
2. Variabel-variabel yang diteliti harus dapat diukur setidaknya dalam skala interval.

Seperti yang disiratkan oleh namanya, statistik nonparametrik tidak didasarkan pada parameter kurva normal. Oleh karena itu, jika data melanggar/tidak memenuhi asumsi dari statistik parametric, maka nonparametrik dapat digunakan dalam melanjutkan penganalisisan data. Meskipun uji statistik nonparametrik memiliki lebih banyak fleksibilitas daripada uji statistik parametrik, tes nonparametrik tidak begitu kuat; oleh karena itu, sebagian besar ahli statistik merekomendasikan bahwa bila diperlukan, statistik parametrik lebih disukai dan sering digunakan.

### **Keunggulan Statistik Nonparametrik**

1. Asumsi dalam uji-uji statistik nonparametrik relatif lebih longgar. Jika pengujian data menunjukkan bahwa salah satu atau beberapa asumsi yang mendasari uji statistik parametrik. (misalnya mengenai sifat distribusi data) tidak terpenuhi, maka statistik nonparametrik lebih sesuai diterapkan dibandingkan statistic parametrik.
2. Perhitungan-perhitungannya dapat dilaksanakan dengan cepat dan mudah, sehingga hasil penelitian segera dapat disampaikan.
3. Untuk memahami konsep-konsep dan metode-metodenya tidak memerlukan dasar matematika serta statistika yang mendalam.
4. Uji-uji pada statistik nonparametrik dapat diterapkan jika kita menghadapi keterbatasan data yang tersedia, misalnya jika data telah diukur menggunakan skala pengukuran yang lemah (nominal atau ordinal).
5. Efisiensi statistik nonparametrik lebih tinggi dibandingkan dengan metode parametrik untuk jumlah sampel yang sedikit.

### **Keterbatasan Statistik Nonparametrik**

1. Beberapa keterbatasan statistik nonparametrik antara lain:
2. Jika asumsi uji statistik parametrik terpenuhi, penggunaan uji nonparametric meskipun lebih cepat dan sederhana, akan menyebabkan pemborosan informasi.
3. Jika jumlah sampel besar, tingkat efisiensi nonparametrik relatif lebih rendah dibandingkan dengan metode parametrik.

## **B. Jenis-jenis Data**

### **1. Nominal**

Data Nominal merupakan data yang paling lemah/rendah di antara data pengukuran yang ada. Data nominal hanya bisa membedakan benda atau peristiwa yang satu dengan yang lainnya berdasarkan nama (predikat). Data pengukuran nominal digunakan untuk mengklasifikasi obyek, individual atau kelompok dalam bentuk kategori. Pemberian angka atau simbol pada data nominal tidak memiliki maksud kuantitatif hanya menunjukkan ada atau tidak adanya atribut atau karakteristik pada objek yang diukur. Misalnya, jenis kelamin diberi kode 1 untuk laki-laki dan kode 2 untuk perempuan. Angka ini hanya berfungsi sebagai label kategori, tanpa memiliki nilai instrinsik dan tidak memiliki arti apa pun. Kita tidak bisa mengatakan perempuan dua kali dari laki-laki. Kita bisa saja mengkode laki-laki menjadi 2 dan perempuan dengan kode 1, atau bilangan apapun asal kodenya berbeda antara laki-laki dan perempuan. Misalnya lagi untuk agama, kita bisa mengkode 1=Islam, 2=Kristen, 3=Hindu, 4=Budha dstnya. Kita bisa menukar angka-angka tersebut, selama suatu karakteristik memiliki angka yang berbeda dengan karakteristik lainnya.

Karena tidak memiliki nilai instrinsik, maka angka-angka (kode-kode) yang kita berikan tersebut tidak memiliki sifat sebagaimana bilangan pada umumnya. Oleh karenanya, pada variabel dengan data nominal tidak dapat diterapkan operasi matematika standar (aritmatik) seperti pengurangan, penjumlahan, perkalian, dan lainnya. Peralatan statistik yang sesuai dengan data nominal adalah peralatan statistik yang berbasiskan (berdasarkan) jumlah dan proporsi seperti modus, distribusi frekuensi, Chi Square dan beberapa peralatan statistik non-parametrik lainnya.

## 2. Ordinal

Data Ordinal ini lebih tinggi daripada data nominal, dan sering juga disebut dengan data peringkat. Hal ini karena dalam data ordinal, lambang-lambang bilangan hasil pengukuran selain menunjukkan perbedaan juga menunjukkan urutan atau tingkatan obyek yang diukur menurut karakteristik tertentu. Misalnya tingkat kepuasan seseorang terhadap produk. Bisa kita beri angka dengan 5=sangat puas, 4=puas, 3=kurang puas, 2=tidak puas dan 1=sangat tidak puas. Atau misalnya dalam suatu lomba, pemenangnya diberi peringkat 1,2,3 dan seterusnya.

Dalam data ordinal, tidak seperti data nominal, ketika kita ingin mengganti angka-angkanya, harus dilakukan secara berurut dari besar ke kecil atau dari kecil ke besar. Jadi, tidak boleh kita buat 1=sangat puas, 2=tidak puas, 3=puas dan seterusnya. Selain itu, yang perlu diperhatikan dari karakteristik data ordinal adalah meskipun nilainya sudah memiliki batas yang jelas tetapi belum memiliki jarak (selisih). Kita tidak tahu berapa jarak kepuasan dari tidak puas ke kurang puas. Dengan kata lain juga, walaupun sangat puas kita beri angka 5 dan sangat tidak puas kita beri angka 1, kita tidak bisa mengatakan bahwa kepuasan yang sangat puas lima kali lebih tinggi dibandingkan yang sangat tidak puas. Sebagaimana halnya pada data nominal, pada data ordinal kita juga tidak dapat menerapkan operasi matematika standar (aritmatik) seperti pengurangan, penjumlahan, perkalian, dan lainnya. Peralatan statistik yang sesuai dengan data ordinal juga adalah peralatan statistik yang berbasiskan (berdasarkan) jumlah dan proporsi seperti modus, distribusi frekuensi, Chi Square dan beberapa peralatan statistik non-parametrik lainnya.

### 3. Interval

Data interval mempunyai karakteristik seperti yang dimiliki oleh data nominal dan ordinal dengan ditambah karakteristik lain, yaitu berupa adanya interval yang tetap. Dengan demikian, data interval sudah memiliki nilai intrinsik, sudah memiliki jarak, tetapi jarak tersebut belum merupakan kelipatan. Pengertian “jarak belum merupakan kelipatan” ini kadang-kadang diartikan bahwa data interval tidak memiliki nilai nol mutlak.

Misalnya pada pengukuran suhu. Kalau ada tiga daerah dengan suhu daerah A =  $10^{\circ}\text{C}$ , daerah B =  $15^{\circ}\text{C}$  dan daerah C =  $20^{\circ}\text{C}$ . Kita bisa mengatakan bahwa selisih suhu daerah B,  $5^{\circ}\text{C}$  lebih panas dibandingkan daerah A, dan selisih suhu daerah C dengan daerah B adalah  $5^{\circ}\text{C}$ . (Ini menunjukkan pengukuran interval sudah memiliki jarak yang tetap). Tetapi, kita tidak bisa mengatakan bahwa suhu daerah C dua kali lebih panas dibandingkan daerah A (artinya tidak bisa jadi kelipatan). Kenapa ? Karena dengan pengukuran yang lain, misalnya dengan Fahrenheit, di daerah A suhunya adalah  $50^{\circ}\text{F}$ , di daerah B =  $59^{\circ}\text{F}$  dan daerah C =  $68^{\circ}\text{F}$ . Artinya, dengan pengukuran Fahrenheit, daerah C tidak dua kali lebih panas dibandingkan daerah A, dan ini terjadi karena dalam derajat Fahrenheit titik nolnya pada 32, sedangkan dalam derajat Celcius titik nolnya pada 0.

### 4. Ratio

Data rasio adalah data data dengan kualitas paling tinggi. Pada data rasio, terdapat semua karakteristik data nominal, ordinal dan data interval ditambah dengan sifat adanya nilai nol yang bersifat mutlak. Nilai nol mutlak ini artinya adalah nilai dasar yang tidak bisa diubah meskipun menggunakan data yang lain. Oleh karenanya, pada data ratio, pengukuran sudah mempunyai nilai perbandingan/rasio. Pengukuran-pengukuran dalam data rasio yang sering digunakan

adalah pengukuran tinggi dan berat. Misalnya berat benda A adalah 30 kg, sedangkan benda B adalah 60 kg. Maka dapat dikatakan bahwa benda B dua kali lebih berat dibandingkan benda A.

### **C. Normalitas dan Homogenitas**

#### **Uji Normalitas**

Uji Normalitas adalah sebuah uji yang dilakukan dengan tujuan untuk menilai sebaran data pada sebuah kelompok data atau variabel, apakah sebaran data tersebut berdistribusi normal atau tidak. Uji statistik normalitas yang dapat digunakan diantaranya Lilliefors, Kolmogorov-Smirnov, Shapiro Wilk, Chi-Square, Jarque Bera.

##### 1. Uji Lilliefors

Ketentuan dalam penggunaan uji lilliefors sebagai berikut:

- a. Uji lilliefors dilakukan apabila data merupakan data tunggal atau data frekuensi tunggal, bukan data distribusi frekuensi kelompok.
- b. Metode Lilliefors menggunakan data dasar yang belum diolah dalam tabel distribusi frekuensi.
- c. Data ditransformasikan dalam nilai Z untuk dapat dihitung luasan kurva normal sebagai probabilitas kumulatif normal.
- d. Probabilitas tersebut dicari bedanya dengan probabilitas kumulatif empiris.
- e. Beda terbesar dibanding dengan tabel Lilliefors pada Tabel Nilai Quantil Statistik Lilliefors Distribusi Normal.

##### 2. Uji Kolmogorov-smirnov

Uji Kolmogorov Smirnov adalah pengujian normalitas yang banyak dipakai, terutama setelah adanya banyak program statistik yang beredar. Kelebihan dari uji ini adalah sederhana

dan tidak menimbulkan perbedaan persepsi di antara satu pengamat dengan pengamat yang lain, yang sering terjadi pada uji normalitas dengan menggunakan grafik. Konsep dasar dari uji normalitas Kolmogorov Smirnov adalah dengan membandingkan distribusi data (yang akan diuji normalitasnya) dengan distribusi normal baku. Distribusi normal baku adalah data yang telah ditransformasikan ke dalam bentuk Z-Score dan diasumsikan normal. Jika signifikansi di atas 0,05 maka berarti tidak terdapat perbedaan yang signifikan antara data yang akan diuji dengan data normal baku.

### 3. Uji Saphiro Wilk

Shapiro Wilk adalah salah satu uji normalitas yang dianjurkan oleh banyak pakar apabila jumlah sampel kecil yaitu kurang dari atau sama dengan 50 sampel. Uji ini sangat sensitif untuk mendeteksi adanya ketidak normalan sebaran data. Hanya saja dalam penghitungannya relatif lebih sulit dibandingkan dengan jenis uji lainnya. Metode Shapiro-Wilk menggunakan data dasar yang belum diolah dalam tabel distribusi frekuensi. Data diurut, kemudian dibagi dalam dua kelompok untuk dikonversi dalam Shapiro Wilk. Dapat juga dilanjutkan transformasi dalam nilai Z untuk dapat dihitung luasan kurva normal. **Persyaratan:** Data berskala interval atau ratio (kuantitatif), Data tunggal / belum dikelompokkan pada tabel distribusi frekuensi, Data dari sampel random.

### **Uji Homogenitas**

Uji homogenitas bertujuan untuk mencari tahu apakah dari beberapa kelompok data penelitian memiliki varians yang sama atau tidak. Dengan kata lain, homogenitas berarti bahwa himpunan data yang kita teliti memiliki karakteristik yang sama. Sebagai contoh, jika kita ingin meneliti sebuah permasalahan misalnya mengukur pemahaman siswa untuk suatu sub materi



dalam pelajaran tertentu di sekolah yang dimaksudkan homogen bisa berarti bahwa kelompok data yang kita jadikan sampel pada penelitian memiliki karakteristik yang sama, misalnya berasal dari tingkat kelas yang sama. Perhitungan uji homogenitas dapat dilakukan dengan berbagai cara dan metode, beberapa yang cukup populer dan sering digunakan adalah : Uji Varians (Uji F), Uji Bartlett, Uji Levene.

### 1. Uji F

Uji F dikenal dengan Uji serentak atau uji Model/Uji Anova, yaitu uji untuk melihat bagaimanakah pengaruh semua variabel bebasnya secara bersama-sama terhadap variabel terikatnya, atau untuk menguji apakah model regresi yang kita buat baik/signifikan atau tidak baik/non signifikan.

Uji F dapat dilakukan dengan membandingkan F hitung dengan F tabel (dapat menggunakan excel), jika  $F_{hitung} > F_{tabel}$ , ( $H_0$  di tolak  $H_a$  diterima) maka model signifikan atau bisa dilihat dalam kolom signifikansi pada Anova (Olahan dengan SPSS, Gunakan Uji Regresi dengan Metode Enter/Full Model). Model signifikan selama kolom signifikansi (%) < Alpha (kesiapan berbuat salah tipe 1, yang menentukan peneliti sendiri, ilmu sosial biasanya paling besar alpha 10%, atau 5% atau 1%). Dan sebaliknya jika  $F_{hitung} < F_{tabel}$ , maka model tidak signifikan, hal ini juga ditandai nilai kolom signifikansi (%) akan lebih besar dari alpha.

### 2. Uji Bartlett

Uji Bartlett digunakan untuk menguji apakah k sampel berasal dari populasi dengan varians yang sama. k sampel bisa berapa saja. karena biasanya uji bartlett digunakan untuk menguji sampel/kelompok yang lebih dari 2. Varians yang sama di seluruh sampel disebut

homoscedasticity atau homogenitas varians. Uji bartlett pertama kali diperkenalkan oleh M. S. Bartlett (1937). Uji bartlett diperlukan dalam beberapa uji statistik seperti analysis of variance (ANOVA) sebagai syarat jika ingin menggunakan Anova. berdasarkan info dari wikipedia uji barlett ini dinamai Maurice Stevenson Bartlett. Selain uji bartlett terdapat uji lavene yang fungsinya sama yaitu mengetahui homogenitas varians. Untuk Kali ini akan dicoba mencoba membahas uji bartlett.

### 3. Uji Levene

Beberapa metoda statistik memerlukan adanya asumsi kesamaan varians sebagai salah satu syarat dapat digunakannya metoda statistik tersebut sebagai metoda analisis seperti penggunaan uji-F pada uji ANOVA satu arah dan uji t pada uji kesamaan rata-rata. Uji levne (*levne 1960*) digunakan untuk menguji kesamaan varians dari beberapa populasi. Uji levne merupakan uji alternatif dari uji Bartlett. Jika ada bukti yang kuat bahwa data berdistribusi normal atau mendekati normal, maka uji Bartlett lebih baik digunakan.

## TUGAS 7

1. Jelaskanlah perbedaan antara statistic paramterik dengan statistic non parametric?
2. Berikanlah dua sampel data untuk : data rasio, ordinal, nominal dan interval?
3. Diketahui data ujian siswa sebagai berikut:

70, 80, 45, 80, 55, 80, 90, 50, 55, 70, 75, 80, 75, 55, 70

apakah data tersebut diatas terdistribusi normal? Lakukan uji normalitas dengan salah satu jenis uji yang telah dipaparkan sebelumnya.

## CHAPTER 8

### UJI T DALAM STATISTIK PARAMETRIK

#### D. Defenisi Uji Z, t, t'

Uji Z adalah salah satu uji statistika yang pengujian hipotesisnya didekati dengan distribusi normal. Menurut teori limit terpusat, data dengan ukuran sampel yang besar akan berdistribusi normal. Oleh karena itu, uji Z dapat digunakan untuk menguji data yang sampelnya berukuran besar. Jumlah sampel 30 atau lebih dianggap sampel berukuran besar. Selain itu, uji Z ini dipakai untuk menganalisis data yang varians populasinya diketahui. Namun, bila varians populasi tidak diketahui, maka varians dari sampel dapat digunakan sebagai penggantinya.

Uji t adalah salah satu tes yang digunakan untuk menguji kebenaran atau penolakan hipotesis nol yang menyatakan bahwa diantara dua buah Mean sampel yang diambil secara random dari populasi yang sama, tidak terdapat perbedaan yang signifikan. Uji t' adalah uji hipotesis yang membandingkan dua kelompok sampel yang berdistribusi normal dan variansi tidak homogen. Perbedaan penerapan antara uji t dan uji z hanya terletak pada jumlah sampel yang digunakan. Untuk uji z, sampel besar dari 30, sedangkan untuk uji t sampelnya kecil dari 30.

#### E. Jenis-jenis Uji t

*One sample t-test* merupakan teknik analisis untuk membandingkan satu variabel bebas. Teknik ini digunakan untuk menguji apakah nilai tertentu berbeda secara signifikan atau tidak dengan rata-rata sebuah sampel. Pada uji hipotesis ini, diambil satu sampel yang kemudian dianalisis apakah ada perbedaan rata-rata dari sampel tersebut.

*Paired-sample t-Test* merupakan prosedur yang digunakan untuk membandingkan rata-rata dua variabel dalam satu group. Artinya analisis ini berguna untuk melakukan pengujian terhadap satu sampel yang mendapatkan suatu treatment yang kemudian akan dibandingkan rata-rata dari sampel tersebut antara sebelum dan sesudah treatment.

*Independent sample t-Test* adalah uji yang digunakan untuk menentukan apakah dua sampel yang tidak berhubungan memiliki rata-rata yang berbeda. Jadi tujuan metode statistik ini adalah membandingkan rata-rata dua grup yang tidak berhubungan satu sama lain.

## F. Persyaratan Uji t

### 1. Uji Z

- a. Data berdistribusi normal
- b. Variance ( $\sigma$ ) diketahui
- c. Ukuran sampel (n) besar ( $n \geq 30$ )

### 2. Uji t

- a. Data berdistribusi normal
- b. Variance ( $\sigma$ ) tidak diketahui
- c. Untuk uji t' variansi data tidak homogen
- d. Ukuran sampel (n) kecil ( $n < 30$ )

Untuk lebih jelas, berikut disajikan perbedaan Uji Z, Uji t, dan Uji t'.

$H_0$	Nilai Uji Statistik	$H_1$
1. $\mu = \mu_0$ sampel besar $n \geq 30$	$z = \frac{\bar{x} - \mu_0}{\sigma / \sqrt{n}}$  $\sigma$ diketahui dan dapat diganti dengan s	$\mu < \mu_0$  $\mu > \mu_0$  $\mu \neq \mu_0$

2. $\mu = \mu_0$  sampel kecil $n < 30$	$t = \frac{\bar{x} - \mu_0}{s / \sqrt{n}}$  $\sigma$ tidak diketahui	$\mu < \mu_0$  $\mu > \mu_0$  $\mu \neq \mu_0$
3. $ \mu_1 - \mu_2  = d_0$  sampel-sampel besar $n_1 \geq 30$ $n_2 \geq 30$	$z = \frac{ \bar{x}_1 - \bar{x}_2  - d_0}{\sqrt{(\sigma_1^2 / n_1) + (\sigma_2^2 / n_2)}}$  Jika $\sigma_1^2$ dan $\sigma_2^2$ tidak diketahui $\rightarrow$ gunakan $s_1^2$ dan $s_2^2$	$ \mu_1 - \mu_2  < d_0$  $ \mu_1 - \mu_2  > d_0$  $ \mu_1 - \mu_2  \neq d_0$
4. $ \mu_1 - \mu_2  = d_0$  sampel - sampel kecil $n_1 < 30$ $n_2 < 30$	$t = \frac{ \bar{x}_1 - \bar{x}_2  - d_0}{s \sqrt{(1/n_1) + (1/n_2)}}$  $s = \sqrt{\frac{(n_1 - 1)S_1^2 + (n_2 - 1)S_2^2}{n_1 + n_2 - 2}}$  $\sigma_1 = \sigma_2$ tetapi tidak diketahui	$ \mu_1 - \mu_2  < d_0$  $ \mu_1 - \mu_2  > d_0$  $ \mu_1 - \mu_2  \neq d_0$
5. $ \mu_1 - \mu_2  = d_0$  sampel - sampel kecil $n_1 < 30$ $n_2 < 30$	$t' = \frac{ \bar{x}_1 - \bar{x}_2  - d_0}{\sqrt{(s_1^2 / n_1) + (s_2^2 / n_2)}}$  $v = \frac{\left(\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}\right)^2}{\frac{\left(\frac{s_1^2}{n_1}\right)^2}{n_1 - 1} + \frac{\left(\frac{s_2^2}{n_2}\right)^2}{n_2 - 1}}$  $\sigma_1 \neq \sigma_2$ dan tidak diketahui	$ \mu_1 - \mu_2  < d_0$  $ \mu_1 - \mu_2  > d_0$  $ \mu_1 - \mu_2  \neq d_0$

Keterangan:

1. Rata-rata dari Sampel Besar
2. Rata-rata dari Sampel Kecil
3. Beda 2 Rata-rata dari Sampel Besar
4. Beda 2 Rata-rata dari Sampel Kecil
5. Beda 2 Rata-rata dari Sampel Kecil (untuk data mempunyai variansi yang tidak homogen)

Catatan: beberapa nilai z yang penting,  $z_{5\%} = z_{0.05} = 1.645$ ,  $z_{2.5\%} = z_{0.025} = 1.96$ ,  $z_{1\%} = z_{0.01} = 2.33$ ,

$$z_{0.5\%} = z_{0.005} = 2.575$$

## **G. Perhitungan Menggunakan SPSS**

### **1. One Sample T-Test**

Adapun langkah-langkah yang ditempuh pada tahap ini dalam SPSS adalah sebagai berikut.

- a. Buka menu SPSS
- b. Pada Data View, masukkan data yang akan di analisis
- c. Klik menu Analyze > pilih sub menu Compare Means, kemudian klik One-Sample T test
- d. Pada One-Sample T test, pindahkan Variable Data yang telah diberi nama ke kotak Test Variable dengan mengklik tanda panah
- e. Pada kotak Test Value masukkan angka nilai yang akan diuji, **SETELAH** itu klik OK
- f. Pada jendela Output akan muncul hasil perhitungan, silahkan di interpretasikan hasil tersebut.

### **2. Independent-Samples T-Test**

Adapun langkah-langkah yang ditempuh pada tahap ini dalam SPSS adalah sebagai berikut.

- a. Buka menu SPSS
- b. Pada Data View, masukkan data yang akan di analisis
- c. Pada Variable View, ketiklah nama variabel di kolom label
- d. Pada kolom Value, untuk VAR0002 klik none
- e. Isi Value Labels dengan Value sebagai simbol 1, label nama simbol 1, begitu juga seterusnya. Kemudian klik Add dan OK
- f. Klik menu Analyze > pilih sub menu Compare Means, kemudian klik Independent-Samples T test
- g. Pada Independent-Samples T test, pindahkan Nama Label Data (1) yang telah diberi nama ke kotak Test Variables dan Nama Label Data (2) ke kotak Grouping Variable dengan mengklik tanda panah

- h. Pada Define Group, Klik Specified Values, masukkan (1) pada group 1 dan (2) pada group 2. Kemudian klik Continue, SETELAH itu klik OK

Pada jendela Output akan muncul hasil perhitungan, silahkan di interpretasikan hasil tersebut.

### 3. Paired-Samples T-Test

Adapun langkah-langkah yang ditempuh pada tahap ini dalam SPSS adalah sebagai berikut.

- a. Buka menu SPSS
- b. Pada Data View, masukkan data yang akan di analisis
- c. Pada Variable View, ketiklah nama variabel di kolom name
- d. Klik menu Analyze > pilih sub menu Compare Means, kemudian klik Paired-Samples T test
- e. Pada Paired-Samples T test, pindahkan Nama variabel(2) yang telah diberi nama pada kolom name tadi ke kotak Variable 1 dan Nama Variabel (2) ke kotak Variable 2 dengan mengklik tanda panah, SETELAH itu klik OK
- f. Pada jendela Output akan muncul hasil perhitungan, silahkan di interpretasikan hasil tersebut.

## TUGAS 8

1. Ujilah data berikut menggunakan uji paired sample t- test, dan nyatakan kesimpulanmu

No.	Kode Siswa	Sebelum (x)	Setelah (y)
1.	S-01	60	80
2.	S-02	70	70
3.	S-03	50	80
4.	S-04	80	90
5.	S-05	70	90
6.	S-06	60	80
7.	S-07	60	80

8.	S-08	60	90
9.	S-09	80	100
10.	S-10	50	70

2. Ujilah data berikut menggunakan uji independent sample t-test, dan nyatakan kesimpulanmu

No.	Kode Siswa	Kontrol (x)	eksperimen (y)
1.	S-01	40	61
2.	S-02	90	48
3.	S-03	10	90
4.	S-04	30	54
5.	S-05	50	75
6.	S-06	60	89
7.	S-07	40	80
8.	S-08	50	80
9.	S-09	90	70
10.	S-10	70	50
11.	S-11	40	80
12.	S-12	76	50
13.	S-13	67	56
14.	S-14	89	93
15.	S-15	60	45
16.	S-16	40	55
17.	S-17	20	48
18.	S-18	70	67
19.	S-19	78	92
20.	S-20	82	57



## CHAPTER 9

### ONE WAY ANOVA/ANOVA 1 ARAH

ANOVA digunakan untuk menguji perbedaan antara sejumlah rata-rata populasi dengan cara membandingkan variansinya. Pembilang pada rumus variansi tidak lain adalah jumlah kuadrat skor simpangan dari rata-ratanya, yang secara sederhana dapat ditulis sebagai  $\sum(X_i - \mu)^2$ . Istilah jumlah kuadrat skor simpangan sering disebut jumlah kuadrat (sum of squares). Jika jumlah kuadrat tersebut dibagi dengan n atau n-1 maka akan diperoleh rata-rata kuadrat yang tidak lain dari variansi suatu distribusi. Rumus untuk menentukan varians sampel yaitu,

$$S^2 = \frac{\sum_{i=1}^n (Y_i - \bar{Y})^2}{n - 1}$$

Seandainya kita mempunyai suatu populasi yang memiliki variansi  $\sigma^2$  dan rata-rata  $\mu$ . Dari populasi tersebut misalkan diambil tiga buah sampel secara independent, masing-masing dengan  $n_1$ ,  $n_2$ , dan  $n_3$ . Dari setiap sampel tersebut dapat ditentukan rata-rata dan variansinya, sehingga akan diperoleh tiga buah rata-rata dan variansi sampel yang masing-masing merupakan statistik (penaksir) yang tidak bias bagi parameternya.

ANOVA digunakan untuk menguji hipotesis nol tentang perbedaan dua buah rata-rata atau lebih. Secara formal, hipotesis tersebut dapat ditulis sebagai berikut.

$$H_0: \mu_1 = \mu_2 = \mu_3 = \dots = \mu_k$$

$H_1$ : Paling tidak salah satu tanda sama dengan (=) tidak berlaku

Hipotesis nol di atas mengatakan bahwa rata-rata populasi pertama sama dengan rata-rata populasi ke dua dan seterusnya yang berarti bahwa seluruh sampel diambil dari populasi

yang sama. Jika demikian maka, rata-ratanya akan mirip satu sama lain. Dalam menguji hipotesis nol tersebut,

ANOVA meakukan perbandingan antara variansi antar kelompok ( $MS_B$ ) dengan variansi dalam kelompok ( $MS_W$ ). Jika ternyata kedua variansi itu sama ( $F=1$ ) maka berarti seluruh sampel yang dianalisis berasal dari populasi yang sama, dan kita tidak memiliki dasar untuk menolak hipotesis nol. Namun, jika ada salah satu nilai rata-rata yang jauh berbeda dengan nilai rata-rata lainnya maka berarti sampel tersebut berasal dari populasi yang berbeda.

Seluruh subjek yang berada dalam satu kelompok memiliki karakteristik yang sama pada peubah bebas yang tengah dikaji. Dalam bahasa eksperimen, mereka seluruhnya menerima perlakuan yang sama, sehingga keragaman mereka pada peubah terikat dipandang sebagai keragaman galat dan tidak berkaitan dengan perbedaan jenis perlakuan atau peubah bebas.

Perbedaan rata-rata antar kelompok terdiri atas dua unsur yaitu keragaman galat dan keragaman yang berkaitan perbedaan pada peubah bebas. Oleh karena keragaman di dalam kelompok ( $MS_W$ ) merupakan penaksir yang tidak bias atas variansi populasi dan keragaman antara kelompok ( $MS_B$ ) terdiri atas  $MS_W$  dan keragaman yang berkaitan dengan perlakuan, maka hubungan antara keduanya dapat dituliskan sebagai berikut:

$$MS_W = \sigma^2$$

$$MS_B = \sigma^2 + \text{dampak perlakuan}$$

Dengan demikian, F dapat juga dituliskan:

$$F = MS_B/MS_W$$

$$F = (\sigma^2 + \text{dampak perlakuan})/\sigma^2$$

Jika dampak perlakuan sama dengan nol, maka

$$F = \frac{\sigma^2}{\sigma^2} = 1$$

Persoalan kita sekarang adalah bagaimana membedakan pengaruh yang sistematis dari pengaruh yang tidak sistematis (acak). ANAVA dan statistika inferensial pada umumnya mendekati persoalan ini dengan menggunakan teori peluang. Statistika inferensial bertugas untuk menjawab suatu pertanyaan yang dapat dirumuskan sebagai berikut: "jika hipotesis nol ternyata benar berapakah peluang memperoleh harga statistik tertentu?"

Misalkan dalam ANAVA, kita memperoleh  $F=3,96$ . Pertanyaan yang harus dijawab adalah "berapa besar peluang memperoleh  $F=3,96$  jika ternyata hipotesis nol itu benar?" Paket analisis statistik pada komputer umumnya memberikan jawaban terhadap pertanyaan tersebut secara langsung dalam bentuk  $p=0,25, 0,01, 0,001$  dan sebagainya. Namun jika dilakukan secara manual maka harga  $F_{hitung}$  harus dibandingkan dengan nilai kritis yang sudah disediakan dalam bentuk  $F_{tabel}$  pada derajat kebebasan dan tingkat keyakinan. Nilai  $p$  yang lebih kecil dari nilai yang ditentukan menunjukkan penolakan terhadap  $H_0$ . Kesimpulan yang sama diperoleh jika ternyata  $F_{hitung} > F_{tabel}$ . Menolak hipotesis nol berarti menyimpulkan bahwa perbedaan antara  $MS_B$  dengan  $MS_W$  berkaitan dengan pengaruh yang sistematis dari faktor atau peubah bebas yang diteliti.

### **A. Anava Satu Arah**

ANAVA satu arah/satu jalur yaitu analisis yang melibatkan hanya satu peubah bebas. Secara rinci, ANAVA satu jalur digunakan dalam suatu penelitian yang memiliki ciri-ciri berikut: Melibatkan hanya satu peubah bebas dengan dua kategori atau lebih yang dipilih dan ditentukan oleh peneliti secara tidak acak. Kategori yang dipilih disebut tidak acak karena peneliti tidak bermaksud menggeneralisasikan hasilnya ke kategori lain di luar yang diteliti pada peubah itu.

Sebagai contoh, peubah jenis kelamin hanya terdiri atas dua ketgori (pria-wanita), atau peneliti hendak membandingkan keberhasilan antara Metode A, B, dan C dalam meningkatkan semangat belajar tanpa bermaksud menggeneralisasikan ke metode lain di luar ketiga metode tersebut.

1. Perbedaan antara kategori atau tingkatan pada peubah bebas dapat bersifat kualitatif atau kuantitatif.
2. Setiap subjek merupakan anggota dari hanya satu kelompok pada peubah bebas, dan dipilih secara acak dari populasi tertentu.

Tujuan dari uji anova satu jalur adalah untuk membandingkan lebih dari dua rata-rata. Sedangkan gunanya untuk menguji kemampuan generalisasi. Maksudnya dari signifikansi hasil penelitian. Jika terbukti berbeda berarti kedua sampel tersebut dapat digeneralisasikan (data sampel dianggap dapat mewakili populasi). Anova satu jalur dapat melihat perbandingan lebih dari dua kelompok data.

Anova pengembangan atau penjabaran lebih lanjut dari uji-t ( $t_{hitung}$ ). Uji-t atau uji-z hanya dapat melihat perbandingan dua kelompok data saja. Sedangkan anova satu jalur lebih dari dua kelompok data. Contoh: Perbedaan prestasi belajar statistika antara mahasiswa tugas belajar ( $X_1$ ), izin belajar ( $X_2$ ) dan umum ( $X_3$ ).

Anova lebih dikenal dengan uji-F (*Fisher Test*), sedangkan arti variasi atau varian itu asalnya dari pengertian konsep “*Mean Square*” atau kuadrat rerata (KR).

Rumusnya :

$$KR = \frac{JK}{ab}$$

Dimana:  $JK$  = jumlah kuadrat (*some of square*)

$db$  = derajat bebas (*degree of freedom*)

Menghitung nilai Anova atau  $F$  ( $F_{hitung}$ ) dengan rumus :

$$F_{hitung} = \frac{V_A}{V_D} = \frac{KR_A}{KR_D} = \frac{JK_A: db_A}{JK_D: db_D} = \frac{\text{varian antar group}}{\text{varian antar group}}$$

Varian dalam group dapat juga disebut Varian Kesalahan (Varian Galat). Dapat dirumuskan :

$$JK_A = \sum \frac{(\sum X_{Ai})^2}{n_{Ai}} - \frac{(\sum X_\tau)^2}{N} \text{ untuk } db_A = A - 1$$

$$JK_D = (\sum X_\tau)^2 - \sum \frac{(\sum X_{Ai})^2}{n_{Ai}} \text{ untuk } db_D = N - A$$

## B. Langkah-langkah dalam Perhitungan Anava Satu Arah

1. Sebelum anova dihitung, asumsikan bahwa data dipilih secara random, berdistribusi normal, dan variannya homogen.
2. Buatlah hipotesis ( $H_a$  dan  $H_0$ ) dalam bentuk kalimat.
3. Buatlah hipotesis ( $H_a$  dan  $H_0$ ) dalam bentuk statistik.
4. Buatlah daftar statistik induk.
5. Hitunglah jumlah kuadrat antar group ( $JK_A$ ) dengan rumus :

$$JK_A = \sum \frac{(\sum X_{Ai})^2}{n_{Ai}} - \frac{(\sum X_\tau)^2}{N} = \left( \frac{(\sum X_{A1})^2}{n_{A1}} + \frac{(\sum X_{A2})^2}{n_{A2}} + \frac{(\sum X_{A3})^2}{n_{A3}} \right) - \frac{(\sum X_\tau)^2}{N}$$

6. Hitunglah derajat bebas antar group dengan rumus :  $db_A = A - 1$
7. Hitunglah kudrat rerata antar group ( $KR_A$ ) dengan rumus :  $KR_A = \frac{JK_A}{db_A}$
8. Hitunglah jumlah kuadrat dalam antar group ( $JK_D$ ) dengan rumus :

$$JK_D = (\sum X_\tau)^2 - \sum \frac{(\sum X_{Ai})^2}{n_{Ai}}$$

$$= \sum X^2_{A1} + \sum X^2_{A2} + \sum X^2_{A3} - \left( \frac{(\sum X_{A1})^2}{n_{A1}} + \frac{(\sum X_{A2})^2}{n_{A2}} + \frac{(\sum X_{A3})^2}{n_{A3}} \right)$$

9. Hitunglah derajat bebas dalam group dengan rumus :  $db_D = N - A$
10. Hitunglah kuadrat rerata dalam antar group ( $KR_D$ ) dengan rumus :  $KR_D = \frac{JK_D}{db_D}$
11. Carilah  $F_{hitung}$  dengan rumus :  $F_{hitung} = \frac{KR_A}{KR_D}$
12. Tentukan taraf signifikansinya, misalnya  $\alpha = 0,05$  atau  $\alpha = 0,01$
13. Cari  $F_{tabel}$  dengan rumus :  $F_{tabel} = F_{(1-\alpha)(db_A, db_D)}$
14. Buat Tabel Ringkasan Anova

TABEL RINGKASSAN ANOVA SATU ARAH

Sumber Varian (SV)	Jumlah Kuadrat (JK)	Derajat bebas (db)	Kuadrat Rerata (KR)	$F_{hitung}$	Taraf Signifikan ( $\rho$ )
Antar group (A)	$\sum \frac{(\sum X_{Ai})^2}{n_{Ai}} - \frac{(\sum X_\tau)^2}{N}$	$A - 1$	$\frac{JK_A}{db_A}$	$\frac{KR_A}{KR_D}$	$\alpha$
Dalam group (D)	$(\sum X_\tau)^2 - \sum \frac{(\sum X_{Ai})^2}{n_{Ai}}$	$N - A$	$\frac{JK_D}{db_D}$	-	-
Total	$(\sum X_\tau)^2 - \frac{(\sum X_\tau)^2}{N}$	$N - 1$	-	-	-

15. Tentukan kriteria pengujian : jika  $F_{hitung} \geq F_{tabel}$  , maka tolak  $H_0$  berarti signifikan dan konsultasikan antara  $F_{hitung}$  dengan  $F_{tabel}$  kemudian bandingkan
16. Buat kesimpulan.

## TUGAS 9

Perhatikanlah data Berikut

Seorang ingin mengetahui perbedaan prestasi belajar untuk mata kuliah dasar-dasar statistika antara mahasiswa tugas belajar, izin belajar dan umum.

Data diambil dari nilai UTS sebagai berikut :

Tugas belajar ( $A_1$ ) = 6 8 5 7 7 6 6 8 7 6 7 = 11 orang

Izin belajar ( $A_2$ ) = 5 6 6 7 5 5 5 6 5 6 8 7 = 12 orang

Umum ( $A_3$ ) = 6 9 8 7 8 9 6 6 9 8 6 8 = 12 orang

Buktikan apakah ada perbedaan atau tidak?

## CHAPTER 10

### ANALISIS REGRESI DAN KORELASI SEDERHANA

#### A. Analisa Regresi

Analisis regresi diartikan sebagai suatu analisis tentang hubungan suatu variable kepada variable lain, yaitu variabel bebas dalam rangka membuat estimasi atau prediksi dari nilai variable terikat dengan diketahuinya nilai variable bebas (Riadi, 2016: 145). Sedangkan (Kadir, 2015: 175) mengatakan bahwa analisis regresi merupakan teknik analisis yang khas untuk jenis penelitian asosiatif. Menurut Kadir analisis regresi bertujuan mempelajari “pengaruh” variable bebas (*predictor*) terhadap variable tak bebas suatu mengetahui bentuk hubungan dua variable atau lebih, digunakan analisis regresi. Bila ingin melihat keeratan hubungan, gunakan analisis korelasi. Dengan demikian, tampak jelas bahwa apabila peneliti melakukan analisis korelasi dan regresi, maka sebenarnya peneliti baru sampai pada tahap analisis hubungan dalam menjawab hipotesis’ seberapa besar keeratan hubungan variable  $y$  dapat diprediksikan oleh variable independen  $x$ . Variable analisis regresi dapat digunakan untuk mempelajari pengaruh antara *predictor* dan *criterion* Apabila memenuhi empat syarat sebagai berikut :

1. Terdapat logika (konseptual) yang menghubungkan antara variable bebas (*predictor*) dan variable tak bebas (*criterion*). Artinya, hubungan *predictor* dan *criterion* mempunyai dasar rasional yang kuat atau didukung oleh teori yang kuat
2. Pada umumnya *predictor* mendahului *criterion*. Artinya, dalam urutan waktu *predictor* terjadi lebih dahulu kemudian *criterion*. Sebagai contoh pemberian sertifikasi guru kejadiannya mendahului pengukuran kinerja guru, sehingga dapat dipelajari kinerja guru sebagai pengaruh dari pemberian sertifikasi guru



3. Terdapat arah pengaruh (*direct effect*), yaitu dari *predictor* ke *criterion* atau dalam representasi simbol ditulis sebagai anak panah berkepala satu. Misalnya  $\text{predictor} = X$  dan  $\text{criterion} = Y$  maka arah pengaruh ditulis sebagai  $X \rightarrow Y$  atau pengaruh X terhadap Y tidak sebaliknya.
4. Terdapat kontrol secara statistik sehingga pengaruh prediktor lain dalam model, terhadap *criterion* diluar *predictor* yang dipelajari dapat dikontrol pengaruhnya. Misalnya dalam analisis regresi ganda Y atas  $X_1$ , dan  $X_2$ .

Analisis regresi sederhana mempelajari apakah antara dua variabel atau lebih mempunyai pengaruh/hubungan atau lebih, mengukur kekuatan pengaruhnya, dan membuat ramalan yang didasarkan kepada kuat lemahnya pengaruh/ hubungan tersebut (Kadir ,2015 : 176). Teknik analisis ini akan bermakna apabila pengaruh antar variable-variabel didasarkan pada kerangka teori yang terkuat. Jika skala pengukuran data dua variable yang akan dianalisis merupakan skala interval atau rasio, maka untuk menjelaskan pengaruh antara kedua variable tersebut dapat dilakukan dengan menggunakan regresi sederhana. Misalnya kedua variabel tersebut adalah X dan Y, maka pengaruh X terhadap Y dianalisis melalui regresi sederhana Y atas X. Variabel X disebut variable bebas (*predictor*) dan variable Y disebut variabel tak bebas /variable terikat (*criterion*). Assosiasi antara variable X dan Y dinyatakan dalam suatu persamaan atau model matematika sebagai berikut:

Model Regresi :  $Y = \alpha + \beta X + e$  (populasi)

Fungsi tafsiran :  $\hat{Y} = a + bX$  (sampel)

Dimana

Y = Variabel Response atau Variabel Akibat (Dependent)

X = Variabel Predictor atau Variabel Faktor Penyebab (Independent)

a = konstanta,

b = koefisien regresi (kemiringan), yang nilainya dapat diperoleh dari data sampel.

Untuk memperoleh nilai a dan b dibutuhkan pasangan dari (X,Y) sebanyak n. Misalnya data variable X dan Y disajikan data berikut :

NO	X	Y
1	$X_1$	$Y_1$
2	$X_2$	$Y_2$
3	$X_3$	$Y_3$
4	$X_4$	$Y_4$
.	.	.
.	.	.
.	.	.
n	$X_n$	$Y_n$

Pada data regresi dari variable X dan Y mensyaratkan data sampel yang dipilih harus random. Berdistribusi normal, dan homogen. Dari perhitungan melalui pasangan dari (X,Y) minimal ditentukan :

1. Persamaan atau model regresi Y atas X
2. Linearitas regresi dan signifikansi regresi Y atas X
3. Koefisien korelasi dan koefisien determinasi.

## **B. Analisis Korelasi**

Secara sederhana, korelasi dapat diartikan sebagai hubungan. Namun ketika dikembangkan lebih jauh, korelasi tidak hanya dapat dipahami sebatas pengertian tersebut. Korelasi merupakan salah satu teknik analisis dalam statistik yang digunakan untuk mencari hubungan antara dua variabel yang bersifat kuantitatif. Hubungan dua variabel tersebut dapat terjadi karena adanya hubungan sebab akibat atau dapat pula terjadi karena kebetulan saja. Dua variabel dikatakan berkorelasi apabila perubahan pada variabel yang satu akan diikuti perubahan pada variabel yang lain secara teratur dengan arah yang sama (korelasi positif) atau berlawanan (korelasi negatif).

Korelasi Sederhana merupakan suatu teknik statistik yang dipergunakan untuk mengukur kekuatan hubungan antara 2 variabel dan juga untuk dapat mengetahui bentuk hubungan keduanya dengan hasil yang bersifat kuantitatif. Kekuatan hubungan antara 2 variabel yang dimaksud adalah apakah hubungan tersebut erat, lemah, ataupun tidak erat. Sedangkan bentuk hubungannya adalah apakah bentuk korelasinya linear positif ataupun linear negatif.

Korelasi Pearson menghitung korelasi dengan menggunakan variasi data. Keragaman data tersebut dapat menunjukkan korelasinya. Korelasi ini menghitung data apa adanya, tidak membuat ranking atas data yang digunakan seperti pada korelasi Rank Spearman. Ketika kita memiliki data numerik seperti nilai tukar rupiah, data rasio keuangan, tingkat pertumbuhan ekonomi, data berat badan dan contoh data numerik lainnya, maka Korelasi Pearson Product Moment cocok digunakan.

Sebaliknya, Koefisien Korelasi Rank Spearman digunakan untuk data diskrit dan kontinu namun untuk statistik nonparametrik. Koefisien korelasi Rank Spearman lebih cocok untuk digunakan pada statistik nonparametrik. Statistik nonparametrik adalah statistik yang digunakan

ketika data tidak memiliki informasi parameter, data tidak berdistribusi normal atau data diukur dalam bentuk ranking. Berbeda dengan Korelasi Pearson, korelasi ini tidak memerlukan asumsi normalitas, maka korelasi Rank Spearman cocok juga digunakan untuk data dengan sampel kecil.

Korelasi Rank Spearman menghitung korelasi dengan menghitung ranking data terlebih dahulu. Artinya korelasi dihitung berdasarkan orde data. Ketika peneliti berhadapan dengan data kategorik seperti kategori pekerjaan, tingkat pendidikan, kelompok usia, dan contoh data kategorik lainnya, maka Korelasi Rank Spearman cocok digunakan. Korelasi Rank Spearman pun cocok digunakan pada kondisi dimana peneliti dihadapkan pada data numerik (kurs rupiah, rasio keuangan, pertumbuhan ekonomi), namun peneliti tidak memiliki cukup banyak data (data kurang dari 30).

### **1. Rumus Pearson Product Moment**

Koefisien Korelasi Sederhana disebut juga dengan Koefisien Korelasi Pearson karena rumus perhitungan Koefisien korelasi sederhana ini dikemukakan oleh Karl Pearson yaitu seorang ahli Matematika yang berasal dari Inggris.

Rumus yang dipergunakan untuk menghitung Koefisien Korelasi Sederhana adalah sebagai berikut : (Rumus ini disebut juga dengan Pearson Product Moment)

$$r = \frac{n\sum xy - (\sum x)(\sum y)}{\sqrt{\{n\sum x^2 - (\sum x)^2\} \{n\sum y^2 - (\sum y)^2\}}}$$

**Dimana :**

$n$  = Banyaknya Pasangan data X dan Y

$\Sigma x$  = Total Jumlah dari Variabel X

$\Sigma y$  = Total Jumlah dari Variabel Y

$\Sigma x^2$  = Kuadrat dari Total Jumlah Variabel X

$\Sigma y^2$  = Kuadrat dari Total Jumlah Variabel Y

$\Sigma xy$  = Hasil Perkalian dari Total Jumlah Variabel X dan Variabel Y

**2. Pola / Bentuk Hubungan antara 2 Variabel :**

**a. Korelasi Linear Positif (+1)**

Perubahan salah satu Nilai Variabel diikuti perubahan Nilai Variabel yang lainnya secara teratur dengan arah yang sama. Jika Nilai Variabel X mengalami kenaikan, maka Variabel Y akan ikut naik. Jika Nilai Variabel X mengalami penurunan, maka Variabel Y akan ikut turun. Apabila Nilai Koefisien Korelasi mendekati +1 (positif Satu) berarti pasangan data Variabel X dan Variabel Y memiliki Korelasi Linear Positif yang kuat/erat.

**b. Korelasi Linear Negatif (-1)**

Perubahan salah satu Nilai Variabel diikuti perubahan Nilai Variabel yang lainnya secara teratur dengan arah yang berlawanan. Jika Nilai Variabel X mengalami kenaikan, maka Variabel Y akan turun. Jika Nilai Variabel X mengalami penurunan, maka Nilai Variabel Y akan naik. Apabila Nilai Koefisien Korelasi mendekati -1 (Negatif Satu) maka hal ini menunjukkan pasangan data Variabel X dan Variabel Y memiliki Korelasi Linear Negatif yang kuat/erat.

### c. Tidak Berkorelasi (0)

Kenaikan Nilai Variabel yang satunya kadang-kadang diikuti dengan penurunan Variabel lainnya atau kadang-kadang diikuti dengan kenaikan Variable yang lainnya. Arah hubungannya tidak teratur, kadang-kadang searah, kadang-kadang berlawanan. Apabila Nilai Koefisien Korelasi mendekati 0 (Nol) berarti pasangan data Variabel X dan Variabel Y memiliki korelasi yang sangat lemah atau berkemungkinan tidak berkorelasi.

### TUGAS 10

Soal :

Judul Penelitian: *Pengaruh Sertifikasi (X) terhadap Kinerja Guru (Y) pada SMA Taman*

Untuk keperluan tersebut telah diambil data 15 orang guru sebagai berikut :

Var	Data Penelitian														
X	40	55	32	55	50	52	61	44	30	22	40	64	58	48	44
Y	4	16	12	24	15	24	22	17	4	14	24	26	20	9	14

Nyatakanlah kesimpulan penelitian dengan menggunakan analisis regresi sederhana

## CHAPTER 11

### ANALISIS FAKTOR

Analisis faktor merupakan studi tentang keterhubungan antar sekumpulan variabel dalam usaha menemukan sekumpulan variabel baru. Variabel baru merupakan persekutuan dari beberapa variabel asli, sehingga banyak variabel dalam kumpulan variabel baru yang tercipta lebih sedikit dibandingkan dengan banyak variabel asli. Jadi istilah analisis faktor merupakan teknik analisis yang digunakan untuk membedakan beberapa jenis varian, sedangkan istilah faktor dimaksudkan sebagai varian persekutuan.

Analisis faktor digunakan dalam dua konteks. Pertama, analisis faktor untuk kepentingan eksplorasi, yaitu penggunaan analisis faktor hanya untuk menemukan variabel persekutuan. Dalam konteks ini, peneliti tidak memiliki cukup teori yang mendukung hipotesisnya. Pada konteks kedua, analisis faktor digunakan untuk kepentingan konfirmasi, yaitu untuk mendukung atau menolak hipotesis yang disusun peneliti tentang variabel persekutuan. Tentunya dalam konteks ini, peneliti memiliki teori yang kuat untuk mendukung hipotesisnya tentang variabel persekutuan.

Untuk kepentingan analisis butir instrumen, analisis faktor digunakan untuk menguji, apakah butir-butir instrumen yang disusun sudah mengukur indikator yang ditetapkan. Variabel analisis faktor untuk kepentingan analisis butir adalah butir itu sendiri. Sementara itu, faktor yang terbentuk merupakan indikator yang diukur dengan butir-butir yang telah disusun. Dalam hal ini, bisa terjadi butir-butir instrumen dari beberapa indikator yang berbeda mengukur hal yang sama, sehingga indikator-indikator itu perlu digabungkan. Hal lain yang bisa terjadi adalah satu butir yang disusun untuk mengukur satu indikator ternyata lebih cocok untuk mengukur indikator yang lain.

Asumsi statistik seperti normalitas, linieritas, dan homokedastisitas dipersyaratkan sebagai asumsi tambahan dalam analisis faktor. Uji asumsi yang lebih penting adalah uji matriks korelasi keseluruhan dengan *Bartlett test of sphericity* dan uji interkorelasi antar variabel yang disebut *measure of sampling adequacy* (MSA). Analisis faktor terdiri dari beberapa langkah yaitu : memilih variabel yang layak, menentukan faktor, dan melakukan rotasi.

Variabel-variabel yang akan dianalisis tidak semuanya layak untuk dimasukkan ke dalam salah satu faktor yang akan dibentuk. Oleh karena itu, pertama kali harus dipilih variabel mana saja yang layak untuk dianalisis lebih lanjut, yaitu dimasukkan ke dalam salah satu faktor. Tentunya variabel yang layak dipilih adalah variabel yang memiliki korelasi cukup tinggi dengan variabel-variabel lainnya, sehingga memiliki kecenderungan untuk membentuk faktor. Sebaliknya, variabel-variabel yang memiliki korelasi rendah harus dieliminasi.

Hipotesis yang akan diuji dalam memilih variabel yang layak adalah sebagai berikut.

H<sub>0</sub> : variabel belum layak untuk dianalisis lebih lanjut.

H<sub>1</sub> : variabel sudah layak untuk dianalisis lebih lanjut.

Kriteria penerimaan atau penolakan hipotesis adalah dengan memperhatikan signifikansi dari koefisien Kaiser-Meyer-Olkin (KMO) dalam tabel *KMO and Bartlett's Test*. Apabila koefisien KMO > 0,5, maka analisis sebetulnya sudah bisa dilanjutkan. Akan tetapi uji signifikansi masih perlu dilakukan dengan membandingkan taraf signifikansi yang tetap ( $\alpha$ ).

- Apabila bilangan signifikansi  $KMO \geq \alpha$ , maka H<sub>0</sub> diterima dan H<sub>1</sub> ditolak, yang artinya variabel-variabel belum layak untuk dianalisis lebih lanjut.
- Apabila bilangan signifikansi  $KMO < \alpha$ , maka H<sub>0</sub> ditolak dan H<sub>1</sub> diterima, yang artinya variabel-variabel sudah layak untuk dianalisis lebih lanjut.



Apabila variabel-variabel dinyatakan belum layak untuk dianalisis lebih lanjut, maka dilakukan eliminasi variabel yang tidak layak untuk dianalisis, dan selanjutnya dilakukan pengulangan analisis faktor.

Selanjutnya, untuk melihat kelayakan masing-masing variabel atau menentukan variabel mana yang patut dieliminasi, digunakan koefisien Measure of Sampling Adequacy (MSA) dalam tabel *Anti-Image*. Koefisien MSA berkisar dari 0 sampai dengan 1, dengan kriteria sebagai berikut.

- Apabila koefisien MSA = 1, maka variabel tersebut dapat diprediksi tanpa ada kesalahan oleh variabel lain.
- Apabila koefisien MSA  $\geq 0,5$ , maka variabel masih bisa diprediksi dan dianalisis lebih lanjut.

Apabila koefisien MSA  $< 0,5$ , maka variabel tidak bisa diprediksi dan tidak bisa dianalisis lebih lanjut.

Faktor-faktor yang baru terbentuk dapat dilihat dari tabel *Total Variance Explained*. Apabila dalam tabel tersebut sudah tampak nilai *eigenvalue* berubah menjadi kurang dari 1, maka banyak faktor yang ditunjukkan merupakan banyak faktor yang paling ideal. Selanjutnya, untuk menentukan butir mana yang masuk ke faktor yang mana, dapat dilihat dari tabel *Component Matrix*. Kriterianya, pada faktor mana variabel tersebut memiliki korelasi tertinggi, pada faktor itulah variabel tersebut seharusnya tergabung.

Apabila satu variabel memiliki korelasi yang berimbang terhadap lebih dari satu faktor, maka peneliti sering menemui kesulitan untuk menentukan ke faktor mana variabel tersebut harus dikelompokkan. Apabila hal ini terjadi, maka perlu dilakukan rotasi. Proses rotasi akan menghasilkan Tabel *Rotated Component Matrix*. Dalam tabel *Rotated Component Matrix*,

koefisien korelasi satu variabel dengan faktor-faktor yang terbentuk akan semakin jelas perbedaannya. Dengan demikian, peneliti akan lebih mudah untuk memasukkan variabel tersebut ke faktor yang paling tepat. Fungsi dari analisis faktor antara lain:

- a. Menentukan himpunan dari dimensi yang tidak mudah diamati dalam himpunan variabel (R faktor analysis)
- b. Mengelompokkan orang-orang (misalnya, responden kuis) ke dalam kelompok-kelompok yang berbeda di dalam populasi (Q factor analysis).
- c. Mengidentifikasi variabel-variabel yang digunakan dalam analisis lanjutan (regresi, korelasi, atau diskriminan)
- d. Membentuk himpunan dari variabel (dengan lebih sedikit) untuk menggantikan (sebagian/ seluruh) himpunan variabel awal.
- e. Menganalisis suatu fenomena dengan data yang sangat besar.
- f. Menjabarkan/ menguraikan suatu kaitan yang kompleks di antara fenomena ke dalam fungsi kesatuan-kesatuan atau ke dalam bagian-bagiannya, dan dapat mengidentifikasi pengaruh dari luar (independen).

## TUGAS 11

Seorang peneliti melakukan penelitian dengan judul analisis faktor yang menentukan faktor-faktor yang mempengaruhi seseorang memilih presiden. Untuk itu peneliti mengajukan beberapa kuesioner kepada responden yang memuat pertanyaan tentang faktor-faktor yang mempengaruhi seseorang memilih presiden. Setelah didapatkan data dari responden, maka dilakukan pengujian analisis faktor. Lakukan analisa untuk data berikut menggunakan analisis factor.

Responden	Berwibawa	Tampakan	Cerdas	Berwawasan luas	Merakyat	Pendidikan	Ikutan-ikutan	Bebas korupsi	Religius
1	4	4	5	5	4	5	6	6	4
2	5	5	5	6	5	6	6	5	5
3	6	5	8	7	6	7	8	6	6
4	6	6	8	8	7	6	6	7	8
5	8	4	9	4	7	3	4	5	6
6	7	5	8	5	5	6	7	6	6
7	7	6	8	6	6	6	5	7	7
8	7	2	9	7	4	5	6	4	6
9	7	5	6	4	6	4	5	5	4
10	7	6	7	6	6	6	7	5	4
11	7	5	6	7	6	5	6	7	5
12	5	4	7	8	4	5	6	7	4
13	7	2	7	6	6	4	6	5	4
14	7	3	8	6	6	4	7	6	5
15	7	6	7	7	6	6	5	6	6
16	6	3	8	5	5	4	6	7	4
17	8	3	8	4	5	6	7	5	3
18	8	4	9	3	6	5	4	8	7
19	2	4	10	2	5	6	5	4	5
20	3	7	9	5	4	6	8	4	6

## CHAPTER 12

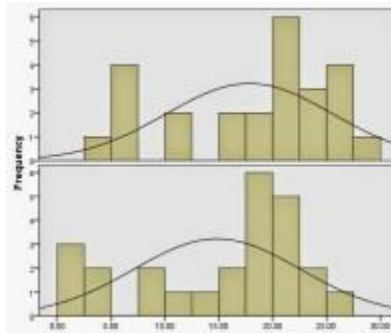
### UJI DALAM STATISTIK NON PARAMETRIK (MANN WHITNEY)

Mann Whitney U Test adalah uji non parametris yang digunakan untuk mengetahui perbedaan median 2 kelompok bebas apabila skala data variabel terikatnya adalah ordinal atau interval/ratio tetapi tidak berdistribusi normal. Berdasarkan definisi di atas, uji Mann Whitney U Test mewajibkan data berskala ordinal, interval atau rasio. Apabila data interval atau rasio, maka distribusinya tidak normal. Sumber data adalah 2 kelompok yang berbeda, misal kelas A dan kelas B di mana individu atau objek yang diteliti adalah objek yang berbeda satu sama lain.

Mann Whitney U Test disebut juga dengan Wilcoxon Rank Sum Test. Merupakan pilihan uji non parametris apabila uji Independent T Test tidak dapat dilakukan oleh karena asumsi normalitas tidak terpenuhi. Tetapi meskipun bentuk non parametris dari uji independent t test, uji Mann Whitney U Test tidak menguji perbedaan Mean (rerata) dua kelompok seperti layaknya uji Independen T Test, melainkan untuk menguji perbedaan Median (nilai tengah) dua kelompok.

Tetapi beberapa ahli tetap menyatakan bahwasanya uji Mann Whitney U Test tidak hanya menguji perbedaan Median, melainkan juga menguji Mean. Mengapa seperti itu? karena dalam berbagai kasus, Median kedua kelompok bisa saja sama, tetapi nilai P Value hasilnya kecil yaitu  $< 0,05$  yang berarti ada perbedaan. Penyebabnya adalah karena Mean kedua kelompok tersebut berbeda secara nyata. Maka dapat disimpulkan bahwa uji ini bukan hanya menguji perbedaan Median, melainkan juga perbedaan Mean.

Seseorang akan melakukan uji Mann Whitney U Test apabila menemui kasus: Diketahui dengan jelas bahwa terdapat perbedaan median, bentuk dan sebaran data sama, tetapi tidak diketahui secara pasti apakah perbedaan median tersebut bermakna atau tidak.



Gambar 4. Histogram Data

Perhatikan dua histogram di atas, di mana bentuk lebar dan ketinggian keduanya sama, yang berarti bentuk dan sebaran data kedua kelompok sama, tetapi median keduanya berbeda. Lihat bahwa histogram yang di atas lebih ke kanan dari pada yang di bawah, yaitu dengan median 18 sedangkan yang di bawah dengan median 15. Maksud dari peneliti melakukan uji Mann Whitney U Test adalah menguji apakah perbedaan median tersebut bermakna atau tidak. Bagaimana jika bentuk dan sebaran dari histogram tidak sama? apakah masih bisa dilakukan uji ini? Jawabannya adalah “Ya”, tetapi peneliti tidak lagi menguji perbedaan Median dan Mean, melainkan menguji perbedaan Mean saja.

Berdasarkan uraian di atas, maka dapat disimpulkan Asumsi yang harus terpenuhi dalam Mann Whitney U Test, yaitu:

1. Skala data variabel terikat adalah ordinal, interval atau rasio. Apabila skala interval atau rasio, asumsi normalitas tidak terpenuhi. (Normalitas dapat diketahui setelah uji normalitas).

2. Data berasal dari 2 kelompok. (Apabila data berasal dari 3 kelompok atau lebih, maka sebaiknya gunakan uji Kruskal Wallis).
3. Variabel independen satu dengan yang lainnya, artinya data berasal dari kelompok yang berbeda atau tidak berpasangan.
4. Varians kedua kelompok sama atau homogen. (Karena distribusi tidak normal, maka uji homogenitas yang tepat dilakukan adalah uji Levene's Test. Di mana uji Fisher F diperuntukkan bila asumsi normalitas terpenuhi).

Langkah-langkah perhitungan:

1. Gabungkan kedua sampel dan beri peringkat atau ranking dari data terkecil sampai terbesar.
2. Bila ada peringkat / ranking yang sama, peringkatnya diambil rata-ratanya
3. Hitung jumlah peringkat sampel 1 dan sampel 2 dan dinotasikan dengan R1 dan R2
4. Kemudian menentukan nilai U:

$$U_1 = n_1 n_2 + \frac{n_1(n_1 + 1)}{2} - R_1 \dots\dots\dots 1)$$

$$U_2 = n_1 n_2 + \frac{n_2(n_2 + 1)}{2} - R_2 \dots\dots\dots 2)$$

Untuk  $n_1 : n_2 \leq 20$      $U \sim$  berdistribusi  $Un_{1:n_2; \alpha}$  (nilai  $U_\alpha \rightarrow$  tabel mann whitney)  
 Untuk  $n_1 : n_2 > 20$      $U \sim$  berdistribusi normal  
    dengan rata-rata     $\mu_u = \frac{n_1 n_2}{2}$   
    dengan standard deviasi     $\sigma_u = \sqrt{\frac{n_1 n_2 (n_1 + n_2 + 1)}{12}}$   
  
    sehingga :     $Z_{hitung} = \frac{W - \mu_u}{\sigma_u} \sim$  berdistribusi normal  
    standard  $N(0; 1)$

Dengan nilai ketentuan

$n_1 : n_2 \leq 20$   
 $H_1 : \mu_1 \neq \mu_2 \rightarrow H_0$  ditolak jika  $U < U_\alpha$   
 $H_1 : \mu_1 > \mu_2 \rightarrow H_0$  ditolak jika  $U_1 < U_\alpha$   
 $H_1 : \mu_1 < \mu_2 \rightarrow H_0$  ditolak jika  $U_2 < U_\alpha$

## TUGAS 12

Diketahui data sebagai berikut, ujilah apakah nilai fakultas ekonomi lebih baik dibandingkan dengan fakultas computer

### Berikut adalah nilai UAS Statistika 2 mahasiswa fakultas Ekonomi dan ilmu komputer

Berikan peringkat (ranking) data dalam tabel berikut ini!

Mahasiswa Fak. Ekonomi		
Nilai	Urutan	Rangking
30	2	2
55	4	4
65	5	5
70	8	7
75	10	9.5
88	16	15.5
90	17	17
95	18	18
98	19	19
100	20	20
$R_1 =$		117

Mahasiswa Fak. Ilmu Komputer		
Nilai	Urutan	Ranking
25	1	1
50	3	3
70	6	7
70	7	7
75	9	9.5
78	11	11
80	12	12
85	13	13.5
85	14	13.5
88	15	15.5
$R_2 =$		93

Urutan	Nilai	Rank
1	25	1
2	30	2
3	50	3
4	55	4
5	65	5
6	70	7
7	70	7
8	70	7
9	75	9.5
10	75	9.5
11	78	11
12	80	12
13	85	13.5
14	85	13.5
15	88	15.5
16	88	15.5
17	90	17
18	95	18
19	98	19
20	100	20

Catatan: jumlah sampel mahasiswa 20

SI 2 - Statistik Non Parametrik





## REFERENSI

- Agresti, Alan & Barbara Finlay. 1999. *Statistical Methods for the sosial sciences* Prentice Hall, Upper Saddle River, New Jersey. Cochran,
- Antoni, 2007. *Statistik Deskriptif*, Padang: Bung Hatta University Press.
- Adhityo Kuncoro, April 2017, Korelasi Penguasaan Kosakata Dengan Keterampilan Berbicara Siswa Dalam Bahasa Inggris, *Jurnal SAP* Vol. 1 No. 3.
- Cindy Viane Bertan, 2016, Pengaruh Pendayagunaan Sumber Daya Manusia (Tenaga Kerja) Terhadap Hasil Pekerjaan (Studi Kasus Perumahan Taman Mapanget Raya (Tamara), **Jurnal Sipil Statik** Vol.4 No.1 (13-20) ISSN: 2337-6732 13
- Djarwanto, 2006. *Statistik Nonparametrik*, Jogjakarta: BPFPE.
- Dawson B, Trapp RG. 4th ed. New York: McGraw Hill; 2004. *Basic and clinical biostatistics*.
- Dewanto & Tarsis Tarmudji. 1995. *Metode Statistika*, Liberty. Yogyakarta. Gulo, W. 2002. *Metodologi Penelitian*, Grasindo. Jakarta.
- Furqon. 2009. *Statistika Terapan untuk Penelitian*. Cetakan ketujuh. ALFABETA: Bandung.
- Gravetter FJ, Wallnau LB. 5th ed. Belmont: Wadsworth – Thomson Learning; 2000. *Statistics for the behavioral sciences*.
- Hermawati, 2010, *Statistik Sosial*, Padang: Hayfa Press.
- James H. Torrie. 1995. *Prinsip dan Prosedur Statistika Suatu Pendekatan Biometrik* Gramedia. Jakarta. Walpole,
- Norman GR, Streiner DL. 2nd ed. Hamilton: B.C. Decker Inc; 2000. *Biostatistics the bare essentials*.
- Riduwan. 2008. *Dasar-dasar Statistika*. Bandung: Alfabeta
- Riduwan, 2009. *Pengantar Statistika Sosial*, Bandung: Alvabeth
- Rasyad, Rashidan. 2003. *Metode Statistik Deskriptif*. Grasindo, Jakarta Steel, Robert GD & Ronald E. 1992. *Pengantar Statistika*, Gramedia, Jakarta.
- Sudjana.1996. *Metoda Statistika*. Bandung: Tarsito Bandung
- Sundar Rao PS, Richard J. 4th ed. New Delhi: Prentice Hall of India Pvt Ltd; 2006. *Introduction to biostatistics and research methods*.
- Sundaram KR, Dwivedi SN, Sreenivas V. 1st ed. New Delhi: B.I Publications Pvt Ltd; 2010. *Medical statistics principles and methods*.
- Swinscow TDV, Campbell MJ. (Indian) 10th ed. New Delhi: Viva Books Private Limited; 2003. *Statistics at square one*.10 th ed (Indian)

Usman, Husaini. 2006. Pengantar Statistika. Jakarta: PT Bumi Aksara

William G. 1991. Teknik Penarikan Sampel. Universitas Indonesia, Jakarta.

Zanten, Wim Van, 1982. Statistika Untuk Ilmu-ilmu Sosial, Gramedia. Jakarta.