
Statistik Terapan - Analisis Data

Oleh: KINTOKO

DAFTAR ISI

KATA PENGANTAR.....	Error! Bookmark not defined.
DAFTAR ISI	iii
DAFTAR TABEL.....	v
DAFTAR GAMBAR.....	vi
DAFTAR LAMPIRAN.....	vii
BAB I. PENDAHULUAN.....	1
Tujuan pembelajaran.....	1
Struktur bahasan.....	1
A. Statistik dan Statistika	1
B. Statistika Murni dan Statistika Terapan	2
C. Statistika Deskriptif dan Statistika Inferensial.....	3
BAB II. PENYAJIAN DATA DALAM BENTUK TABEL.....	1
Tujuan Pembelajaran.....	1
Susunan Materi.....	1
A. Pengantar	1
B. Tabel Distribusi Frekuensi Tunggal dan Bergolong.....	2
1. Tabel Distribusi Frekuensi Tunggal.....	2
2. Tabel Distribusi Frekuensi Bergolong.....	3
C. Tabel Distribusi Frekuensi Relatif.....	5
D. Tabel Distribusi Frekuensi Kumulatif.....	5
LATIHAN	6
BAB III. PENYAJIAN DATA DALAM BENTUK GRAFIK.....	8
Tujuan Pembelajaran.....	8
Susunan materi.....	8
A. Pengantar	8
B. Grafik Histogram	8
C. Grafik Poligon	10
D. Grafik Ogive.....	11
BAB IV. UKURAN PEMUSATAN DATA (<i>TENDENSI SENTRAL</i>)	13
TUJUAN PEMBELAJARAN.....	Error! Bookmark not defined.
SUSUNAN MATERI	Error! Bookmark not defined.
A. PENGANTAR	13
B. Mean (rata-rata).....	14

C. Modus	16
D. Median.....	17
LATIHAN	Error! Bookmark not defined.
BAB V. UKURAN VARIABILITAS	19
tujuan pembelajaran	Error! Bookmark not defined.
Susunan materi.....	20
A. Pengantar	20
B. Kuartil, Desil, dan Persentil.....	21
1. Kuartil.....	21
2. Desil dan Persentil.....	21
C. Range, Deviasi Kuartil, dan Standar Deviasi	23
1) Range.....	23
2) Deviasi Kuartil.....	24
3. Standar Deviasi (SD).....	24
LATIHAN	28
BAB VI. UJI - t.....	30
BAB VII. ANALISIS VARIAN	38
BAB VIII. TEKNIK KORELASI	53
BAB IX. ANALISIS REGRESI	58
DAFTAR PUSTAKA.....	69

DAFTAR TABEL

Tabel 2- 1. Distribusi Frekuensi Skor Kemampuan Membaca _____	2
Tabel 2- 2. Distribusi Frekuensi Skor Kemampuan Membaca Pemahaman _____	4
Tabel 2- 3. Distribusi Frekuensi Skor Kemampuan Membaca Pemahaman _____	5
Tabel 2- 4. Distribusi Frekuensi Relatif Skor Kemampuan Membaca _____	5
Tabel 2- 5. Distribusi Frekuensi Kumulatif "Kurang Dari" _____	6
Tabel 2- 6. Distribusi Frekuensi Kumulatif "Atau Lebih" _____	6
Tabel 3- 1. Kelas interval, batas kelas dan frekuensi hasil tes kemampuan _____	9
Tabel 3- 2. Kelas interval, batas kelas, dan frekuensi skor membaca _____	10
Tabel 3- 3. <i>Distribusi Frekuensi Kumulatif "Kurang Dari"</i> _____	11
Tabel 3- 4. <i>Tabel Distribusi Frekuensi Kumulatif "Atau Lebih"</i> _____	12
Tabel 4- 1. <i>Tabel Distribusi Frekuensi Hasil Tes Statistik</i> _____	14
Tabel 4- 2. Data yang telah dikelaskan _____	15
Tabel 4- 3. Perhitungan dengan metode mean kerja (MK) _____	16

DAFTAR GAMBAR

Gambar 3- 1. Histogram skor kemampuan membaca pemahaman _____	9
Gambar 3- 2. Poligon skor kemampuan membaca pemahaman _____	10
Gambar 3- 3. Ogive skor kemampuan membaca _____	12

DAFTAR LAMPIRAN

Lampiran 1. Tabel nilai F	70
Lampiran 2. Tabel nilai r	71

BAB I

PENDAHULUAN

TUJUAN PEMBELAJARAN

Mahasiswa dapat

1. Mendefinisikan pengertian statistik, parameter, statistika dan lainnya
2. Membedakan berbagai jenis statistik
3. Memberikan contoh penggunaan statistika sehari-hari

STRUKTUR BAHASAN

- a) Statistik dan Statistika
- b) Statistika Murni dan Statistika Terapan
- c) Statistika Deskriptif dan Statistika Inferensial

A. STATISTIK DAN STATISTIKA

Kata statistik dipakai untuk menyatakan kumpulan fakta, umumnya berbentuk angka yang disusun dalam tabel atau diagram yang menggambarkan suatu persoalan. Dalam pengertian ini kita mengenal nama-nama statistik, seperti: *statistik penduduk, statistik kelahiran, statistik pendidikan, statistik pertanian, statistik kesehatan*, dan lain-lain.

Kata statistik juga diartikan sebagai suatu ukuran yang dihitung dari sekumpulan data dan merupakan wakil dari data itu.

Misalnya: a. 90% mahasiswa yang mengikuti kuliah ini berasal dari kota "A"

b. Kecelakaan lalu lintas itu kebanyakan diakibatkan karena kecerobohan pengemudi angkutan kota.

c. Rata-rata berat badan dari mahasiswa yang mengikuti lomba ini adalah 50 kg.

Dalam hal ini 90%, kebanyakan, rata-rata termasuk ke dalam statistik.

Pengertian statistik yang ketiga dikaitkan dengan ilmu pengetahuan atau metode ilmiah yang sering disebut statistika. Statistika adalah pengetahuan yang berhubungan dengan cara-cara pengumpulan fakta, pengolahan serta penganalisisannya, penarikan kesimpulan serta pembuatan keputusan yang cukup beralasan berdasarkan fakta dan penganalisisan yang dilakukan.

Pengetahuan mengenai statistika perlu dimiliki oleh mereka yang sedang mengajukan usulan penelitian. Dengan pengetahuan statistika mereka dapat menganalisis dan menafsirkan data mereka, serta dapat menyampaikan hasil penelitiannya kepada orang lain.

Para pendidik yang harus selalu mengikuti perkembangan dan memanfaatkan hasil-hasil penelitian seyogyanya juga memahami statistika sehingga mereka dapat mengerti dan menilai penelitian yang dilakukan orang lain. Selain itu, pelaksanaan dan penafsiran tes-tes yang dipakai di sekolah pun menuntut adanya pengertian tentang statistika. Guru-guru yang tidak memahami statistika mungkin akan menemui kesulitan dalam menilai kemampuan dan prestasi belajar murid-muridnya. Mereka juga akan menemui kesulitan dalam membaca laporan penelitian dalam bidang spesialisasinya maupun dalam memperoleh informasi umum.

Hinkle (1979:3) menyatakan bahwa peneliti menggunakan statistik sebagai metode atau prosedur yang diterapkan dalam usahanya untuk mengerti atau memahami data penelitian. Data terdiri atas informasi dalam bentuk angka yang mewakili/mencerminkan sebuah karakteristik atau fenomena.

Peran statistik dalam ilmu sosial dan pendidikan juga penting. Tingkah laku individu sering dipantau dan dijelaskan dengan data yang mewakili tingkah laku-tingkah laku tersebut. Statistik digunakan untuk memberikan bukti-bukti empiris yang mendukung atau menolak teori-teori yang digunakan dalam ilmu-ilmu sosial atau pendidikan.

B. STATISTIKA MURNI DAN STATISTIKA TERAPAN

Ada dua macam statistika, yaitu statistika murni dan statistika terapan. Statistika murni membicarakan teori-teori dan rumus-rumus secara terinci, sedangkan statistika terapan lebih menekankan pembahasannya pada penerapan atau penggunaan rumus-rumus tersebut.

Statistika Pendidikan tergolong dalam statistika terapan, yaitu statistika yang diterapkan dalam bidang pendidikan. Dalam memahami statistika pendidikan para mahasiswa cukup memahami bagaimana penerapan statistika di bidang pendidikan. Karena itu, tidak benar anggapan yang menyatakan bahwa untuk mempelajari statistika pendidikan diperlukan dasar matematika yang kuat. Untuk mempelajari

statistika pendidikan cukup diperlukan dasar-dasar matematika yang sederhana yang telah diperoleh sewaktu mereka masih di sekolah lanjutan.

C. STATISTIKA DESKRIPTIF DAN STATISTIKA INFERENSIAL

Statistika pendidikan menurut fungsinya terbagi menjadi dua bagian, yaitu statistika deskriptif dan statistika inferensial. Statistika deskriptif adalah statistika yang tugasnya untuk mendeskripsikan dan merangkum data hasil pengamatan yang telah dilakukan, sedang statistika inferensial adalah statistika yang tugasnya untuk menarik kesimpulan dari data yang telah dikumpulkan.

Statistika Pendidikan yang dibahas dalam perkuliahan ini adalah statistik yang digunakan untuk menganalisis data bidang pendidikan, meliputi statistika deskriptif dan statistika inferensial.

BAB II

PENYAJIAN DATA DALAM BENTUK TABEL

TUJUAN PEMBELAJARAN

Mahasiswa dapat:

1. membaca data
2. menyajikan data dalam grafik

SUSUNAN MATERI

- a) Pengantar
- b) Tabel Distribusi Frekuensi Tunggal dan Bergolong
- c) Tabel Distribusi Frekuensi Relatif
- d) Tabel Distribusi Frekuensi Kumulatif

A. PENGANTAR

Jika seseorang melakukan suatu studi atau suatu penelitian, mereka akan memperoleh data. Data yang telah dikumpulkan, baik berasal dari populasi maupun dari sampel, untuk keperluan laporan atau analisis selanjutnya, perlu diatur, disusun, disajikan dalam bentuk yang jelas dan baik. Statistik deskriptif berkenaan dengan tugas ini. Dua cara penyajian data yang sering digunakan ialah: (1) tabel distribusi frekuensi, dan (2) grafik.

Pada bagian berikut akan dipaparkan tentang bagaimana data penelitian disajikan ke dalam tabel distribusi frekuensi

Sebuah distribusi frekuensi dimaksudkan sebagai rangkaian angka yang menunjukkan berapa kali skor atau sekelompok skor muncul. Mengapa seorang peneliti perlu membuat distribusi frekuensi? Pengembangan sebuah distribusi frekuensi secara sistematis mengorganisir data dan memberikan peneliti beberapa indikasi mengenai keadaan data. Sebuah distribusi frekuensi memang tidak mencerminkan semua informasi dari sekumpulan data, namun penyajian data dalam distribusi frekuensi merupakan permulaan mengerti sesuatu mengenai data.

B. TABEL DISTRIBUSI FREKUENSI TUNGGAL DAN BERGOLONG

Distribusi frekuensi dapat dibedakan menjadi dua macam, yaitu (1) distribusi frekuensi tunggal, dan (2) distribusi frekuensi bergolong. Jika jarak sebaran angka (*range*) kecil kita dapat mempergunakan distribusi frekuensi tunggal. Sebaliknya, jika jarak sebaran angka yang diperoleh cukup tinggi, lebih tepat kiranya jika dipergunakan distribusi frekuensi bergolong. *Range* atau *rentang* dari data terbesar dikurangi data terkecil.

1. TABEL DISTRIBUSI FREKUENSI TUNGGAL

Contoh:

Skor hasil Tes Kemampuan Membaca Pemahaman dari 40 orang siswa dengan kemungkinan skor tertinggi 75 adalah sebagai berikut:

55 60 57 60 55 53 62 60 55 58
54 62 52 60 54 60 61 61 55 64
58 58 60 63 62 59 59 61 56 54
53 53 55 56 60 57 58 62 63 59

Kisaran (*range*) data tersebut = $(64 - 52) = 12$ (tidak lebih dari 15) karena itu data tersebut lebih tepat disajikan dalam tabel distribusi tunggal seperti berikut.

Tabel 2- 1. Distribusi Frekuensi Skor Kemampuan Membaca

No.	Skor	Cacahan/Tally	Frekuensi (f)
1	64	/	1
2	63	//	2
3	62	////	4
4	61	///	3
5	60	///// //	7
6	59	///	3
7	58	////	4
8	57	//	2
9	56	//	2
10	55	////	5
11	54	///	3
12	53	///	3
13	52	/	1
		Jumlah (N)	40

Dari Tabel 2-1 distribusi frekuensi dapat dilihat “bentuk” umum sebaran skor tersebut. Dengan skor yang telah tersusun kita dapat menetapkan penyebarannya, apakah skor tersebar secara merata ataukah cenderung mengelompok, dan di mana pengelompokan itu terjadi. Sebagai contoh, dengan melihat tabel distribusi frekuensi skor kemampuan membaca pemahaman di atas kita dengan mudah melihat bahwa skor itu terentang antara 52 sampai 64, bahwa skor 60 adalah skor dengan frekuensi terbanyak, dan bahwa ada kecenderungan skor tersebut mengelompok di tengah. Semua itu tidak akan tampak seandainya skor tersebut belum tersusun. Penyusunan data ke dalam distribusi frekuensi juga akan memudahkan penghitungan berbagai statistik yang diperlukan.

2. TABEL DISTRIBUSI FREKUENSI BERGOLONG

Contoh:

Skor hasil tes Kemampuan Membaca dari 50 orang siswa dengan kemungkinan skor tertinggi 75 adalah sebagai berikut:

48	56	44	60	34	41	55	62	53	37
58	63	28	52	69	40	38	44	49	60
23	25	18	54	46	15	22	38	34	54
24	40	44	39	47	44	42	34	37	32
26	36	42	58	65	40	29	53	40	28

Range data tersebut = $(69 - 15) = 54$ relatif besar (lebih dari 15) karena itu data tersebut lebih tepat disajikan dalam tabel distribusi frekuensi bergolong. Untuk membuat tabel distribusi frekuensi bergolong, kita lakukan langkah-langkah berikut.

- a. Menentukan **banyak kelas interval** (K) = 7

Tidak ada ketentuan yang harus diikuti secara mutlak. Sebagai acuan-ancuan yaitu antara 5 sampai dengan 15.

Dalam hal ini ada sebuah aturan untuk menentukan banyak kelas yang digunakan, yaitu **Aturan Sturges** dengan rumusnya sebagai berikut.

$$K = 1 + (3,3 \log n)$$

- b. Menentukan panjang kelas interval, disingkat p , dengan rumus:

$$p = \frac{\text{range}}{\text{kelas}} = \frac{54}{7} = 8$$

- c. Pilih ujung bawah kelas interval pertama. Untuk ini dapat diambil sama dengan data terkecil atau nilai data yang lebih kecil dari data terkecil tetapi selisihnya harus kurang dari panjang kelas yang telah ditentukan. Selanjutnya tabel diselesaikan dengan menggunakan harga-harga yang telah dihitung.

Sebelum tabel sebenarnya dituliskan, ada baiknya dibuat tabel penolong yang berisikan *kolom tabulasi*. Kolom ini merupakan kumpulan deretan garis-garis miring yang banyaknya sesuai dengan banyak data yang terdapat dalam kelas interval yang bersangkutan. Dengan mengambil banyak kelas 7, panjang kelas 8 dan dimulai dengan ujung bawah kelas pertama sama dengan 15, tabel penolong akan seperti di bawah ini.

Tabel 2- 2. Distribusi Frekuensi Skor Kemampuan Membaca Pemahaman

No	Kelas interval	Tabulasi	Frekuensi (f)
1	15-22	///	3
2	23-30	//// //	7
3	31-38	//// ////	9
4	39-46	//// //// ////	13
5	47-54	//// ////	8
6	55-62	//// //	7
7	63-70	///	3
		Jumlah (N)	50

No	innterval	f absolud	f relatife (%)
1	62-69	4	8
2	54-61	8	16
3	46-53	7	14
4	38-45	14	28
5	30-37	7	14
6	22-29	8	16
7	14-21	2	4
		50	100

Setelah dituliskan dalam bentuk tabel yang lazim dipakai, hasilnya seperti berikut.

Tabel 2- 3. Distribusi Frekuensi Skor Kemampuan Membaca Pemahaman

No.	Kelas interval	Frekuensi
1	15-22	3
2	23-30	7
3	31-38	9
4	39-46	13
5	47-54	8
6	55-62	7
7	63-70	3
		50

Dari Tabel tersebut kita dapat membuat penafsiran sebagai berikut.

- a. Siswa yang mendapat skor 15 sampai 22 ada 3 orang
- b. Siswa yang mendapat skor antara 23-30 ada 7 orang
- c. Siswa yang mendapat skor antara 31-38 ada 9 orang dan seterusnya.

C. TABEL DISTRIBUSI FREKUENSI RELATIF

Dalam Tabel 2-3 di atas frekuensi dinyatakan dengan banyak data yang terdapat dalam tiap kelas. Jadi dalam bentuk absolut. Jika frekuensi dinyatakan dalam bentuk persen, maka diperoleh tabel distribusi frekuensi relatif. Tentu saja, absolut dan relatif dapat disajikan dalam sebuah tabel. Skor Kemampuan Membaca dalam tabel di atas dapat kita sajikan ke dalam tabel sebagai berikut.

Tabel 2- 4. Distribusi Frekuensi Relatif Skor Kemampuan Membaca

No.	Kelas interval	Frekuensi Absolut	Frekuensi Relatif (%)
1	15-22	3	6
2	23-30	7	14
3	31-38	9	18
4	39-46	13	26
5	47-54	8	16
6	55-62	7	14
7	63-70	3	6
		50	100

D. TABEL DISTRIBUSI FREKUENSI KUMULATIF

Tabel distribusi frekuensi kumulatif dapat dibentuk dari tabel di atas dengan jalan menjumlahkan frekuensi demi frekuensi. Ada dua macam distribusi frekuensi kumulatif yang dikenal, yaitu distribusi frekuensi kumulatif *kurang dari* dan frakuensi komulatif *atau lebih*.

Tabel 2- 5. Distribusi Frekuensi Kumulatif “Kurang Dari”

No.	Skor Kemampuan Membaca	Frekuensi Kumulatif	Frekuensi Kum (%)
1	Kurang dari 15	0	0
2	Kurang dari 23	3	6
3	Kurang dari 31	10	20
4	Kurang dari 39	19	38
5	Kurang dari 47	32	64
6	Kurang dari 55	40	80
7	Kurang dari 63	47	94
8	Kurang dari 71	50	100

Tabel 2- 6. Distribusi Frekuensi Kumulatif “Atau Lebih”

No.	Skor Kemampuan Membaca	Frekuensi Kumulatif	Frekuensi Kum (%)
1	15 atau lebih	50	100
2	23 atau lebih	47	94
3	31 atau lebih	40	80
4	39 atau lebih	31	62
5	47 atau lebih	18	36
6	55 atau lebih	10	20
7	63 atau lebih	3	6
8	71 atau lebih	0	0

LATIHAN

1. Untuk mendapatkan gambaran yang lebih jelas mengenai nilai-nilai Bahasa Indonesia

20 orang siswa kelas III suatu Sekolah Dasar di DIY berikut sajikanlah ke dalam sebuah tabel distribusi frekuensi.

7, 6, 6, 5, 4, 7, 9,8, 9, 10

7, 10, 6, 8, 5, 7, 8, 9, 7, 8

2. Sajikanlah nilai-nilai Bahasa Indonesia tersebut dalam tabel distribusi frekuensi relatif.

3. Sajikanlah IQ 100 orang mahasiswa berikut ke dalam tabel distribusi frekuensi.

127, **129**, 120, 122, 124, 115, 116, 117, 118, 119
118, 119, 110, 111, 112, 113, 114, 114, 113, 112
110, 111, 112, 113, 105, 106, 107, 108, 109, 108
107, 106, 105, 106, 107, 108, 109, 108, 107, 106
105, 106, 107, 108, 109, 100, 101, 102, 103, 104
103, 102, 101, 95, 96, 97, 98, 99, 98, 97
101, 100, 101, 95, 96, 97, 98, 99, 98, 97
96, 95, 96, 97, 98, 99, 98, 97, 96, 95
96, 97, 98, 90, 91, 92, 93, 94, 93, 92
91, 90, 92, 91, 85, 86, 87, 88, 89, **84**.

4. Sajikan IQ 100 orang mahasiswa di atas ke dalam tabel distribusi frekuensi kumulatif.

5. Sajikan juga ke dalam tabel distribusi frekuensi kumulatif dalam persen

BAB III

PENYAJIAN DATA DALAM BENTUK GRAFIK

TUJUAN PEMBELAJARAN

Mahasiswa dapat:

- 1) Memilih bentuk grafik yang tepat
- 2) Menyajikan data dalam bentuk grafik tersebut

SUSUNAN MATERI

- a) Pengantar
- b) Grafik histogram
- c) Grafik poligon
- d) Grafik Ogive

A. PENGANTAR

Penyajian data penelitian dalam bentuk grafik adalah salah satu cara yang banyak digunakan peneliti. Penyajian dalam bentuk grafik sering dirasa tepat dan ada manfaatnya. Grafik lebih cepat ditangkap mata, mudah dicerna jika ada pola kecenderungan dan perbedaan, lebih artistik dan menarik dalam presentasi, Dalam statistika grafik yang sering digunakan ada tiga, yaitu grafik histogram, grafik poligon, dan grafik ogive.

B. GRAFIK HISTOGRAM

Grafik histogram biasa juga disebut *Diagram Bar*, yaitu grafik yang terdiri dari segi empat-segi empat panjang pada sebuah absis. Lebar segi empat menunjukkan lebar kelas, dan tinggi segi empat menunjukkan frekuensi.

Misalnya, kita akan membuat grafik histogram tentang hasil tes kemampuan membaca 50 orang siswa. Untuk membuat grafik histogram terlebih dahulu datanya harus disajikan dalam tabel distribusi frekuensi dengan menambahkan kolom *batas kelas*. *Batas bawah kelas* sama dengan ujung bawah kelas dikurangi 0,5, sedangkan *batas atas kelas* sama dengan ujung atas ditambah 0,5. Contoh: Kelas interval pertama

adalah 15–22. Ujung bawah kelas = 15, ujung atas = 22. Adapun batas bawah kelas = 14,5 dan batas atas kelas = 22,5.

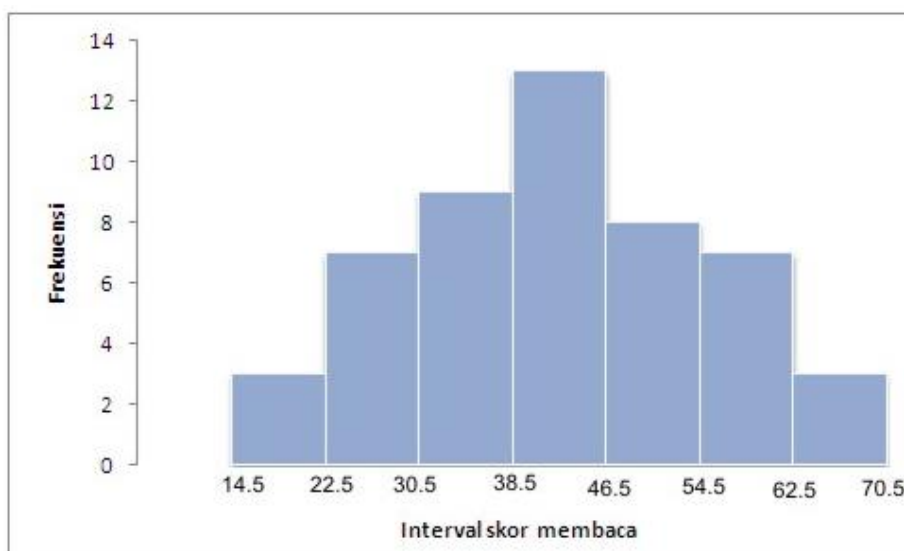
Tabel 3- 1. Kelas interval, batas kelas dan frekuensi hasil tes kemampuan membaca 50 siswa

No.	Kelas interval	Batas Kelas	Frekuensi
1	15-22	14,5-22,5	3
2	23-30	22,5-30,5	7
3	31-38	30,5-38,5	9
4	39-46	38,5-46,5	13
5	47-54	46,5-54,5	8
6	55-62	54,5-62,5	7
7	63-70	62,5-70,5	3
			50

Langkah-langkah membuat grafik histogram ialah:

1. membuat absis (sumbu mendatar) dan ordinat (sumbu tegak)
2. absis memuat batas nyata tiap-tiap kelas, dan ordinat menyatakan frekuensi.
3. membuat skala pada absis dan pada ordinat.
4. membuat segi empat-segi empat pada absis yang tingginya menunjukkan frekuensi.

Dengan langkah-langkah di atas, sajikan skor kemampuan membaca pada Tabel 3-1 di atas ke dalam grafik histogram (Gambar 3-1).



Gambar 3- 1. Histogram skor kemampuan membaca pemahaman

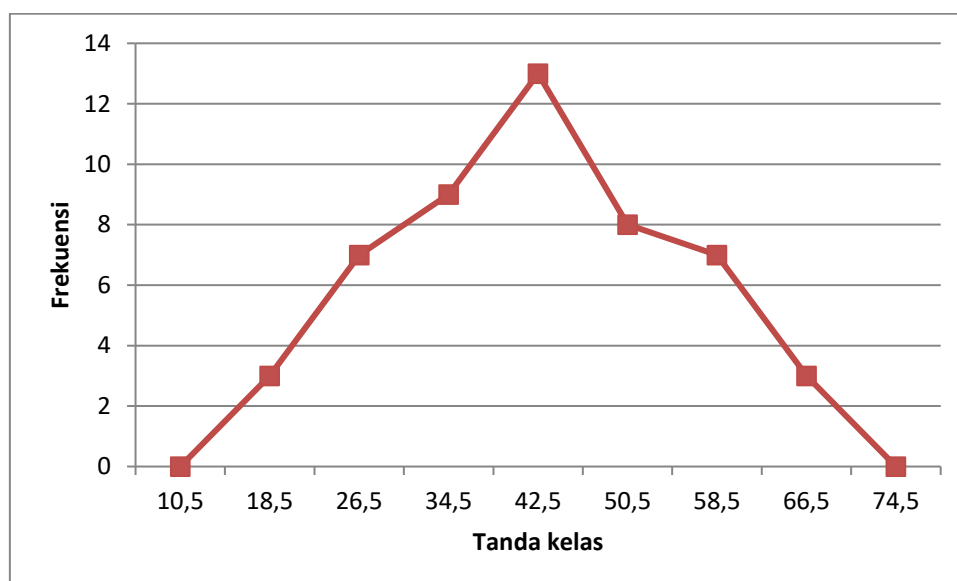
C. GRAFIK POLIGON

Pada dasarnya pembuatan poligon tidak banyak berbeda dengan pembuatan histogram. Hanya, perlu diingat, jika pada pembuatan histogram digunakan batas kelas, pada pembuatan poligon digunakan tanda kelas atau titik tengah. Untuk membuat grafik poligon terlebih dahulu datanya harus dibuat tabel distribusi frekuensi yang memuat kolom tanda kelas atau titik tengah kelas, Tanda kelas (ujung bawah + ujung atas) dibagi dua. Contoh: Kelas interval pertama adalah 15 – 22. Tanda kelasnya = $(15 + 22) : 2 = 18,5$.

Tabel 3- 2. Kelas interval, batas kelas, dan frekuensi skor membaca

No.	Kelas interval	Tanda Kelas	Frekuensi
1	15-22	18,5	3
2	23-30	26,5	7
3	31-38	34,5	9
4	39-46	42,5	13
5	47-54	49,5	8
6	55-62	58,5	7
7	63-70	66,5	3
			50

Melalui langkah-langkah yang sama dengan pembuatan histogram, sajikan skor statistik pada Tabel 3-2 di atas ke dalam poligon (Gambar 3-2).



Gambar 3- 2. Poligon skor kemampuan membaca pemahaman

D. GRAFIK OGIVE

Ogive sering disebut poligon kumulatif atau poligon meningkat. Misalnya, kita mempunyai sekumpulan data yang telah disusun dalam tabel distribusi frekuensi kumulatif, maka kita dapat menggambarkan grafik ogive. Karena tabel distribusi frekuensi kumulatif ada dua macam, yaitu tabel distribusi frekuensi kumulatif "*kurang dari*" dan tabel distribusi frekuensi kumulatif "*atau lebih*", maka grafiknya juga ada dua macam, yaitu *ogive positif* dan *ogive negatif*.

Grafik yang dibuat berdasarkan data yang sudah disusun dalam tabel distribusi frekuensi kumulatif "*kurang dari*" disebut *ogive positif*, sedang grafik yang dibuat berdasarkan data yang sudah disusun dalam tabel distribusi frekuensi "*atau lebih*" disebut *ogive negatif*.

Contoh:

Salin kembali data mengenai skor Kemampuan Membaca dari 50 siswa yang telah disusun dalam tabel distribusi frekuensi kumulatif "*kurang dari*" dan tabel distribusi frekuensi kumulatif "*atau lebih*" seperti nampak pada Bab II.

Tabel 3- 3. Distribusi Frekuensi Kumulatif "Kurang Dari"

No.	Skor Kemampuan Membaca	Frekuensi Kumulatif	Frekuensi Kum (%)
1	Kurang dari 15	0	0
2	Kurang dari 23	3	6,0
3	Kurang dari 31	10	20,0
4	Kurang dari 39	19	38,0
5	Kurang dari 47	32	64,0
6	Kurang dari 55	40	80,0
7	Kurang dari 63	47	94,0
8	Kurang dari 71	50	100,0

Tabel 3- 4. Tabel Distribusi Frekuensi Kumulatif "Atau Lebih"

No.	Skor Kemampuan Membaca	Frekuensi Kumulatif	Frekuensi Kum (%)
1	15 atau lebih	50	100,0
2	23 atau lebih	47	94,0
3	31 atau lebih	40	80,0
4	39 atau lebih	31	62,0
5	47 atau lebih	18	36,0
6	55 atau lebih	10	20,0
7	63 atau lebih	3	6,0
10	71 atau lebih	0	0,0

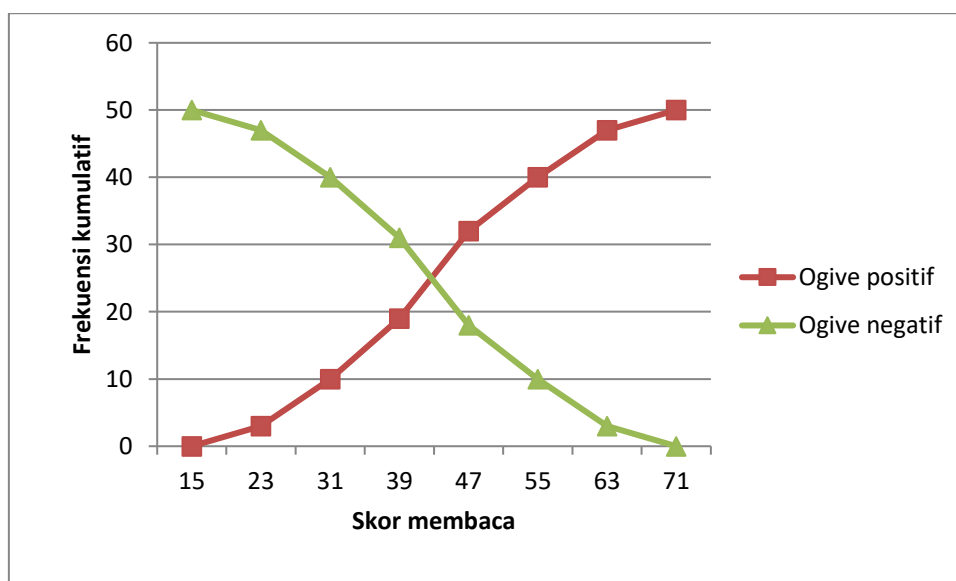
Gambarkan ogivenya!

Penyelesaian:

Sumbu datarnya berisi bilangan-bilangan 15, 23, 31, 39, 47, 55, 63, 71

Sumbu tegaknya berisi bilangan-bilangan 0, 10, 20, 30, 40, 50.

Hasil grafiknya sebagai berikut!



Gambar 3- 3. Ogive skor kemampuan membaca

BAB IV

UKURAN PEMUSATAN DATA (*TENDENSI SENTRAL*)

TUJUAN PEMBELAJARAN

Mahasiswa dapat

- 1) Menghitung nilai rata-rata
- 2) Menetapkan modus
- 3) Menentukan nilai median
- 4) Membandingkan kondisi pemakaian ketiga ukuran pemusatan

SUSUNAN MATERI

- a) Pengantar
- b) Mean (rata-rata)
- c) Modus
- d) Median

A. PENGANTAR

Dalam Bab II dan III sebelumnya telah dikemukakan bermacam-macam cara mengorganisasikan data ke dalam kategori-kategori yang bermakna, sehingga data yang dikumpulkan tidaklah terserak-serak secara individual tetapi telah terkelompokkan serta disajikan ke dalam bentuk gambar. Itu semuanya merupakan langkah awal dari usaha untuk memahami data secara menyeluruh.

Selanjutnya, kita perlu mengerti ciri-ciri tertentu dari data tersebut. Salah satu ukuran yang dapat menunjukkan nilai khusus dari data adalah *tendensi sentral* sebagai suatu numerik tunggal yang dapat mewakili gambaran data yang diberikan. Untuk itu, sering pula orang menggunakan istilah ukuran pemusatan data atau kecenderungan memusat. Ada tiga indeks kecenderungan memusat yang dapat berfungsi untuk mewakili kelompok data secara keseluruhan, yaitu *Mean, Modus, dan Median*. Ketiga memiliki sifat, ciri dan kondisi tertentu dalam penggunaannya. Asumsi dasar yang harus diketahui yaitu sebaran normal dari data pengamatan.

B. MEAN (RATA-RATA)

Ukuran kecenderungan memusat yang paling banyak dipakai adalah *mean*, yang dikenal dengan nama rata-rata atau rata-rata hitung. *Mean* adalah nilai rata-rata dari seluruh nilai yang dijumpai. Istilah rata-rata adalah istilah yang digunakan dalam bahasa sehari-hari, sedangkan istilah statistiknya adalah *mean*. Dengan demikian, konotasi sehari-hari *mean* adalah rata-rata.

a. Menghitung *Mean* dari data kasar

Untuk menghitung *mean* dari data kasar, digunakan rumus sebagai berikut.

$$\bar{X} = \frac{\sum X_i}{N}$$

di mana :

X = jumlah nilai

N = jumlah individu

Contoh : jika ada lima mahasiswa mempunyai nilai ujian Statistik masing-masing yaitu 70, 69, 45, 80, 56, maka nilai rata-ratanya adalah:

$$\bar{X} = \frac{70 + 69 + 45 + 80 + 56}{5} = 64$$

b. Menghitung *mean* untuk data yang telah disajikan dalam tabel distribusi tunggal

$$\bar{X} = \frac{\sum f_i X_i}{N}$$

Contoh Penghitungan :

Tabel 4- 1. Tabel Distribusi Frekuensi Hasil Tes Statistik

Nilai (X)	F	fX
70	5	350
69	6	414
45	3	135
80	1	80
56	1	56
Jumlah	16	1035

Dari tabel distribusi frekuensi tersebut, dengan tambahan kolom fX dapat dihitung nilai rata-rata keenam belas siswa tersebut sebagai berikut:

$$\bar{X} = \frac{1035}{16} = 64,6$$

- c. Menghitung *mean* untuk data yang telah disajikan dalam tabel distribusi frekuensi bergolong.

Rumus yang digunakan sama dengan rumus (b) di atas, hanya di sini X = tanda kelas interval, dan N = frekuensi yang sesuai dengan tanda kelas.

Contoh: Marilah kita hitung nilai rata-rata hasil tes kemampuan membaca yang telah disajikan dalam Tabel 4-2 berikut.

Tabel 4- 2. Data yang telah dikelaskan

No.	Kelas interval	f	Tanda kelas (X)	fX
1	15-22	3	18,5	55,5
2	23-30	7	26,5	185,5
3	31-38	9	34,5	310,5
4	39-46	13	42,5	552,5
5	47-54	8	50,5	404
6	55-62	7	58,5	409,9
7	63-70	3	66,5	199,5
		50		21170,5

Dari tabel tersebut, dengan tambahan kolom tanda kelas dan fX dapat dihitung nilai rata-rata kelimpuluh puluh siswa dengan menerapkan rumus sebagai berikut:

$$\bar{X} = \frac{21170,5}{50} = 42,34$$

- d. Menghitung *Mean* dengan Metode MK (*mean kerja*)

Untuk menghindari angka-angka yang lebih besar dan juga untuk mengurangi kemungkinan melakukan kesalahan, *Mean* dapat dihitung dengan menggunakan "*mean kerja*" disingkat MK. Rumus untuk menghitung *mean* tersebut sebagai berikut.

$$\bar{X} = MK + p \left(\frac{\sum fx}{N} \right)$$

Keterangan :

MK = Mean Kerja, yaitu titik tengah kelas yang digunakan sebagai landasan kerja

x = deviasi atau penyimpangan dari MK

p = panjang kelas atau panjang interval

Langkah-langkah perhitungan :

- 1) Tentukan salah satu kelas sebagai landasan kerja (usahakan jangan terlalu di ujung-ujung distribusi), dan tanda kelas sebagai MK.
- 2) Isi kolom x, yaitu kolom penyimpangan dengan angka nol (0) pada baris MK.
- 3) Teruskan mengisi kolom x itu ke atas dan ke bawah. Tanda kelas yang lebih kecil dari x_0 berturut-turut diberi harga $x = -1, -2, -3$, dst; dan tanda kelas yang lebih besar dari x_0 berturut-turut diberi harga $x = +1, +2, +3$, dst.
- 4) Selesaikan pengisian kolom fx dan jumlahkan kolom fx itu. Dengan demikian semua unsure rumus mean kerja telah diketahui.

Data dalam Tabel 4-2 dapat dihitung nilai rata-ratanya menggunakan rumus MK dengan langkah-langkah seperti tersebut hasilnya sebagai berikut.

Tabel 4- 3. Perhitungan dengan metode mean kerja (MK)

Kelas	Frekuensi (f)	x	fX
15-22	3	-3	-9
23-30	7	-2	-14
31-38	9	-1	-9
39-46	13	0	0
47-54	8	+1	8
55-62	7	+2	14
63-70	3	+3	9
Jumlah	50		-1

$$\bar{X} = 42,5 + (8) \frac{-1}{50} = 42,34$$

Kesimpulannya, kedua cara menghitung mean tersebut menghasilkan nilai yang sama.

C. MODUS

Modus adalah suatu titik nilai yang menunjukkan frekuensi yang terbanyak pada suatu distribusi frekuensi. Dengan demikian, ada kemungkinan pada suatu distribusi mempunyai dua modus. Distribusi semacam ini disebut distribusi bimodal.

Contoh : Dari Tabel 4-1 di atas frekuensi yang terbanyak adalah 6, terletak pada nilai 69. jadi modusnya adalah 69.

Jika data telah disusun dalam tabel distribusi frekuensi, modusnya dapat ditentukan dengan rumus sebagai berikut:

$$Mo = Bb + p \left(\frac{b_1}{b_1 + b_2} \right)$$

dengan :

Bb = batas bawah kelas modus, ialah kelas interval dengan frekuensi terbanyak

p = panjang kelas

b₁ = frekuensi kelas modus dikurangi frekuensi kelas interval terdekat sebelumnya

b₂ = frekuensi kelas modus dikurangi frekuensi kelas interval terdekat berikutnya.

Contoh :

Dari Tabel 4-2 distribusi frekuensi dapat dicari modusnya sebagai berikut.

No.	Kelas interval	f
1	15-22	3
2	23-30	7
3	31-38	9
4	39-46	13
5	47-54	8
6	55-62	7
7	63-70	3
		50

1. kelas modal = kelas keempat
2. b = 38,5
3. b₁ = 13 - 9 = 4
4. b₂ = 13 - 8 = 5
5. p = 8

$$Mo = 38,5 + (8) \left(\frac{4}{4 + 5} \right) = 42,05$$

D. MEDIAN

Median adalah suatu titik nilai yang menunjukkan 50% dari frekuensi berada di bagian atasnya dan 50% frekuensi lainnya berada pada bagian bawahnya.

Contoh : Dari data : 4, 12, 5, 7, 8, 10, 10 setelah disusun menurut nilainya menjadi : 4, 5, 7, 8, 10, 10, 12. Data paling tengah bernilai 8. jadi mediannya adalah 8.

Jika jumlah nilainya genap, setelah nilai atau data disusun menurut urutan nilainya medianya sama dengan rata-rata dihitung dua data tengah.

Contoh : Dari data : 12, 7, 8, 14, 16, 19, 10, 8 setelah disusun menurut nilainya menjadi : 7, 8, 8, 10, 12, 14, 16, 19. Data tengahnya ialah 10 dan 12 sehingga medianya sama dengan $\frac{1}{2}(10+12) = 11$.

Untuk data yang disusun dalam daftar distribusi frekuensi bergolong, mediannya dihitung dengan rumus :

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

dengan :

b = batas bawah kelas median, ialah kelas dimana median akan terletak

p = panjang kelas median

n = ukuran sample atau banyak data

f = frekuensi kelas median

F = jumlah semua frekuensi sebelum kelas median

Contoh :

Dari Tabel 4-2 dihitung mediannya sebagai berikut :

No.	Kelas interval	f
1	15-22	3
2	23-30	7
3	31-38	9
4	39-46	13
5	47-54	8
6	55-62	7
7	63-70	3
		50

Setengah dari seluruh data ada 25 buah. Jadi median terletak di kelas interval keempat, karena sampai dengan ini jumlah frekuensi sudah lebih dari 25.

Dari kelas median ini didapat :

b = 38,5; p = 8 dan f = 13.

Adapun $F = 3 + 7 + 9 = 19$, sehingga

$$Me = 38,5 + (8) \left(\frac{25-19}{13} \right) = 42,19$$

Ada 50% dari data yang bernilai paling rendah 42,19 dan setengahnya lagi bernilai paling besar 42,19

Dari data dalam Tabel 4-2 tentang nilai kemampuan membaca pemahaman wacana argumentasi 50 orang siswa telah didapat $\bar{X} = 42,34$; $Mo = 40,02$ dan $Me = 42,19$. kita lihat bahwa harga-harga statistik tersebut berlainan.

LATIHAN

1. Tentukan mean dari data berikut menggunakan rumus tanda kelas dan mean duga.

50 40 30 25 22 32 38 35 38 55

48 35 28 24 25 35 35 30 30 30

45 30 26 20 28 40 30 32 40 20

2. Hitunglah nilai Mo data tersebut.

3. Hitunglah nilai Me data tersebut.

BAB V UKURAN VARIABILITAS

TUJUAN PEMBELAJARAN

Mahasiswa dapat:

- 1) Menjelaskan beberapa ukuran variabilitas data
- 2) Menghitung variabilitas data
- 3) Menafsirkan hasil pengukuran berdasarkan variabilitasnya

SUSUNAN MATERI

- a) Pengantar
- b) Kuartil, Desil, dan Persentil
- c) Range, Deviasi Kuartil, dan Standar Deviasi

A. PENGANTAR

Variabilitas adalah derajat penyebaran nilai-nilai variabel dari suatu tendensi sentral dalam distribusi. Meskipun indeks tendensi sentral dapat membantu kita melukiskan data berdasarkan nilai rata-rata atau ukuran yang khas, indeks ini tidak dapat memberikan gambaran yang menyeluruh tentang sebaran. Nilai mean dari dua sebaran bisa jadi sama, tetapi keragaman (variability) skor kedua sebaran tersebut mungkin sangat berbeda. Secagai contoh, perhatikan sebaran skor berikut.

$$(A) \quad 24 \quad 24 \quad 25 \quad 25 \quad 25 \quad 26 \quad 26 \quad X = \frac{175}{7} = 25$$

$$(B) \quad 16 \quad 19 \quad 22 \quad 25 \quad 28 \quad 30 \quad 35 \quad X = \frac{175}{7} = 25$$

Nilai *mean* dalam kedua sebaran ini sama, yaitu 25, tetapi derajat sebaran skornya (*dispersion*) sangat berbeda. Skor dalam sebaran (A) jelas lebih homogen daripada skor dalam sebaran (B). Jelas di sini diperlukan suatu indeks yang dapat memberikan gambaran berdasarkan keragaman skor. Dalam statistik ada beberapa indeks untuk keperluan ini. Ukuran dispersi yang paling banyak dipakai Range, Rentang Antar Kuartil (RAK), dan Standar Deviasi (SD).

Untuk mempelajari hal ini perlu dipahami terlebih dahulu pengukuran kuartil, desil, dan persentil, yang kesemuanya merupakan pengembangan dari ukuran median.

B. KUARTIL, DESIL, DAN PERSENTIL

1. KUARTIL

Kalau median dapat dikatakan sebagai ukuran *perdua*, maka kuartil dapat dikatakan sebagai ukuran *perempat*, artinya nilai-nilai kuartil akan membagi 4 sama banyak terhadap banyak data. Dengan demikian, kita kenal *kuartil pertama* (K_1), *kuartil kedua* (K_2), *kuartil ketiga* (K_3).

Untuk nilai kuartil tergantung dari $\frac{i}{n}$, ($i = 1, 2, \text{ dan } 3$). Artinya untuk K_1 tergantung dari nilai $\frac{1}{4}n$; untuk nilai K_2 tergantung dari nilai $\frac{2}{4}n$, dan untuk K_3 tergantung dari $\frac{3}{4}n$. Dengan demikian rumus umum untuk kuartil adalah sebagai berikut.

$$K_i = Bb + p \left[\frac{\frac{i}{4}n - F}{f_{K_i}} \right], i = 1, 2, 3.$$

Bb = Batas bawah kelas interval yang mengandung K_i

P = panjang kelas interval

n = banyak data

F = Frekuensi kumulatif sebelum K_i

f_{K_i} = frekuensi kelas interval yang mengandung K_i

2. DESIL DAN PERSENTIL

Jika kumpulan data dibagi menjadi 10 bagian yang sama, maka tiap bagian disebut *persepuluhan* atau disebut *desil*. Artinya, nilai-nilai desil akan membagi 10 sama banyak terhadap banyak data. Dengan demikian, kita kenal desil 1 (D_1), desil dua (D_2), desil tiga (D_3), desil empat (D_4)..... desil sembilan (D_9).

Persentil (P_i) merupakan ukuran lokasi yang paling halus karena pembagiannya 1 s.d 99. Rumus umum untuk desil adalah sebagai berikut.

$$D_i = Bb + p \left[\frac{\frac{i}{10}n - F}{f_{D_i}} \right], i = 1, 2, 3, \dots, 9$$

Bb = Batas bawah kelas interval yang mengandung D_i

P = panjang kelas interval

n = banyak data

F = Frekuensi kumulatif sebelum D_i

f_{Di} = frekuensi kelas interval yang mengandung D_i

Sedangkan untuk persentil rumusnyasebagai berikut.

$$P_i = Bb + p \left[\frac{\frac{i}{10}n - F}{f_p} \right], i = 1, 2, 3, \dots, 99$$

Bb = Batas bawah kelas interval yang mengandung P_i

P = panjang kelas interval

n = banyak data

F = Frekuensi kumulatif sebelum P_i

f_{Di} = frekuensi kelas interval yang mengandung P_i

Contoh:

Pada data yang berdistribusi sebagai berikut

No.	Kelas interval	Frekuensi Absolut	Frekuensi Kumulatif
1	15-22	3	3
2	23-30	7	10
3	31-38	9	19
4	39-46	13	32
5	47-54	8	40
6	55-62	7	47
7	63-70	3	50
		50	

1) Hitung nilai K_3

Jawab:

K_3 terletak pada kelas interval no 5 sebab $\frac{3}{4}n$ paling sedikit harus sama

dengan 37.5

Bb = 46,5

p = 8

$$F = 32$$

$$f_{k3} = 8$$

$$\text{Maka nilai } K_3 = 46,5 + 8 \left[\frac{37,5 - 32}{8} \right] = 52$$

2) Hitung D_4

Jawab:

D_4 terletak pada kelas interval nomor 4 sebab $\frac{4}{10}n$ paling sedikit harus 20.

$$Bb = 38,5$$

$$p = 8$$

$$F = 19$$

$$f_{D4} = 13$$

$$\text{Maka nilai } D_4 = 38,5 + 8 \left[\frac{20 - 19}{13} \right] = 39,11$$

3) Hitung P_{45}

Jawab:

P_{45} terletak pada kelas interval no 4 sebab $\frac{45}{100}n$ paling sedikit harus 22.5

$$Bb = 38,5$$

$$P = 8$$

$$F = 19$$

$$f_{p45} = 13$$

$$\text{Maka } P_{45} = 38,5 + 8 \left[\frac{22,5 - 19}{13} \right] = 40,65$$

C. RANGE, DEVIASI KUARTIL, DAN STANDAR DEVIASI

1) RANGE

Kisaran (range) menggambarkan rentang nilai terkecil hingga terbesar dan dapat dinyatakan sebagai $R = \text{Nilai Maksimum} - \text{Nilai Minimum}$. Nilai ini memberikan gambaran lebar data pengamatan yang penting untuk diketahui saat menggolongkan data dalam kelas berfrekuensi.

Diketahui data tersebar sebagai berikut.

25 37 28 29 35 30 21 31

Maka nilai $R = 37 - 21 = 16$

2) DEVIASI KUARTIL

Nilai deviasi kuartil menunjukkan besaran rerata selisih nilai Kuartil 3 dan Kuartil 1 dan dapat dirumuskan sebagai berikut.

$$DK = \frac{K_3 - K_1}{2}$$

a. Mengurutkan data.

Contoh data di atas, diurutkan

21 25 28 29 30 31 35 37

b. Menghitung nilai K_1

Letak $K_1 = \frac{1}{4}(n + 1) = 2,25$. Artinya, nilai K_1 terletak antara data ke-2 dan data

Ke-3. Besarnya = $25 + 0,25$ (nilai data ke-3 - nilai data ke-2) = $25 + 0,25$ (28-25) = 25,75.

c. Menghitung nilai K_3

Letak $K_3 = \frac{3}{4}(n + 1) = 6,75$. Artinya, nilai K_3 terletak antara data ke-6 dan data

ke-7. Besarnya = $31 + 0,75$ (nilai data ke-7 - nilai data ke-6) = $31 + 3 = 34$.

d. $DK = \frac{34 - 25,75}{2} = 4,125$

3. STANDAR DEVIASI (SD)

Standar Deviasi (SD) dalam beberapa buku ada yang menyebutnya Simpangan Baku (S). Kuadrat dari Simpangan Baku disebut Varian (S^2) dan dapat dirumuskan sebagai berikut. Misalnya, untuk gugus data yang sama untuk contoh sebelumnya dapat dihitung nilai ragamnya yaitu:

$$\begin{aligned}
S^2 &= \frac{\sum (X_i - \bar{X})^2}{n-1} \\
&= \frac{(25 - 29,5)^2 + (37 - 29,5)^2 + (28 - 29,5)^2 + (29 - 29,5)^2}{8-1} \\
&\quad + \frac{(35 - 29,5)^2 + (30 - 29,5)^2 + (21 - 29,5)^2 + (31 - 29,5)^2}{8-1} = \frac{184}{7} = 26,285 \\
S &= \sqrt{26,285} \\
S &= 5,13.
\end{aligned}$$

Rumus Simpangan Baku (S) untuk data berbobot.

Ada kalanya simpangan baku ditentukan dari data yang berkelas atau terbobot. Untuk itu, rumus asal harus disesuaikan menjadi:

$$S = \sqrt{\frac{\sum f(X_i - \bar{X})^2}{n-1}}$$

Atau, dapat dihitung dengan rumus S^2

$$S^2 = \frac{n \sum fx^2 - (\sum fx)^2}{n(n-1)}$$

Atau, dapat dihitung dengan cara coding

$$S = p \sqrt{\frac{\sum fd^2}{n} - \left[\frac{\sum fd}{n} \right]^2}$$

Rumus-rumus tersebut dapat berlaku untuk data tersebar maupun data terkelompok, asal diingat bahwa yang dimaksud dengan X untuk data terkelompok adalah tanda kelas tiap-tiap kelas interval.

Contoh 1

1. Penghitungan menggunakan rumus S

$$S = \sqrt{\frac{\sum f(X_i - \bar{X})^2}{n-1}}$$

x	f	fx	$(x - \bar{x})$	$(x - \bar{x})^2$	$f(x - \bar{x})^2$
35	2	70	6,15	37,82	75,64

30	6	180	1,15	1,32	7,92
29	4	116	0,15	0,0225	0,09
28	2	56	-0,85	0,7225	1,445
27	2	54	-1,85	3,4225	6,845
26	1	26	-2,85	8,225	8,125
25	3	75	-3,85	14,8225	44,47
	20	577			144,555

$$\bar{X} = \frac{577}{20} = 28,85$$

$$S = \sqrt{\frac{\sum f(X_i - \bar{X})^2}{n-1}}$$

$$S = \sqrt{\frac{144,5}{19}} = 2,75$$

2. Penghitungan menggunakan rumus Varian (S²)

$$S^2 = \frac{n \sum fx^2 - (\sum fx)^2}{n(n-1)}$$

x	f	x ²	fx	fx ²
35	2	1225	70	2450
30	6	900	180	5400
29	4	841	116	3364
28	2	784	56	1568
27	2	729	54	1458
26	1	676	26	676
25	3	625	75	1875
Total	20	5780	577	16791

$$S^2 = \frac{n \sum fx^2 - (\sum fx)^2}{n(n-1)}$$

$$S^2 = \frac{20(16791) - (577)^2}{20(19)} = 7,6$$

$$S = \sqrt{7,6} = 2,75$$

Contoh 2

Diketahui 40 data berikut.

63 81 67 85 77 78 57 80 53 74 85 97 62 71 75 95 61 78 83 71
75 68 60 62 87 79 63 93 93 73 84 85 70 90 82 80 76 68 75 75

a. Penghitungan menggunakan rumus S

$$S = \sqrt{\frac{\sum f(X_i - \bar{X})^2}{n-1}}$$

Kelas Interval	Interval Kelas	f	X _i	fX _i	(X _i - \bar{X})	(X _i - \bar{X}) ²	f(X _i - \bar{X}) ²
1	53-59	2	56	112	20,15	406,02	812,045
2	60-66	6	63	378	13,12	172,13	1032,87
3	67-73	7	70	490	6,15	37,82	264,75
4	74-80	12	77	825	0,85	0,72	8,67
5	81-87	8	84	672	7,85	61,62	492,98
6	88-94	3	91	273	14,85	220,52	661,567
7	95-101	2	98	198	21,85	477,42	954,85
Total		40		3046			4227,73

$$\bar{X} = \frac{3046}{40} = 76,13$$

$$S = \sqrt{\frac{\sum f(X_i - \bar{X})^2}{n-1}}$$

$$S = \sqrt{\frac{4227,73}{40-1}}$$

$$= 10,41$$

b. Penghitungan menggunakan rumus Varian (S²)

$$S^2 = \frac{n \sum fx^2 - (\sum fx)^2}{n(n-1)}$$

Kelas Interval	Interval kelas	f	x	x ²	fx	fx ²
1	53-59	2	56	3126	112	6252
2	60-66	6	63	3969	378	23814

3	67-73	7	70	4900	490	34300
4	74-80	12	77	5929	924	71148
5	81-87	8	84	7056	672	56448
6	88-94	3	91	8281	273	24843
7	95-101	2	98	9604	196	19208
Total		40			3045	236013

$$S = \sqrt{\frac{(40)(236013) - (3045)^2}{(40)(39)}}$$

$$S = \sqrt{\frac{(40)(236013) - (3045)^2}{(40)(39)}}$$

$$S = \sqrt{\frac{9440520 - 9272025}{1560}}$$

$$\sqrt{\quad} = 10,41$$

c. Penghitungan dengan Coding

$$S = p \sqrt{\frac{\sum fd^2}{n} - \left[\frac{\sum fd}{n} \right]^2}$$

Kelas Interval	Interval kelas	f	d	d ²	fd	fd ²
1	53-59	2	-3	9	-6	18
2	60-66	6	-2	4	-12	24
3	67-73	7	-1	1	-7	7
4	74-80	12	0	0	0	0
5	81-87	8	1	1	8	8
6	88-94	3	2	4	6	12
7	95-101	2	3	9	6	18
Total		40			-5	87

$$S = p \sqrt{\frac{87}{40} - \left(\frac{-5}{40} \right)^2}$$

$$S = 10,38$$

Latihan

1. Jelaskan arti K1 = 8; D7 = 8; P55 65

2. Hitung nilai K2, K3, D4, D5, P50, P67 data dalam Tabel berikut.

No	Kelas interval	Frekuensi (f)
1	65 - 69	2
2	60 - 64	4
3	55 - 59	5
4	50 - 54	3
5	45 - 49	5
6	40 - 44	9
7	35 - 39	6
8	30 - 34	6
9	25 - 29	5
10	20 - 24	3
11	15 - 19	2
		50

3. Hitung Rentang dari data tersebut.

4. Hitung Rentang Antar Kuartil dari data tersebut.

5. Hitung Simpangan Baku dari data tersebut

- a. Menggunakan Rumus S
- b. Menggunakan Rumus Varian
- c. Menggunakan Rumus Coding.

BAB VI

UJI - T

TUJUAN PEMBELAJARAN

Mahasiswa dapat:

- 1) Melakukan pengujian
- 2) Membandingkan
- 3) Menjelaskan arti

SUSUNAN MATERI

- a) Pengantar
- b) Uji-t untuk sampel bebas
- c) Uji-t untuk dua mean yang berkorelasi

PENGANTAR

Nilai statistik dari dua gugus data seringkali harus dibandingkan untuk mengetahui apakah ada perbedaan di antara mereka. Jika ukuran contoh asal relatif kecil asalkan data aslinya menyebar normal atau diasumsi menyebar normal, maka beda diantara dua gugus data dapat diuji dengan menggunakan Uji-t (Student's test).

A. UJI-T UNTUK DUA KELOMPOK SAMPEL YANG BERLAINAN (INDEPENDENT SAMPLE)

Untuk mengetahui apakah ciri statistik (mean comparison) dari dua kelompok yang diamati berbeda atau mirip dan memenuhi kaidah sampel kecil dan asumsi menyebar normal dapat digunakan persamaan berikut. Ukuran contoh umumnya kurang dari 30.

Rumus:

$$t = \frac{X_1 - X_2}{\sqrt{\left[\frac{\sum X_1^2 - \frac{(\sum X_1)^2}{N_1} + \sum X_2^2 - \frac{(\sum X_2)^2}{N_2}}{(N_1 + N_2) - 2} \right] \left[\frac{1}{N_1} + \frac{1}{N_2} \right]}}$$

Contoh Kasus:

Seorang peneliti ingin mengetahui perbedaan motivasi kerja antara pegawai pria dan wanita pada suatu kantor. Dengan menggunakan angket (atau instrumen lain) peneliti mengumpulkan data dan hasilnya sebagai berikut:

Pegawai pria		Pegawai Wanita	
1	107	1	100
2	96	2	94
3	88	3	127
4	88	4	76
5	109	5	115
6	84	6	121
7	79	7	87
8	105	8	92
9	108	9	91
10	92	10	98
11	96	11	104
12	96	12	96
		13	110
		14	108

Langkah-langkah perhitungan :

X ₁	X ₁ ²	X ₂	X ₂ ²
107	-	109	
96	-	94	
88	-	127	
131	-	76	
109	-	115	
84	-	121	
79	Dst.	87	
105		92	
108		91	
92		98	
96		104	
101		96	
		110	
		108	
1196	121.318	1428	148.202

Langkah 1 Jumlahkan skor masing-masing kelompok pria (X₁) dan wanita (X₂).

Langkah 2 Kuadratkan setiap skor pada masing-masing kelompok dan kemudian jumlahkan hasil pengkuadratan tersebut pada masing-masing kelompok (hasilnya pada lajur 2 dan 4)

Langkah 3 Kuadratkan semua skor kelompok pria, dan demikian juga jumlah skor kelompok wanita, kemudian hasilnya dibagi n pada kelompok masing-masing.

Kelompok Pria (X_1)

$$\frac{(\sum X_1)^2}{n_1} = \frac{(1196)^2}{12} = \frac{1.430.416}{12} = 119.201$$

Kelompok Wanita (X_2)

$$\frac{(\sum X_2)^2}{n_2} = \frac{(1428)^2}{14} = \frac{1.039.416}{14} = 145.656$$

Langkah 4 Kurangkan hasil langkah ke 3 pada hasil langkah ke 2, sesuai pengurangan pada kedua kelompok masing-masing, kemudian hasil pengurangan pada kedua kelompok tersebut dijumlahkan.

Kelompok Pria (X_1)

$$121.318 - 119.201 = 2117$$

Kelompok Wanita (X_2)

$$148.202 - 145.656 = \underline{2546}$$

$$\text{Jumlah} = 4663$$

Langkah 5 Bagilah hasil pada langkah 4 dengan $(N_1 + N_2) - 2$
 $= 12 + 14 - 2 = 24$

$$\frac{4663}{24} = 194$$

Langkah 6 Hasil dari langkah ke 5 dikalikan dengan $\frac{1}{12} + \frac{1}{14}$

$$= 194 \times \frac{13}{84} = 30$$

$$\sqrt{30} = 5,48$$

Langkah 7 Hitunglah skor rata-rata untuk masing-masing kelompok dengan cara membagi jumlah skor kelompok dengan n kelompok.

$$\text{Kelompok pria : } Mean = \frac{1196}{12} = 99,67(\bar{X}_1)$$

$$\text{Kelompok wanita : } Mean = \frac{1428}{14} = 102(\bar{X}_2)$$

Langkah 8 Menghitung harga t; dengan jalan mengurangkan mean kelompok pria dari mean kelompok wanita, kemudian hasilnya dibagi dengan hasil langkah ke 6.

$$t = \frac{102 - 99,67}{5,48} = \frac{2,33}{5,48} = 0,43$$

Langkah 9 Konsultasikan hasilnya dengan nilai pada tabel t dengan ketentuan:

$$df = n_1 + n_2 - 2 = 12 + 14 - 2 = 24$$

$$ts = 5\% \text{ atau } 0,05.$$

Dari tabel t diperoleh harga t dengan $df = 24$ dan $ts = 0,05$, $t = 2,064$.

Karena harga t hitung lebih kecil dari harga t tabel, kesimpulannya tidak ada perbedaan motivasi kerja antara pegawai pria dengan pegawai wanita (dalam kasus ini).

Catatan:

1. Untuk uji t jenis ini, jumlah anggota sampel pada dua kelompok tidak perlu sama.
2. Derajat kebebasan untuk menentukan signifikan atau tidaknya harga t adalah $n_1 + n_2 - 2$

B. UJI T UNTUK DUA MEAN YANG BERKORELASI

Adakalanya pengujian dilakukan terhadap data **berpasangan**. Misalnya dalam evaluasi hasil belajar kita membandingkan efektivitas suatu metode pengajaran dengan melihat hasil **pre-test dan post-test**. Untuk pengujian seperti ini dapat digunakan persamaan berikut.

Rumus:

$$t = \frac{X_1 - X_2}{\sqrt{\frac{\sum D^2 - \frac{(\sum D)^2}{N}}{N(N-1)}}}, \text{ dengan } D = X_2 - X_1$$

Kasus 1:

Seorang peneliti tertarik untuk mengetahui efek penyuluhan terhadap cara hidup sehat para ibu di pedesaan. Mula-mula peneliti mengukur/mengambil data tentang cara ibu-ibu mencuci, memasak, mengatur rumah tangga, dan sebagainya, yang terkait dengan kebersihan. Kemudian peneliti memberi penyuluhan tentang segala sesuatu yang berkaitan dengan cara hidup sehat. Beberapa minggu kemudian peneliti kembali melakukan pengukuran/mengambil data tentang cara-cara ibu-ibu hidup sehat seperti sebelum penyuluhan. Hasil pengukuran sebelum dan sesudah penyuluhan dibandingkan untuk mengetahui ada/tidaknya perbedaan cara hidup mereka.

Data:Sebelum Perlakuan				Data : Sesudah Perlakuan			
1	89	8	93	1	94	8	96
2	86	9	87	2	94	9	90
3	96	10	89	3	101	10	88
4	100	11	110	4	105	11	115
5	94	12	95	5	100	12	100
6	86	13	107	6	84	13	110
7	81	14	96	7	81	14	102

Langkah-langkah penghitungan

Langkah 1 Kurangkan skor sebelum perlakuan pada skor setelah perlakuan

1	$94 - 89 =$	5	8	$96 - 93 =$	3
2	$94 - 86 =$	8	9	$90 - 97 =$	3
3	$101 - 96 =$	5	10	$88 - 89 =$	-1
4	$105 - 100 =$	5	11	$115 - 110 =$	5
5	$100 - 94 =$	6	12	$100 - 95 =$	5
6	$84 - 86 =$	-1	13	$110 - 107 =$	3
7	$81 - 81 =$	0	14	$102 - 96 =$	6

Langkah 2 Kuadratkan hasil pengurangan masing-masing skor pada langkah 1 dan jumlahkan hasil pengkuadratannya.

$$(5)^2 + (8)^2 + \dots + (-2)^2 + \dots (-1)^2 + \dots (6)^2 = 293$$

Langkah 3 Jumlahkan hasil pengurangan tiap pasangan pada langkah 1, kemudian kuadratkan jumlah tersebut, dan hasilnya bagilah dengan n.

$$(5 + 8 + \dots 6 - 2 \dots -1 \dots + 6 = 51$$

$$\frac{(51)^2}{14} = \frac{2001}{14} = 186$$

Langkah 4 Kurangkan hasil langkah 3 pada hasil langkah 2, kemudian bagilah dengan n-1.

$$\frac{293 - 186}{14 - 1} = \frac{107}{13} = 8,23$$

Langkah 5 Tariklah akar dari hasil pada langkah 4, kemudian bagilah hasilnya dengan \sqrt{n} .

$$\sqrt{8,23} = 2,87$$

$$\frac{2,87}{\sqrt{14}} = \frac{2,87}{3,74} = 0,767$$

Langkah 6 Hitunglah skor rata-rata, masing-masing dari kelompok skor sebelum perlakuan dan setelah perlakuan.

$$I. \quad 89 + 86 + \dots + 96 = 1309 ; \text{mean} = \frac{1309}{14} = 93,50$$

$$II. \quad 94 + 94 + \dots + 102 = 1360 ; \text{mean} = \frac{1360}{14} = 97,14$$

Langkah 7 Hitunglah harga t dengan cara menghitung selisih mean I dan mean II kemudian dibagi hasil pada langkah ke 5.

$$t = \frac{97,14 - 93,50}{0,767} = \frac{3,64}{0,767} = 4,74$$

Langkah 8 Konsultasikan harga t pada langkah 7 dengan tabel t, dengan derajat kebebasan n-1 (dalam hal ini n= jumlah pasangan; yaitu 14). Dari tabel t diperoleh bahwa harga t dengan taraf signifikan 5% pada derajat bebas 13, adalah 2,16. Karena harga t hitung lebih besar dari t tabel, maka kesimpulannya penyuluhan dapat meningkatkan cara hidup sehat.

Kasus 2:

t-test jenis di atas dapat juga diterapkan pada dua kelompok subyek yang berlainan yang dipasangkan (match) atas dasar pertimbangan rasional seperti usia, tingkat

pendidikan, ras, seks, dsb. untuk menyamakan antara kelompok yang satu dengan yang lain. Penugasan setiap individu pada masing-masing kelompok harus secara acak, sedangkan pasangan harus ditentukan sebelum perlakuan dimulai.

Analisis untuk rancangan penelitian jenis ini dapat dilakukan dengan cara seperti pada kasus I di atas.

LATIHAN

1. Seorang peneliti ingin mengetahui perbedaan antara hasil belajar siswa yang berintelegensi tinggi dan yang berintelegensi rendah dalam belajar Membaca Pemahaman dengan menggunakan pertanyaan bacaan bentuk objektif. Dengan menggunakan tes peneliti mengumpulkan data, hasilnya sebagai berikut:

Siswa dengan IQ tinggi:

17 17 18 14 17 14 14 17 10 13 12 15 15 9 13 14 16 17 18 17
16 10 9 12 13 14 12 10 9 15 14 12 11 13 11 10 12 11 16 12

Siswa dengan IQ rendah:

16 14 14 14 15 12 13 12 13 12 14 15 13 12 11 10 9 9 8 10 9
10 9 11 10 9 9 8 8 9 10 10 9 9 10 11

Analisislah data tersebut untuk menguji apakah ada perbedaan yang signifikan antara hasil belajar siswa yang berintelegensi tinggi dan yang berintelegensi rendah dalam belajar MP dengan menggunakan pertanyaan bacaan bentuk objektif.

2. Seorang peneliti melaksanakan eksperimen untuk menguji efektivitas latihan bentuk meringkas dalam meningkatkan hasil belajar MP di SD. Eksperimen diawali dengan pretest untuk mengetahui kemampuan awal siswa. Setelah 6 x beturut-turut pembelajaran membaca pemahaman dilaksanakan dengan melatih

siswa meringkas isi bacaan, posttest diselenggarakan. Hasil pretest dan posttest sbb:

Skor Pretest:

30 35 29 29 30 40 40 38 37 38 35 40 40 30 39 40 34 36 28 30
26 30 38 38 35 29 34 38 35 34 32 3 4 26 27 34 45 34 34 35 29

Skor Posttest:

43 38 39 45 46 50 52 56 38 45 46 56 54 53 56 48 49 56 58 39

39 42 48 54 56 58 58 60 45 46 56 54 54 48 56 56 58 60 45 56

Analisislah data tersebut, buatlah kesimpulan tentang efektivitas bentuk meringkas

dalam meningkatkan hasil belajar MP.

BAB VII

ANALISIS VARIAN

TUJUAN PEMBELAJARAN

Mahasiswa dapat:

- 1) Melakukan pengujian
- 2) Membandingkan
- 3) Menjelaskan arti

SUSUNAN MATERI

- a) Pengantar
- b) Anava satu jalur
- c) Uji-t untuk dua mean yang berkorelasi

PENGANTAR

Uji-t seperti dicontohkan di muka hanya sesuai untuk mengetahui perbedaan antara dua kelompok data. Untuk menentukan perbedaan yang melibatkan lebih dari dua kelompok secara simultan digunakan cara lain, yaitu analisis varian, karena penerapan uji-t akan terlalu banyak perhitungan yang harus dilakukan. Uji-t bukan untuk rancangan yang demikian. Ambil contoh. Peneliti ingin membandingkan 4 kelompok data yang diambil dari kelompok sampel. Jika peneliti menerapkan uji-t untuk mengetahui ada tidaknya perbedaan pada keempat kelompok tersebut, peneliti harus melakukan

$$\frac{4(4-1)}{2} = 6x\text{perhitungan}$$

Dengan analisis varian perhitungan cukup singkat dan oleh banyak ahli dianggap lebih sesuai dan akurat.

A. ANAVA SATU JALUR

Contoh Kasus:

Seorang peneliti bermaksud mengetahui efek suatu metode terhadap hasil. Subjek penelitian diambil secara acak untuk ditugaskan ke dalam empat kelompok eksperimen, masing-masing dengan menggunakan metoda A, B, C, dan D. Setelah eksperimen dilakukan data diambil dari masing-masing subjek, dan hasilnya sbb.

Metode A	Metode B	Metode C	Metode D
10	3	19	23
7	8	12	14
9	7	16	16
8	5	14	18
15	6	7	12
3	10	8	13
8	12	14	16
9	4	10	17
11	7	19	19
9	6	9	14
5	5	15	16
17	15	14	17

Yang ingin diketahui peneliti adalah:

Apakah terdapat perbedaan hasil belajar antara siswa yang diajar dengan metode A, B, C, dan D?

Perumusan Hipotesis

H_0 : Tidak terdapat perbedaan hasil belajar siswa yang diajar dengan metode A, B, C, dan D.

H_1 : Paling sedikit salah satu rata-rata hasil belajar siswa yang dilatih dengan metode A, B, C, dan D tidak sama.

Langkah-langkah dalam perhitungan :

1. Jumlahkan skor pada masing-masing kelompok (Metode)

Kelompok1	Kelompok2	Kelompok3	Kelompok4
10	3	19	23
7	8	12	14
-	-	-	-
-	-	-	-
-	-	-	-
T ₁ 111	T ₂ 88	T ₃ 156	T ₄ 195

2. Kuadratkan setiap skor, kemudian jumlahkan hasil pengkuadratan ini secara keseluruhan untuk empat kelompok.

$$10^2 + 7^2 + \dots + 17^2 + 3^2 + 8^2 + \dots + 15^2 + 19^2 + 12^2 + \dots + 14^2 + 23^2 + 14^2 + \dots + 17^2 = 7434$$

3. Jumlahkan kelompok jumlah skor dari tiap-tiap kelompok menjadi jumlah total (grand total)

$$111 + 88 + 156 + 159 = 550$$

4. Kuadratkan hasil jumlah grand total tersebut kemudian dibagi dengan n total

$$\frac{550^2}{48} = 6302 \text{ (faktor koreksi)}$$

5. Kurangkan hasil langkah 4 pada hasil langkah 2 : $7434 - 6302 = 1132$

(SSt: Sum Square total)

6. Kuadratkan jumlah skor tiap kelompok pada langka 1, kemudian masing-masing jumlah dibagi dengan n pada masing-masing kelompok, dan jumlahkan hasil tersebut dari semua (keempat) kelompok.

$$\frac{111^2}{12} + \frac{88^2}{12} + \frac{156^2}{12} + \frac{195^2}{12} = 6869$$

7. Kurangkan hasil langkah 4 pada langkah 6, dan hasilnya adalah Sum of Squares between (SSb)

$$6869 - 6302 = 567$$

8. Kuarangkan hasil pada langkah 7 pada hasil langkah 5, dan hasilnya merupakan Sum of Square within groups (SSw)

$$1132 - 567 = 565$$

9. Menghitung *Means Square*. Sebagai komponennya df (degree of freedom) untuk SSt, SSb, dan SSw perlu dihitung lebih dulu.

$$\text{df untuk SSt} = 48 - 1 = 47$$

$$\text{df untuk SSb} = 4 - 1 = 3$$

$$\text{df untuk SSw} = 47 - 3 = 44 \text{ (df SSt - df SSb).}$$

maka :

$$M .St = \frac{SSt}{df} \rightarrow \text{nilai ini tidak digunakan}$$

$$M .Sb = \frac{SSb}{dfb} = \frac{567}{3} = 189$$

$$M.Sw = \frac{SSw}{dfw} = \frac{565}{44} = 12,84$$

10. Menghitung harga F

$$F = \frac{M.Sb}{M.Sw} = \frac{189}{12,84} = 14,71$$

Tabel Ringkasan Anava

Source	Df	SS	Ms	F	P
Total	47	1132	-	-	-
Between groups	3	567	189	14,71	-
Within groups	44	565	12,84	-	-

11. Konsultasikan dengan nilai pada tabel F.

Harga F yang diperoleh dikonsultasikan dengan tabel F dengan df 3 dan 44 untuk ts 5 % adalah 2,84. Jadi harga F yang diperoleh dari perhitungan lebih tinggi, maka kita menolak hipotesis nol. Dengan kata lain, terdapat perbedaan yang signifikan pada keempat kelompok tersebut sebagai efek dari penerapan metode.

ANALISIS LANJUTAN (POST-HOC)

Hasil perhitungan tersebut di atas barulah sampai informasi bahwa terdapat perbedaan yang signifikan pada keempat kelompok eksperimen tersebut. Pertanyaan selanjutnya, di mana letak perbedaan itu? Untuk mengetahui letak perbedaan, perlu dilakukan analisis *post-hoc*.

Berikut ini langkah-langkahnya : **(Duncan's Multiple Range)**

1. Hitung mean untuk masing-masing kelompok.

$$\text{Mean kelompok 1} = \frac{111}{12} = 9,25$$

$$\text{Mean kelompok 2} = \frac{88}{12} = 7,33$$

$$\text{Mean kelompok 3} = \frac{156}{12} = 13,00$$

$$\text{Mean kelompok 4} = \frac{195}{12} = 16,25$$

2. Hitung standar kesalahan dari mean dengan rumus

$$MSe = \sqrt{\frac{MSw}{n(\text{perkelompok})}} = \sqrt{\frac{12,84}{12}} = \sqrt{1,07} = 1,03$$

Catatan : Jika jumlah subyek (n) untuk masing-masing kelompok tidak sama satu sama lain, khususnya jika n terbesar tidak lebih dari dua kali (2x) n terkecil, dapat digunakan = (n rata-rata) (Bruning, hal. 15)

Misalnya : $n_1=12, n_2=8, n_3=10, n_4=6$.

$$n = \frac{4}{\frac{1}{12} + \frac{1}{8} + \frac{1}{10} + \frac{1}{6}} = 8,42$$

3. Hasil langkah 2 konsultasikan dengan Tabel Duncan Multiple range. Dengan mengambil $\alpha = 0,05$, dengan $df = 44$, (df untuk MSw) diperoleh tiga angka berturut-turut, selanjutnya disebut k (*significant studentized range*)

$$k = 2 : 2,858$$

$$k = 3 : 3,006$$

$$k = 4 : 3,102$$

4. Tentukan harga kritik setiap pasangan dengan mengalikan harga k dengan standar kesalahan mean pada langkah 2.

$$k = 2 : 2,858 \times 1,03 = 2,944$$

$$k = 3 : 3,006 \times 1,03 = 3,096$$

$$k = 4 : 3,102 \times 1,03 = 3,195$$

5. Buatlah urutan mean keempat kelompok dari yang terkecil sampai yang terbesar, hasilnya :

$$7,33 \rightarrow \text{Kelompok 2}$$

$$9,25 \rightarrow \text{Kelompok 1}$$

$$13,00 \rightarrow \text{Kelompok 3}$$

$$16,25 \rightarrow \text{Kelompok 4}$$

6. Bandingkan perbedaan mean setiap pasangan dengan batas signifikan yang dihitung pada langkah 4.

Caranya :

a. Kel. 2 Vs Kel. 4 (batas sig = 3,195) $\rightarrow 16,26 - 7,33 = 8,92$ (signifikan)

b. Kel. 2 Vs Kel. 3 (batas sig = 3,096) $\rightarrow 13,00 - 7,33 = 5,67$ (signifikan)

c. Kel. 2 Vs Kel. 1 (batas sig = 2,944) $\rightarrow 9,25 - 7,33 = 1,92$ (**tidak** signifikan)

d. Kel. 1 Vs Kel. 4 (batas sig = 3,195) $\rightarrow 16,25 - 9,25 = 7,00$ (signifikan)

e. Kel. 1 Vs Kel. 3 (batas sig = 2,096) $\rightarrow 13,00 - 9,25 = 3,75$ (signifikan)

f. Kel. 3 Vs Kel. 4 (batas sig = 3,195) $\rightarrow 16,25 - 13,00 = 3,25$ (signifikan)

Cara kedua dengan menggunakan T - Metode (Tukey's = hsd = honest significant difference)

$$HSD = q \sqrt{\frac{MS_w}{n}}$$

Tulis kembali skor rata-rata tiap kelompok :

Kel 1; X = 9,25

Kel 2; X = 7,33

Kel 3; X = 13,00

Kel 4; X = 16,25

MSw = 12,84, dfw = 44, k = 4 (lihat tabel ringkasan).

Langkah 1 : Hitunglah selisih skor rata-rata untuk setiap pasangan,

	X ₁	X ₂	X ₃
X ₂	1,92	-	-
X ₃	3,75	5,67	-
X ₄	7	8,92	3,25

Langkah 2 : Hitunglah hsd (*honest significant difference*).

Misalnya diambil ts = 0,5, dari *Tabel Studentized Range Statistic* dengan dfw = 44 dan k = 4. diperoleh (dalam tabel diambil dfw terdekat, yaitu 40):

$$q_{0,5} = 3,79$$

$$\begin{aligned} hsd &= q \sqrt{\frac{MS_w}{n}} \\ &= 3,79 \sqrt{\frac{12,84}{12}} = 3,92 \end{aligned}$$

Langkah 3 : Pengambilan kesimpulan

Setiap nilai pada tabel langkah 1 yang sama atau melebihi harga hsd pada langkah 2, dinyatakan perbedaannya signifikan.

Dari Tabel di atas dapat disimpulkan bahwa: **Tampaknya ada yang tidak tepat dalam kesimpulan ini.**

$$X_1 = X_2$$

$$X_1 = X_3$$

$$X_1 = X_4$$

$$X_2 = X_3$$

$$X_2 = X_4$$

$$X_3 = X_4$$

Rata-rata kelompok paling tinggi adalah kelompok 4 maka metode 4 adalah metode yang paling tepat dibandingkan dengan ketiga metode lainnya (1, 2, dan 3). **Metode 4 tidak berbeda nyata dengan metode 3.**

B. ANAVA DUA ARAH

Apabila design yang dikembangkan untuk mencari ada tidaknya perbedaan dari 2 variabel bebas, dan masing-masing variabel bebas dibagi dalam beberapa kelompok, maka design yang dikembangkan tersebut sering disebut dengan *two factorial design*. Dalam kasus ini peneliti akan menghadapi kelompok sebanyak hasil kali banyak kelompok variabel bebas pertama dan banyak kelompok variabel bebas kedua.

Sebagai ilustrasi, perhatikan contoh design factorial berikut:

		METODE MENGAJAR		
		A	B	C
Jenis kelamin	Laki-laki			
	Perempuan			

Hipotesis dalam Anava 2 arah (2 variabel bebas) terdiri dari (untuk ilustrasi di atas):

1) Yang berkaitan dengan pengaruh faktor A atau efek baris.

H_0 : Tidak terdapat perbedaan hasil belajar siswa antara laki-laki dan perempuan

H_1 : Terdapat perbedaan hasil belajar siswa antara laki-laki dan perempuan.

2) Yang berkaitan dengan pengaruh faktor kedua (B), atau efek kolom.

H_0 : Tidak terdapat perbedaan hasil belajar siswa yang diajar dengan metode A, B, dan C

H_1 : Paling sedikit salah satu rata-rata hasil belajar siswa yang dilatih dengan metode A, B, dan C tidak sama.

3) Interaksi A x B

H_0 : Efek faktor jenis kelamin tidak tergantung pada faktor metode mengajar atau efek metode mengajar tidak tergantung pada faktor jenis kelamin terhadap hasil belajar.

H_1 : Efek faktor jenis kelamin tergantung pada faktor metode mengajar atau efek metode mengajar tergantung pada faktor jenis kelamin terhadap hasil belajar.

Contoh Kasus:

Suatu eksperimen metode mengajar yang terdiri dari tiga macam metode (A, B, dan C) diterapkan untuk siswa SLTP dengan memperhatikan intelegensi siswa. Dari hasil tes setelah eksperimen selesai, penyebaran skornya sbb.:

		METODE MENGAJAR			
		A (B ₁)	B (B ₂)	C (B ₃)	
Intelegensi	Rendah (A ₁)	40	60	60	
		30	70	75	
		50	70	75	
		70	65	85	
		50	50	90	
			A ₁ B ₁ = 240	A ₁ B ₂ = 315	A ₁ B ₃ = 385
	Tinggi (A ₂)	50	45	55	
		60	75	80	
		75	80	90	
		65	90	95	
60		70	80		
		A ₂ B ₁ = 310	A ₂ B ₂ = 360	A ₂ B ₃ = 400	

Yang ingin diketahui peneliti adalah:

1. Apakah intelegensi (tinggi rendah) mempunyai efek terhadap hasil belajar yang berbeda?
2. Apakah metode mengajar (A, B, dan C) mempunyai efek terhadap hasil belajar yang berbeda?
3. Apakah intelegensi (tinggi rendah) berinteraksi dengan metode mengajar (A, B, dan C)?

Perumusan Hipotesis:

a. Pengaruh faktor intelegensi

H_0 : Tidak terdapat perbedaan hasil belajar siswa yang berintelegensi tinggi dan rendah.

H_1 : Terdapat perbedaan antara hasil belajar siswa yang berintelegensi tinggi dan rendah

b. Pengaruh faktor metode mengajar

H_0 : Tidak terdapat perbedaan hasil belajar siswa yang diajar dengan metode A, B, dan C

H_1 : Paling sedikit salah satu rata-rata hasil belajar siswa yang dilatih dengan metode A, B, dan C tidak sama.

c. Interaksi intelegensi dan metode mengajar

H_0 : Efek faktor intelegensi tidak tergantung pada faktor metode mengajar atau efek metode mengajar tidak tergantung pada faktor intelegensi terhadap hasil belajar..

H_1 : Efek faktor intelegensi tergantung pada faktor metode mengajar atau efek metode mengajar tergantung pada faktor intelegensi terhadap hasil belajar.

Langkah-langkah Penghitungan:

1. Inventarisasi hal yang diketahui:

$$A_1 = 940 \quad A_2 = 1070 \quad (A = \text{Jumlah skor faktor A})$$

$$B_1 = 550 \quad B_2 = 675 \quad B_3 = 785 \quad (\text{Jumlah skor faktor B})$$

$$G = 2010 \quad (\text{skor total})$$

$$\sum X^2 = 142250$$

$$p = 2 \quad (\text{banyak kelompok pada faktor A})$$

$$q = 3 \quad (\text{banyak kelompok pada faktor B})$$

$$n = 5 \quad (\text{banyak sampel masing-masing sel})$$

$$N = 30 \quad (\text{banyak sampel keseluruhan})$$

2. Penghitungan derajat kebebasan

$$dk \text{ SS}_t = N - 1$$

$$= 30 - 1$$

$$= 29$$

$$dk \text{ SS}_b = pq - 1$$

$$= (2 \times 3) - 1$$

$$= 5$$

$$\text{dk } SS_w = N - pq$$

$$= 30 - (2 \times 3)$$

$$= 24$$

$$\text{dk } SS_A = p - 1$$

$$= 2 - 1$$

$$= 1$$

$$\text{dk } SS_B = q - 1$$

$$= 3 - 1$$

$$= 2$$

$$\text{dk } SS_{AB} = \text{dk } SS_A \times \text{dk } SS_B$$

$$= 1 \times 2$$

$$= 2$$

3. Penghitungan Sum Squares (SS):

$$SS_t = \sum X^2 - \frac{G^2}{N}$$

$$= 142250 - \frac{2010^2}{30}$$

$$= 7580$$

$$SS_b = \sum \frac{AB^2}{n} - \frac{G^2}{N}$$

$$= \frac{240^2}{5} + \frac{310^2}{5} + \frac{315^2}{5} + \frac{365^2}{5} + \frac{385^2}{5} + \frac{400^2}{5} - \frac{2015^2}{30}$$

$$= 138875 - 13467$$

$$= 4205$$

$$SS_w = SS_t - SS_b$$

$$= 7580 - 4205$$

$$= 3375$$

$$SS_A = \sum \frac{A^2}{qn} - \frac{G^2}{N}$$

$$= \frac{940^2}{3 \times 4} + \frac{1070^2}{3 \times 5} - \frac{2010^2}{30}$$

$$= 135233,33 - 134670$$

$$= 563,33$$

$$\begin{aligned}SS_B &= \sum \frac{B^2}{pn} - \frac{G^2}{N} \\&= \frac{550^2}{2 \times 5} + \frac{675^2}{2 \times 5} + \frac{785^2}{2 \times 5} - \frac{2010^2}{30} \\&= 137435 - 134670 \\&= 2765\end{aligned}$$

$$\begin{aligned}SS_{AB} &= SS_b - SS_A - SS_B \\&= 4205 - 563,33 - 2765 \\&= 875,67\end{aligned}$$

4. Penghitungan Mean Squares (MS)

$$\begin{aligned}MS_w &= \frac{SS_w}{dkSS_w} \\&= 3375 : 24 \\&= 140,625\end{aligned}$$

$$\begin{aligned}MS_A &= \frac{SS_A}{dkSS_A} \\&= 563,33 : 1 \\&= 563,33\end{aligned}$$

$$\begin{aligned}MS_B &= \frac{SS_B}{dkSS_B} \\&= 2765 : 2 \\&= 1382,5\end{aligned}$$

$$\begin{aligned}MS_{AB} &= \frac{SS_{AB}}{dkSS_{AB}} \\&= 875,67 : 2 \\&= 438,335\end{aligned}$$

5. Penghitungan F ratio

a. Faktor Tingkat Intelegensi (faktor A)

$$\begin{aligned}F_A &= \frac{MS_A}{MS_w} \\&= 563,33 : 140,625\end{aligned}$$

$$= 4,00590222 \text{ (dibulatkan menjadi 4,00)} < F_{0,05 (1, 24)} = 4,26$$

b. Faktor Metode Mengajar (faktor B)

$$F_B = \frac{MS_B}{MS_w}$$

$$= 1382,5 : 140,625$$

$$= 9,831111111 \text{ (dibulatkan menjadi 9,83)} > F_{0,05 (2,24)} = 3,40$$

c. Interaksi Faktor Intelegensi dan Metode Mengajar (interaksi A x B)

$$F_{AB} = \frac{MS_{AB}}{MS_w}$$

$$= 438,335 : 140,625$$

$$= 3,117048889 \text{ (dibulatkan menjadi 3,12)} < F_{0,05 (2, 24)} = 3,40.$$

6. Tabel ANOVA.

Untuk lebih jelasnya penghitungan-penghitungan di atas dapat disatukan dalam satu tabel ANOVA (jika menggunakan komputer merupakan print out analisis of variance) sbb.

Sumber variance	dk	SS	MS	F
Baris (A)	1	563,33	563,33	4,00ns
Kolom (B)	2	2765	1382,5	9,83*
Interaksi (AB)	2	876,67	438,335	3,12ns
Dalam sel	24	3375	140.625	
	29	7580		

7. Pengambilan keputusan

- 1) Untuk faktor A, kita menerima hipotesis nol: tidak terdapat perbedaan hasil belajar antara siswa yang mempunyai intelegensi tinggi dan rendah.
Ini berarti bahwa intelegensi tidak mempunyai peranan yang cukup signifikan terhadap hasil belajar sehingga perbedaan hasil belajar siswa yang berintelegensi tinggi tidak berbeda dengan hasil belajar siswa yang berintelegensi rendah.
- 2) Untuk faktor B, kita menolak hipotesis nol: paling tidak salah satu rata-rata hasil belajar siswa yang diajar dengan metode berbeda akan berbeda dengan yang lainnya. Ini berarti bahwa dari ketiga metode mengajar paling tidak salah satu

mempunyai efek yang berbeda dengan yang lainnya. Sampai pada tahap ini kita belum memperoleh informasi yang jelas tentang metode mana yang benar-benar mempunyai efek berbeda dengan yang lainnya. Untuk mempengaruhi secara pasti rata-rata mana yang berbeda dengan yang lainnya perlu penghitungan pasca-ANOVA (ingat pertemuan lalu)

- 3) Untuk interaksi A x B, kita menerima hipotesis nol: efek faktor metode mengajar terhadap hasil belajar tidak tergantung pada faktor intelegensi. Ini berarti bahwa kombinasi jenis metode mengajar dengan tingkat intelegensi mana saja tidak mengakibatkan perbedaan hasil belajar.

RINGKASAN

ANOVA merupakan analisis statistik yang dapat memberikan informasi tentang perbedaan antarkelompok satu dengan kelompok lain dalam satu populasi maupun antarpopulasi. ANOVA mengandung kesalahan yang lebih kecil dan lebih efisien daripada pengujian perbedaan dengan t-test.

Jenis ANOVA ada dua, yaitu ANOVA satu arah dan ANOVA dua arah. Penghitungan ANOVA didasarkan pada varian walaupun tujuannya menguji perbedaan rata-rata. Variabilitas dalam ANOVA dikelompokkan menjadi 3 bagian, yaitu:

- (1) Variabilitas antarkelompok (SSb) dengan $dk = k - 1$ atau $(N - 1) - (N - k)$
- (2) Variabilitas dalam kelompok (SSw) dengan $dk = N - k$
- (3) Variabilitas total (SSt) dengan $dk = N - 1$.

Pengujian signifikansi perbedaan dalam ANOVA dengan F tes. ANOVA satu arah dengan jumlah sampel per sel tidak sama, analisisnya tidak berbeda dengan jumlah sampel yang sama tiap sel, asal jumlah sampel cukup besar dan perbedaan jumlah sampel antar sel tidak mencolok.

Signifikansi perbedaan harus dilanjutkan dengan analisis lanjutan, di antaranya dengan Tukey's HSD untuk memperoleh informasi tentang kelompok mana yang berbeda dengan kelompok lainnya.

ANOVA dua arah digunakan untuk mengatasi perbedaan nilai variabel terikat yang dikategorikan berdasarkan variabel bebas yang banyak dan masing-masing variabel terdiri dari beberapa kelompok.

Hipotesis ANOVA dua arah terdiri dari:

- (1) Perbedaan yang dipengaruhi oleh variabel bebas (yang jumlahnya tergantung jumlah variabel bebas)
- (2) Interaksi antarvariabel bebas.

LATIHAN

1. Sebuah penelitian terhadap kemampuan membaca pemahaman siswa ditinjau dari perbedaan ragam bacaan, diperoleh data sebagai berikut :

NARASI	DESKRIPSI	EKSPOSISI	ARGUMENTASI
70	80	75	65
60	85	80	70
75	75	85	75
65	90	80	60
65	85	70	60
90	75	75	75
85	85	70	75
90	95	60	70
75	90	75	65
70	100	95	85

- a. Apakah ada perbedaan yang signifikan antara kemampuan membaca pemahaman siswa ditinjau dari ragam bacaan yang dibaca? (Anava 1 jalur dg Uji F)
 - b. Lanjutkan ANOVA pada soal no 1 di atas sampai Saudara dapat mengambil kesimpulan yang lebih terinci. Kemampuan membaca ragam bacaan manakah yang nyata-nyata berbeda dengan kemampuan membaca ragam bacaan yang lain? (Tukey HSD)
2. Suatu eksperimen dari sebuah penelitian tentang dampak metode mengajar ditinjau dari banyak siswa dalam satu kelas terhadap hasil belajar, menghasilkan data sebagai berikut.

Metode A				Metode B				Metode C			
<20		>30		<20		>30		<20		>30	
80	70	75	70	75	60	65	60	90	80	80	75
70	70	85	80	65	60	70	65	95	80	75	65
85	65	85	80	60	60	75	60	85	70	65	75
90	70	80	70	85	70	80	75	85	70	70	60
80	60	90	75	80	75	85	75	70	60	80	75
75	60	95	75	85	75	80	70	75	60	70	80
80	70	70	60	80	70	80	70	80	80	75	70
90	75	80	70	90	65	70	60	85	80	80	70
70	70	75	65	75	65	85	70	85	75	80	75
60	70	70	60	85	65	75	60	85	75	75	70

- a. Rumuskan hipotesis yang mungkin bisa diuji dengan ANOVA.
- b. Uji hipotesis yang sudah Saudara kembangkan.

BAB VIII

TEKNIK KORELASI

TUJUAN PEMBELAJARAN

Mahasiswa dapat:

- 1) Melakukan pengujian
- 2) Membandingkan
- 3) Menjelaskan arti

SUSUNAN MATERI

- a) Pengantar
- b) Korelasi product moment
- c) Korelasi Spearman

PENGANTAR

Pada uraian terdahulu kita telah membahas perbedaan rata-rata dua kelompok. Pembahasan berikut akan menitikberatkan pada hubungan antara dua kelompok nilai. (korelasi).

Untuk mengukur besarnya hubungan antara sekelompok nilai satu (X) dengan sekelompok nilai yang lainnya (Y) rumus-rumus korelasi yang sering dipakai di antaranya adalah Product Moment Correlation dan Spearman Correlation.

A. KORELASI PRODUCT MOMENT

Korelasi yang sering digunakan oleh peneliti, terutama kalau datanya interval, adalah korelasi *product moment* atau korelasi *Pearson*. Korelasi product moment dapat dihitung dengan rumus:

$$r_{1.2} = \frac{N\Sigma XY - (\Sigma X)(\Sigma Y)}{\sqrt{(N\Sigma X^2 - (\Sigma X)^2)(N\Sigma Y^2 - (\Sigma Y)^2)}}$$

Contoh Kasus:

Suatu penelitian ingin melihat apakah ada hubungan antara banyaknya kredit yang diambil dengan indeks prestasi yang dicapai mahasiswa dalam satu semester. Setelah dilakukan pengumpulan data dari 10 mahasiswa, penyebaran datanya sbb:

Mahasiswa ke	Jumlah kredit yang diambil	Indeks Prestasi.
.1.	20	3,1
2.	18	3,0
3.	15	2,2
4.	20	3,0
5.	22	3,6
6.	22	3,0
7.	18	2,0
8.	16	2,8
9.	14	2,2
10.	12	3,2

Untuk keperluan penghitungan korelasi, sebaiknya data di atas disusun dalam suatu tabel yang mengandung unsur-unsur yang diperlukan dalam penghitungan korelasi, sebagai berikut.

X	Y	X ²	Y ²	XY
20	3,1	400	9,61	62
18	3,0	324	9,00	54
15	2,2	225	4,84	33
20	3,0	400	9,99	60
22	3,6	484	12,96	79,2
22	3,0	484	9,00	66
18	2,0	324	4,00	36
16	2,8	256	7,84	44,8
14	2,2	196	4,84	30,8
12	3,2	144	10,24	38,4
177	28,1	3237	81,33	504,2

Langkah-langkah Penghitungan:

1) Inventarisasi hal yang diketahui berdasarkan tabel di atas.

$$n = 10$$

$$\sum X = 177$$

$$\sum Y = 28,1$$

$$\sum X^2 = 3237$$

$$\sum XY = 504,2$$

2) Substitusikan faktor yang diperlukan dalam rumus Korelasi Product Moment dan menghitungnya, sbb:

$$\begin{aligned} r_{1.2} &= \frac{N\sum XY - (\sum X)(\sum Y)}{\sqrt{(N\sum X^2 - (\sum X)^2)(N\sum Y^2 - (\sum Y)^2)}} \\ &= \frac{(10 \times 504,2) - (177)(28,1)}{\sqrt{((10 \times 3237) - (177)^2)((10 \times 81,33) - (28,1)^2)}} \\ &= \frac{5042 - 4973,7}{\sqrt{(32370 - 31329)(813,3 - 789,61)}} \\ &= \frac{68,3}{\sqrt{(1041)(23,69)}} \\ &= \frac{68,3}{\sqrt{24661,29}} \\ &= \frac{68,3}{157,04} \\ &= 0,43 \end{aligned}$$

3) Pengambilan Kesimpulan

Konsultasikan harga r pada langkah 2 dengan tabel r. Dari tabel r diperoleh bahwa harga r dengan taraf signifikan 5% adalah 0,632. Karena harga r hitung lebih kecil dari r tabel, maka kesimpulannya adalah tidak ada hubungan yang signifikan antara banyaknya kredit yang diambil dengan indeks prestasi yang dicapai mahasiswa dalam satu semester.

Hasil penghitungan korelasi pada dasarnya dapat dikelompokkan menjadi tiga kelompok.

1. *Korelasi positif kuat*, apabila hasil penghitungan mendekati +1 atau sama dengan +1. Ini berarti bahwa setiap *kenaikan* skor/nilai pada variabel X akan diikuti *kenaikan* skor/nilai variabel Y. Sebaliknya, jika variabel X mengalami *penurunan* maka akan diikuti dengan *penurunan* variabel Y.
2. *Korelasi negatif kuat*, apabila hasil penghitungan mendekati -1 atau sama dengan -1. Ini berarti bahwa setiap *kenaikan* skor/nilai pada variabel X akan diikuti dengan *penurunan* nilai/skor variabel Y. Sebaliknya apabila skor/nilai dari variabel X

turun, maka skor/nilai variabel Y akan *naik*.

3. *Tidak ada korelasi*, apabila hasil penghitungan korelasi mendekati 0 atau sama dengan 0. Hal ini berarti bahwa naik turunnya skor/nilai satu variabel tidak mempunyai kaitan dengan naik turunnya skor/nilai variabel yang lainnya. Apabila skor/nilai variabel X naik tidak selalu diikuti dengan naik turunnya skor/nilai variabel Y, demikian juga sebaliknya.
4. Besar kecilnya indeks korelasi menentukan kuat lemahnya hubungan kedua variabel. Kriterianya sebagai berikut.
 - 0 – 25 : Korelasi sangat lemah
 - >0,25 – 0,5 : Korelasi cukup
 - >0,5 – 0,75 : Korelasi kuat
 - >0,75 – 1 : Korelasi sangat kuat.

B. KORELASI SPEARMAN

Apabila korelasi product moment digunakan untuk menganalisis data berskala interval, korelasi Spearman digunakan untuk menganalisis data berskala ordinal. Korelasi Spearman dihitung dengan rumus sbb.

$$r_s = 1 - \frac{6 \sum D^2}{n(n^2 - 1)}$$

Keterangan:

D = selisih antara X dan Y

6 = angka konstan

Contoh kasus:

Suatu penelitian terhadap hubungan antara ranking tes masuk mahasiswa dengan di dalam kelas sesudah ikut kuliah. Dari sepuluh mahasiswa yang terambil sebagai sampel ternyata penyebaran datanya sebagai berikut:

Mahasiswa ke	Rangking tes masuk	Rangking kelas
1.	1	10
2.	2	7
3.	3	8
4.	4	6
5.	5	5
6.	6	3
7.	7	4

8.	8	2
9.	9	9
10.	10	1

Yang ingin diketahui peneliti adalah:

Berapa tingkat hubungan antara ranking tes masuk dengan ranking kelas sesudah kuliah?

Jawab:

Untuk keperluan penghitungan, sebaiknya data di atas disusun dalam suatu tabel yang mengandung unsur-unsur yang diperlukan dalam penghitungan korelasi, sebagai berikut.

X	Y	D	D ²
1	10	9	81
2	7	5	25
3	8	5	25
4	6	2	4
5	5	0	0
6	3	3	9
7	4	3	9
8	2	6	36
9	9	0	0
10	1	9	81
	.	42.	270

1) Inventarisasi hal yang diketahui berdasarkan tabel di atas.

$$n = 10$$

$$\sum D = 42$$

$$\sum D^2 = 270$$

2) Mensubstitusikan harga-harga tersebut ke dalam rumus, dan menghitungnya, sbb:

$$\begin{aligned}
 r_s &= 1 - \frac{6 \sum D^2}{n(n^2 - 1)} \\
 &= 1 - \frac{(6 \times 270)}{10(10 - 1)} \\
 &= 1 - 1,636363636
 \end{aligned}$$

$$= -0,636363636$$

$$= -0,63$$

3) Pengambilan kesimpulan

Dari Tabel r diperoleh bahwa harga r untuk $n = 10$ taraf signifikansi 5% adalah $r = 0,632$. Karena harga r hitung lebih kecil dari r tabel, **maka kesimpulannya adalah tidak ada hubungan yang signifikan antara banyaknya kredit yang diambil dengan indeks prestasi yang dicapai mahasiswa dalam satu semester. Tolong dicek lagi !!!**

Ada kalanya dua kelompok data yang kita hadapi tidak mempunyai skala sama. Di satu pihak berskala ordinal dan di lain pihak berskala interval. Untuk kondisi seperti ini besarnya korelasi tidak dapat dihitung dengan Product moment, tetapi harus digunakan korelasi Spearman.

LATIHAN

1. Dalam pengumpulan nilai mata kuliah Evaluasi Pendidikan (X) dengan nilai Statistik Pendidikan dari 20 orang mahasiswa sebagai berikut:

X: 80 80 70 75 85 90 85 85 90 75 70 75 80 85 85 95 90 90 85 70

Y: 70 75 60 70 80 90 80 85 85 70 65 70 80 80 85 90 90 90 85 70

- Hitung korelasi antara nilai mata kuliah Evaluasi Pendidikan dan nilai mata kuliah Statistik Pendidikan
- Apakah terdapat korelasi yang signifikan antara nilai mata kuliah Evaluasi Pendidikan dan Statistik Pendidikan?

2. Berikut adalah data status anak dalam keluarga, anak ke-1, ke-2, ke-3, dst. (X) dan prestasi belajarnya (Y) dari 25 siswa.

X: 1 4 3 1 4 4 2 1 3 5 3 4 2 4 3 6 4 2 3 4 5 6 4 1 1

Y: 7 8 7 8 8 8 9 7 8 7 6 6 7 8 8 9 7 6 7 8 7 8 9 5 4

Adakah hubungan antara status anak dengan prestasi belajar?

BAB IX

ANALISIS REGRESI

TUJUAN PEMBELAJARAN

Mahasiswa dapat:

- 1) Melakukan pengujian
- 2) Membandingkan
- 3) Menjelaskan arti

SUSUNAN MATERI

- a) Pengantar
- b) Analisis regresi linier: satu prediktor
- c) Analisis regresi linier: dua prediktor

PENGANTAR

Teknik analisis regresi pada pokoknya adalah teknik untuk mendapatkan bentuk hubungan-hubungan di antara variabel-variabel yang menjadi fokus perhatian. Salah satu manfaat dari hasil analisis regresi adalah dapat memberikan dasar untuk mengadakan prediksi. Suatu variabel dapat diprediksikan dari yang lain (*prediktor*) jika variabel tersebut (dinamakan *kriterium*) terdapat korelasi yang signifikan.

Korelasi atau hubungan antara kedua variabel tersebut dapat digambarkan dalam bentuk sebuah garis (disebut garis regresi). Bentuk garis tersebut ada yang linier dan ada yang nonlinier. Garis regresi tersebut dinyatakan dalam bentuk persamaan garis regresi;

Tugas pokok dari analisis regresi adalah:

1. mencari korelasi antara kriterium dengan prediktor
2. menguji apakah korelasi tersebut signifikan atau tidak.
3. mencari persamaan regresinya.
4. mencari besarnya sumbangan relatif dan sumbangan efektif di antara prediktor terhadap kriterium (apabila jumlah prediktornya lebih dari satu)

A. ANALISIS REGRESI LINIER: SATU PREDIKTOR

Contoh Kasus:

Suatu penelitian tentang **berat badan** orang pada kelompok umur tertentu dapat diramalkan dari **tinggi badan**. Kemudian data dikumpulkan dari 10 orang sebagai berikut.

Subj, no.	Tinggi Badan (cm) (X)	Berat Badan (kg) (Y)
1.	168	68
2.	173	81
3.	162	54
4.	157	49
5.	160	52
6.	165	62
7.	163	56
8.	170	78
9.	168	64
10	164	62

Dengan menggunakan kalkulator statistik, diperoleh :

$$N = 10$$

$$\sum X = 1650$$

$$\sum Y =$$

$$\sum X^2 = 271,460$$

$$\sum Y^2 = 39,432$$

$$\sum XY = 102,732$$

1) Mencari korelasi antara kreterium dan prediktor

$$r_{xy} = \frac{\sum xy}{(\sum x^2)(\sum y^2)}$$

$$\begin{aligned}\sum xy &= \sum XY - \frac{(\sum X)(\sum Y)}{N} \\ &= 102,732 - \frac{(1650)(620)}{10} = 432\end{aligned}$$

$$\begin{aligned}\sum x^2 &= \sum X^2 - \frac{(\sum X)^2}{N} \\ &= 271,469 - \frac{(1650)^2}{10} = 210\end{aligned}$$

$$\begin{aligned}\sum y^2 &= \sum Y^2 - \frac{(\sum Y)^2}{N} \\ &= 39,432 - \frac{(620)^2}{10} = 992 \\ r_{xy} &= \frac{432}{(210)(992)} \\ &= 0,946\end{aligned}$$

2) Uji Signifikansi harga r_{xy}

$r_{xy} = 0,946$ dikonsultasikan dengan Tabel r. Dari Tabel r dengan $N = 10$ taraf signifikansi 1% didapatkan harga $r = 0,765$. Dengan demikian, dapat disimpulkan bahwa korelasi antara variabel X dan Y, yaitu antara tinggi badan dan berat badan sangat signifikan.

3) Mencari persamaan garis regresi

Dengan korelasi antara tinggi badan dan berat badan yang sangat signifikan tsb., kita mempunyai landasan untuk memprediksikan berat badan dari tinggi badan, dan karenanya kita dapat membuat garis regresi untuk prediksi, dengan rumus persamaan garis regresi dengan satu prediktor sbb.:

$$Y = aX + K$$

Untuk mengisi persamaan garis regresi tersebut, harga koefisien prediktor (yaitu a) dan harga bilangan konstan (yaitu K) antara lain dapat kita temukan dengan metode angka kasar. Dengan metode angka kasar harga a dan K dapat dicari dari persamaan berikut.

$$(1) \sum XY = a \sum X^2 + K \sum X$$

$$(2) \sum Y = a \sum X + NK$$

Data yang sudah kita ketahui kita masukkan ke dalam rumus-rumus tersebut, sbb:

$$(1) 102,732 = 272,460 a + 1,650 K$$

$$(2) 620 = 1.650 a + 10 K$$

Dengan penyelesaian persamaan secara simultan (dengan membagi persamaan 1 dengan 1650 dan persamaan 2 dengan 10) akan kita temukan:

$$(3) 62,26 = 165,13 a + K$$

$$(4) \underline{62} = \underline{165} a + K$$

$$(5) 0,26 = 0,13 a \quad a = 0,26 : 0,13 = 2$$

$$(4) 62 = (165) (2) + K \quad K = 62 - 330 = - 268$$

Dengan harga $a = 2$ dan $K = -262$, persamaan garis regresinya adalah:

$Y = 2 X - 268$. Persamaan garis regresi digunakan untuk memprediksikan berat badan seseorang (berdasarkan tinggi badannya), misalnya:

$$\text{Untuk } X = 175 \quad Y = 2 (175) - 268 = 82.$$

$$X = 174 \quad Y = 2 (174) - 268 = 80$$

$$X = 150 \quad Y = 2 (150) - 268 = 32, \text{ dan seterusnya.}$$

Pekerjaan analisis regresi apabila hanya satu prediktor telah selesai. Analisis dikatakan selesai karena korelasi antara prediktor dengan kreterium telah ditemukan, uji signifikansi telah dilakukan, dan persamaan garis regresi telah diperoleh.

B. ANALISIS REGRESI LINIER: DUA PREDIKTOR

Contoh Kasus:

Seorang dosen Statistik ingin memastikan apakah skor Statistik Pendidikan (Y) dapat diprediksikan dari nilai Matematika (X1) dan Indeks Prestasi SMA (X2). Dari pengambilan sampel secara acak diperoleh 10 mahasiswa dengan data sbb:

(X ₁)	(X ₂)	(Y)
57	3,00	27
93	2,85	34
79	3,20	27
26	2,49	24
69	3,07	35
24	2,38	18
76	3,74	33
68	2,62	39
82	2,53	35
29	3,17	25

Dengan menggunakan kalkulator statistik, diperoleh :

$$N = 10$$

$$\sum X_1 = 596$$

$$\sum X_1^2 = 41214$$

$$\sum X_2 = 29,05$$

$$\sum X_2^2 = 85,9537$$

$$\sum Y = 297$$

$$\sum Y^2 = 9199$$

$$\sum X_1 X_2 = 1765,99$$

$$\sum X_1 Y = 18787$$

$$\sum X_2 Y = 867,75$$

1. Mencari persamaan garis regresi

Rumus persamaan garis regresi dua prediktor adalah sbb.

$$Y = a_1 X_1 + a_2 X_2 + K$$

Dalam skor deviasi persamaan itu dapat dituliskan sbb:

$$y = a_1 x_1 + a_2 x_2$$

Untuk mengisi persamaan garis regresi tersebut, dalam contoh ini akan digunakan *metode skor deviasi*

1) Mencari data berikut, dengan rumus masing-masing

$$\begin{aligned}\sum x_1^2 &= \sum X_1^2 - \frac{(\sum X_1)^2}{N} \\ &= 41214 - \frac{(596)^2}{10} = 5692,4\end{aligned}$$

$$\begin{aligned}\sum x_2^2 &= \sum X_2^2 - \frac{(\sum X_2)^2}{N} \\ &= 85,95,37 - \frac{(29,05)^2}{10} = 1,556345\end{aligned}$$

$$\begin{aligned}\sum y^2 &= \sum Y^2 - \frac{(\sum Y)^2}{N} \\ &= 9199 - \frac{(297)^2}{10} = 378,1\end{aligned}$$

$$\begin{aligned}\sum x_1 x_2 &= \sum X_1 X_2 - \frac{(\sum X_1)(\sum X_2)}{N} \\ &= 1765,99 - \frac{(596)(29,05)}{10} = 34,61\end{aligned}$$

$$\sum x_1 y = \sum X_1 Y - \frac{(\sum X_1)(\sum Y)}{N}$$

$$= 18787 - \frac{(596)(297)}{10} = 1085,8$$

$$\begin{aligned} \sum x_2 y &= \sum X_2 Y - \frac{(\sum X_2)(\sum Y)}{N} \\ &= 867,75 - \frac{(29,05)(297)}{10} = 4,965 \end{aligned}$$

2) Mencari a_1 dan a_2 dengan persamaan simultan

$$(1) \sum x_1 y = a_1 \sum x_1^2 + a_2 \sum x_1 x_2$$

$$(2) \sum x_2 y = a_1 \sum x_1 x_2 + a_2 \sum x_2^2$$

Diisi dan dihitung:

$$(1) 1085,8 = 5692,4 a_1 + 34,61 a_2$$

$$(2) 4,965 = 34,61 a_1 + 1,56345 a_2$$

$$(1) : 34,61 \rightarrow (3) 31,37243571 = 164,4726958 a_1 + a_2$$

$$(2) : 1,56345 \rightarrow (4) \underline{3,175669193} = \underline{22,1376487 a_1 + a_2}$$

$$(5) 28,19676652 = 142,335755 a_1$$

$$a_1 = \frac{28,19676652}{142,335755} = 0,198100375$$

$$(4) 3,175669193 = 22,1376487 a_1 + a_2$$

$$3,175669193 = (22,1376487)(0,198100375) + a_2$$

$$a_2 = -1,209807316$$

Jadi, persamaan garis regresi dalam skor deviasi yang kita cari adalah:

$$y = a_1 x_1 + a_2 x_2$$

$$Y - Y = a_1 (X_1 - X_1) + a_2 (X_2 - X_2)$$

$$Y = a_1 (X_1 - X_1) + a_2 (X_2 - X_2) + Y$$

Dari data di muka dapat diketemukan:

$$X_1 = \frac{596}{10} = 59,6$$

$$X_2 = \frac{29,05}{10} = 2,905$$

$$Y = \frac{297}{10} = 29,7$$

$$a_1 = 0,198100375$$

$$a_2 = -1,209807316$$

$$\begin{aligned}
\text{Jadi, } Y &= (0,198100375)(X_1 - 59,6) + (-1,209807316) (X_2 - 2,905) + 29,7 \\
&= 0,198100375 X_1 - 11,80678235 - 1,209667071X_2 - 3,514082841 \\
&\quad + 29,7 \\
&= 0,198100375 X_1 - 1,209667071X_2 + 21,40730049, \text{ dibulatkan} \\
Y &= 0,2 X_1 - 1,2 X_2 + 21,4
\end{aligned}$$

3) Mencari koefisien korelasi antara kriterium (Y) dengan prediktor X₁ dan X₂.

$$\text{Rumus: } R_{y(1,2)} = \sqrt{\frac{a_1 \sum x_1 y + a_2 \sum x_2 y}{\sum y^2}}$$

Jika hasil penghitungan di atas dimasukkan ke dalam rumus:

$$\begin{aligned}
R_{y(1,2)} &= \sqrt{\frac{(0,198100375)(1085,8) + (-1,209807316)(4,965)}{378,1}} \\
&= 0,744
\end{aligned}$$

$$R_{y(1,2)}^2 = 0,553005527$$

4) Uji signifikansi

Untuk menguji apakah $R_{y(1,2)}$ itu signifikan atau tidak, dilakukan analisis regresi. Analisis regresi menghasilkan harga "F", yang kemudian harga F ini signifikan atau tidak, kita lakukan uji terhadap harga F ini.

Harga F regresi dihitung dengan rumus:

$$F_{reg} = \frac{R^2(N - m - 1)}{m(1 - R^2)}$$

dengan: F_{reg} = harga F garis regresi

m = jumlah prediktor

N = jumlah kasus

R = koefisien korelasi antara kriterium dengan prediktor.

Derajat kebebasan untuk menguji F ini adalah besarnya m melawan N-m-1.

Jika harga-harga yang telah ditung di muka dimasukkan, akan diperoleh:

$$F_{reg} = \frac{(0,553005527)(10 - 2 - 1)}{2(1 - 0,553005527)} = 4,330$$

$$F_{reg} = 4,330 < F_{\text{tabel } 5\%} = 4,74$$

Kesimpulan : **tidak ada korelasi** antara Y dengan X₁ dan X₂. Oleh karena itu, (dengan t.s 5%) kita tidak berani memprediksikan nilai statistik dengan menggunakan nilai matematika dan indeks prestasi SMA.

Rumus F_{reg} di atas diperoleh dari proses analisis varian garis regresi yang agak

panjang. Keseluruhan proses tersebut dapat dilihat dalam Tabel Rangkuman Analisis sebagai berikut.

Sumber Varian	d.b	JK	RK
Regresi (Reg)	m	$R(\sum y^2)$	$\frac{R^2(\sum y^2)}{m}$
Residu (Res)	N-m-1	$(1 - R^2)(\sum y^2)$	$\frac{(1 - R^2)(\sum y^2)}{N - m - 1}$
Total	N-1	$\sum y^2$	

Jadi, jika seluruh proses tersebut kita ikuti, maka:

$$Jk_{reg} = (0,553005527)(378,1) = 209,0913897$$

$$db_{reg} = 2$$

$$RK_{reg} = \frac{209,0913897}{2} = 104,5456948$$

$$Jk_{res} = (1 - 0,553005527) (378,1) = 169,0086102$$

$$db_{res} = 10 - 2 - 1 = 7$$

$$RK_{res} = \frac{169,0086102}{7} = 24,14408717$$

$$\text{Jadi, } F_{reg} = \frac{RK_{reg}}{RK_{res}} = \frac{104,5456946}{24,14408717} = 4,330$$

Tabel Ringkasan Analisis Regresi

Sumber Varian	d.b	JK	RK	F
Regresi (Reg)	2	209,0913897	104,5456948.	4,33
Residu (Res)	7	169,0086102.	24,14408717	
Total	9	378,1		

5) Mencari Sumbangan Relatif dan Sumbangan Efektif

Mencari sumbangan relatif dan sumbangan efektif menjadi tugas keempat analisis regresi.

Rumus:

$$SR_{x_1} = \frac{a_1 \sum x_1 y}{JK_{reg}} \times 100\%$$

$$SR_{x_2} = \frac{a_2 \sum x_2 y}{JK_{reg}} \times 100\%$$

Dari pekerjaan kita tentang ANAREG dua prediktor kita temukan:

$$a_1 \sum x_1 y = (0,198100375)(1085,8) = 215,097387$$

$$a_2 \sum x_2 y = (-1,209667070)(4,965) = -6,005997$$

JK_{reg} dapat pula ditulis:

$$\begin{aligned} JK_{reg} &= a_1 \sum x_1 y + a_2 \sum x_2 y \\ &= 215,097387 + (-6,005997) \end{aligned}$$

Catatan: Jika harga komponen-komponen dalam JK_{reg} tidak semuanya bertanda positif (seperti dalam contoh ini) tanda negatif ditiadakan (dihitung dengan harga mutlak).

$$\text{Jadi, } JK_{reg} = 215,097387 + 6,005997 = 221,103384$$

$$SR_{x_1} = \frac{215,097387}{221,103384} \times 100\% = 97,29\%$$

$$SR_{x_2} = \frac{6,005997}{221,103384} \times 100\% = 2,72\%$$

Sumbangan Efektif dihitung dengan rumus:

$$\begin{aligned} SE_{x_1} &= SR_{x_1} \times R^2 \\ &= 97,29\% \times 0,553 = 53,8\% \end{aligned}$$

$$\begin{aligned} SE_{x_2} &= SR_{x_2} \times R^2 \\ &= 2,72\% \times 0,553 = 1,5\% \end{aligned}$$

LATIHAN

1 Hitung persamaan regresi untuk data di bawah ini

Intelegensi (X) : 90 100 100 95 105 110 105 105 115 120

Nilai Statistik (Y) : 70 75 80 75 85 85 85 90 95 100

2. Hitunglah koefisien regresi untuk data di bawah ini.

X1	X2	Y
70	90	85
80	75	85
90	80	90
80	85	85
75	70	70
90	85	80
90	95	95
80	90	80
75	75	70
70	70	80

X1 : Kemampuan Membaca

X2 : Nilai Bahasa Indonesia

Y : Kemampuan Menulis Karya Ilmiah

- Hitunglah koefisien regresi untuk data tersebut.
- Uji signifikansi regresi ganda untuk data tersebut.
- Berapa besarnya sumbangan variabel bebas terhadap variabel terikatnya.

DAFTAR PUSTAKA

- Agus Irianto. 1988. *Statistik Pendidikan*. Jakarta: Departemen Pendidikan dan Kebudayaan, Dirjen Dikti, Proyek LPTK
- Dergibson Siagian Sugiarto. 2002. *Metode Statistika*. Jakarta: Gramedia Pustaka Utama
- Neter, John. *Applied Linear Statistical Model*. Otario: Irwin-Dorsey
- Netter, John, William Wasserman, Michael H. Kurtner. 1983. *Applied Linier Regression Models*. Illinois: Irwin home wood.
- Spegel, Murray R, Stephaens Larry J. 1999. *Schaum's Outlines of Theory and Problems of Statistics*, Terjemahan. Bandung: Eralangga.
- Singgih Santoso. 2005. *Mengatasi berbagai masalah statistik dengan SPSS versi 11.5*. Jakarta: Elex Media Computindo
- Sudjana. 2005. *Metode Statistika*. Bandung : Tarsito.
- Sugiyono. 2009. *Statistika untuk Penelitian*. Bandung: Alfabeta
- Walpole, E. Ronald. 1995. *Pengantar Statistika*. Jakarta: Gramedia Pustaka Utama.

Lampiran 1. Tabel nilai F

Lampiran 2. Tabel nilai r