



DASAR-DASAR STATISTIKA UNTUK PENELITIAN

Anisa Fitri ■ Rani Rahim ■ Nurhayati ■ Azis
Sadrack Luden Pagiling ■ Irmawaty Natsir
Anis Munfarikhatin ■ Daniel Nicson Simanjuntak
Kartini Hutagaol ■ Nanda Eska Anugrah



DASAR-DASAR STATISTIKA UNTUK PENELITIAN

UU 28 tahun 2014 tentang Hak Cipta

Fungsi dan sifat hak cipta Pasal 4

Hak Cipta sebagaimana dimaksud dalam Pasal 3 huruf a merupakan hak eksklusif yang terdiri atas hak moral dan hak ekonomi.

Pembatasan Perlindungan Pasal 26

Ketentuan sebagaimana dimaksud dalam Pasal 23, Pasal 24, dan Pasal 25 tidak berlaku terhadap:

- a. penggunaan kutipan singkat Ciptaan dan/atau produk Hak Terkait untuk pelaporan peristiwa aktual yang ditujukan hanya untuk keperluan penyediaan informasi aktual;
- b. Penggandaan Ciptaan dan/atau produk Hak Terkait hanya untuk kepentingan penelitian ilmu pengetahuan;
- c. Penggandaan Ciptaan dan/atau produk Hak Terkait hanya untuk keperluan pengajaran, kecuali pertunjukan dan Fonogram yang telah dilakukan Pengumuman sebagai bahan ajar; dan
- d. penggunaan untuk kepentingan pendidikan dan pengembangan ilmu pengetahuan yang memungkinkan suatu Ciptaan dan/atau produk Hak Terkait dapat digunakan tanpa izin Pelaku Pertunjukan, Produser Fonogram, atau Lembaga Penyiaran.

Sanksi Pelanggaran Pasal 113

1. Setiap Orang yang dengan tanpa hak dan/atau tanpa izin Pencipta atau pemegang Hak Cipta melakukan pelanggaran hak ekonomi Pencipta sebagaimana dimaksud dalam Pasal 9 ayat (1) huruf c, huruf d, huruf f, dan/atau huruf h untuk Penggunaan Secara Komersial dipidana dengan pidana penjara paling lama 3 (tiga) tahun dan/atau pidana denda paling banyak Rp500.000.000,00 (lima ratus juta rupiah).
2. Setiap Orang yang dengan tanpa hak dan/atau tanpa izin Pencipta atau pemegang Hak Cipta melakukan pelanggaran hak ekonomi Pencipta sebagaimana dimaksud dalam Pasal 9 ayat (1) huruf a, huruf b, huruf e, dan/atau huruf g untuk Penggunaan Secara Komersial dipidana dengan pidana penjara paling lama 4 (empat) tahun dan/atau pidana denda paling banyak Rp1.000.000.000,00 (satu miliar rupiah).

Dasar-dasar Statistika untuk Penelitian

Anisa Fitri, Rani Rahim, Nurhayati, Azis
Sadrack Luden Pagiling, Irmawaty Natsir, Anis Munfarikhatin
Daniel Nicson Simanjuntak, Kartini Hutagaol, Nanda Eska Anugrah



Penerbit Yayasan Kita Menulis

Dasar-dasar Statistika untuk Penelitian

Copyright © Yayasan Kita Menulis, 2023

Penulis:

Anisa Fitri, Rani Rahim, Nurhayati, Azis
Sadrack Luden Pagiling, Irmawaty Natsir
Anis Munfarikhatin, Daniel Nicson Simanjuntak
Kartini Hutagaol, Nanda Eska Anugrah

Editor: Ronal Watrianthos

Desain Sampul: Devy Dian Pratama, S.Kom.

Penerbit

Yayasan Kita Menulis

Web: kitamenulis.id

e-mail: press@kitamenulis.id

WA: 0821-6453-7176

IKAPI: 044/SUT/2021

Anisa Fitri., dkk.

Dasar-dasar Statistika untuk Penelitian

Yayasan Kita Menulis, 2023

xii 130 hlm; 16 x 23 cm

ISBN: 978-623-342-945-0

Cetakan 1, September 2023

- I. Dasar-dasar Statistika untuk Penelitian
- II. Yayasan Kita Menulis

Katalog Dalam Terbitan

Hak cipta dilindungi undang-undang

Dilarang memperbanyak maupun mengedarkan buku tanpa
izin tertulis dari penerbit maupun penulis

Kata Pengantar

Puji Syukur ke hadirat Tuhan Yang Maha Esa yang telah memberikan rahmat-Nya, sehingga kami mampu menyelesaikan buku Dasar-Dasar Statistika Untuk Penelitian. Buku Dasar-Dasar Statistika Untuk Penelitian membahas topik-topik penting yang perlu diketahui oleh periset sebagai dasar ketika melakukan penelitian. Semoga dengan dibuatkan buku ini para pendidik, akademisi, periset, serta setiap pihak yang melakukan studi dan riset dapat memahami tentang analisis-analisis dalam Statistika yang lebih berorientasi pada Penelitian.

Buku ini berisi penjabaran mengenai Dasar-Dasar Statistika Untuk Penelitian yang terdiri dari 11 (sebelas) bab, yaitu:

Bab 1 Statistik dan Statistika

Bab 2 Penyajian Data Penelitian

Bab 3 Distribusi Frekuensi

Bab 4 Kemencengan dan Kurtosis

Bab 5 Statistik Inferensi

Bab 6 Uji Normalitas dan Uji Homogenitas Data

Bab 7 Uji T-Test

Bab 8 Uji Chi Square

Bab 9 Analisis Varians

Bab 10 Analisis Korelasi dan Regresi

Akhir kata, penulis mengucapkan banyak terima kasih kepada teman-teman sejawat yang telah berkolaborasi dan memberikan masukan positif selama penulisan buku ini. Kiranya buku ini bermanfaat bagi pembaca khususnya mahasiswa dan para peneliti dan dapat dijadikan sebagai bahan pembelajaran dan berguna bagi pembaca.

Penulis

Daftar Isi

Kata Pengantar	v
Daftar Isi	vii
Daftar Gambar	ix
Daftar Tabel	xi

Bab 1 Statistik dan Statistika

1.1 Sejarah dan Peranan Statistika	1
1.2 Data Statistik	4
1.3 Variabel Penelitian	5
1.4 Populasi dan Sampel	9

Bab 2 Penyajian Data Penelitian

2.1 Pendahuluan	11
2.2 Penyajian Data Dalam Bentuk Tabel	12
2.3 Penyajian Data dalam Bentuk Diagram/Grafik	18
2.3.1 Jenis Diagram/Grafik	19
2.3.2 Grafik Tabel Distribusi Frekuensi	21

Bab 3 Distribusi Frekuensi

3.1 Distribusi Frekuensi Data Tunggal	25
3.2 Distribusi Frekuensi Data Kelompok	26
3.3 Diagram Distribusi Frekuensi	28

Bab 4 Kemencengan dan Kurtosis

4.1 Kemencengan	33
4.2 Kurtosis	39
4.3 Menghitung Normalitas	45

Bab 5 Statistik Inferensi

5.1 Pendahuluan	47
5.2 Pengujian Hipotesis	48
5.3 Uji Mean Populasi	52
5.4 Metode Pengujian Pada Statistik Inferensi	55

Bab 6 Uji Normalitas dan Uji Homogenitas Data

6.1 Pendahuluan.....	57
6.2 Uji Homogenitas.....	61

Bab 7 Uji T-Test

7.1 Pengertian Uji T.....	69
7.2 Macam-Macam Uji T.....	70
7.2.1 Uji Beda Satu Mean.....	71
7.2.2 Uji Beda Dua Mean.....	73
7.3 Uji T Menggunakan Ms. Excel.....	79

Bab 8 Uji Chi Square

8.1 Pendahuluan.....	83
8.2 Contoh Kasus 1.....	85
8.3 Contoh Kasus 2.....	87

Bab 9 Analisis Varians

9.1 Pendahuluan.....	91
9.2 Ciri dan Prinsip ANOVA.....	93
9.3 Jenis ANOVA.....	94
9.3.1 One Way Analysis of Variance.....	94
9.3.2 Two Way Analysis of Variance.....	95

Bab 10 Analisis Korelasi dan Regresi

10.1 Analisis Korelasi.....	99
10.2 Analisis Regresi.....	109
Daftar Pustaka.....	119
Biodata Penulis.....	125

Daftar Gambar

Gambar 1.1: Hubungan Variabel X dengan Variabel Y	6
Gambar 1.2: Hubungan Variabel X1 dan X2 Secara Bersama-sama Dengan Variabel Y	7
Gambar 1.3: Hubungan Variabel X1 Dengan Variabel Y dan X2 Sebagai Variabel Moderator	7
Gambar 1.4: Hubungan Variabel X1 Dengan Variabel Y dan X2 Sebagai Variabel Intervening.....	8
Gambar 1.5: Hubungan Variabel X1 Dengan Variabel Y dan X2 Sebagai Variabel Kontrol	8
Gambar 2.1: Diagram/Grafik Batang Mengenai Jumlah Dosen Universitas Dharmawangsa Berdasarkan Jenis Kelamin dan Status Kepegawaiannya Pada Tahun 2020	19
Gambar 2.2: Diagram/Grafik Garis Mengenai Kunjungan Poliklinik RS ABC pada Tahun 2020	20
Gambar 2.3: Persentase Dosen Universitas Dharmawangsa Berdasarkan Jenis Kelamin Pada Tahun 2020	20
Gambar 2.4: Histogram	21
Gambar 2.5: Poligon Frekuensi	22
Gambar 2.6: Ogive Kurang Dari	23
Gambar 2.7: Ogive Lebih Dari	24
Gambar 3.1: Grafik Histogram Skor Kemampuan Matematika	29
Gambar 3.2: Grafik Poligon skor kemampuan matematika	30
Gambar 3.3: Gabungan Grafik Poligon dan Histogram	30
Gambar 3.4: Grafik Ogive Frekuensi Kumulatif.....	31
Gambar 3.5: Grafik Ogive Frekuensi Relatif.....	32
Gambar 4.1: Kemencengan	34
Gambar 4.2: Kurtosis.....	41
Gambar 4.3: Kurtosis.....	41
Gambar 6.1: Tahapan Analyze Descriptive.....	59
Gambar 6.2: Tahapan Analyze Descriptive.....	59
Gambar 6.3: Normal Q-Q Plot Hasil Belajar.....	61

Gambar 6.4: Tahapan Uji Levene	65
Gambar 6.5: Tahapan Uji Levene	66
Gambar 6.6: Tahapan Uji Levene	66

Daftar Tabel

Tabel 2.1: Jumlah Siswa di Kota Medan Berdasarkan Jenis Kelamin dan Tingkat Pendidikan	13
Tabel 2.2: Tabel Distribusi Frekuensi Nilai MID Mata Kuliah Statistik dari 60 Mahasiswa	16
Tabel 2.3: Tabel Distribusi Frekuensi Relatif.....	17
Tabel 2.4: Tabel Distribusi Frekuensi Kumulatif Kurang Dari.....	18
Tabel 2.5: Tabel Distribusi Frekuensi Kumulatif Lebih Dari.....	18
Tabel 3.1: Distribusi Frekuensi Skor Kemampuan Berbahasa Indonesia	26
Tabel 3.2: Distribusi Frekuensi Skor Kemampuan Matematika.....	28
Tabel 3.3: Tabel Distribusi Frekuensi	29
Tabel 3.4: Data Frekuensi Kumulatif.....	31
Tabel 4.1: Ketinggian Mahasiswa Laki-Laki di Perguruan Tinggi	35
Tabel 5.1: Pengambilan Keputusan.....	50
Tabel 6.1: Case Processing Summary.....	60
Tabel 6.2: Descriptive.....	60
Tabel 6.3: Test of Normality.....	60
Tabel 6.4: Nilai Varians Tiap Kelompok.....	64
Tabel 6.5: Test of Homogeneity of Variance	67
Tabel 7.1: Nilai Siswa Kelas F Dan G	79
Tabel 8.1: Proporsi Asupan Vitamin C	85
Tabel 9.1: Anova Dua Arah.....	95
Tabel 9.2: Tabel ANOVA Dua Arah	96
Tabel 9.3: Pengamatan Unit Penjualan Dalam 1 Bulan.....	96
Tabel 9.4: Pengaruh Bentuk Kemasan Pada Tingkat Penjualan.....	97
Tabel 9.5: ANOVA	97
Tabel 10.1: Tipe Data dan Teknik Korelasinya.....	101
Tabel 10.2: Nilai Koefisien Korelasi (r).....	101
Tabel 10.3: Rata-rata Nilai Ulangan dan UAS siswa.....	102
Tabel 10.4: Tabel Pembantu Nilai Ulangan dan UAS Siswa.....	103
Tabel 10.5: Tabel r.....	105
Tabel 10.6: Kategori Versi/Kualitas Smartphone dan Motivasi Belajar	107

Tabel 10.7: Data Versi/Kualitas Smartphone dan Motivasi Belajar Siswa...107

Tabel 10.8: Data Motivasi Belajar Siswa Terhadap Hasil Belajar Siswa.....111

Tabel 10.9: Data Motivasi Belajar Siswa Terhadap Hasil Belajar Siswa.....114

Bab 1

Statistik dan Statistika

1.1 Sejarah dan Peranan Statistika

Ilmu statistika mempunyai sejarah yang sangat panjang bersamaan dengan perkembangan sejarah manusia. Statistik telah banyak digunakan oleh bangsa-bangsa di Mesopotamia (Babilonia), Mesir, dan Cina Pada zaman sebelum Masehi. Mereka mengumpulkan data statistik untuk memperoleh informasi tentang besar pajak yang harus dibayar oleh penduduk, jumlah hasil pertanian yang mampu diproduksi, dan lain sebagainya.

Statistika pertama kali ditemukan oleh Aristoteles dalam bukunya yang berjudul “politea”, dalam buku tersebut ia menjelaskan data tentang keadaan 158 negara yang disebut sebagai statistika. Pada abad ke-17 di Inggris, statistika disebut sebagai *political arithmetic*. Selanjutnya istilah statistika dipopulerkan oleh Sir John Sinclair dalam bukunya berjudul “*Statistical Account Of Scotland* (1791-1799) pada abad ke-18,” yang sebelumnya dikemukakan oleh seorang ahli hitung asal Jerman yang bernama Gottfried Achenwall (1719-1772).

Awal perkembangan statistika induktif dipelopori oleh Karl Pearson (1857-1936) yang terjadi pada peralihan tahun 1900 ke 2000 an. Pada abad ini itu menjadi titik awal untuk pengembangan statistik modern Pada tahun 1900 an, Karl Pearson memberlakukan statistik pada biologi tentang proses evolusi masalah pewarisan. Pada tahun 1893 hingga 1912, Karl Pearson telah menulis

18 karya berjudul *Kontribusi Matematika untuk Teori Evolusi Berdasarkan Analisis Regresi dan Koefisien Korelasi*.

Statistika merupakan pokok bahasan dalam matematika dimana berkaitan dengan pengumpulan data, pengolahan data, dan pengambilan kesimpulan dari data hasil pengamatan di lapangan. Sejak dahulu, statistika telah menjadi dasar bagi penelitian, riset ataupun pengamatan di berbagai bidang ilmu (Sunaryo et al., 2003).

Dalam penelitian, terutama penelitian kuantitatif statistika berperan dalam penyusunan model, perumusan hipotesis, dalam pengembangan alat dan instrumen pengumpulan data, dalam penyusunan desain penelitian, dalam penentuan sampel dan dalam analisis data. Pengujian hipotesis mengharuskan kita untuk menarik kesimpulan yang bersifat umum dari kasus-kasus yang bersifat individual. Misalnya kita ingin mengetahui berapa proporsi mahasiswa di Jawa Timur yang putus sekolah, maka nilai proporsi siswa di Jawa Timur yang putus sekolah tersebut merupakan kesimpulan umum (tentunya yang berlaku di Jawa Timur) yang ditarik dari kasus-kasus individual di Jawa Timur.

Pengertian Statistik dan Statistika

Pada umumnya istilah statistik dan statistika sering kali dianggap memiliki makna yang sama. Statistik berasal dari kata “state” yang berarti negara. Awalnya, statistik mengacu pada ilmu angka untuk atas nama pemimpin atau raja di negara yang ingin mengetahui kekayaan tanah, populasi, produk pertanian dan perkebunan, serta hewan peliharaannya dari Kaisar Augustus, yang membuat pernyataan bahwa setiap orang mengenakan pajak. jadi, setiap orang harus melapor ke ahli statistik atau pemungut pajak terdekat merupakan contoh tertua mengenai statistika di masa lalu (Spiegel, 1961).

Banyak permasalahan, apakah itu hasil penelitian, riset atau pengamatan dinyatakan dan dicatat dalam bentuk bilangan atau angka-angka. Kumpulan angka-angka tersebut biasanya disusun, diatur, atau disajikan dalam bentuk daftar atau tabel. Sering pula daftar atau tabel tersebut disertai dengan gambar-gambar yang biasa disebut diagram atau grafik supaya lebih dapat menjelaskan persoalan yang sedang dipelajari.

Pada awalnya kata statistik diartikan sebagai keterangan-keterangan yang dibutuhkan oleh Negara dan berguna bagi Negara. Semula, statistik hanya diartikan sebagai sekumpulan angka-angka yang menggambarkan kejadian masa lalu sampai saat gambaran itu dibuat. Misalnya gambaran tentang kondisi

penduduk, pendapatan masyarakat, tingkat produksi pertanian, dan lain sebagainya (Mundir, 2012).

Statistik merupakan kumpulan fakta, informasi, data, bilangan maupun non bilangan, yang disusun dalam tabel dan atau diagram yang menggambarkan suatu persoalan. Statistik sering juga disebut parameter, namun dua istilah ini memiliki perbedaan. Statistik merupakan sembarangan nilai yang menjelaskan nilai dari sampel, sedangkan parameter merupakan sembarangan nilai yang mewakili atau menjelaskan nilai dari populasi.

Dari hasil penelitian, riset maupun pengamatan tersebut disajikan dalam bentuk tabel/diagram, selanjutnya sering kali diminta suatu uraian, penjelasan, hingga dapat ditarik kesimpulan tentang persoalan yang diteliti. Sebelum kesimpulan dibuat, informasi atau data yang telah terkumpul terlebih dahulu dipelajari, diolah dan dianalisis sehingga nantinya dapat ditarik sebuah kesimpulan.

Ini semua ternyata merupakan pengetahuan tersendiri yang disebut statistika. Jadi statistika adalah ilmu pengetahuan yang mempelajari tentang cara-cara pengumpulan data, penyajian, pengolahan, dan penganalisisan sampai pada penarikan kesimpulan dari suatu keseluruhan data hasil pengamatan di lapangan.

Berdasarkan pengertian diatas, statistika dibedakan menjadi dua jenis, yaitu statistika deskriptif dan statistika induktif atau biasa disebut statistika inferensia:

1. Statistika deskriptif adalah bagian statistika yang mempelajari cara penyusunan dan penyajian data untuk mendeskripsikan atau memberikan gambaran yang diteliti sebagaimana adanya tanpa menarik sebuah kesimpulan atau generalisasi. Cara-cara penyajian data pada statistika deskriptif ini dikemukakan dalam bentuk tabel maupun diagram, penentuan rata-rata (mean), modus, median, rentang serta simpangan baku.
2. Statistika inferensia (induktif) adalah bagian statistika yang tidak hanya sebatas menyajikan data melainkan mempelajari cara penarikan kesimpulan (inference) mengenai populasi berdasarkan data pada sampel, yang sebelumnya telah dilakukan suatu dugaan yang diperoleh dari statistika deskriptif. Kesimpulan ini lazimnya dilakukan dalam rangka menguji hipotesis penelitian yang telah dirumuskan. Dengan

kata lain untuk mempelajari statistik inferensial seseorang diperlukan mempelajari statistik deskriptif terlebih dahulu.

1.2 Data Statistik

Setiap kegiatan yang berkaitan dengan statistik, selalu berhubungan dengan data. Data adalah segala informasi, keterangan, tentang suatu keadaan, gejala atau peristiwa. Keterangan atau informasi tersebut dapat berbentuk kategori (non-angka), misalnya: laki-laki, perempuan, berhasil, gagal, setuju, baik, senang, dan lain sebagainya, atau dapat berbentuk bilangan (angka). Data yang berbentuk angka disebut data kuantitatif, fakta dan fenomena dalam data ini tidak dinyatakan dalam bahasa alami, melainkan dalam numerik (Priadana, Sidik., Sunarsi, 2021).

Sedangkan data yang berbentuk kategori disebut data kualitatif. Agar diperoleh suatu kesimpulan yang tepat dan benar, maka data yang dikumpulkan sebaiknya memenuhi empat syarat yaitu: obyektif (sesuai dengan keadaan sebenarnya), dapat mewakili (representatif), update, dan relevan. Data angka (kuantitatif) ada yang berjenis nominal (diskrit) dan ada yang berjenis kontinum. Data kontinum terdiri dari data ordinal, interval, atau rasio.

Dengan demikian, data penelitian dapat dikelompokkan sebagai berikut:

1. Data kualitatif (non-angka)
2. Data kuantitatif (angka)
 - a. Data Nominal (diskret)
 - b. Data Kontinum:
 - Ordinal
 - Interval
 - Rasio

Data nominal (diskrit) data yang diperoleh dari menghitung (membilang). Oleh karena itu angka data nominal tidak mungkin berbentuk pecahan (desimal), dan tidak mungkin bersambung (kontinum). Data nominal adalah data statistik yang memuat angka yang tidak mempunyai arti apa-apa (Malik and Chusni, 2018). Artinya angka ini hanya sebagai label atau identitas yang membedakan satu

objek/subjek dengan objek/subjek yang lain. Contoh; jumlah siswa, jumlah hari libur, jenis kelamin, jenis pekerjaan, agama, dan lain-lain.

Data kontinum, yakni data yang diperoleh dengan cara mengukur sehingga angkanya dimungkinkan berbentuk pecahan (desimal), dan bersambung (kontinu).

Data ordinal, yaitu data yang mempunyai daya berjenjang, peringkat, atau ranking. Angka pada data ini berfungsi untuk menunjukkan adanya penjenjangan kualitatif atau secara sederhana angka ini menunjukkan sebuah peringkat (Abigail Soesana, Hani Subakti, dkk 2023) Sebagai contoh jenjang pendidikan, pemberian ranking di kelas, dan lain-lain.

Data Interval, pada dasarnya data interval merupakan hasil pengukuran ordinal yang memiliki jarak antar jenjang yang tetap (selalu sama) dan memiliki nilai nol (0) relatif. Artinya nilai nol tetap berarti, bukan berarti kosong sama sekali.

Data Rasio, yaitu data yang memiliki sifat paling lengkap, yaitu membedakan, mengurutkan, menjumlahkan dan mengalikan (Sukardi, 2018). Data ini memiliki jarak sama dan nol (0) mutlak. Artinya bila suatu hasil pengukuran menunjukkan nilai nol (0), maka berarti objek terukur tersebut memang kosong (zero). Misal hasil suatu pengukuran 0 kg berarti tidak ditemukan berat sama sekali.

1.3 Variabel Penelitian

Secara etimologi variabel berasal dari bahasa Inggris “variable” yang berarti peubah, faktor tak tetap, gejala yang dapat diubah-ubah, sesuatu yang bervariasi, warna-warni, tidak sama, atau tidak satu jenis. Sedang secara terminologi, variabel dapat diartikan sebagai suatu konsep yang mempunyai keragaman atau variasi yang dapat diberi nilai atau bilangan. Variabel penelitian adalah segala sesuatu yang berbentuk apa saja yang ditetapkan oleh peneliti untuk dipelajari sehingga diperoleh informasi tentang hal tersebut, kemudian ditarik kesimpulannya (Sugiyono, 2009).

Menurut Kerlinger (2006), variabel adalah konstruk atau sifat yang akan dipelajari, yang mempunyai nilai bervariasi. Variabel adalah simbol atau lambang yang padanya kita letakkan sembarang nilai atau bilangan. Dengan demikian, variabel penelitian adalah segala sesuatu yang akan menjadi objek

pengamatan penelitian, dengan kata lain variabel merupakan fenomena yang menjadi pusat perhatian penelitian untuk di observasi atau diukur.

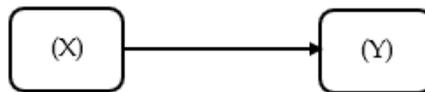
Dalam suatu penelitian, jumlah variabel bisa lebih dari satu. Variabel-variabel tersebut saling berhubungan. Jika ditinjau dari konteks hubungannya, maka setidaknya terdapat lima macam variabel yakni, variabel bebas, variabel terikat, variabel moderator, variabel intervening, dan variabel kontrol.

Variabel Bebas (Independent Variable) Dan Variabel Terikat (Dependent Variable)

Variabel bebas adalah variabel yang nilainya mempengaruhi variabel lainnya yaitu variabel terikat. Dengan kata lain variabel bebas adalah variabel yang menjadi penyebab munculnya perubahan variabel lain. Dengan demikian, variabel yang terpengaruh (dipengaruhi) atau yang menjadi akibat dari adanya variabel bebas tersebut disebut variabel terikat (variabel dependent, respon, output, kriteria).

Sebagai contoh dalam penelitian yang berjudul: Pengaruh motivasi belajar terhadap prestasi siswa. Dalam hal ini motivasi belajar diduga dapat mempengaruhi prestasi belajar. Maka motivasi belajar sebagai variabel bebas, dikatakan bebas sebab nilainya dapat berubah-ubah dan setiap perubahan itu mempengaruhi nilai variabel terikatnya yakni prestasi belajar.

Pengaruh atau hubungan variabel X (bebas) dengan variabel Y (terikat) dapat digambarkan sebagai berikut:



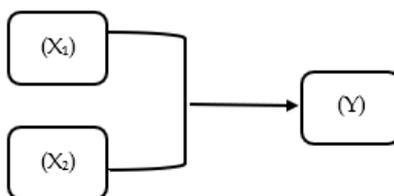
Gambar 1.1: Hubungan Variabel X dengan Variabel Y

Variabel Moderator

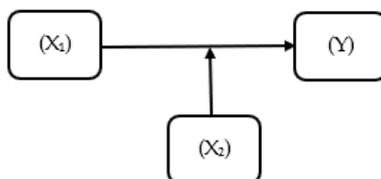
Variabel moderator yaitu merupakan variabel yang juga mempengaruhi variabel terikat, dalam hal ini variabel moderator dapat memperkuat atau memperlemah hubungan antara variabel bebas (X1) dan variabel terikat (Y). Variabel ini sering disebut sebagai variabel bebas kedua (X2). Jadi di sini ada tiga variabel penelitian, yaitu dua variabel bebas dan satu variabel terikat. Sebagai contoh dalam penelitian yang berjudul: Pengaruh strategi pembelajaran dan motivasi belajar terhadap prestasi belajar.

Dalam mempengaruhi variabel terikat (Y) yaitu prestasi belajar, variabel bebas pertama (X_1) yaitu strategi pembelajaran, dapat bersama-sama dengan variabel bebas kedua (X_2) yaitu motivasi, sehingga hasil pengaruh variabel bebas (X_1 dan X_2) terhadap variabel terikat (Y) akan lebih besar. Sebaliknya dapat pula variabel bebas kedua (X_2) dijadikan sebagai pengontrol terhadap kemurnian pengaruh variabel bebas pertama (X_1) terhadap variabel terikat (Y).

Hal tersebut dapat digambarkan sebagai berikut:



Gambar 1.2: Hubungan Variabel X_1 dan X_2 Secara Bersama-sama Dengan Variabel Y



Gambar 1.3: Hubungan Variabel X_1 Dengan Variabel Y dan X_2 Sebagai Variabel Moderator

Variabel Intervening

Variabel *Intervening* merupakan variabel yang secara teoritis mempengaruhi hubungan antara variabel bebas dan variabel terikat, namun ia tidak terukur (sulit diukur). Variabel ini disebut juga sebagai variabel mediasi atau penghubung.

Sebagai contoh kecerdasan seseorang akan berpengaruh terhadap peningkatan prestasi belajarnya. Namun kecerdasan itu tidak akan terlalu berpengaruh manakala ia sedang sakit, atau sedang dalam kondisi frustrasi, atau pikiran kacau. Kondisi sakit, frustrasi, atau pikiran kacau ini disebut variabel *intervening* (variabel yang memiliki intervensi, memiliki pengaruh juga). Namun tingkat sakit, frustrasi, atau pikiran kacau ini sulit diukur (Mundir, 2012).



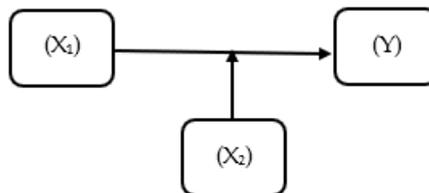
Gambar 1.4: Hubungan Variabel X1 Dengan Variabel Y dan X2 Sebagai Variabel Intervening

Variabel Kontrol

Variabel kontrol yaitu variabel yang dikendalikan atau dinetralkan pengaruhnya oleh peneliti, sehingga tidak akan mempengaruhi hubungan variabel bebas dengan variabel terikat dengan tujuan untuk mencegah adanya hasil perhitungan bias. Variabel ini sering digunakan oleh peneliti saat akan melakukan penelitian yang bersifat komparatif atau membandingkan variabel satu dengan variabel yang lain melalui penelitian eksperimen. Sebagai contoh, seorang peneliti ingin membandingkan kemampuan praktik mengajar mahasiswa Jurusan Pendidikan Matematika dari Institut Keguruan dan Ilmu Pendidikan (IKIP) dan Universitas A.

Sebelumnya peneliti telah memberikan pelatihan (Microteaching), peneliti ingin melihat pengaruh pelatihan yang diberikan terhadap keterampilan mengajar disekolah mahasiswa antara mahasiswa dari Institut Keguruan dan Ilmu Pendidikan (IKIP) dan Universitas A. Sebagai variabel kontrol, ditentukan misalnya kelas tempat mengajar, dan media yang digunakan.

Berikut ini gambaran visualnya:



Gambar 1.5: Hubungan Variabel X1 Dengan Variabel Y dan X2 Sebagai Variabel Kontrol

Untuk dapat menentukan kedudukan variabel apakah ia sebagai variabel bebas, terikat, moderator, intervening, atau kontrol maka peneliti harus melihat konteksnya dengan memperhatikan konsep teoritis yang mendasari maupun hasil pengamatan empiris di tempat penelitian. Sebelum peneliti memilih variabel apa yang akan diteliti, maka terlebih dahulu perlu ia melakukan kajian

teoritis dan melakukan studi pendahuluan pada objek yang akan diteliti, sehingga nantinya diperoleh hasil yang sesuai dengan tujuan peneliti.

1.4 Populasi dan Sampel

Salah satu fungsi statistik inferensial yaitu menarik kesimpulan tentang suatu variabel diteliti berdasarkan data sampel yang kemudian digeneralisasikan pada populasi. Populasi adalah seluruh objek yang menjadi sasaran penelitian yang memiliki sifat yang sama. Sedangkan sampel adalah bagian dari populasi yang diambil untuk dijadikan objek pengamatan dan dijadikan dasar dalam pengambilan keputusan. Sampel yang baik adalah sampel yang dapat mewakili populasi, artinya sampel harus memiliki sifat-sifat, karakteristik, dan ciri-ciri yang dimiliki oleh populasi, sehingga dapat mewakili populasi (representatif).

Jika sampel tidak representatif, maka secara ilmiah peneliti tidak diperkenankan melakukan generalisasi. Karena nantinya generalisasi akan melenceng dari realitas sebenarnya. Sehingga pada kasus sampel yang tidak representatif kesimpulan yang diambil peneliti hanya dapat berlaku bagi sampel itu sendiri. Oleh karena itu, peneliti perlu memiliki dasar pemikiran yang tepat dalam penentuan sampel.

Terdapat beberapa teknik pengambilan sampel, diantaranya adalah sebagai berikut:

1. Sampel acak sederhana (sample random sampling)

Pengambilan sampel pada teknik ini dilakukan dengan mengambil sampel secara acak melalui cara yang sederhana, seperti pengundian atau menggunakan pendekatan bilangan acak. Pengambilan sampel dari semua anggota populasi dilakukan secara acak tanpa memperhatikan strata yang ada dalam anggota populasi. Hal ini dapat dilakukan jika anggota populasi dianggap homogen.

2. Sampel sistematis (systematic random sampling)

Teknik sampling ini merupakan teknik penarikan sampel dengan cara menetapkan sampel awal secara acak kemudian sampel selanjutnya dipilih secara sistematis berdasarkan pola tertentu. Teknik penarikan sampel ini mengambil setiap unsur ke- k dalam populasi, untuk

dijadikan contoh dengan titik awal ditentukan secara acak diantara k unsur yang pertama. Sebagai contoh jumlah anggota populasi sebanyak 150 orang. Anggota populasi diberi nomor urut dari nomor 1 sampai nomor 150. Selanjutnya pengambilan sampel dilakukan dengan memilih nomor urut ganjil, atau genap saja, atau kelipatan dari bilangan tertentu, seperti bilangan 5 dan lainnya.

3. Sampel stratifikasi (stratified random sampling)

Teknik pengambilan sampel ini melakukan penentuan sampel penelitian dengan menetapkan pengelompokan anggota populasi dalam kelompok-kelompok tingkatan tertentu seperti tingkat tinggi, sedang, dan rendah. Teknik ini digunakan apabila populasi mempunyai anggota/karakteristik yang tidak homogen dan berstrata secara proporsional. Sebagai contoh suatu organisasi mempunyai personil yang terdiri dari latar belakang pendidikan yang berbeda yaitu: SMP, SMA, S1, dan S2 dengan jumlah setiap kelas pendidikan juga berbeda. Jumlah anggota populasi untuk setiap strata pendidikan tidak sama atau bervariasi. Jumlah sampel yang harus diambil harus meliputi strata pendidikan yang ada yang diambil secara proporsional (Nuryadi, 2017)

4. Sampel kelompok (cluster random sampling)

Teknik pengambilan sampel ini menentukan sampel berdasar kelompok wilayah dari anggota populasi penelitian. Pada teknik ini subyek penelitian akan dikelompokkan menurut area atau wilayah anggota populasi. Sebagai contoh Indonesia terdiri dari 38 provinsi, sampel yang akan diambil sebanyak 8 provinsi, maka pengambilan 8 provinsi dari 38 provinsi dilakukan secara random. Suatu hal yang perlu diingat adalah bahwa karena provinsi yang ada di Indonesia juga berstrata, maka pengambilan sampel untuk 8 provinsi juga dilakukan dengan menggunakan teknik *stratified random sampling*. Teknik cluster sampling dilakukan dalam dua tahap yaitu: (1) menentukan sampel daerah, dan (2) menentukan orang-orang yang ada pada daerah dengan cara sampling juga.

Bab 2

Penyajian Data Penelitian

2.1 Pendahuluan

Pada bab ini akan dibahas mengenai penyajian data dalam bentuk laporan penelitian. Selain sebagai informasi kepada pembaca mengenai teknik penyajian data sebagai bentuk visualisasi dari hasil penelitian yang akan kita lakukan, pada bab ini juga akan dijelaskan cara menyusun sistematika penelitian, yaitu mengenai susunan isi dari laporan penelitian dimulai dari bagian awal hingga bagian akhir penulisan. Diharapkan dengan membaca pada bab ini, pembaca dapat lebih memahami bagaimana cara membuat laporan penelitian sesuai dengan sistematika penulisan yang baik dan benar serta berlaku secara umum.

Penyajian data adalah suatu kegiatan dalam pembuatan laporan hasil penelitian sehingga dapat dipahami dan dianalisis sesuai dengan tujuan yang diinginkan. Data yang disajikan harus sederhana dan jelas agar mudah dibaca oleh pembaca.

Tujuan penyajian data adalah:

1. memberikan sebuah gambaran mengenai kejadian yang merupakan hasil dari penelitian atau observasi yang telah dilakukan;
2. data yang disajikan lebih mudah dimengerti;
3. memudahkan peneliti dalam membuat analisis data;

4. mendapatkan proses pengambilan keputusan dan kesimpulan yang lebih tepat, cepat dan akurat.

Pada bab ini akan dijelaskan bentuk penyajian data sesuai dengan hasil penelitian menggunakan pendekatan kuantitatif maupun kualitatif.

2.2 Penyajian Data Dalam Bentuk Tabel

Tabel merupakan kumpulan data yang disusun berdasarkan baris dan kolom. Penyajian data dalam bentuk tabel merupakan susunan angka yang singkat dan jelas dalam baris dan kolom, sehingga memberikan deskripsi atau perbandingan.

Suatu tabel dikatakan lengkap, jika memiliki persyaratan sebagai berikut:

1. Kepala tabel
Kepala tabel terdiri dari nomor tabel dan judul tabel. Nomor tabel menunjukkan urutan tabel tersebut. Sedangkan judul tabel diletakkan di atas tabel dan berisi keterangan “apa”, “dimana” dan “kapan” mengenai isi tabel tersebut.
2. Leher tabel
Leher tabel berisi keterangan atau judul kolom yang harus ditulis dengan singkat dan jelas.
3. Badan tabel
Badan tabel berisi penjelasan mengenai keterangan baris dan kolom. Keterangan baris dan kolom harus mudah dipahami oleh pembaca. Keterangan yang tertulis pada baris di dalam tabel dapat disusun menurut:
 - a. Abjad, misalnya nama-nama orang.
 - b. Geografis, misalnya nama-nama daerah yang ada di suatu wilayah.
 - c. Perkembangan waktu, misalnya urutan hari, bulan atau tahun.
 - d. Tingkatan pendidikan misalnya SD, SMP, SMA atau Perguruan Tinggi.

Pada dasarnya terdapat dua jenis tabel yang sering digunakan dalam deskripsi data, yaitu tabel kontingensi dan tabel distribusi frekuensi.

Tabel Kontingensi (Tabel Faktorial)

Tabel kontingensi sering dikatakan tabel faktorial, yaitu tabel yang mendeskripsikan dua variabel (faktor) atau lebih dalam satu perpaduan baris dan kolom (Mundir, 2012).

Berikut contoh tabel kontingensi (tabel faktorial) $2 \times 4 =$ (baris = 2 sel, kolom = 4 sel):

Tabel 2.1: Jumlah Siswa di Kota Medan Berdasarkan Jenis Kelamin dan Tingkat Pendidikan

Jenis Kelamin	Tingkat Pendidikan				Jumlah
	TK	SD	SMP	SMA/SMK	
Laki-laki	1.000	1.750	2.500	5.500	10.750
Perempuan	1.500	2.000	2.750	6.250	12.500
Jumlah	2.500	3.750	5.250	11.750	23.250

Tabel Distribusi Frekuensi

Tabel distribusi frekuensi adalah metode statistik untuk menyusun data dengan cara membagi nilai-nilai observasi data ke dalam kelas-kelas dengan interval tertentu.

Perhatikan nilai MID Mata Kuliah Statistik dari 60 mahasiswa berikut ini:

23 60 79 32 57 74 52 70 82 36

80 77 81 95 41 65 92 85 55 76

52 10 64 75 78 25 80 98 81 67

41 71 83 54 64 72 78 62 74 43

60 78 89 76 84 48 84 90 15 79

34 67 17 82 69 74 63 80 85 61

Pada tabel distribusi frekuensi, data dikelompokkan ke dalam kelas-kelas tertentu. Pada umumnya, banyak kelas pada suatu data antara 5 sampai dengan 15 kelas. Setiap kelas memiliki batas kelas yaitu nilai yang terdapat pada ujung-ujung suatu kelas.

Batas kelas terdiri atas batas bawah kelas dan batas atas kelas:

1. batas bawah kelas yaitu nilai ujung bawah pada suatu kelas, dan;
2. batas atas kelas yaitu nilai ujung atas pada suatu kelas.

Selain memiliki batas kelas, setiap kelas juga memiliki panjang kelas. Panjang kelas adalah selisih tepi atas dengan tepi bawah pada kelas tersebut. Jika data dicatat teliti sampai satuan maka tepi atas kelas sama dengan batas atas kelas ditambah 0,5 dan tepi bawah sama dengan batas bawah kelas dikurangi 0,5.

Jika data dicatat teliti sampai satu satuan desimal maka tepi atas kelas sama dengan batas atas kelas ditambah 0,05 dan tepi bawah sama dengan batas bawah kelas dikurangi 0,05 dan seterusnya. Setiap data yang telah disusun ke dalam kelas-kelas tertentu harus memiliki panjang kelas yang sama. Selain batas kelas, terdapat juga titik tengah kelas yaitu setengah dari jumlah batas bawah dan batas atas. Titik tengah kelas merupakan nilai yang dianggap mewakili kelas tersebut.

Untuk menyusun sebuah tabel distribusi berkelompok, lakukanlah langkah-langkah berikut:

- Langkah ke-1: Menentukan jangkauan data (J), yaitu nilai data terbesar dikurangi data terkecil.

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

- Langkah ke-2: Menentukan banyak kelas interval (k). Kelas interval adalah pembagian data ke dalam interval tertentu. Untuk menentukan banyak kelas, gunakanlah aturan Sturges, yaitu:

$$k = 1 + 3,3 \log n$$

dengan: k = bilangan bulat

n = banyaknya data

- Langkah ke-3: Menentukan panjang kelas interval (p)

$$p = \frac{\text{Jangkauan } (J)}{\text{Banyaknya kelas } (k)}$$

- Langkah ke-4: Menentukan batas kelas interval
- Langkah ke-5: Menentukan titik tengah interval

$$\text{Titik tengah} = \frac{1}{2} (\text{batas bawah} + \text{batas atas})$$

- Langkah ke-6: Menentukan tabel distribusi frekuensi berkelompok disertai nilai tepi atas dan tepi bawah.

Agar Anda dapat membuat tabel distribusi frekuensi berkelompok, pelajari contoh berikut.

Contoh Soal:

Buatlah tabel distribusi frekuensi untuk data nilai MID Mata Kuliah Statistik dari 60 mahasiswa berikut ini:

23 60 79 32 57 74 52 70 82 36

80 77 81 95 41 65 92 85 55 76

52 **10** 64 75 78 25 80 **98** 81 67

41 71 83 54 64 72 78 62 74 43

60 78 89 76 84 48 84 90 15 79

34 67 17 82 69 74 63 80 85 61

- Langkah ke-1: Menentukan jangkauan (J)

Nilai terbesar (X_{\max}) = 98

Nilai terkecil (X_{\min}) = 10

$$J = X_{\max} - X_{\min} = 98 - 10 = 88$$

- Langkah ke-2: Menentukan banyak kelas interval (k)

$k = 1 + 3,3 \log n = 1 + 3,3 \log (60) = 1 + 3,3 (1,78) = 6,86 = 7$. Oleh karena itu, banyak kelas interval (k) dibulatkan menjadi 7.

- Langkah ke-3: Menentukan panjang kelas (p)

$$p = \frac{J}{k} = \frac{88}{7} = 12,57 = 13$$

Oleh karena itu, panjang kelas (p) dibulatkan menjadi 13.

- Langkah ke-4: Menentukan batas kelas interval.

Batas bawah kelas pertama adalah 10.

Kelas I : $10 + 12 = 22$ sehingga kelas pertama memiliki interval 10 – 22

Kelas II : $23 + 12 = 35$ sehingga kelas kedua memiliki interval 23 – 35

Kelas III : $36 + 12 = 48$ sehingga kelas ketiga memiliki interval 36 – 48

Kelas IV : $49 + 12 = 61$ sehingga kelas kedua memiliki interval 49 – 61

Kelas V : $62 + 12 = 74$ sehingga kelas kedua memiliki interval 62 – 74

Kelas VI : $75 + 12 = 87$ sehingga kelas kedua memiliki interval 75 – 87

Kelas VII : $88 + 12 = 100$ sehingga kelas kedua memiliki interval 88 – 100

- Langkah ke-5: Menentukan titik tengah kelas

$$\text{Titik tengah kelas I} = \frac{10+22}{2} = 16$$

$$\text{Titik tengah kelas II} = \frac{23+35}{2} = 29$$

$$\text{Titik tengah kelas III} = \frac{36+48}{2} = 42$$

$$\text{Titik tengah kelas IV} = \frac{49+61}{2} = 55$$

$$\text{Titik tengah kelas V} = \frac{62+74}{2} = 68$$

$$\text{Titik tengah kelas VI} = \frac{75+87}{2} = 81$$

$$\text{Titik tengah kelas VII} = \frac{88+100}{2} = 94$$

- Langkah ke-6: Menentukan tabel distribusi frekuensi berkelompok disertai nilai tepi atas dan bawah.

Tabel 2.2: Tabel Distribusi Frekuensi Nilai MID Mata Kuliah Statistik dari 60 Mahasiswa

Nilai	Nilai Tengah	Tepi Bawah	Tepi Atas	Frekuensi
10 – 22	16	9,5	22,5	3
23 – 35	29	22,5	35,5	4
36 – 48	42	35,5	48,5	5
49 – 61	55	48,5	61,5	8
62 – 74	68	61,5	74,5	14
75 – 87	81	74,5	87,5	21
88 – 100	94	87,5	100,5	5
Jumlah				60

Tabel Distribusi Frekuensi Relatif

Distribusi frekuensi dapat diubah menjadi distribusi frekuensi relatif. Perubahan ini dapat dilakukan dengan membagi setiap frekuensi kelas dengan n (jumlah frekuensi seluruhnya). Tabel distribusi frekuensi relatif merupakan tabel distribusi frekuensi yang dinyatakan dalam bentuk persen (Somantri, 2006). Tabel distribusi frekuensi relatif juga dinamakan tabel persentase. Dikatakan “frekuensi relatif” sebab frekuensi yang disajikan disini bukanlah frekuensi yang sebenarnya, melainkan frekuensi yang dituangkan dalam bentuk angka persen (Sudijono, 2009).

Nilai frekuensinya **tidak** dinyatakan dalam bentuk **angka mutlak** tapi dalam bentuk **angka persentase (%)** atau **angka relatif**.

Rumus mencari frekuensi relatif adalah:

$$f_{\text{relatif kelas-}i} = \frac{f_{\text{mutlak kelas-}i}}{n} \times 100\%$$

Contoh menggunakan soal dan tabel distribusi frekuensi mutlak dari contoh soal sebelumnya sebagai acuan).

Tabel 2.3: Tabel Distribusi Frekuensi Relatif

Nilai	Frekuensi	Frekuensi Relatif (%)
10 – 22	3	5%
23 – 35	4	6,67%
36 – 48	5	8,33%
49 – 61	8	13,33%
62 – 74	14	23,33%
75 – 87	21	35%
88 – 100	5	8,33%
Jumlah	60	100%

Tabel Distribusi Frekuensi Kumulatif

Tabel distribusi frekuensi kumulatif disusun dengan cara menjumlahkan frekuensi. Tabel distribusi frekuensi kumulatif terdiri atas tabel distribusi frekuensi kumulatif kurang dari dan tabel distribusi frekuensi kumulatif lebih dari. Frekuensi kumulatif kurang dari menyatakan jumlah frekuensi semua nilai data yang kurang dari atau sama dengan tepi bawah kelasnya. Adapun frekuensi kumulatif lebih dari menyatakan jumlah frekuensi semua nilai data yang lebih dari atau sama dengan tepi bawah kelasnya.

Dari data nilai MID Mata Kuliah Statistik dari 60 mahasiswa dapat dibuat dalam tabel distribusi frekuensi kumulatif sebagai berikut:

Tabel 2.4: Tabel Distribusi Frekuensi Kumulatif Kurang Dari

Nilai	Frekuensi Kumulatif	Frekuensi Kumulatif Relatif (%)
kurang dari 10	0	0%
kurang dari 23	3	5%
kurang dari 36	$3 + 4 = 7$	11,67%
kurang dari 49	$7 + 5 = 12$	20%
kurang dari 62	$12 + 8 = 20$	33,33%
kurang dari 75	$20 + 14 = 34$	56,67%
kurang dari 88	$34 + 21 = 55$	91,67%
kurang dari 101	$55 + 5 = 60$	100%

Tabel 2.5: Tabel Distribusi Frekuensi Kumulatif Lebih Dari

Nilai	Frekuensi Kumulatif	Frekuensi Kumulatif Relatif (%)
10 atau lebih	60	60%
23 atau lebih	$60 - 3 = 57$	95%
36 atau lebih	$57 - 4 = 53$	88,33%
49 atau lebih	$53 - 5 = 48$	80%
62 atau lebih	$48 - 8 = 40$	66,67%
75 atau lebih	$40 - 14 = 26$	43,33%
88 atau lebih	$26 - 21 = 5$	8,33%
101 atau lebih	$5 - 5 = 0$	0%

2.3 Penyajian Data dalam Bentuk Diagram/Grafik

Data statistik dapat disajikan dalam bentuk tabel atau grafik. Penyajian data dalam bentuk grafik umumnya lebih menarik perhatian dan mengesankan. Penyajian data statistik secara grafis mempunyai berbagai fungsi. Pada hakikatnya grafik dan tabel seyogyanya digunakan secara bersama-sama. Grafik statistik lebih mudah dan menarik dibanding tabel statistik.

Selain itu, grafik dapat melukiskan suatu peristiwa secara lebih mengesankan dan tidak membosankan. Namun demikian, penyajian secara grafis hanyalah

bersifat aprosimatif. Angka-angka yang pasti dan rinci tentang suatu peristiwa dimuat dalam tabel. Oleh karena itu, analisis dan interpretasi data umumnya dilakukan terhadap data yang terdapat dalam tabel statistik (Nuryadi, dkk, 2017).

Diagram atau grafik adalah gambar yang menunjukkan data secara visual dan didasarkan pada nilai-nilai pengamatan aslinya maupun tabel yang telah dibuat sebelumnya (Somantri, 2006). Adapun tujuan menyajikan data dalam bentuk diagram/grafik adalah membantu visualisasi data yang besar dan kompleks menjadi lebih sederhana.

2.3.1 Jenis Diagram/Grafik

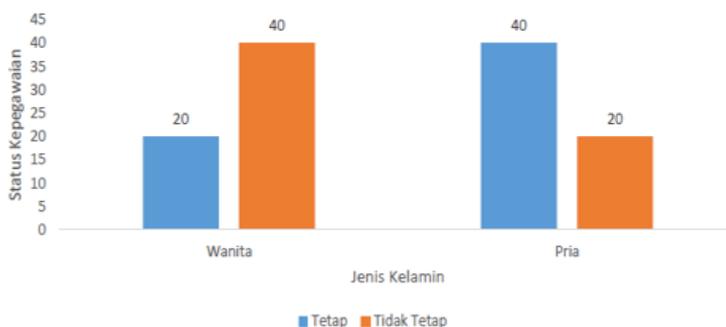
Jenis diagram/ grafik adalah sebagai berikut:

1. Diagram Batang

Hasan (2011) menyatakan grafik batang atau balok adalah grafik data berbentuk persegi panjang yang lebarnya sama dan dilengkapi dengan skala ukuran data. Diagram batang digunakan untuk data yang berbentuk kategori.

Contoh:

Berikut ini adalah diagram/grafik batang mengenai jumlah dosen Universitas Dharmawangsa berdasarkan jenis kelamin dan status kepegawaiannya pada tahun 2020.



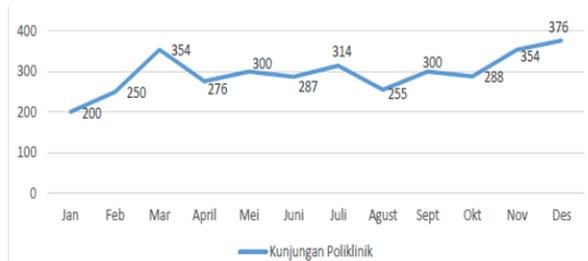
Gambar 2.1: Diagram/Grafik Batang Mengenai Jumlah Dosen Universitas Dharmawangsa Berdasarkan Jenis Kelamin dan Status Kepegawaiannya Pada Tahun 2020

Diagram Garis

Hasan (2011) menyatakan grafik garis adalah grafik data yang berupa garis, diperoleh dari beberapa ruas garis yang menghubungkan titik-titik pada bilangan.

Contoh:

Berikut ini adalah diagram/grafik garis mengenai kunjungan poliklinik RS ABC pada Tahun 2020.



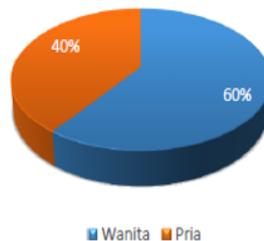
Gambar 2.2: Diagram/Grafik Garis Mengenai Kunjungan Poliklinik RS ABC pada Tahun 2020

Diagram Lingkaran

Hasan (2011) menyatakan grafik lingkaran (pie diagram) adalah grafik data berupa lingkaran yang telah dibagi menjadi bagian sesuai data tersebut.

Contoh:

Berikut ini adalah diagram/grafik lingkaran mengenai persentase dosen Universitas Dharmawangsa berdasarkan jenis kelamin pada tahun 2020.



Gambar 2.3: Persentase Dosen Universitas Dharmawangsa Berdasarkan Jenis Kelamin Pada Tahun 2020

2.3.2 Grafik Tabel Distribusi Frekuensi

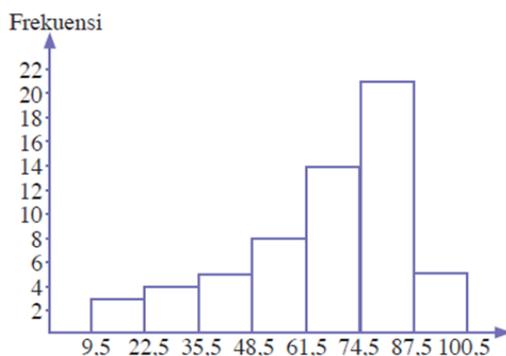
Ada beberapa grafik tabel distribusi frekuensi:

Histogram

Histogram sangat berhubungan dengan tabel distribusi frekuensi. Histogram adalah grafik yang menyajikan data dari tabel distribusi frekuensi. Sumbu horizontal pada histogram menyatakan suatu kelas dan sumbu vertikal menyatakan frekuensi. Jika data yang disajikan dalam bentuk histogram adalah data dari tabel distribusi frekuensi maka sumbu horizontal menyatakan kelas-kelas, sedangkan sumbu vertikal menyatakan frekuensi dari kelas-kelas tersebut.

Histogram digambarkan dengan persegi panjang dimana antara persegi panjang satu dengan lainnya tidak terdapat jarak. Untuk histogram dari data pada tabel distribusi frekuensi berkelompok, setiap persegi panjang mewakili kelas tertentu. Lebar persegi panjang menunjukkan panjang kelas, sedangkan tinggi persegi panjang menyatakan frekuensi kelas.

Untuk lebih jelasnya, perhatikan histogram untuk data nilai MID Mata Kuliah Statistik sebagai berikut:



Gambar 2.4: Histogram

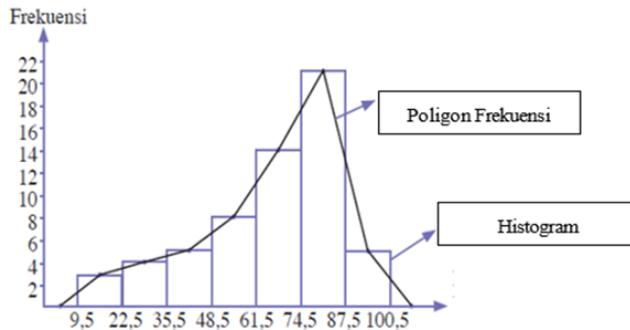
Langkah-langkah untuk membuat histogram adalah sebagai berikut:

1. Buatlah sumbu horizontal dan sumbu vertikal yang saling berpotongan tegak lurus.
2. Buatlah skala yang sama besar untuk kedua sumbu. Skala pada sumbu horizontal tidak perlu sama dengan skala pada sumbu vertikal.

3. Untuk histogram yang menyajikan data dari tabel distribusi frekuensi, tuliskan tepi bawah kelas dan tepi atas kelas pada sumbu horizontal dan frekuensi setiap kelas pada sumbu vertikal.
4. Gambarlah persegi panjang untuk setiap kelas pada data yang berasal dari tabel distribusi frekuensi berkelompok. Tinggi persegi panjang menunjukkan frekuensi dari nilai dan kelas pada data.

Poligon Frekuensi

Poligon frekuensi adalah grafik yang dihasilkan dengan menghubungkan puncak dari masing-masing nilai tengah kelas histogram. Poligon frekuensi dari data nilai MID Mata Kuliah Statistik disajikan sebagai berikut:



Gambar 2.5: Poligon Frekuensi

Ogive

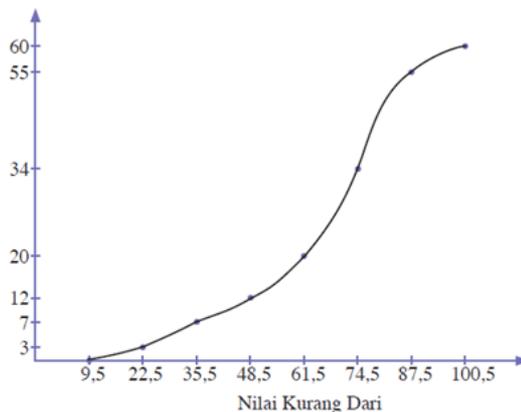
Ogive adalah grafik yang menyajikan data dari tabel distribusi frekuensi kumulatif. Titik-titik pada ogive adalah pasangan tepi kelas dengan nilai frekuensi kumulatif. Titik-titik ini dihubungkan dengan kurva mulus. Ogive terdiri atas ogive positif dan ogive negatif. Ogive positif menyajikan data yang berasal dari tabel distribusi frekuensi kumulatif kurang dari. Adapun ogive negatif menampilkan data yang berasal dari tabel distribusi frekuensi kumulatif lebih dari.

Langkah-langkah membuat ogive sebagai berikut:

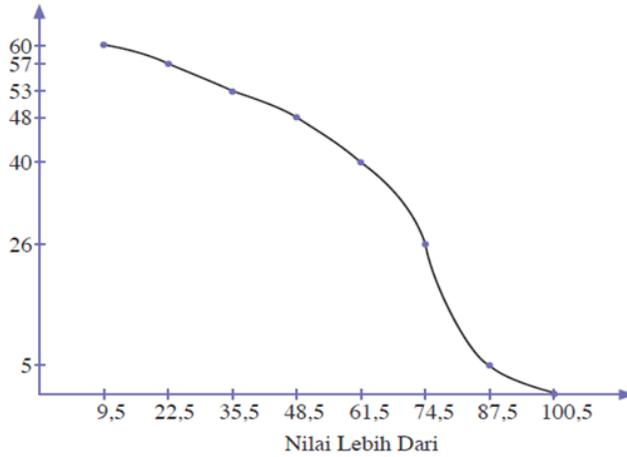
1. Buatlah sumbu horizontal dan sumbu vertikal yang saling berpotongan tegak lurus.

2. Buat skala untuk kedua sumbu. Skala untuk sumbu horizontal tidak perlu sama dengan sumbu vertikal.
3. Pada sumbu horizontal, tuliskanlah bilangan-bilangan yang menyatakan tepi bawah kelas dan tepi atas kelas. Pada sumbu vertikal, tuliskanlah bilangan-bilangan yang menyatakan frekuensi kumulatif.
4. Gambarlah titik atau noktah yang menyatakan pasangan tepi bawah kelas dengan frekuensi kumulatifnya. Cara menggambarkan titik ini serupa dengan menggambar pasangan berurutan (x,y) pada bidang koordinat Cartesius. Untuk ogive positif, tepi bawah kelas menunjukkan sumbu-x dan frekuensi kumulatif kurang dari menunjukkan sumbu-y. Untuk ogive negatif, tepi bawah kelas menunjukkan sumbu-x dan frekuensi kumulatif lebih dari menunjukkan sumbu-y.
5. Hubungkanlah titik-titik yang diperoleh dengan kurva mulus.

Dengan melakukan langkah-langkah untuk membuat ogive akan diperoleh kurva ogive positif dan kurva ogive negatif untuk data nilai MID Mata Kuliah Statistik dari 60 mahasiswa. Gambar berikut menunjukkan kurva ogive positif dan ogive negatif untuk data nilai MID Mata Kuliah Statistik dari 60 mahasiswa.



Gambar 2.6: Ogive Kurang Dari



Gambar 2.7: Ogive Lebih Dari

Bab 3

Distribusi Frekuensi

3.1 Distribusi Frekuensi Data Tunggal

Data hasil penelitian yang telah diperoleh pertama kali masih berbentuk data mentah yang belum dapat memberikan informasi yang berarti, hal tersebut dikarenakan data masih tersaji secara acak. Istilah “Distribusi” digunakan dalam statistik untuk menunjukkan adanya penyebaran data sedangkan “frekuensi” menunjukkan banyaknya data yang muncul. Distribusi frekuensi adalah data yang tersusun menurut kelas-kelas interval tertentu atau menurut kategori tertentu dalam sebuah tabel.

Distribusi frekuensi data tunggal, disajikan frekuensi dari data angka dimana angka yang ada tidak dikelompokkan. Atau dengan kata lain tidak ada pengelompokan data pada kolom pertama di tabel distribusi frekuensi.

Misalkan, diperoleh data skor hasil pengukuran terhadap kemampuan berbahasa Indonesia 40 orang mahasiswa:

65 61 60 68 63 63 60 67 67 70
68 62 63 70 69 67 66 66 69 61
61 64 64 62 62 59 64 65 65 67
69 65 71 65 64 64 65 66 64 66

Data disajikan dalam bentuk tabel distribusi frekuensi secara tunggal dengan cara diurutkan terlebih dahulu dimulai dari data yang terbesar hingga data terkecil, seperti terlihat pada tabel 3.1 berikut:

Tabel 3.1: Distribusi Frekuensi Skor Kemampuan Berbahasa Indonesia

Skor	Cacahan	Frekuensi
71	/	1
70	//	2
69	///	3
68	///	3
67	///	3
66	////	4
65	###/	6
64	###/	6
63	///	3
62	///	3
61	///	3
60	//	2
59	/	1
Jumlah		40

3.2 Distribusi Frekuensi Data Kelompok

Komponen-komponen yang diperlukan untuk menyusun distribusi frekuensi data kelompok yaitu sebagai berikut:

1. Frekuensi (N)
Jumlah atau banyaknya data dari hasil penelitian.
2. Interval Kelas
Sejumlah nilai variabel yang ada dalam batas kelas tertentu. Setiap interval kelas terdapat batas bawah kelas dan batas atas kelas.
3. Range (R)
Rentang kelas yang diperoleh dengan cara data terbesar dikurangi data terkecil.

4. Titik tengah kelas (M)
Nilai yang terletak di tengah interval kelas yaitu batas bawah kelas ditambah batas atas kelas dibagi 2.
5. Tepi kelas
Batas kelas sesungguhnya yang memuat nilai terkecil dan terbesar yang masuk dalam interval kelas tertentu. Terdapat dua tepi kelas yaitu tepi kelas bawah dan tepi kelas atas.

Langkah-langkah menyusun distribusi frekuensi data kelompok yaitu sebagai berikut:

1. Tentukan data terbesar dan data terkecil kemudian temukan nilai range (R).

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

2. Tentukan banyaknya kelas (k) dengan aturan Sturges, yaitu:

$$k = 1 + 3,3 \log n$$

Dimana n adalah jumlah atau banyaknya data.

Dapat dipakai cara konvensional, banyak kelas yang ideal antara 9 – 12 kelas.

3. Tentukan lebar kelas

$$d = R/k$$

Hasilnya dapat dibulatkan.

4. Susunlah batas bawah kelas pertama hingga batas atas kelas akhir yang bersesuaian.

Misalkan terdapat data skor pengukuran kemampuan matematika dari 50 siswa

68 62 54 70 48 45 55 54 47 63

58 79 38 66 73 48 44 64 59 67

35 33 28 64 56 25 33 48 43 70

50 34 54 48 57 54 56 68 40 38

52 46 36 68 75 50 39 44 50 42

Berdasarkan langkah-langkah menyusun distribusi frekuensi data kelompok maka dapat dilakukan untuk data di atas adalah sebagai berikut:

1. Data terbesar 79 dan data terkecil 25

$$R = 79 - 25 = 54$$

2. Banyak kelas

$$k = 1 + 3,350$$

$$= 1 + 3,3 \cdot 1,69897$$

$$= 6,61 \approx 7$$

3. Lebar kelas

$$\begin{aligned} d &= \frac{R}{k} \\ &= \frac{54}{7} \\ &= 7,7 \\ &\approx 8 \end{aligned}$$

4. Selanjutnya kelompokkan data dalam bentuk tabel distribusi frekuensi.

Tabel 3.2: Distribusi Frekuensi Skor Kemampuan Matematika

Kelas Interval	Cacahan	Frekuensi
25 - 32	//	2
33 - 40	### ////	9
41 - 48	### ####/	11
49 - 56	### ####/	11
57 - 64	### //	7
65 - 72	### //	7
73 - 80	///	3
Jumlah		50

3.3 Diagram Distribusi Frekuensi

Data yang telah disajikan dalam bentuk tabel distribusi frekuensi dapat pula disajikan dalam bentuk diagram. Adapun jenis diagram yang digunakan untuk menggambarkan data berdasarkan tabel distribusi frekuensi yaitu:

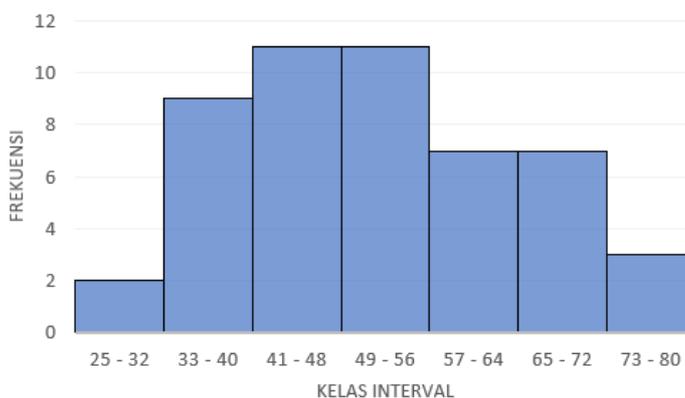
Grafik Histogram

Histogram merupakan grafik berbentuk persegi panjang atau batang. Dalam menggambar grafik histogram digunakan grafik dengan sumbu X dan sumbu Y. Sumbu mendatar/horizontal (sumbu X) menyatakan interval kelas (batas bawah dan batas atas masing-masing kelas) dan sumbu tegak/vertikal (sumbu Y) menyatakan frekuensi.

Perhatikan di bawah ini:

Tabel 3.3: Tabel Distribusi Frekuensi

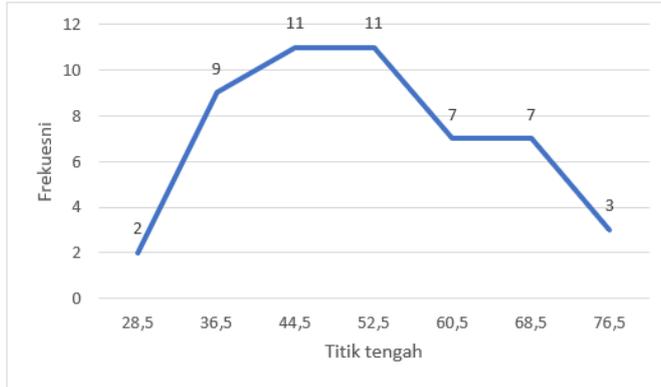
Kelas Interval	Titik Tengah	Frekuensi
25 - 32	28,5	2
33 - 40	36,5	9
41 - 48	44,5	11
49 - 56	52,5	11
57 - 64	60,5	7
65 - 72	68,5	7
73 - 80	76,5	3
Jumlah		50



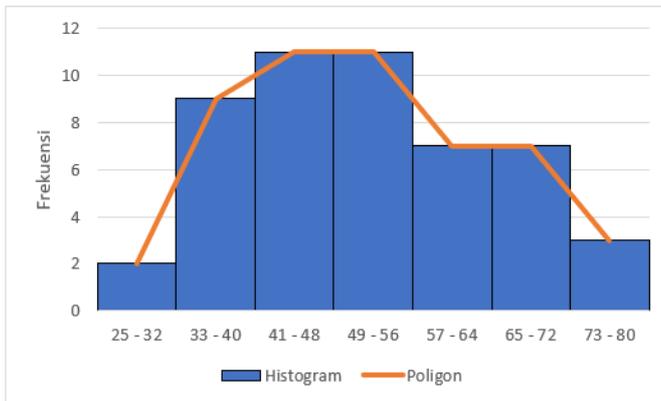
Gambar 3.1: Grafik Histogram Skor Kemampuan Matematika

Grafik Poligon

Grafik poligon disebut juga grafik poligon frekuensi, dibuat dengan menghubungkan-hubungkan titik-titik koordinat (pertemuan titik tengah dengan frekuensi tiap kelas) secara berturut-turut.



Gambar 3.2: Grafik Poligon skor kemampuan matematika



Gambar 3.3: Gabungan Grafik Poligon dan Histogram

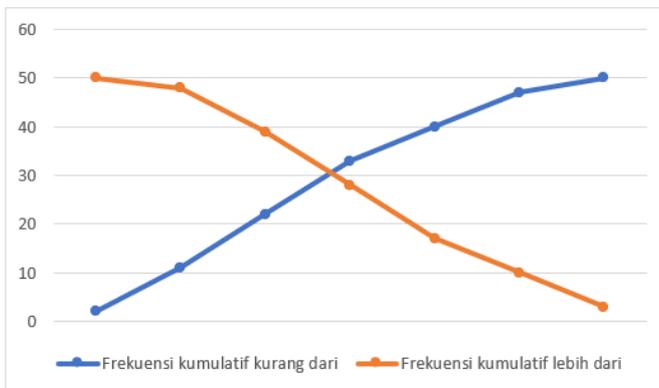
Grafik Ogive

Pembuatan grafik Ogive dengan cara yang sama dengan membuat grafik lainnya. Hal yang perlu diketahui bahwa grafik Ogive dibuat dengan menggunakan batas kelas dan bukan titik tengah sebagaimana grafik poligon. Berikut ini diberikan contoh untuk membuat grafik Ogive dari distribusi data kelompok.

Tabel 3.4: Data Frekuensi Kumulatif

Kelas Interval	Titik Tengah	Frekuensi	Frekuensi Kumulatif Kurang Dari	Frekuensi Kumulatif Lebih Dari
25 - 32	28,5	2	2	50
33 - 40	36,5	9	11	48
41 - 48	44,5	11	22	39
49 - 56	52,5	11	33	28
57 - 64	60,5	7	40	17
65 - 72	68,5	7	47	10
73 - 80	76,5	3	50	3
Jumlah		50		

Frekuensi kumulatif kurang dari adalah jenis frekuensi kumulatif yang dihitung dengan menjumlahkan frekuensi kelas pertama dengan frekuensi kelas berikutnya hingga akhir, dimana nilai kelas akhir sama dengan jumlah data. Sedangkan frekuensi kumulatif lebih dari adalah jenis frekuensi kumulatif yang dihitung dari mengurangi jumlah data dengan frekuensi kelas pertama kemudian frekuensi kelas berikutnya hingga frekuensi kelas paling akhir.

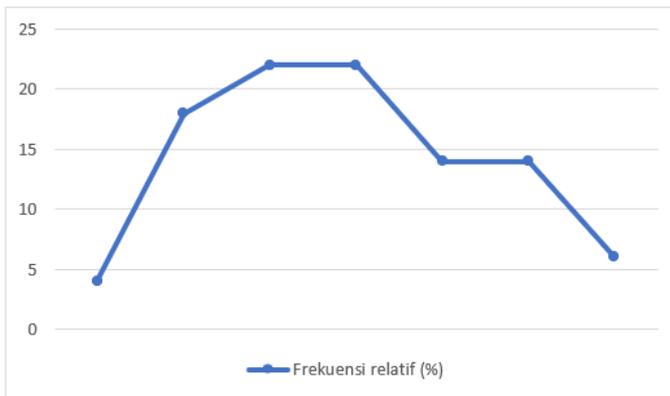
**Gambar 3.4:** Grafik Ogive Frekuensi Kumulatif

Frekuensi kumulatif kurang dari grafiknya berupa ogive positif dan Frekuensi kumulatif lebih dari grafiknya berupa ogive negatif.

Distribusi frekuensi relatif merupakan frekuensi kelas interval relatif terhadap total frekuensi. Formula untuk distribusi frekuensi relatif diberikan oleh:

$$\text{Frekuensi relatif} = \frac{\text{frekuensi kelas interval}}{\text{jumlah data}}$$

Kelas interval	Titik Tengah	Frekuensi	Frekuensi relatif	Frekuensi relatif (%)
25 - 32	28,5	2	2/50	4
33 - 40	36,5	9	9/50	18
41 - 48	44,5	11	11/50	22
49 - 56	52,5	11	11/50	22
57 - 64	60,5	7	7/50	14
65 - 72	68,5	7	7/50	14
73 - 80	76,5	3	3/50	6
Jumlah		50	1.00	100



Gambar 3.5: Grafik Ogive Frekuensi Relatif

Bab 4

Kemencengan dan Kurtosis

4.1 Kemencengan

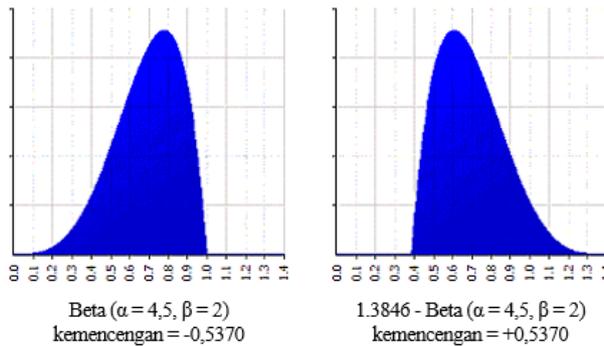
Setelah mempelajari ukuran numerik pusat, penyebaran, dan pencilan, tetapi bagaimana dengan ukuran bentuk? Histogram dapat memberikan gambaran umum tentang bentuk, tetapi dua ukuran numerik dari bentuk memberikan evaluasi yang lebih tepat: kemencengan atau kemiringan menunjukkan jumlah dan arah kemiringan (penyimpangan dari simetri horizontal), dan kurtosis menunjukkan seberapa tinggi dan tajamnya puncak pusat, relatif terhadap kurva lonceng standar.

Salah satu aplikasinya adalah menguji *normalitas*; banyak kesimpulan statistik yang mengharuskan sebuah distribusi normal atau hampir normal. Distribusi normal memiliki kemencengan dan kelebihan kurtosis sebesar 0, jadi jika distribusi mendekati nilai tersebut maka kemungkinan besar distribusi tersebut mendekati normal.

Hal pertama yang biasanya jadi perhatian tentang bentuk distribusi adalah apakah distribusi tersebut memiliki satu modus (puncak) atau lebih dari satu. Jika *unimodal* (hanya memiliki satu puncak), seperti kebanyakan kumpulan data, hal berikutnya yang perlu jadi perhatian adalah apakah distribusi tersebut simetris atau miring ke satu sisi. Jika sebagian besar data berada di sebelah kiri dan ekor kanan lebih panjang, kita mengatakan bahwa distribusinya miring ke

kanan atau miring positif; jika puncaknya berada di sebelah kanan dan ekor kirinya lebih panjang, kita mengatakan bahwa distribusinya miring ke kiri atau miring negatif.

Lihatlah dua grafik di bawah ini. Keduanya memiliki $\mu = 0,6923$ dan $\sigma = 0,1685$, tetapi bentuknya berbeda. Distribusi beta adalah salah satu dari sekian banyak distribusi miring yang digunakan dalam pemodelan matematika.



Gambar 4.1: Kemencengan

Yang pertama agak miring ke kiri: ekor kiri lebih panjang dan sebagian besar distribusinya ada di sebelah kanan. Sebaliknya, distribusi kedua agak miring ke kanan: ekor kanan lebih panjang dan sebagian besar distribusi berada di sebelah kiri. Ingat bahwa rata-rata dan standar deviasi memiliki satuan yang sama dengan data asli, dan varians memiliki kuadrat dari satuan tersebut. Namun, kemencengan tidak memiliki satuan: ini adalah angka murni, seperti skor-z.

Komputasi

Koefisien kemencengan momen dari kumpulan data adalah:

$$\text{kemencengan: } g_1 = \frac{m_3}{m_2^{3/2}}$$

di mana:

$$m_3 = \frac{\sum(x-\bar{x})^3}{n} \text{ dan } m_2 = \frac{\sum(x-\bar{x})^2}{n} \quad (1)$$

\bar{x} adalah rata-rata dan n adalah ukuran sampel, seperti biasa. m_3 disebut momen ketiga dari kumpulan data. m_2 adalah varians, kuadrat dari standar deviasi.

Kemencengan juga dapat dihitung sebagai $g_1 =$ nilai rata-rata z^3 , di mana z adalah skor-z yang sudah dikenal, $z = (x - \bar{x})/\sigma$. Tentu saja nilai rata-rata z selalu nol, tetapi bagaimana dengan rata-rata z^3 ? Misalkan anda memiliki beberapa titik yang berada jauh di sebelah kiri rata-rata, dan banyak titik yang berada di sebelah kanan rata-rata. Karena mengkuadratkan deviasi akan memberikan bobot yang lebih besar pada deviasi yang besar, maka anda akan mendapatkan kemencengan yang negatif. Hal ini berlaku sebaliknya jika anda memiliki deviasi yang besar di sebelah kanan rata-rata.

Anda akan ingat bahwa anda harus menghitung varians dan standar deviasi dengan cara yang sedikit berbeda, tergantung pada apakah anda memiliki data untuk seluruh populasi atau hanya sampel. Hal yang sama juga berlaku untuk kemencengan. Jika anda memiliki seluruh populasi, maka g_1 di atas adalah ukuran kemencengan. Namun jika anda hanya memiliki sampel, anda memerlukan kemiringan sampel:

$$\text{kemiringan sampel: } G_1 = \frac{\sqrt{n(n-1)}}{n-2} g_1 \quad (2)$$

(Joanes and Gill, 1998)

Contoh:

Tabel 4.1: Ketinggian Mahasiswa Laki-Laki di Perguruan Tinggi

Tinggi Badan (cm)	Nilai Tengah (x)	Frekuensi (f)
150-159	154,5	8
160-169	164,5	22
170-179	174,5	54
180-189	184,5	14
190-199	194,5	2



Berikut ini adalah data yang dikelompokkan untuk tinggi badan 100 mahasiswa laki-laki yang dipilih secara acak. Histogram menunjukkan bahwa data condong ke kiri, tidak simetris. Namun, seberapa besar kemiringannya, dibandingkan dengan kumpulan data lainnya? Untuk menjawab pertanyaan ini, anda harus menghitung kemencengan.

Mulailah dengan ukuran sampel dan rata-rata sampel. (Ukuran sampel sudah diberikan, tetapi tidak ada salahnya untuk memeriksanya).

$$n = 8 + 22 + 54 + 14 + 2 = 100$$

$$\bar{x} = (154,5 \times 8 + 164,5 \times 22 + 174,5 \times 54 + 184,5 \times 14 + 194,5 \times 2) \div 100$$

$$\bar{x} = 1236 + 3619 + 9423 + 2583 + 389 \div 100$$

$$\bar{x} = 17250 \div 100 = 172,5$$

Sekarang, dengan nilai rata-rata di tangan, anda bisa menghitung kemencengan. (Tentu saja dalam kehidupan nyata, anda mungkin akan menggunakan Excel atau paket statistik, tetapi akan lebih baik jika anda mengetahui dari mana angka-angka tersebut berasal).

Nilai Tengah (x)	Frekuensi (f)	x. f	(x - \bar{x})	(x - \bar{x}) ² . f	(x - \bar{x}) ³ . f
154,5	8	1236	-18	2592	-46656
164,5	22	3619	-8	1408	-11264
174,5	54	9423	2	216	432
184,5	14	2583	12	2016	24192
194,5	2	389	22	968	21296
Jumlah		17250		7200	-12000
Rataan		172,5		72	-120

Kemiringannya adalah:

$$g_1 = \frac{m_3}{m_2^{3/2}} = \frac{-120}{(72)^{3/2}} = -0,1964$$

Itu akan menjadi kemencengan jika anda memiliki data untuk seluruh populasi.

Namun, jelas ada lebih dari 100 siswa laki-laki di dunia, atau bahkan di hampir semua sekolah, jadi yang anda miliki di sini adalah sampel, bukan populasi. Anda harus menghitung kemencengan sampel:

$$G_1 = \frac{\sqrt{n(n-1)}}{n-2} g_1 = \left[\frac{\sqrt{100(99)}}{98} \right] \left[\frac{-120}{[(72)^{3/2}]} \right] = -0,1994$$

Menafsirkan

Jika kemencengan positif, data condong ke kanan atau miring ke kanan, yang berarti ekor kanan distribusi lebih panjang daripada ekor kiri. Jika skewness negatif, data condong ke kiri atau miring ke kiri, yang berarti ekor kiri lebih panjang.

Jika kemencengan = 0, data adalah simetris sempurna. Tetapi kemencengan yang tepat nol sangat tidak mungkin terjadi pada data dunia nyata, jadi bagaimana anda dapat menginterpretasikan angka kemencengan? Tidak ada interpretasi yang disepakati, namun (Bulmer, 1979) - seorang ahli klasik - menyarankan aturan praktis ini:

1. Jika kemencengan kurang dari -1 atau lebih besar dari +1, distribusi dapat disebut sangat miring;
2. Jika kemencengan berada di antara -1 dan $-\frac{1}{2}$ atau di antara $+\frac{1}{2}$ dan +1, distribusi dapat disebut agak miring;
3. Jika *skewness* berada di antara $-\frac{1}{2}$ dan $+\frac{1}{2}$, distribusi dapat disebut kurang lebih simetris.

Dengan kemiringan -0,1994, data sampel untuk tinggi badan siswa kurang lebih simetris. Ini adalah interpretasi dari data yang anda miliki. Ketika anda memiliki data untuk seluruh populasi, tidak masalah. Namun, ketika anda memiliki sampel, kemencengan sampel tidak selalu berlaku untuk seluruh populasi.

Dalam hal ini pertanyaannya adalah, dari kemencengan sampel, dapatkah anda menyimpulkan apa pun tentang kemencengan populasi? Untuk menjawab pertanyaan tersebut, lihat bagian selanjutnya.

Menyimpulkan

Kumpulan data anda hanyalah satu sampel yang diambil dari sebuah populasi. Mungkin, dari variabilitas sampel biasa, sampel anda miring meskipun populasinya simetris. Tetapi jika sampel terlalu condong sehingga peluang acak tidak dapat menjelaskannya, maka anda dapat menyimpulkan bahwa ada kemiringan dalam populasi.

Namun, apa yang saya maksud dengan "terlalu banyak peluang acak untuk menjadi penjelasan"? Untuk menjawabnya, anda perlu membagi kemencengan sampel G_1 dengan kesalahan standar kemencengan (SES) untuk mendapatkan statistik uji, yang mengukur berapa banyak kesalahan standar yang memisahkan kemencengan sampel dari nol:

$$\text{statistik uji: } Z_{g_1} = \frac{G_1}{SES} \text{ di mana } SES = \sqrt{\frac{6n(n-1)}{(n-2)(n+1)(n+3)}} \quad (3)$$

Rumus ini diadaptasi dari (Cramer, 2002). Beberapa penulis menyarankan $\sqrt{6/n}$, tetapi untuk sampel kecil, itu adalah perkiraan yang buruk. Lagi pula, kita semua memiliki kalkulator, jadi sebaiknya Anda melakukannya dengan benar).

Nilai kritis dari Z_{g_1} adalah sekitar 2 (Ini adalah uji dua ekor kemencengan $\neq 0$ pada tingkat signifikansi 0,05).

1. Jika $Z_{g_1} < -2$, kemungkinan besar populasinya condong ke arah negatif (meskipun Anda tidak tahu seberapa besar).
2. Jika Z_{g_1} berada di antara -2 dan $+2$, Anda tidak dapat mengambil kesimpulan apa pun tentang kemiringan populasi: mungkin populasi tersebut simetris, atau mungkin miring ke salah satu arah.
3. Jika $Z_{g_1} > 2$, kemungkinan besar populasi condong ke arah positif (meskipun Anda tidak tahu seberapa besar).

Jangan mencampuradukkan arti dari statistik uji ini dan jumlah kemiringan. Jumlah kemencengan memberitahu anda seberapa besar kemiringan sampel anda: semakin besar angkanya, semakin besar kemiringannya. Statistik uji memberi tahu anda apakah seluruh populasi mungkin condong, tetapi tidak seberapa besar: semakin besar angkanya, semakin tinggi kemungkinannya.

Memperkirakan

Interval kepercayaan untuk kemiringan:

$$\text{Interval kepercayaan 95\% dari kemiringan populasi} = G_1 \pm 2 SES \quad (4)$$

Joanes dan Gill (1998) menunjukkan bahwa kemencengan sampel adalah penaksir kemencengan populasi yang tidak bias untuk distribusi normal, tetapi tidak untuk distribusi lainnya. Jadi, hitunglah interval kepercayaan tersebut,

tetapi ambillah dengan beberapa butir garam - dan semakin jauh kemencengan sampel dari nol, Anda harus semakin skeptis.

Untuk tinggi badan mahasiswa laki-laki, ingatlah bahwa kemiringan sampel adalah $G_1 = -0,1994$. Ukuran sampel adalah $n = 100$ dan oleh karena itu kesalahan standar kemencengan adalah:

$$SES = \sqrt{\frac{6n(n-1)}{(n-2)(n+1)(n+3)}} = \sqrt{\frac{600(99)}{(98)(101)(103)}} = 0,2414$$

Statistik pengujiannya adalah:

$$Z_{g_1} = \frac{G_1}{SES} = \frac{-0,1994}{0,2414} = -0,8260$$

Nilai ini berada di antara -2 dan +2 (lihat di atas), sehingga dari sampel ini tidak mungkin untuk mengatakan apakah populasi tersebut simetris atau miring. Karena kemencengan sampel kecil, interval kepercayaan mungkin masuk akal:

$$G_1 \pm 2 SES = -0,1994 \pm 2 \times 0,2414 = -0,1994 \pm 0,4828 = -0,6822 \text{ sampai } +0,2834$$

Anda dapat memberikan interval kepercayaan 95% untuk kemencengan sekitar -0,6822 sampai +0,2834, kurang lebih.

4.2 Kurtosis

Ukuran bentuk umum lainnya disebut kurtosis. Karena kemencengan melibatkan momen ketiga dari distribusi, kurtosis melibatkan momen keempat. Oleh karena itu, pencilan dalam sampel memiliki efek yang lebih besar pada kurtosis daripada kemencengan dan dalam distribusi simetris, kedua ekor meningkatkan kurtosis, tidak seperti kemencengan yang saling mengimbangi.

Anda mungkin ingat bahwa rata-rata dan standar deviasi memiliki satuan yang sama dengan data asli, dan varians memiliki kuadrat dari satuan-satuan tersebut. Namun, kurtosis, seperti kemencengan, tidak memiliki satuan: kurtosis adalah angka murni, seperti skor-z.

Secara tradisional, kurtosis telah dijelaskan dalam istilah puncak pusat. Anda akan melihat pernyataan seperti ini: Nilai yang lebih tinggi menunjukkan

puncak yang lebih tinggi dan lebih tajam; nilai yang lebih rendah menunjukkan puncak yang lebih rendah dan kurang jelas. Balanda dan Macgillivray (1988) juga menyebutkan ekor: peningkatan kurtosis dikaitkan dengan "pergerakan massa probabilitas dari bahu distribusi ke pusat dan ekornya."

Namun Westfall (2014) telah melakukan sedikit perang salib untuk mengubah persepsi ini, dan dia membuat argumen yang bagus. Kita dapat mengatakan, bahwa "kurtosis yang lebih tinggi berarti lebih banyak varians yang merupakan hasil dari penyimpangan ekstrem yang jarang terjadi, dibandingkan dengan penyimpangan yang sering terjadi." Dengan kata lain, ekorlah yang sebagian besar menjelaskan kurtosis, bukan puncak pusat.

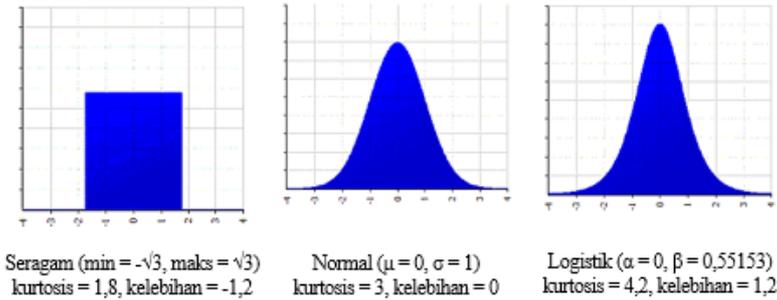
Standar referensi adalah distribusi normal, yang memiliki kurtosis 3. Untuk menunjukkan hal ini, sering kali kurtosis berlebih disajikan: kurtosis berlebih hanyalah kurtosis-3. Sebagai contoh, "kurtosis" yang dilaporkan oleh Excel sebenarnya adalah kurtosis berlebih.

1. Distribusi normal memiliki kurtosis tepat 3 (kelebihan kurtosis tepat 0). Distribusi apa pun dengan kurtosis ≈ 3 (kelebihan kurtosis ≈ 0) disebut *mesokurtik*.
2. Distribusi dengan kurtosis < 3 (kurtosis berlebih < 0) disebut *platykurtik*. Dibandingkan dengan distribusi normal, ekornya lebih pendek dan lebih tipis, dan sering kali puncak pusatnya lebih rendah dan lebih luas.
3. Distribusi dengan kurtosis > 3 (kelebihan kurtosis > 0) disebut *leptokurtik*. Dibandingkan dengan distribusi normal, ekornya lebih panjang dan lebih gemuk, dan sering kali puncak tengahnya lebih tinggi dan lebih tajam.

Perhatikan kata "sering" dalam menggambarkan perubahan pada puncak pusat karena perubahan pada ekor. Westfall (2014) memberikan beberapa ilustrasi contoh tandingan.

Memvisualisasikan

Kurtosis sayangnya lebih sulit untuk digambarkan daripada kemencengan, tetapi ilustrasi berikut ini, seharusnya dapat membantu. Ketiga distribusi ini memiliki rata-rata 0, standar deviasi 1, dan kemencengan 0, dan semuanya diplot pada skala horizontal dan vertikal yang sama. Lihatlah perkembangan dari kiri ke kanan, seiring dengan meningkatnya kurtosis.

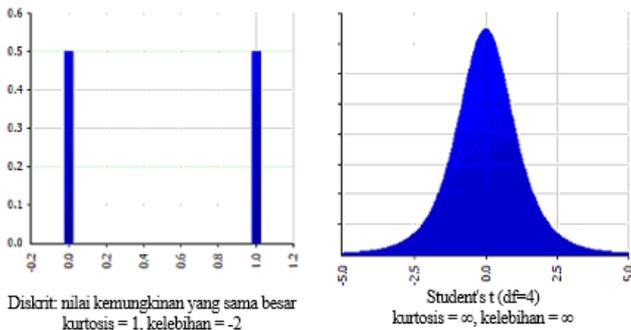


Gambar 4.2: Kurtosis

Beranjak dari distribusi seragam yang diilustrasikan ke distribusi normal, anda dapat melihat bahwa "bahu" telah memindahkan sebagian massanya ke bagian tengah dan ekor. Dengan kata lain, nilai tengah menjadi lebih kecil kemungkinannya dan nilai tengah dan ekstrem menjadi lebih besar kemungkinannya. Kurtosis meningkat sementara deviasi standar tetap sama, karena lebih banyak variasi disebabkan oleh nilai ekstrem.

Beranjak dari distribusi normal ke distribusi logistik yang diilustrasikan, tren terus berlanjut. Bahkan ada lebih sedikit di bagian bahu dan lebih banyak lagi di bagian ekor, dan puncak tengahnya lebih tinggi dan lebih sempit.

Seberapa jauh hal ini bisa terjadi? Berapa nilai kurtosis terkecil dan terbesar yang mungkin terjadi? Kurtosis terkecil yang mungkin adalah 1 (kelebihan kurtosis -2), dan yang terbesar adalah ∞ , seperti yang ditunjukkan di sini:



Gambar 4.3: Kurtosis

Distribusi diskrit dengan dua hasil yang sama-sama mungkin terjadi, seperti menang atau kalah dalam pelemparan koin, memiliki kurtosis serendah

ungkinan. Distribusi ini tidak memiliki puncak pusat dan tidak memiliki ekor yang nyata, dan Anda bisa mengatakan bahwa distribusi ini "semua bahu" - distribusi ini sama seperti distribusi plato.

Di sisi lain, distribusi *t Student* dengan empat derajat kebebasan memiliki kurtosis tak terbatas. Sebuah distribusi tidak bisa lebih *leptokurtik* dari ini. Anda mungkin ingin melihat Gambar 5.3 dari (Westfall, 2014) untuk tiga distribusi yang sangat berbeda dengan kurtosis yang sama.

Komputasi

Koefisien momen kurtosis dari kumpulan data dihitung dengan cara yang hampir sama dengan koefisien kemencengan: cukup ubah eksponen 3 menjadi 4 dalam rumus:

$$\text{kurtosis: } a_4 = \frac{m_4}{m_2^2} \text{ dan kurtosis berlebih: } g_2 = a_4 - 3$$

di mana

(5)

$$m_4 = \frac{\sum(x - \bar{x})^4}{n} \text{ dan } m_2 = \frac{\sum(x - \bar{x})^2}{n}$$

Sekali lagi, kurtosis berlebih umumnya digunakan karena kurtosis berlebih dari distribusi normal adalah 0. \bar{x} adalah rata-rata dan n adalah ukuran sampel, seperti biasa. m_4 disebut momen keempat dari kumpulan data. m_2 adalah varians, kuadrat dari standar deviasi.

Kurtosis juga dapat dihitung sebagai $a_4 = \text{nilai rata-rata } z^4$, di mana z adalah nilai- z yang sudah dikenal, $z = (x - \bar{x})/\sigma$. Tentu saja nilai rata-rata z selalu nol, tetapi nilai rata-rata z^4 selalu ≥ 1 , dan lebih besar ketika anda memiliki beberapa deviasi besar di kedua sisi rata-rata daripada ketika anda memiliki banyak deviasi kecil.

Seperti halnya varians, standar deviasi, dan kemencengan, di atas adalah penghitungan akhir kurtosis jika anda memiliki data untuk seluruh populasi. Namun jika Anda hanya memiliki data untuk sampel, anda harus menghitung kurtosis berlebih sampel menggunakan rumus ini, yang berasal dari (Joanes and Gill, 1998):

$$\text{sampel kelebihan kurtosis: } G_2 = \frac{n-1}{(n-2)(n-3)} [(n+1)g_2 + 6] \quad (6)$$

Contoh:

Mari lanjutkan dengan contoh tinggi badan mahasiswa, dan hitung kurtosis dari kumpulan data tersebut. $n = 100$, $\bar{x} = 172,5 \text{ cm}$, dan varians $m_2 = 72$ telah dihitung sebelumnya.

Nilai Tengah (x)	Frekuensi (f)	$(x - \bar{x})$	$(x - \bar{x})^4 \cdot f$
154,5	8	-18	839808
164,5	22	-8	90112
174,5	54	2	864
184,5	14	12	290304
194,5	2	22	468512
Jumlah			1689600
Rataan			16896

Kurtosisnya adalah

$$a_4 = \frac{m_4}{m_2^2} = \frac{16896}{72^2} = 3,2593$$

dan kelebihan kurtosisnya adalah: $g_2 = 3,2593 - 3 = 0,2593$

Tetapi ini adalah sampel, bukan populasi, jadi Anda harus menghitung kurtosis berlebih sampel:

$$G_2 = [99/(98 \times 97)][101 \times (0,2593) + 6] = 0,3352$$

Sampel ini sedikit bersifat leptokurtik: puncaknya sedikit lebih dangkal daripada puncak distribusi normal.

Menyimpulkan

Kumpulan data anda hanyalah satu sampel yang diambil dari sebuah populasi. Seberapa jauh kelebihan kurtosis harus dari 0, sebelum anda dapat mengatakan bahwa populasi juga memiliki kelebihan kurtosis yang tidak nol?

Pertanyaannya mirip dengan pertanyaan tentang kemencengan, dan jawabannya juga serupa. Anda membagi kurtosis berlebih sampel dengan kesalahan standar kurtosis (SEK) untuk mendapatkan statistik uji, yang memberi tahu Anda berapa banyak kesalahan standar kurtosis berlebih sampel dari nol:

$$\text{statistik uji: } Z_{g_2} = \frac{G_2}{SEK} \text{ dimana } SEK = 2(SES) \sqrt{\frac{n^2-1}{(n-3)(n+5)}} \quad (7)$$

Rumus ini diadaptasi dari Cramer (2002). (Beberapa penulis menyarankan $24/n$, tetapi untuk sampel kecil, ini adalah perkiraan yang buruk. Lagi pula, kita semua memiliki kalkulator, jadi sebaiknya anda melakukannya dengan benar).

Nilai kritis dari Z_{g_2} adalah sekitar 2. (Ini adalah uji dua ekor untuk kurtosis berlebih $\neq 0$ pada tingkat signifikansi 0,05).

1. Jika $Z_{g_2} < -2$, populasi kemungkinan besar memiliki kurtosis berlebih negatif (kurtosis < 3 , platykurtik), meskipun Anda tidak tahu berapa banyak.
2. Jika Z_{g_2} berada di antara -2 dan $+2$, Anda tidak dapat mengambil kesimpulan apa pun tentang kurtosis: kelebihan kurtosis mungkin positif, negatif, atau nol.
3. Jika $Z_{g_2} > +2$, populasi kemungkinan besar memiliki kurtosis berlebih yang positif (kurtosis > 3 , leptokurtik), meskipun Anda tidak tahu berapa banyak.

Untuk sampel tinggi badan mahasiswa pria ($n = 100$), Anda menemukan kurtosis berlebih sebesar $G_2 = 0,3352$. Sampel tersebut berbentuk lonceng, tetapi apakah ini cukup untuk membuat Anda mengatakan bahwa seluruh populasi berbentuk lonceng (memiliki kurtosis yang lebih rendah daripada kurva lonceng)?

Pertama-tama, hitunglah kesalahan standar kurtosis:

$$SEK = 2(SES) \sqrt{\frac{n^2 - 1}{(n - 3)(n + 5)}}$$

$n = 100$, dan SES sebelumnya dihitung sebagai 0,2414.

$$SEK = 2 \times 0,2414 \times \sqrt{\frac{100^2 - 1}{(97)(105)}} = 0,4784$$

Statistik pengujiannya adalah

$$Z_{g_2} = \frac{G_2}{SEK} = \frac{0,3352}{0,4784} = 0,7008$$

Anda tidak dapat mengatakan apakah kurtosis populasi sama atau berbeda dengan kurtosis distribusi normal.

4.3 Menghitung Normalitas

Ada banyak cara untuk menilai normalitas, dan sayangnya, tidak ada satu pun yang tanpa masalah. Salah satu pengujiannya adalah uji omnibus D'Agostino-Pearson (Bowman, D'Agostino and Stephens, 1988). Uji D'Agostino-Pearson jika diterapkan pada Excel di Pemeriksaan Normalitas dan Menemukan Pencilan di Excel. Ini disebut uji omnibus karena menggunakan statistik uji untuk kemencengan dan kurtosis untuk menghasilkan nilai p tunggal yang menilai apakah bentuk kumpulan data ini terlalu berbeda dari normal.

Statistik uji tersebut adalah:

$$DP = Z_{g_1}^2 + Z_{g_2}^2 \text{ mengikuti } \chi^2 \text{ dengan } df = 2 \quad (8)$$

Anda dapat mencari nilai-p dalam tabel, atau menggunakan χ^2cdf pada TI-83 atau TI-84.

Uji χ^2 ini selalu memiliki 2 derajat kebebasan, berapa pun ukuran sampelnya. D'Agostino tidak mengatakan alasannya secara eksplisit, tetapi ada yang mengatakan bahwa itu adalah kecocokan empiris.

$\chi^2cdf(2; 5,991464546) = 0,95$, jadi jika statistik uji lebih dari sekitar 6, anda akan menolak hipotesis normalitas pada tingkat 0,05. Uji D'Agostino-Pearson memiliki kecenderungan untuk melakukan kesalahan dalam menolak normalitas, terutama dengan ukuran sampel yang kecil.

Untuk tinggi badan mahasiswa, anda memiliki statistik uji $Z_{g_1} = -0,8260$ untuk kemencengan dan $Z_{g_2} = 0,7008$ untuk kurtosis. Statistik uji omnibus adalah:

$$DP = Z_{g_1}^2 + Z_{g_2}^2 = (-0,8260)^2 + (0,7008)^2 = 1,1734$$

dan nilai p-value untuk χ^2 (df=2) < 1,1734, dari tabel atau kalkulator statistik, adalah 0,8203. Anda dapat menolak asumsi normalitas. (Ingat, anda pernah menerima hipotesis nol, jadi anda dapat mengatakan dari pengujian ini bahwa distribusi normal). Histogram menunjukkan normalitas, dan tes ini memberi anda alasan untuk menolak kesan tersebut.

Bab 5

Statistik Inferensi

5.1 Pendahuluan

Bab ini dituliskan untuk membantu pembaca khususnya peneliti pemula untuk mampu memahami, memilih, menggunakan, dan mendemonstrasikan teknik statistik inferensial dalam penelitian. Secara garis besar bab ini menguraikan pengujian hipotesis dan penggunaan statistika inferensial dalam penelitian kuantitatif.

Statistika inferensial meliputi semua teknik yang berkaitan dengan analisis data untuk penarikan kesimpulan mengenai keseluruhan data. Statistika inferensial merupakan pengetahuan statistika yang berhubungan dengan penarikan kesimpulan yang bersifat umum dari data yang telah disusun dan diolah dengan menggunakan konsep dan prinsip probabilitas (Walpole et al., 2016).

Lebih jauh lagi, statistika inferensial berkaitan dengan tidak sekadar deskripsi data tetapi juga teknik-teknik mendapatkan sampel dan selanjutnya membuat inferensi atau kesimpulan-kesimpulan ataupun prediksi tentang populasi atas dasar sampel data tersebut (Cohen, 2013). Secara singkat, statistika inferensial adalah statistika yang digunakan untuk menggeneralisasikan data sampel terhadap populasi.

Statistik inferensial selanjutnya dibagi ke dalam dua cakupan, yaitu statistik parametrik dan statistik non-parametrik. Statistik parametrik didasarkan pada

asumsi bahwa data yang dianalisis terdistribusi normal (Spiegel, Schiller and Srinivasan, 2013). Statistika parametrik umumnya diterapkan pada data berjenis interval maupun rasio. Statistika non-parametrik digunakan ketika data tidak terdistribusi secara normal. Jenis data yang diolah dalam statistika nonparametrik biasanya bersifat nominal maupun ordinal. Namun, bahasan pada buku ini masih berfokus pada statistik parametrik.

5.2 Pengujian Hipotesis

Sebelumnya telah dibahas cara-cara menaksir parameter untuk mengambil kesimpulan tentang besarnya nilai parameter. Cara pengambilan kesimpulan yang kedua akan disajikan adalah melalui pengujian hipotesis. Hipotesis adalah asumsi atau dugaan mengenai sesuatu hal yang dibuat untuk menjelaskan hal tersebut yang sering dituntut untuk melakukan pengecekannya. Jika asumsi atau dugaan tersebut dikhususkan mengenai populasi, umumnya mengenai nilai-nilai parameter populasi, maka hipotesis disebut hipotesis statistik.

Dengan demikian, pengujian hipotesis adalah proses untuk mengambil keputusan tentang klaim atau hipotesis yang diajukan tentang parameter populasi berdasarkan pada bukti dari data sampel. Sering kali, masalah yang dihadapi ilmuwan dan peneliti bukannya untuk mengestimasi parameter populasi tetapi lebih kepada pembentukan prosedur pengambilan keputusan berdasarkan data.

Sebagai contoh, peneliti akan memutuskan berdasarkan data eksperimen pada 2 kasus berikut:

1. Apakah penggunaan teknologi dapat meningkatkan keterlibatan peserta didik pada pembelajaran matematika?
2. Peneliti memutuskan berdasar data apakah pemberian masalah kontekstual sebelum aktivitas pembelajaran memberi dampak pada pemecahan masalah peserta didik?

Pada 2 (dua) kasus tersebut, peneliti membuat suatu pernyataan atau konjektur tertentu tentang sebuah sistem, pernyataan tersebut didasarkan pada data eksperimen dan merupakan keputusan berdasar data yang diperoleh.

Pada setiap kasus, konjektur dapat ditulis dalam bentuk hipotesis statistik. Prosedur yang dikerjakan untuk menuju pada diterima atau ditolaknya hipotesis statistik merupakan bagian penting dari statistika inferensial, yang disebut uji hipotesis. Benar atau tidaknya suatu hipotesis statistika tidak pernah diketahui secara mutlak kecuali mengambil data dari keseluruhan populasi, yang tentunya tidak praktis atau kadang tidak mungkin dapat dilakukan.

Sebagai gantinya, diambil sampel yang representatif dari populasi dan menggunakan data yang diperoleh dari sampel untuk memutuskan mendukung atau tidak mendukung hipotesis. Petunjuk dari sampel yang tidak konsisten atau kontradiksi dengan pernyataan hipotesis membawa pada penolakan hipotesis.

Ketika seorang peneliti menginterpretasikan hasil dari eksperimen, pertanyaan alami yang muncul adalah apakah hasil yang diperoleh bersifat kebetulan? Oleh karena itu, prosedur pengambilan keputusan seorang peneliti harus memperhatikan peluang dari kesimpulan yang salah.

Hal esensial yang perlu dipahami dan dibedakan oleh peneliti pemula adalah harus mampu membedakan antara hipotesis penelitian dan hipotesis statistik. Hipotesis penelitian adalah pernyataan sementara yang merupakan jawaban dari pertanyaan penelitian yang didasarkan pada teori dan hasil penelitian terdahulu yang relevan, sementara hipotesis statistik adalah pernyataan yang berkaitan dengan parameter populasi.

Sejak awal, hipotesis dirumuskan peneliti dengan harapan akan ditolak atau dibuat nol, sehingga membawa kepada istilah hipotesis nol. Akan tetapi sekarang, bentuk ini dapat diaplikasikan pada sebarang hipotesis yang akan diuji dan dinotasikan dengan H_0 . Sebarang hipotesis yang berbeda dengan hipotesis yang diberikan (H_0) disebut hipotesis alternatif dan dinotasikan dengan H_a atau H_1 .

Kita dapat memahami bahwa penolakan dari H_0 membawa pada penerimaan H_1 . Hipotesis nol selalu dinyatakan dalam nilai parameter eksak sementara hipotesis alternatif bisa dengan bermacam nilai. Jadi pada H_0 digunakan notasi “=” (sama dengan); dan pada H_1 digunakan notasi “>” (lebih dari), “<” (kurang dari), atau “≠” (tidak sama dengan).

Misalkan θ adalah suatu parameter populasi dan θ_0 merupakan suatu bilangan real, maka rumusan hipotesis statistik dapat dijabarkan sebagai berikut:

1. Uji satu pihak

a. Uji pihak kanan

$$H_0 : \theta = \theta_0$$

$$H_1 : \theta > \theta_0$$

b. Uji pihak kiri

$$H_0 : \theta = \theta_0$$

$$H_1 : \theta < \theta_0$$

2. Uji dua pihak

$$H_0 : \theta = \theta_0$$

$$H_1 : \theta \neq \theta_0$$

Untuk menentukan uji hipotesis yang akan digunakan, peneliti harus memerhatikan pertanyaan penelitian dan hipotesis penelitian. Hipotesis penelitian harus terletak tepat di salah satu dari hipotesis statistik. Dua tipe kesalahan sering dilakukan peneliti meskipun dalam penelitian hipotesis telah diterima atau ditolak, tidak berarti bahwa telah dibuktikan kebenaran hipotesis karena yang diperlihatkan adalah hanya menerima atau menolak hipotesis saja.

Dalam melakukan pengujian hipotesis untuk menuntun pada pengambilan keputusan, ada 4 (empat) macam situasi yang dapat terjadi yang dapat disajikan pada tabel 5.1

Tabel 5.1: Pengambilan Keputusan

Keputusan	H_0 Benar	H_0 Salah
Menerima H_0	Keputusan yang benar	Kesalahan tipe II
Menolak H_0	Kesalahan tipe I	Keputusan yang benar

Secara verbal, penolakan H_0 ketika H_0 benar disebut dengan kesalahan tipe I dan penerimaan H_0 ketika H_0 salah disebut dengan kesalahan tipe II.

Peluang dari melakukan kesalahan tipe I disebut dengan taraf signifikansi dari uji statistik, dinotasikan dengan α . Terkadang taraf signifikansi disebut dengan ukuran daerah kritis. Daerah kritis dengan ukuran 0,05 atau $\alpha = 0,05$ secara sederhana berarti kriteria keputusan dibuat sehingga peluang melakukan

kesalahan tipe I adalah sebesar 0,05. Peluang melakukan kesalahan tipe II dinotasikan dengan β . Jika α diperkecil, maka β menjadi besar dan demikian sebaliknya. Idealnya adalah menggunakan prosedur dimana α dan β adalah kecil. Dalam statistika, diasumsikan bahwa H_0 selalu benar sehingga kesalahan tipe I (α) yang diperhatikan.

Meskipun semakin kecil α semakin baik, tetapi kadang tidak bisa menggunakan α sekecil mungkin dengan alasan tidak bisa mengontrol variabel lain yang mempengaruhi. Nilai α yang biasa digunakan dalam penelitian adalah $\alpha=0,01$ atau $\alpha = 0,05$. Misalnya, dengan $\alpha = 0,05$ atau sering disebut taraf signifikansi 5%, artinya kira-kira 5 dari tiap 100 kesimpulan bahwa akan menolak hipotesis yang harusnya diterima.

Dengan kata lain, kira-kira 95% yakin bahwa telah dibuat kesimpulan yang benar. Dalam hal demikian dikatakan bahwa hipotesis telah ditolak pada taraf nyata 0,05 yang berarti mungkin salah dengan peluang 0,05. Dalam pengujian hipotesis, kita membandingkan data sampel dengan hipotesis nol (H_0) yang menyatakan tidak ada efek atau perbedaan tertentu.

Adapun langkah-langkah pengujian hipotesis statistik dapat diuraikan berikut ini (Walpole et al., 2016; Coolican, 2019):

1. Menentukan H_0 dengan nilai parameter yang sesuai.
2. Menentukan apakah hipotesis penelitian terletak pada H_0 .
3. Merumuskan H_1 dengan menggunakan notasi “<”, “>”, atau “ \neq ” disesuaikan dengan hipotesis penelitian, jika hipotesis penelitian di H_0 maka H_1 menggunakan “ \neq ”.
4. Menetapkan nilai taraf signifikansi atau α yang digunakan 0,05 atau 0,01 atau 0,10.
5. Memilih uji statistik yang digunakan (uji t, uji F, dll) dan menentukan daerah kritis berdasar α yang digunakan (menentukan kriteria penolakan H_0).
6. Menentukan nilai uji statistik dari data sampel.
7. Mengambil keputusan: menolak H_0 jika nilai statistik data sampel berada pada daerah kritis dan selain itu H_0 diterima.

5.3 Uji Mean Populasi

Misalkan akan diuji hipotesis rata-rata populasi μ adalah μ_0 maka:

$$H_0: \mu = \mu_0$$

Kita dapat menentukan uji statistik yang tepat digunakan berdasarkan kasus atau data yang tersedia.

Kasus 1: Varians populasi (σ^2) diketahui

Jika varians populasi diketahui dan rata-rata sampel (\bar{x}) dari populasi dengan n data, maka dalam menguji H_0 peneliti menggunakan uji z yang diformulasikan berikut.

$$z = \frac{\bar{x} - \mu_0}{\sigma} \cdot \sqrt{n}$$

Kita perlu memperhatikan daerah kritis:

1. Jika $H_1: \mu > \mu_0$, maka tolak H_0 jika $z > z_\alpha$
2. Jika $H_1: \mu < \mu_0$, maka tolak H_0 jika $z < -z_\alpha$
3. Jika $H_1: \mu \neq \mu_0$, maka tolak H_0 jika $z > z_{\frac{\alpha}{2};v}$ atau $z < -z_{\frac{\alpha}{2};v}$

Contoh:

Dalam contoh ini, kita akan menguji rata-rata usia penduduk Indonesia pada tahun 2021. Sampel random dari 100 data kematian yang tercatat di Indonesia pada tahun 2021 menunjukkan bahwa rata-rata penduduk Indonesia berusia 66,5 tahun. Jika standar deviasi dari masa hidup orang Indonesia adalah 8 tahun, apakah hal ini mengindikasikan bahwa rata-rata masa hidup orang Indonesia sekarang adalah lebih besar dari 65 tahun? Gunakan taraf signifikansi 0,05.

Solusi

Diketahui:

$$n = 100$$

$$\bar{x} = 66.5$$

$$\sigma = 8$$

$$\mu_0 = 65$$

Langkah pertama merumuskan hipotesis statistik

$$H_0: \mu = 65$$

$H_1: \mu > 65$ (*) hipotesis penelitian terletak di H_1

$$\alpha = 0.05$$

Selanjutnya kita menentukan daerah kritis yaitu nilai z pada tabel:

$$z > z_{\alpha} = z_{0.05} = 1.65$$

Langkah selanjutnya menentukan nilai z hitung.

$$z = \frac{66.5 - 65}{8} \cdot \sqrt{100} = \frac{1.5 \times 10}{8} = 1.875$$

Penarikan kesimpulan:

Karena $z = 1.875 > 1.65$, maka H_0 ditolak sehingga hipotesis penelitian diterima artinya rata-rata hidup orang Indonesia sekarang lebih dari 65 tahun

Kasus 2: Varians populasi (σ^2) tidak diketahui

Sering dijumpai pada kebanyakan kasus, varians populasi (σ^2) tidak diketahui. Dalam hal ini, varians populasi diestimasi dari varians sampel. Oleh karena itu, σ diestimasi dengan s . Untuk menguji H_0 pada kasus σ tidak diketahui atau sampelnya kecil ($n < 30$), digunakan uji statistik t dengan derajat kebebasan $v = n - 1$, yang diformulasikan berikut:

$$t = \frac{\bar{x} - \mu_0}{s} \cdot \sqrt{n}$$

Kita perlu memperhatikan daerah kritis:

1. Jika $H_1: \mu > \mu_0$, maka tolak H_0 jika $t > t_{\alpha}$
2. Jika $H_1: \mu < \mu_0$, maka tolak H_0 jika $t < -t_{\alpha}$
3. Jika $H_1: \mu \neq \mu_0$, maka tolak H_0 jika $t > \frac{t_{\alpha}}{2}, v$ atau $t < -\frac{t_{\alpha}}{2}, v$

Contoh:

Sampel acak dari 20 rokok merek tertentu mempunyai kandungan nikotin rata-rata 4.2 mgr dan standar deviasi 1.4 mgr. Apakah hal ini sesuai dengan klaim perusahaan bahwa rokok mereka kandungan nikotinnya tidak melebihi 3.5 mgr?

Gunakan taraf signifikansi 0.01 dan asumsikan distribusi kandungan nikotin dalam rokok merek tersebut terdistribusi normal.

Solusi:

Diketahui

$$n = 20$$

$$\bar{x} = 4.2$$

$$s = 1.4$$

$$\mu_0 = 3.5$$

Dari soal kita dapat merumuskan klaim perusahaan yaitu " $\mu < 3.5$ ".

1. Merumuskan hipotesis statistik dan menentukan letak hipotesis penelitian

$$H_0: \mu = 3.5 \text{ (*) hipotesis penelitian terletak di } H_0$$

$$H_1: \mu > 3.5$$

2. Menuliskan taraf signifikansi $\alpha = 0.05$
3. Menentukan daerah kritis: $t > t_{\alpha;v} = t_{0.01;19} = 2.539$
4. Menghitung nilai t yaitu:

$$t = \frac{4.2 - 3.5}{1.4} \cdot \sqrt{20}$$

$$t = 2.236$$

5. Kesimpulan

Karena $t_{hitung} < t_{tabel}$ maka H_0 diterima sehingga hipotesis penelitian diterima artinya benar bahwa rata-rata kandungan nikotin rokok tersebut tidak melebihi 3.5 mgr.

5.4 Metode Pengujian Pada Statistik Inferensi

Beberapa metode pengujian yang digunakan dalam analisis inferensial pada domain statistik parametrik yang akan dibahas pada bab-bab selanjutnya yaitu sebagai berikut:

1. Analisis regresi linier digunakan untuk memahami hubungan antara dua variabel dari kumpulan data. Analisis ini dipilih Ketika salah satu variabelnya merupakan variabel dependen dan yang lainnya adalah variabel independen yang dapat digunakan lebih dari satu.
2. Analisis varians digunakan untuk menguji dan menganalisis perbedaan antara dua atau lebih *mean* dari kumpulan data.
3. Uji-T digunakan untuk membandingkan rata-rata dua kelompok dan memahami jika satu dengan lainnya berbeda.
4. Analisis korelasi digunakan untuk memahami sejauh mana dua variabel saling bergantung. Kita dapat melihat apakah variabel memiliki korelasi yang kuat atau lemah. Korelasi bisa negatif atau positif tergantung pada variabelnya.

Bab 6

Uji Normalitas dan Uji Homogenitas Data

6.1 Pendahuluan

Pengujian normalitas merupakan sebuah asumsi yang menjadi prasyarat untuk menentukan uji statistik yang akan digunakan dalam penelitian. Uji normalitas sering digunakan untuk mengukur data yang memiliki skala interval, ordinal, maupun rasio. Dalam penelitian kuantitatif jika analisis menggunakan metode parametrik, salah satu syarat yang harus terpenuhi yaitu data terdistribusi normal (Sugiyono, 2004).

Uji normalitas digunakan untuk mengetahui apakah data yang diperoleh dari hasil penelitian berdistribusi normal atau tidak. Dasar pengambilan keputusan adalah jika nilai $L_{hitung} > L_{tabel}$ maka H_0 ditolak dan H_1 diterima, karena data tidak berdistribusi normal. Sebaliknya, jika nilai $L_{hitung} < L_{tabel}$ maka H_0 diterima dan H_1 ditolak, karena data berdistribusi normal (Murwani, 2001); (Nuryadi, dkk, 2017).

Menentukan suatu data penelitian berdistribusi normal dapat digunakan dengan beberapa cara meliputi Kolmogorov-Smirnov, Lilliefors, Chi-Square, Shapiro-Wilk, dan lain sebagainya. Namun dalam buku ini, contoh yang diberikan untuk

menentukan normalitas dari suatu data berfokus pada penggunaan Kolmogorov-Smirnov (Hamdi, 2004).

Uji Kolmogorov-Smirnov

Kolmogorov-Smirnov bertujuan untuk menguji *goodness of fit* dengan memperhatikan tingkat kesesuaian antara distribusi sampel dan distribusi lainnya serta membandingkan serangkaian data pada sampel terhadap distribusi normal, serangkaian nilai dengan rerata dan standar deviasi yang sama (Siregar, 2015).

Uji Kolmogorov-Smirnov dilakukan dengan membandingkan D_{hit} dan D_{tabel} (Purwanto, 2011).

$$D_{hitung} = \max |F_0(x) - S_n(x)|$$

Keterangan:

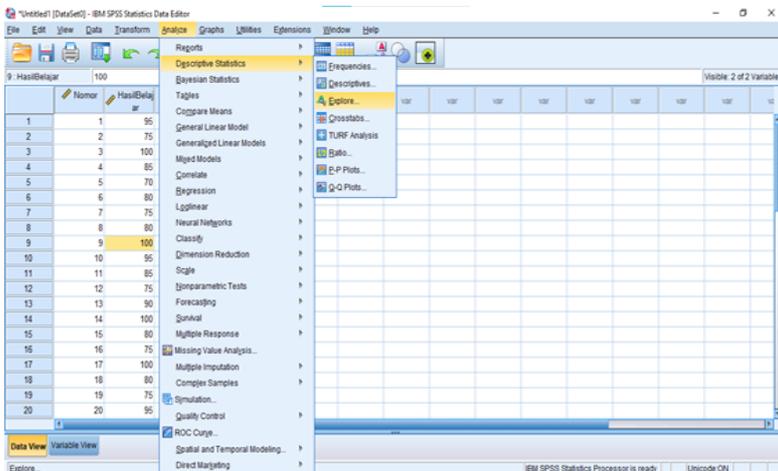
$F_0(x)$ = distribusi frekuensi kumulatif teoritis

$S_n(x)$ = distribusi frekuensi skor observasi

Uji Kolmogorov-Smirnov satu sampel dapat digunakan untuk menguji apakah sampel berasal dari distribusi tertentu (Razali, dkk, 2011). Prosedur ini dapat kita gunakan untuk menentukan apakah sampel berasal dari populasi yang berdistribusi normal (Real Statistic, 2019).

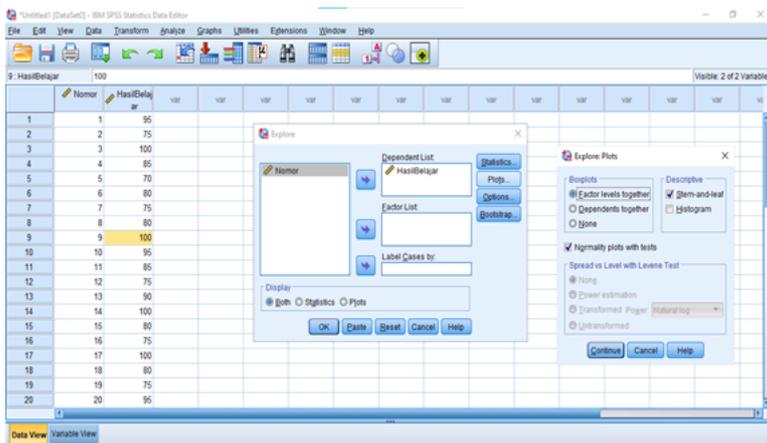
Berikut ini merupakan tahapan pengujian normalitas data menggunakan bantuan program SPSS:

1. Masuk pada program SPSS
2. Klik variabel view pada program SPSS, pada kolom Name baris pertama ketik "Nomor" dan pada kolom Name baris kedua ketik "HasilBelajar". Selanjutnya, pada kolom Type pilih numeric untuk "Nomor dan HasilBelajar", kemudian pada kolom Decimals pilih "0" untuk "Nomor dan HasilBelajar".
3. Buka data view pada program SPSS data editor maka diperoleh kolom variabel "Nomor dan HasilBelajar", import/memasukkan data sesuai dengan variabelnya, selanjutnya klik Analyze, pilih Descriptive Statistics, kemudian pilih Explore. Seperti yang tampak pada gambar berikut:



Gambar 6.1: Tahapan Analyze Descriptive

4. Klik variabel "HasilBelajar" dan masukkan ke dalam kotak Dependent List, kemudian klik Plots. Selanjutnya klik Normality Plots With Test, kemudian klik Continue dan klik Ok. Seperti yang tampak pada gambar berikut:



Gambar 6.2: Tahapan Analyze Descriptive

Setelah seluruh proses analisis selesai, akan muncul Output hasil dari pengolahan data "HasilBelajar".

Berikut adalah hasil dari Analyze Descriptive:

Tabel 6.1: Case Processing Summary

	Cases					
	Valid		Missing		Total	
	N	Percent	N	Percent	N	Percent
HasilBelajar	25	100,0%	0	0,0%	25	100,0%

Berdasarkan tabel di atas terlihat bahwa terdapat 25 data hasil belajar siswa yang valid dengan persentase 100% dan tidak terdapat data yang *missing* dari 25 data hasil belajar siswa.

Tabel 6.2: Descriptive

		Statistic	Std.Error
HasilBelajar	Mean	84,20	2,249
	Median	80,00	
	Variance	126,147	
	Std. Deviation	11,244	
	Minimum	60	
	Maximum	100	
	Range	40	
	Skewness	-0,094	0,464
	Kurtosis	-0,844	0,902

Dari tabel di atas terlihat bahwa rerata nilai hasil belajar siswa yaitu 84,20, dengan standar deviasi sebesar 11,244. Adapun nilai minimum dan maksimum yang diperoleh siswa yaitu 60 dan 100 dengan rentang nilai 40.

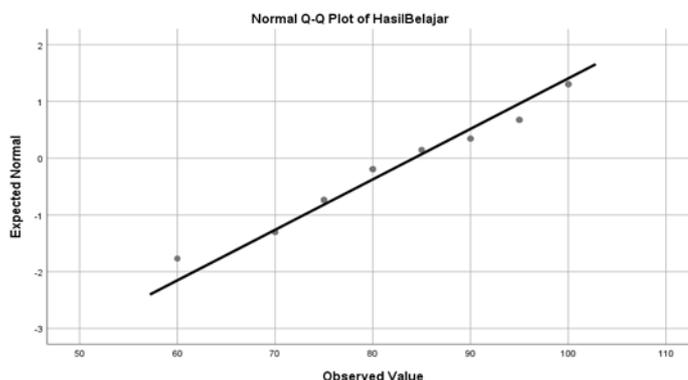
Tabel 6.3: Test of Normality

	Kolmogorov-Smirnov ^a			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
HasilBelajar	0,166	25	0,075	0,931	25	0,092

Berdasarkan tabel di atas hasil uji Kolmogorov-Smirnov menunjukkan bahwa data hasil belajar siswa berdistribusi normal. Hal ini terlihat dari tingkat signifikan yaitu $0,075 > 0,05$.

Data ini diperoleh berdasarkan pedoman pengambilan keputusan (Nuryadi, dkk, 2017) yaitu:

1. Nilai signifikan atau nilai probabilitas $< 0,05$ yang berarti data tidak berdistribusi normal.
2. Nilai signifikan atau nilai probabilitas $> 0,05$ yang berarti data berdistribusi normal.



Gambar 6.3: Normal Q-Q Plot Hasil Belajar

Jika suatu data berdistribusi normal, maka data akan tersebar di sekeliling garis. Berdasarkan gambar di atas terlihat bahwa pola data tersebar di sekeliling garis, yang berarti data berdistribusi normal.

6.2 Uji Homogenitas

Uji homogenitas merupakan uji perbedaan antara dua atau lebih populasi. Semua karakteristik populasi dapat bervariasi antara satu populasi dengan yang lain. Dua di antaranya adalah *mean* dan varian (selain itu masih ada bentuk distribusi, median, modus, range, dll) (Johnson, R. A 2010). Uji homogenitas menjadi prosedur uji statistik yang dapat memperlihatkan dua atau lebih kelompok data sampel yang berasal dari populasi yang memiliki variansi yang sama (Sudjana, 2005). .

Uji homogenitas bertujuan untuk membuktikan apakah dari beberapa kelompok data penelitian memiliki variansi yang sama atau tidak. Sehingga, dapat

diasumsikan bahwa homogenitas berarti kumpulan data yang diteliti memiliki karakteristik yang sama (Santoso, S, 2002).

Pengujian homogenitas dilakukan untuk memberikan keyakinan bahwa sekumpulan data yang dianalisis berasal dari populasi yang sama atau tidak jauh berbeda dari keberagamannya (Riduwan, 2008), misalnya mengukur pemahaman siswa untuk suatu topik pada mata pelajaran tertentu di sekolah yang dimaksudkan homogen bisa berarti bahwa kelompok data yang dijadikan sampel pada penelitian memiliki karakteristik yang sama, misalnya tingkat kelas yang sama, tingkat kesiapan belajar siswa dan lain sebagainya.

Terdapat beberapa faktor yang menyebabkan sampel atau populasi tidak homogen meliputi penyebaran data yang kurang baik, proses sampling yang salah, bahan yang sulit untuk homogen dan lain sebagainya (Nuryadi, dkk, 2017); (Quraisy, A 2020). Jika sampel uji tidak homogen maka sampel tidak dapat digunakan dan perlu dievaluasi kembali mulai dari proses sampling sampai penyebaran data, bahkan jika memungkinkan seluruh proses harus diulangi untuk mendapatkan sampel yang homogen (Sudijono, 2015).

Perhitungan uji homogenitas dapat dilakukan dengan berbagai cara dan metode diantaranya uji Harley Pearson, uji Bartlett, uji Levene, dan uji Cochran (Aprilina, dkk, 2018). Dalam buku ini, akan dijelaskan lebih mendalam uji homogenitas menggunakan uji Bartlett dan uji Levene.

Uji Bartlett

Uji Bartlett digunakan untuk menguji homogenitas varians lebih dari dua kelompok data (Utami, 2020). Tahapan uji homogenitas menggunakan uji Bartlett meliputi:

1. Menghitung derajat kebebasan (dk) masing-masing kelompok.
2. Menghitung varians (s) masing-masing kelompok.
3. Menghitung besarnya $\log S^2$ untuk masing-masing kelompok.
4. Menghitung besarnya dk. $\log S^2$ untuk masing-masing kelompok.
5. Menghitung nilai varians gabungan semua kelompok dengan rumus

sebagai berikut: $S_{gab}^2 = \frac{(\sum dk S_i^2)}{\sum dk}$

Keterangan:

S_{gab}^2 = varians gabungan

6. Menghitung nilai B (nilai Bartlett) dengan rumus sebagai berikut.

B = nilai Bartlett = $\sum dk (\log S_{gab}^2)$

7. Menghitung nilai X^2 dengan rumusan sebagai berikut:

$$X^2 = (\ln 10) \left[B \left(\sum dk \log S_i^2 \right) \right]$$

Keterangan:

S_i^2 = varians tiap kelompok data

$dk = n - 1$ = derajat kebebasan tiap kelompok

$B =$ nilai Bartlett $= \left(\sum db \right) \left(\log S_{gab}^2 \right)$

8. Setelah nilai Chi-Kuadrat hitung diperoleh, maka nilai Chi-Kuadrat tersebut dibandingkan dengan Chi-Kuadrat tabel. Kriteria Homogen ditentukan jika Chi-Kuadrat hitung $<$ Chi-Kuadrat tabel.

Hipotesis pengujian:

$$H_0: \sigma_1^2 = \sigma_2^2 = \sigma_3^2 = \dots = \sigma_n^2$$

H_a : paling sedikit salah satu tanda tidak sama

Kriteria pengujian:

Jika X^2 hitung $\geq X^2$ tabel $(1 - \alpha ; db = n - 1)$ maka tolak H_0

Jika X^2 hitung $< X^2$ tabel $(1 - \alpha ; db = n - 1)$ maka terima H_0

Contoh Perhitungan dengan Uji Bartlett:

Suatu penelitian tentang perbedaan hasil belajar siswa akibat dari suatu perlakuan (eksperimen) dengan menggunakan media geogebra. Adapun perlakuan yang diberikan adalah perbedaan strategi/ metode pembelajaran pada siswa yaitu:

Kelas Eksperimen: metode ceramah dan penggunaan media geogebra.

Kelas Kontrol: metode ceramah dan tanpa penggunaan media geogebra.

Untuk menguji homogenitas varians dari kedua kelompok digunakan teknik Bartlett sebagai berikut:

1. Menghitung derajat kebebasan (dk) masing-masing kelompok
2. Menghitung varians (s) masing-masing kelompok
3. Menghitung besarnya $\log S^2$ untuk masing-masing kelompok
4. Menghitung besarnya $dk \cdot \log S^2$ untuk masing-masing kelompok

Untuk langkah 1-4 dinyatakan dalam tabel di bawah ini yang telah dihitung sebelumnya dalam Excel.

Tabel 6.4: Nilai Varians Tiap Kelompok

Sampel	dk=N-1	1/dk	S	S ²	log(S ²)	dk*(S ²)	dk*log(S ²)
a	23	0,044	188,405797	35496,7444	4,550189	816425,121	104,654336
b	23	0,043	188,405797	35496,7444	4,550189	816425,121	104,654336
Jumlah	46	0,087	376,811594	70993,4888	9,100377	1632850,24	209,308672

5. Menghitung nilai varians gabungan antara semua kelompok

$$S_{gab}^2 = \frac{(\sum db \cdot S_i^2)}{\sum dk}$$

$$S_{gab}^2 = \frac{(46)(70993,49)}{0,087}$$

$$S_{gab}^2 = 37,536$$

6. Menghitung nilai B

$$\begin{aligned} B &= \text{nilai Bartlett} = (\sum db) (\log S_{gab}^2) \\ &= (\sum db) (\log S_{gab}^2) \\ &= (46) (4,574) \\ &= 210,404 \end{aligned}$$

7. Menghitung harga Chi-Kuadrat

$$X^2 = (\ln 10)[B(\sum dk \log S_i^2)]$$

$$X^2 = (2,3)[210,4(3,265)]$$

$$X^2 = 1,579 = 1,58$$

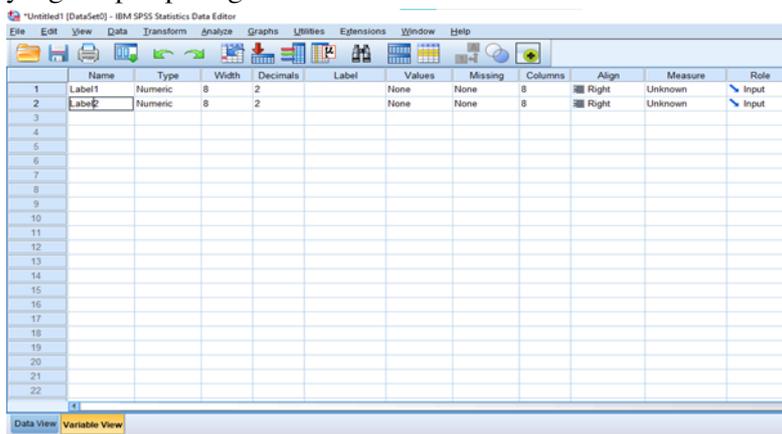
Kesimpulan:

Nilai chi-square hitung dibandingkan dengan nilai chi-square tabel, dengan $dk=1$ dan $\alpha=5\%$ yaitu chi-square tabel $(0,05;1) = 3,84$. Karena chi-square hitung $<$ chi-square tabel yaitu $1,58 < 3,84$, maka H_0 diterima. Ini menunjukkan bahwa varians skor pretes hasil belajar siswa kelas kontrol (menggunakan metode ceramah dan tanpa media geogebra) dan kelas eksperimen (menggunakan metode ceramah dan media geogebra) homogen pada taraf signifikansi 95%.

Uji Levene

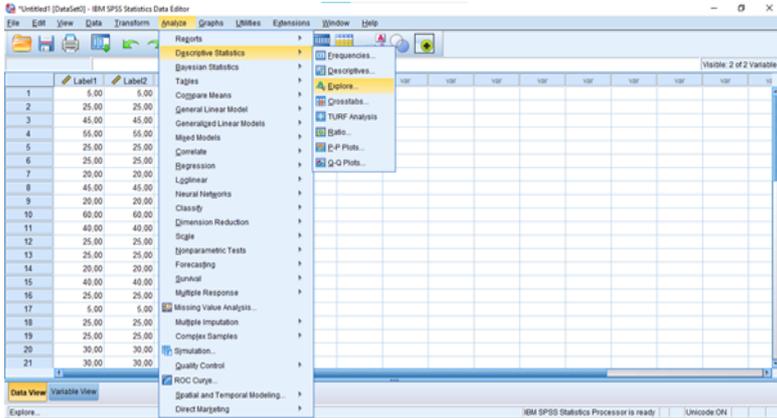
Perhitungan uji homogenitas dengan uji Levene dilakukan menggunakan bantuan program SPSS. Adapun Tahapan melakukan uji Levene dengan bantuan program SPSS sebagai berikut:

1. Memasukkan data variabel yang disusun dalam satu kolom. Setelah variabel pertama dimasukkan, dilanjutkan dengan variabel kedua mulai dari baris kosong setelah variabel pertama.
2. Membuat pengkodean kelas dengan cara membuat variabel baru yang telah diberi “Label 1” untuk variabel pertama dan “Label 2” untuk variabel kedua mulai dari baris kosong setelah variabel kedua. Seperti yang tampak pada gambar berikut:



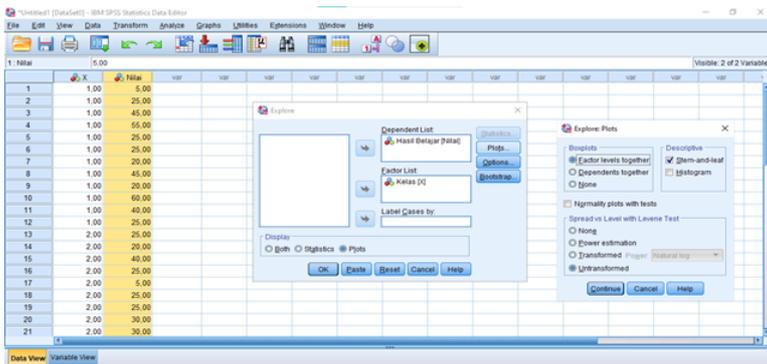
Gambar 6.4: Tahapan Uji Levene

3. Cara menghitung uji Levene dengan program SPSS adalah memilih menu Analyze, Descriptive Statistics kemudian Explore seperti yang tampak pada gambar berikut:



Gambar 6.5: Tahapan Uji Levene

4. Pada jendela yang terbuka memasukkan variabel yang akan dihitung homogenitasnya pada bagian dependent list, dan kode kelas pada bagian factor list, kemudian pilih tombol Plots hingga muncul tampilan sebagai berikut:



Gambar 6.6: Tahapan Uji Levene

5. Selanjutnya, pilih Levene Test untuk Untransformed, klik continue, dan klik ok. Uji homogenitas menghasilkan output yang cukup banyak. Sehingga, untuk keperluan penelitian umumnya hanya memerlukan output dari Homogeneity of Variance, yaitu output yang terdapat pada menu Options.

6. Cara menafsirkan uji Levene ini adalah, jika nilai Levene Statistic $> 0,05$ maka dapat dikatakan bahwa variasi data adalah homogen.

Tabel 6.5: Test of Homogeneity of Variance

		Levene Statistic	df1	df2	Sig.
Hasil Belajar	Based on Mean	5,014	1	22	0,036
	Based on Median	1,918	1	22	0,180
	Based on Median and with adjusted df	1,918	1	17,974	0,183
	Based on trimmed mean	4,990	1	22	0,036

Test Results		
Box's M		2,349
F	Approx	2,349
	df1	1
	df2	1452,000
	Sig.	0,134

Kesimpulan:

Hasil uji Levene menggunakan bantuan program SPSS dihasilkan nilai signifikan sebesar $0,134 > 0,05$ yang berarti kedua kelas tidak berbeda secara signifikan yang bermakna varians kedua kelas yang dibandingkan adalah sama atau homogen.

Bab 7

Uji T-Test

7.1 Pengertian Uji T

Uji beda dua *mean* adalah uji statistik yang membandingkan *mean* dua kelompok data, atau bila seorang peneliti ingin mengetahui apakah parameter dua populasi berbeda atau tidak, maka uji statistik yang digunakan disebut uji beda dua *mean*. Umumnya, pendekatan yang dilakukan bisa dengan distribusi Z (uji Z), ataupun distribusi t (uji t). Uji Z dapat digunakan bila; (1) standar deviasi populasi (σ) diketahui, dan (2) jumlah sampelnya besar (> 30). Bila kedua syarat tersebut tidak terpenuhi, maka jenis uji yang digunakan adalah uji t dua sampel (*two sample t-test*) (Munfarikhatin *et al.*, 2021).

Sebelum kita melakukan uji statistik dua kelompok data, kita perlu perhatikan apakah dua kelompok data tersebut berasal dari dua kelompok yang independen atau berasal dari dua kelompok yang dependen / pasangan (Amir, 2014). Dikatakan kedua kelompok data dependen bila data kelompok yang satu tidak tergantung dari data kelompok kedua, misalnya membandingkan *mean* tekanan darah sistolik orang desa dengan orang kota atau tekanan darah orang kota independen (tidak tergantung) dengan orang desa.

Pengujian hipotesis perbandingan rata-rata dilakukan untuk melihat kesesuaian dugaan peneliti terhadap suatu objek yang diteliti dengan kenyataannya. Misalnya seorang peneliti ingin mengetahui bahwa nilai rata-rata ujian nasional

siswa kelas XII SMA Se-Kabupaten X pada bidang studi Matematika adalah 80. Kemudian peneliti tersebut melakukan penelitian terhadap siswa kelas XII pada salah satu SMA di Kabupaten X yang hasilnya akan digeneralisasikan dengan pertanyaan penelitian dari peneliti tersebut.

Konsep Uji T

Uji beda rata-rata dikenal juga dengan nama uji-t (*t-test*). Konsep dari uji beda rata-rata adalah membandingkan nilai rata-rata beserta selang kepercayaan tertentu (*confidence interval*) dari dua populasi. Prinsip pengujian dua rata-rata adalah melihat perbedaan variasi kedua kelompok data (Sukestiyarno, 2013).

Oleh karena itu dalam pengujian ini diperlukan informasi apakah varian kedua kelompok yang diuji sama atau tidak. Varian kedua kelompok data akan berpengaruh pada nilai standar eror yang akhirnya akan membedakan rumus pengujiannya (Birhan *et al.*, 2021).

Dalam menggunakan uji-t ada beberapa syarat yang harus dipenuhi. Syarat/asumsi utama yang harus dipenuhi dalam menggunakan uji-t adalah data harus berdistribusi normal, Jika data tidak berdistribusi normal, maka harus dilakukan transformasi data terlebih dahulu untuk menormalkan distribusinya (Sugiyono, 2010). Jika transformasi yang dilakukan tidak mampu menormalkan distribusi data tersebut, maka uji-t tidak valid untuk dipakai, sehingga disarankan untuk melakukan uji non-parametrik seperti Wilcoxon (data berpasangan) atau Mann-Whitney U (data independen).

Berdasarkan karakteristik datanya maka uji beda dua rata-rata dibagi dalam dua kelompok, yaitu: uji beda rata-rata independen dan uji beda rata-rata berpasangan (Kurniadi and Purwaningrum, 2018).

7.2 Macam-Macam Uji T

Uji T digunakan untuk menguji hipotesis komparatif (uji perbedaan), digunakan untuk sampel kecil dan varian populasi tidak diketahui serta merupakan salah satu teknik statistik parametrik yang digunakan untuk membedakan *mean* kelompok.

Untuk uji T dapat dikelompokkan menjadi tiga macam bentuk. Diantaranya *One sample t-test*, *Independent sample t-test*, dan *Paired sample t-test*.

7.2.1 Uji Beda Satu Mean

Uji ini digunakan untuk satu sampel, prinsipnya menguji apakah suatu nilai tertentu (yang diberikan sebagai pembandingan) berbeda secara nyata ataukah tidak dengan rata-rata sebuah sampel. Nilai yang dimaksud pada umumnya adalah nilai parameter untuk mengukur suatu populasi, yang dirumuskan sebagai berikut:

$$t = \frac{\bar{x} - \mu}{\left(\frac{SD}{\sqrt{n}}\right)}$$

Keterangan:

t= Nilai t hitung

x= Rata-rata sampel

μ = Nilai Parameter

SD= Standar Deviasi Sampel

n= Jumlah Sampel

Untuk menginterpretasikan hasil uji tersebut, maka langkah-langkahnya adalah sebagai berikut:

1. Tentukan nilai.
2. df (*degree of freedom*) atau derajat kebebasan = N-k.
3. df (*degree of freedom*) untuk one sample t-test df=N.
4. Bandingkan nilai t-hitung dengan nilai t-tabel.
 - a. Apabila nilai t hitung > t tabel maka terdapat perbedaan secara signifikan (H_0 ditolak dan H_a diterima)
 - b. Apabila nilai t hitung < t tabel maka terdapat perbedaan secara signifikan (H_0 diterima dan H_a ditolak).

Contoh Soal (Uji Pihak Kanan)

Proses pembuatan barang rata-rata menghasilkan 15,7 unit per jam. Hasil produksi memiliki varians 2,3. metode baru diusulkan untuk mengganti metode lama jika rata-ratanya per jam menghasilkan paling sedikit 16 buah. Untuk menentukan apakah metode akan diganti atau tidak, metode baru dicoba 20 kali dan ternyata rata-rata per jam menghasilkan 16,9 buah.

Pengusaha bermaksud mengambil risiko 5% untuk menggunakan metode baru apabila metode ini rata-rata menghasilkan lebih dari 16 buah. Apakah keputusan yang akan diambil pengusaha?

Penyelesaian:

Diketahui $\bar{x} = 16,9$ $n = 20$ $\sigma = \sqrt{2,3}$ dan $\mu_0 = 16$

Langkah pengujian hipotesis

1. $H_0 = \mu = \mu_0 \rightarrow H_0 = \mu = 16$

$H_1 = \mu \neq \mu_0 \rightarrow H_1 = \mu \neq 16$

2. Taraf signifikansi $\alpha = 5\%$

3. Kriteria pengujian

Tolak H_0 jika $z \geq z_{0,5-\alpha} \rightarrow z_{0,5-\alpha} = z_{0,5-0,05} = 1,64$

4. Statistik hitung berdasarkan data penelitian yang diambil

$$z = \frac{\bar{x} - \mu_0}{\frac{\sigma}{\sqrt{n}}} = \frac{16,9 - 16}{\sqrt{\frac{2,3}{20}}} = 2,65$$

5. Kesimpulan: karena $z_{\text{hitung}} = 2,65 > z_{0,5-\alpha} = 1,64$ terletak pada daerah kritis maka H_0 ditolak, Jadi $\mu > 16$, sehingga dapat disimpulkan bahwa dengan risiko 5% metode baru dapat menggantikan metode lama.

Contoh Soal (Uji pihak kiri dalam hal varians tidak diketahui)

Masyarakat mengeluh dan mengatakan bahwa isi bersih makanan kaleng tidak sesuai dengan yang tertera pada kemasannya sebesar 5 ons. Untuk meneliti hal ini, 23 kaleng makanan diteliti secara acak. Dari sampel tersebut diperoleh berat rata-rata 4,9 ons dan simpangan baku 0,2 ons. Dengan taraf nyata 5%, bagaimanakah pendapat anda mengenai keluhan masyarakat tersebut.

Penyelesaian

Diketahui $x = 4,9$; $n = 23$; $s = 0,2$; $\mu_0 = 5$

Langkah pengujian hipotesis jika varians populasi tidak diketahui:

1. Hipotesis pengujian

$H_0 = \mu = \mu_0 \rightarrow H_0 = \mu = 5$

$$H_1 = \mu \neq \mu_0 \rightarrow H_1 = \mu < 5$$

Jika rata-rata berat makanan kaleng tidak kurang dari 5 ons tentu masyarakat tidak akan mengeluh.

2. Taraf signifikansi $\alpha = 5\%$
3. Kriteria pengujian

Tolak H_0 jika $t \geq -t_{1-\alpha} \rightarrow -t_{1-0,05} = -1,72$ dengan $dk = 23-1 = 22$

4. Statistik hitung berdasarkan data hasil penelitian (sampel) yang diambil

$$t = \frac{\bar{x} - \mu_0}{\frac{s}{\sqrt{n}}} = \frac{4,9 - 5}{\frac{0,2}{\sqrt{23}}} = -2,398$$

5. Kesimpulan

Karena $t_{\text{hitung}} = -2,398 < -t_{1-\alpha} = -1,72$ terletak pada daerah kritis maka H_0 ditolak, sehingga $\mu < 5$. Sehingga dapat disimpulkan penelitian tersebut menguatkan keluhan masyarakat mengenai berat makanan kaleng yang kurang dari berat yang tertera pada kemasan yaitu 5 ons.

7.2.2 Uji Beda Dua Mean

Uji bertujuan untuk mengetahui perbedaan *mean* dua kelompok data independen syarat atau asumsi yang harus dipenuhi adalah sebagai berikut:

1. Data berdistribusi normal / simetris
Jika 10 sampel Tinggi Badan diambil dari populasi 5000 Mahasiswa sebuah Perguruan Tinggi, maka data Tinggi Badan 5000 Mahasiswa tersebut haruslah berdistribusi normal atau bisa dianggap normal.
2. Kedua kelompok data independen.
3. Variabel yang dihubungkan berbentuk numerik dan kategori (dengan hanya dua kelompok) dan Variabel (data) yang diuji haruslah data bertipe interval atau rasio, yang tingkatnya lebih tinggi dari data tipe nominal atau ordinal. Tinggi Badan Pria atau Wanita (centimeter) jelas bertipe rasio, karena didapat dari proses mengukur. Namun Pendapat atau Sikap Pria dan Wanita (Suka atau Tidak Suka yang diukur dengan skala Likert) bukanlah data interval atau rasio, namun data Ordinal. Pada uji t dan uji F untuk dua sampel atau lebih, kedua sampel diambil

dari dua populasi yang mempunyai varians sama. Jadi jika diambil sampel 10 Tinggi Badan Pria dan 10 Tinggi Badan Wanita dari 3000 Pria dan 2000 Wanita, maka varians 3000 Tinggi Badan Pria dan varians 2000 Tinggi Badan Wanita haruslah sama atau bisa dianggap sama.

Prinsip pengujian dua *mean* adalah melihat perbedaan variasi kedua kelompok data. Oleh karena itu dalam pengujian ini diperlukan informasi apakah varian kedua kelompok yang diuji sama atau tidak. Bentuk varian kedua kelompok data akan berpengaruh pada nilai standar eror yang akhirnya akan membedakan rumus pengujiannya.

Uji Beda Dua Mean (Dua Pihak)

Banyak penelitian yang memerlukan perbandingan antara dua populasi. Misalnya membandingkan hasil belajar, daya kerja suatu obat, dsb. Maka akan digunakan dasar distribusi sampling mengenai selisih statistik, misalnya selisih rata-rata dan selisih proporsi (Munfarikhatin, 2019).

Misalkan dipunyai dua buah populasi, keduanya berdistribusi normal dengan rata-rata dan simpangan baku masing-masing μ_1 dan σ_2 untuk populasi kedua. Secara independen diambil sebuah sampel acak dengan ukuran n_1 dan n_2 dari masing-masing populasi. Rata-rata dan simpangan baku dari sampel itu berturut-turut \bar{x}_1 , s_1 dan \bar{x}_2 , s_2 . Akan diuji tentang rata-rata μ_1 dan μ_2 .

1. Jika $\sigma_1 = \sigma_2 = \sigma$ dan σ diketahui

Langkah-langkah pengujian hipotesis:

- a. Hipotesis pengujian

$$H_0 = \mu_1 = \mu_2$$

$$H_1 = \mu_1 \neq \mu_2$$

- b. Tentukan besarnya taraf signifikansi

- c. Kriteria pengujian

Terima H_0 jika $-z_{\frac{1}{2}(1-\alpha)} < z < z_{\frac{1}{2}(1-\alpha)}$, selainnya H_0 ditolak.

Dengan $z_{\frac{1}{2}(1-\alpha)}$ diperoleh dari daftar distribusi normal baku dengan peluang $\frac{1}{2}(1 - \alpha)$.

- d. Statistik hitung berdasarkan data penelitian (sampel) yang diambil.

$$z = \frac{\bar{x}_1 - \bar{x}_2}{\sigma \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}$$

- e. Menarik kesimpulan berdasarkan hasil 3 dan 4.
2. Jika $\sigma_1 = \sigma_2 = \sigma$ dan σ tidak diketahui
Langkah-langkah pengujian hipotesis:
- Hipotesis pengujian

$$H_0 = \mu_1 = \mu_2$$

$$H_1 = \mu_1 \neq \mu_2$$
 - Tentukan besarnya taraf signifikansi
 - Kriteria pengujian
Terima H_0 jika $-t_{1-\frac{1}{2}\alpha} < t < t_{1-\frac{1}{2}\alpha}$, selainnya H_0 ditolak.
Dengan $t_{1-\frac{1}{2}\alpha}$ diperoleh dari daftar distribusi normal baku dengan peluang $1 - \frac{1}{2}\alpha$ dan $dk = n_1 + n_2 - 2$.
 - Statistik hitung berdasarkan data penelitian (sampel yang diambil) adalah.

$$t = \frac{\bar{x}_1 - \bar{x}_2}{s \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}$$

$$s^2 = \frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2}$$

- e. Menarik kesimpulan berdasarkan hasil 3 dan 4.

Contoh Soal:

Dua macam makanan A dan B diberikan kepada ayam secara terpisah untuk jangka waktu tertentu. Ingin diketahui makanan mana yang lebih baik bagi ayam. Sampel acak yang terdiri atas 11 ayam diberi makanan A dan 10 ayam diberi makanan B.

Hasil percobaan pertambahan berat badan ayam (ons) sebagai berikut.

Makanan A	3,1	3,0	3,3	2,9	2,6	3,0	3,6	2,7	3,8	4,0	3,4
Makanan B	2,7	2,9	3,4	3,2	3,3	2,9	3,0	3,0	2,6	3,7	

Penyelesaian:

Diketahui dari data di atas $\bar{x}_A = 3,22$; $\bar{x}_B = 3,07$; $s_A^2 = 0,1996$; $s_B^2 = 0,1112$. Pada kasus ini populasi dianggap normal dan variansnya tidak diketahui namun sama besar. Langkah pengujian hipotesis dalam hal $\sigma_1 = \sigma_2 = \sigma$ tetapi σ tidak diketahui.

a. Hipotesis Pengujian

$$H_0 = \mu_1 = \mu_2$$

$$H_1 = \mu_1 \neq \mu_2$$

b. Taraf signifikansi 5%

c. Kriteria pengujian

Terima H_0 jika $-t_{1-\frac{1}{2}\alpha} < t < t_{1-\frac{1}{2}\alpha}$, dengan $dk = n_1 + n_2 - 2 = 11 + 10 - 2 = 19$

$$-t_{1-\frac{1}{2}\alpha} < t < t_{1-\frac{1}{2}\alpha} \rightarrow -t_{1-\frac{1}{2}0,05} < t < t_{1-\frac{1}{2}0,05} \rightarrow -2,09 < t < 2,09$$

d. Statistik hitung berdasarkan data penelitian (sampel) yang diambil.

Simpangan baku gabungan $s^2 = \frac{(n_1-1)s_1^2 + (n_2-1)s_2^2}{n_1+n_2-2}$ diperoleh $s = 0,397$.

$$t = \frac{\bar{x}_1 - \bar{x}_2}{s \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} = \frac{3,22 - 3,07}{(0,397) \sqrt{\frac{1}{11} + \frac{1}{10}}} = 0,862$$

e. Kesimpulan: karena $-2,09 < t \text{ hitung} = 0,862 < 2,09$ terletak dalam daerah penerimaan H_0 maka H_0 diterima. Jadi, $\mu_1 = \mu_2$. Artinya, kedua macam makanan tersebut memberikan pertambahan berat badan ayam yang sama, sehingga kedua makanan tersebut sama baiknya.

3. Jika $\sigma_1 \neq \sigma_2$ dan keduanya tidak diketahui

Untuk kasus ini belum ada statistik yang tepat yang dapat digunakan. Pendekatan yang cukup memuaskan adalah dengan menggunakan statistik t .

Langkah pengujian hipotesis dalam hal $\sigma_1 \neq \sigma_2$ dan keduanya tidak diketahui.

a. Hipotesis pengujian

$$H_0 = \mu_1 = \mu_2$$

$$H_1 = \mu_1 \neq \mu_2$$

b. Tentukan besarnya taraf signifikansi

c. Kriteria Pengujian.

Terima H_0 jika $-\frac{w_1 t_1 + w_2 t_2}{w_1 + w_2} < t' < \frac{w_1 t_1 + w_2 t_2}{w_1 + w_2}$ untuk harga t yang lain H_0 ditolak.

$$\text{Dengan } w_1 = \frac{s_1^2}{n_1}; w_2 = \frac{s_2^2}{n_2}$$

$$t_1 = t_{(1-\frac{1}{2}\alpha), (n_1-1)} \text{ dan } t_2 = t_{(1-\frac{1}{2}\alpha), (n_2-1)}$$

$t_{\beta, m}$ diperoleh dari daftar distribusi Student dengan peluang β dan dk = m.

d. Statistik hitung berdasarkan penelitian (sampel) yang diambil.

$$t' = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}$$

e. Menarik kesimpulan berdasarkan hasil 3 dan 4.

Contoh Soal:

Suatu barang dihasilkan dengan menggunakan dua proses. Ingin diketahui apakah kedua proses itu menghasilkan barang yang sama kualitasnya ditinjau dari rata-rata daya tekannya. Maka diadakan percobaan sebanyak 20 kali masing-masing dari hasil proses pertama maupun kedua. Diperoleh informasi $\bar{x}_1 = 9,25$ kg; $\bar{x}_2 = 10,4$ kg; $s_1 = 2,24$ kg; $s_2 = 3,12$ kg. Bila populasinya

dianggap normal dengan varians kedua populasi tidak sama, dengan taraf nyata 5%, ujlilah bagaimana hasilnya

Penyelesaian:

Diketahui $\bar{x}_1 = 9,25$ kg; $\bar{x}_2 = 10,4$ kg; $s_1 = 2,24$ kg ; $s_2 = 3,12$ kg. Pada kasus ini populasi dianggap normal dan variansnya tidak diketahui namun sama besar. Langkah pengujian hipotesis dalam hal $\sigma_1 \neq \sigma_2$ dan keduanya tidak diketahui

a. Hipotesis pengujian

$H_0 = \mu_1 = \mu_2$: kedua proses menghasilkan barang dengan kualitas rata-rata daya tekan yang sama.

$H_1 = \mu_1 \neq \mu_2$: kedua proses menghasilkan barang dengan kualitas rata-rata daya tekan yang berbeda.

b. Taraf signifikansi 5%

c. Kriteria pengujian

Terima H_0 jika $-\frac{w_1 t_1 + w_2 t_2}{w_1 + w_2} < t' < \frac{w_1 t_1 + w_2 t_2}{w_1 + w_2}$

$$w_1 = \frac{s_1^2}{n_1} = \frac{5,0176}{20} = 0,2509; w_2 = \frac{s_2^2}{n_2} = \frac{9,7344}{20} = 0,4867$$

$$t_1 = t_{(1-\frac{1}{2}\alpha), (n_2-1)} = t_{(1-\frac{1}{2}0,05), (20-1)} = t_{0,975,19} = 2,09$$

$$t_2 = t_{(1-\frac{1}{2}\alpha), (n_2-1)} = t_{(1-\frac{1}{2}0,05), (20-1)} = t_{0,975,19} = 2,09$$

Sehingga

$$-\frac{w_1 t_1 + w_2 t_2}{w_1 + w_2} < t' < \frac{w_1 t_1 + w_2 t_2}{w_1 + w_2}$$

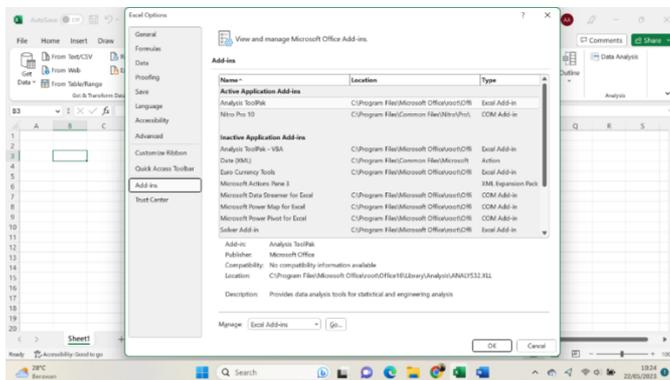
$$-\frac{(0,2509)(2,09) + (0,4867)(2,09)}{(0,2509) + (0,4867)} < t' < \frac{(0,2509)(2,09) + (0,4867)(2,09)}{(0,2509) + (0,4867)}$$

$$-2,09 < t' < 2,09$$

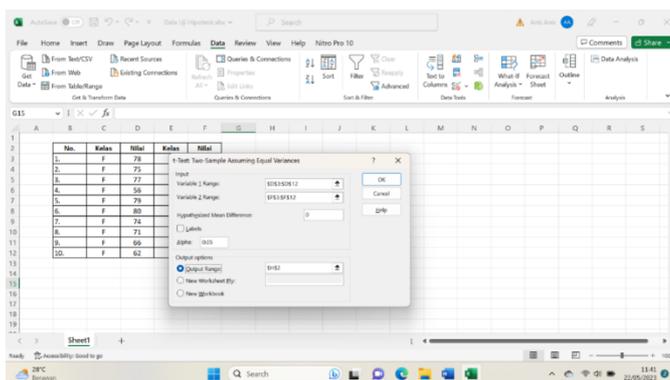
d. Statistik hitung berdasarkan data penelitian (sampel) yang diambil.

$$t' = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}} = \frac{9,25 - 10,4}{\sqrt{\frac{5,0167}{20} + \frac{9,7344}{20}}} = 1,339$$

2. Jika pada menu Data belum muncul data Analisis maka klik menu **File** → **Option** → pilih **Ads In** → **Analysis ToolPak** → **OK**. Maka Opsi Data Analisis akan muncul di sudut kiri atas pada menu Data.



- c. Pada menu Data Analysis inputkan Variabel 1 Range dan Variabel 2 Range. Pada kolom Hypothesized mean differences isikan 0 dengan nilai alpha 0,05. Pada kolom Output option ada tiga pilihan diantaranya **Output Range**: jika hasil akan ditampilkan pada sheet yang sama, **New Worksheet Ply**: jika hasil akan ditampilkan pada sheet yang berbeda, **New workbook**: jika hasil akan ditampilkan pada lembar kerja baru.



d. Diperoleh data output sebagai berikut:

No.	Kelas	Nilai	Kelas	Nilai
1	F	78	G	65
2	F	75	G	70
3	F	77	G	85
4	F	77	G	65
5	F	79	G	56
6	F	80	G	78
7	F	74	G	71
8	F	71	G	73
9	F	68	G	77
10	F	62	G	60
11	Rata-rata	71,8		67,7

t-Test: Two-Sample Assuming Equal Variances	
Variable 1	Variable 2
Mean	71,8 67,7
Variance	64,4 67,57
Observations	10 10
Pooled Variance	65,98
Hypothesized Mean Difference	0
df	18
t Stat	1,83
P(T<=t) one-tail	0,34
t Critical one-tail	1,73
P(T<=t) two-tail	0,27
t Critical two-tail	2,30

e. Pada nilai $P (T \leq t)$ two tail diperoleh nilai $0,27 >$ nilai alfa $0,05$. Maka H_0 diterima dan H_1 ditolak. Kesimpulannya adalah tidak terdapat perbedaan rata-rata nilai siswa kelas F dan kelas G.

Bab 8

Uji Chi Square

8.1 Pendahuluan

Uji Chi Square adalah salah satu distribusi teoritis yang sering digunakan dalam statistika inferensial, (Furqon, 2004a). Uji Chi Square salah satu distribusi yang berasal dari distribusi normal baku. Uji Chi Square atau Uji kai kuadrat (dilambangkan dengan " χ^2 " dari huruf Yunani "*Chi*" dilafalkan "Kai") digunakan untuk menguji dua kelompok data baik variabel independen maupun dependen yang berbentuk data kategori (Freund & Walpole (Furqon, 2004a); (Riduwan, 2009).

Uji Chi Square dapat juga dikatakan sebagai uji proporsi untuk dua peristiwa atau lebih, sehingga datanya bersifat diskrit. Misalnya ingin mengetahui hubungan antara status gizi ibu (baik atau kurang) dengan kejadian Berat Badan Lahir Rendah (ya atau tidak).

Dasar uji Chi Square itu sendiri adalah membandingkan perbedaan frekuensi hasil observasi (O) dengan frekuensi yang diharapkan (E). Perbedaan tersebut meyakinkan jika harga dari Kai Kuadrat (Chi Square) sama atau lebih besar dari suatu harga yang ditetapkan pada taraf signifikan tertentu (dari tabel χ).

Uji (Chi Square) Kai Kuadrat dapat digunakan untuk menguji:

1. Uji χ^2 untuk ada tidaknya hubungan antara dua variabel (Independency test).
2. Uji χ^2 untuk homogenitas antar sub kelompok (Homogeneity test).
3. Uji χ^2 untuk Bentuk Distribusi (Goodness of Fit)

Sebagai rumus dasar dari uji Kai Kuadrat adalah:

$$X^2 = \sum \frac{(f_o - f_e)^2}{f_e}$$

Keterangan:

f_o = Frekuensi hasil observasi. f_e = Frekuensi yang diharapkan.

X^2 = Nilai Chi Square (Riduwan, 2009)

$$f_e = \frac{(\sum fk) \cdot (\sum fb)}{\sum T}$$

Keterangan:

f_e = Frekuensi yang diharapkan

$\sum fk$ = Jumlah frekuensi pada kolom

$\sum fb$ = Jumlah frekuensi pada baris

$\sum T$ = Jumlah keseluruhan baris atau kolom

Dalam melakukan Uji Chi Square, harus memenuhi syarat: Sampel dipilih secara acak

1. Semua pengamatan dilakukan dengan independen.
2. Setiap kolom paling sedikit berisi frekuensi harapan sebesar 1 (satu). Frekuensi harapan kurang dari 5 tidak melebihi 20% dari total kolom.
3. Besar sampel sebaiknya > 40 .

Keterbatasan penggunaan uji Kai Kuadrat (Uji Chi Square adalah memakai data yang diskrit dengan pendekatan distribusi kontinu. Dekatnya pendekatan yang dihasilkan tergantung pada ukuran pada berbagai kolom dari tabel kontingensi.

Untuk menjamin pendekatan yang memadai digunakan aturan dasar “frekuensi harapan tidak boleh terlalu kecil” secara umum dengan ketentuan:

1. Tidak boleh ada kolom yang mempunyai nilai harapan lebih kecil dari 1 (satu).
2. Tidak lebih dari 20% kolom mempunyai nilai harapan lebih kecil dari 5 (lima).

Bila hal ini ditemukan dalam suatu tabel kontingensi, cara untuk menanggulangnya adalah dengan menggabungkan nilai dari kolom yang kecil ke kolom lainnya (penggabungan), artinya kategori dari variabel dikurangi sehingga kategori yang nilai harapannya kecil dapat digabung ke kategori lain. Khusus untuk tabel 2x2 hal ini tidak dapat dilakukan, maka solusinya adalah melakukan uji “*Fisher Exact atau Koreksi Yates*” (Riduwan, 2009); (Furqon, 2004b).

8.2 Contoh Kasus 1

Suatu survei ingin mengetahui apakah ada perbedaan Asupan vitamin C dengan Kesehatan kulit pada penduduk desa X, kemudian diambil sampel sebanyak 100 orang yang terdiri dari 70 orang asupan Vitamin C baik dan 30 orang asupan Vitamin C kurang.

Setelah dilakukan pemeriksaan kulit ternyata dari 70 orang yang asupan Vitamin C baik, ada 5 orang yang dinyatakan kulitnya tampak kurang sehat seperti bersisik. Sedangkan dari 30 orang yang asupan Vitamin C kurang ada 20 orang yang kulitnya seperti bersisik. Ujilah apakah ada perbedaan proporsi pada kedua kelompok tersebut.

Tabel 8.1: Proporsi Asupan Vitamin C

Asupan Vitamin C	Kulit Sehat		Jumlah
	Ya	Tidak	
Baik	65	5	70
Kurang	10	20	30
	75	25	100

Langkah-langkah:

1. Membuat Hipotesis.
2. Menentukan frekuensi (f) yang diharapkan.
3. Mencari nilai X^2 hitung X^2 tabel.
4. Mencari nilai X^2 tabel, $\alpha = 0,05$, dk = (2-1). (2-1).

Langkah 1:

H0: P1= P2

Tidak ada perbedaan proporsi kulit kurang sehat pada asupan vitamin yang baik dan yang kurang baik.

H0: P1≠P2

Ada perbedaan proporsi kulit kurang sehat pada asupan vitamin yang baik dan yang kurang baik.

Langkah 2:

Kemudian tentukan nilai observasi (f) dan nilai ekspektasi (f):

$$f_{e1} = \frac{75.70}{100} = 52,50$$

$$f_{e2} = \frac{75.30}{100} = 22,50$$

$$f_{e3} = \frac{25.70}{100} = 17,50$$

$$f_{o4} = \frac{25.30}{100} = 7,50$$

Langkah 3.

Selanjutnya masukan dalam rumus:

$$X^2 = \frac{(65 - 52,50)^2}{52,50} + \frac{(10 - 22,50)^2}{22,50} + \frac{(5 - 17,50)^2}{17,50} + \frac{(20 - 7,50)^2}{7,50}$$

$$X^2 = \frac{(-12,5)^2}{52,50} + \frac{(-12,5)^2}{22,50} + \frac{(-12,5)^2}{17,50} + \frac{(12,5)^2}{7,50}$$

$$X^2 = \frac{156,25}{52,50} + \frac{156,25}{22,50} + \frac{156,25}{17,50} + \frac{156,25}{7,5}$$

$$X^2 = 39,66$$

Syarat keputusan penolakan atau penerimaan H_0 :

Jika nilai X^2 hitung kurang (lebih kecil) dari nilai X^2 tabel, maka H_0 tidak ditolak, artinya H_0 diterima. Hal tersebut menunjukkan bahwa tidak ada perbedaan antara proporsi 1 dan proporsi 2 sehingga $P_1 = P_2$, Tidak ada perbedaan proporsi kulit kurang sehat pada kedua kelompok tersebut. Sebaliknya bila nilai X^2 hitung lebih besar atau sama dengan nilai X^2 tabel, maka H_0 ditolak, artinya ada perbedaan proporsi kulit kurang sehat pada kedua kelompok tersebut.

Dari perhitungan di atas menunjukkan bahwa χ^2 hitung $39,66 > \chi^2$ tabel $= 3,84$ (df = k-1, dengan $\alpha = 0,05$), sehingga H_0 ditolak, hal berarti bahwa ada perbedaan yang signifikan terhadap proporsi asupan vitamin C dengan Kesehatan kulit. Sebagai kesimpulan bahwa ada perbedaan yang bermakna/signifikan proporsi antara asupan vitamin C dengan Kesehatan kulit (Murti; Bhisma, 1996); (Sabri, 2008).

8.3 Contoh Kasus 2

Diadakan penelitian di Perguruan Tinggi Advent Surya Nusantara, yang bertujuan untuk mengetahui Pelaksanaan Disiplin Kampus (PDK) antara Mahasiswa STFT yang berjumlah 150 orang, Mahasiswa AKPER berjumlah 145 orang dan Mahasiswa STIE berjumlah 148 orang.

Frekuensi Observasi dari 443 orang PDN PTASN:

Mahasiswa	Pelaksanaan Disiplin Kampus			Total
	85-100	66-84	0-65	
STFT	75	50	25	150
AKPER	60	45	40	145
STIE	58	40	50	148
JUMLAH	193	135	115	443

Langkah-langkah:

1. Buat H_0 dan H_a dalam bentuk kalimat.
2. Tentukan frekuensi yang diharapkan (f_e).
3. Cari nilai Chi Square (X^2 hitung).

4. Cari nilai Chi Square (X^2 tabel).

Jawab:

Langkah 1:

H_0 : Melaksanakan disiplin kampus dengan baik (Tidak ada perbedaan yang signifikan) antara Mahasiswa STFT, Mahasiswa AKPER, dan Mahasiswa STIE dalam pelaksanaan disiplin Kampus

H_1 : Ada perbedaan yang signifikan antara Mahasiswa STFT, Mahasiswa AKPER dan Mahasiswa STIE dalam pelaksanaan disiplin kampus

Langkah 2:

Mencari frekuensi yang diharapkan (f) dengan rumus:

$$f_e = \frac{(\sum fk) \cdot (\sum fb)}{\sum T}$$

$$f_{e1} = \frac{(193 \cdot 150)}{443} = 65,34$$

$$f_{e2} = \frac{(193 \cdot 145)}{443} = 63,17$$

$$f_{e3} = \frac{(193 \cdot 148)}{443} = 64,47$$

$$f_{e4} = \frac{(135 \cdot 150)}{443} = 45,71$$

$$f_{e5} = \frac{(135 \cdot 145)}{443} = 44,18$$

$$f_{e6} = \frac{(135 \cdot 148)}{443} = 45,10$$

$$f_{e7} = \frac{(115 \cdot 150)}{443} = 38,93$$

$$f_{e8} = \frac{(115 \cdot 145)}{443} = 37,64$$

$$f_{e9} = \frac{(115 \cdot 148)}{443} = 38,41$$

Langkah 3.

Mencari Chi Square (X^2)

$$X^2 = \sum \frac{(f_o - f_e)^2}{f_e}$$

$$X^2 = \frac{(75 - 65,34)^2}{65,34} + \frac{(60 - 63,17)^2}{63,17} + \frac{(58 - 64,47)^2}{64,47} +$$

$$\frac{(50 - 45,71)^2}{45,71} + \frac{(45 - 44,18)^2}{44,18} + \frac{(40 - 45,10)^2}{45,10} +$$

$$\frac{(25 - 38,93)^2}{38,93} + \frac{(40 - 37,64)^2}{37,64} + \frac{(50 - 38,41)^2}{38,41}$$

$$X^2 = 1,42 + 0,15 + 2,69 + 0,28 + 0,01 + 0,01 + 4,98 + 0,14 + 3,49$$

$$= 14,17$$

Langkah 4:

Mencari X^2 tabel dengan rumus $Dk = (k-1) \cdot (b-1) = 2 \times 2$, X^2 tabel untuk $\alpha = 0,05 = 9,48$

Syarat Jika X^2 hitung $\leq X^2$ tabel, maka H_0 diterima.

Jika X^2 hitung $\geq X^2$ tabel, maka H_0 tidak diterima.

Dari hasil perhitungan ternyata bahwa X^2 hitung sebesar $14,17 \geq X^2$ tabel sebesar $9,48$, hal ini berarti ada perbedaan yang signifikan antara mahasiswa STFT, Mahasiswa AKPER dan Mahasiswa STIE dalam pelaksanaan Disiplin Kampus (PDK) di lingkungan Perguruan Advent Surya Nusantara.

Mahasiswa STFT, Mahasiswa AKPER, Mahasiswa STIE cenderung melaksanakan PDK dengan predikat rendah karena kurang disiplin, kurang patuh, kurang mengindahkan peraturan yang dibuat pihak kampus.

Soal Latihan:

Berikut ini adalah tabel frekuensi observasi dari 280 orang pengguna gadget.
Tabel Frekuensi dari 300 Orang Pengguna Gadget

Gadget	Gadget Paling Digemari			Total
Samsung	90	25	10	125
Oppo	35	40	25	100
IPhone	30	10	15	55
Jumlah	155	75	50	280

Jika diketahui uji hipotesis adalah sebagai berikut:

H₀: Pengguna gadget memiliki minat yang sama terhadap tiga merek gadget.

H_a: Pengguna gadget memiliki minat yang berbeda terhadap tiga merek gadget.

Bab 9

Analisis Varians

9.1 Pendahuluan

Analisis ANOVA dan analisis regresi merupakan dua metode dalam teori statistik untuk menganalisis perilaku salah satu variabel, namun ada perbedaan yg membedakan kedua analisis tersebut yaitu analisis ANOVA hanya untuk menguji ada tidaknya perbedaan dua sampel atau lebih. Sementara analisis regresi hanya untuk menguji ada tidaknya hubungan dua variabel atau lebih.

Diketahui suatu data hasil pengamatan seperti: skor hasil belajar mahasiswa, berat bayi yang baru lahir, pembayaran pajak orang pribadi, tentulah nilai data-data tersebut akan bervariasi dari yang satu dengan yang lain. Karena adanya variasi ini untuk sekumpulan data, maka ditentukan alat ukurnya, yaitu varians.

Varians untuk sekumpulan data melukiskan derajat perbedaan atau variasi nilai data individu yang ada dalam kelompok data tersebut. Secara umum varians dapat digolongkan ke dalam varians sistematis dan varians galat. Varians sistematis adalah pengukuran karena adanya pengaruh yang menyebabkan skor atau nilai data lebih condong ke satu arah tertentu dibandingkan ke arah lain. Salah satu jenis varians sistematis dalam kumpulan data hasil penelitian adalah varians antar kelompok atau disebut juga varians eksperimental. Varians ini menggambarkan adanya perbedaan antara kelompok-kelompok hasil pengukuran.

Dengan demikian varians ini terjadi karena adanya perbedaan antara kelompok-kelompok individu. Jika uji kesamaan dua rata-rata atau uji t digunakan, untuk perbedaan atau persamaan dua rata-rata, maka uji beberapa rata-rata digunakan untuk mencari perbedaan atau kesamaan beberapa rata-rata. Uji ini disebut dengan nama *Analysis of Variance* (ANOVA atau ANAVA).

Konsep Dasar ANOVA

Analysis of Variance (Anova) adalah sebuah analisis statistik yang menguji perbedaan rata-rata antar grup. Grup di sini bisa berarti kelompok atau jenis perlakuan. Anova ditemukan dan diperkenalkan oleh seorang ahli statistik bernama Ronald Fisher. Anova merupakan singkatan dari *Analysis of variance*. Merupakan prosedur uji statistik yang mirip dengan t test. Namun kelebihan dari Anova adalah dapat menguji perbedaan lebih dari dua kelompok. Berbeda dengan *independent sample t test* yang hanya bisa menguji perbedaan rata-rata dari dua kelompok saja.

Kegunaan dan Asumsi ANOVA

Anova digunakan sebagai alat analisis untuk menguji hipotesis penelitian untuk melihat apakah ada perbedaan rata-rata antara kelompok. Hasil akhir dari analisis ANOVA adalah nilai F test atau F hitung. Nilai F Hitung ini yang nantinya akan dibandingkan dengan nilai pada tabel f . Jika nilai f hitung lebih dari f tabel, maka dapat disimpulkan bahwa menerima H_1 dan menolak H_0 atau yang berarti ada perbedaan bermakna rata-rata pada semua kelompok.

Anova bisa juga digunakan untuk menguji rata-rata tiga atau lebih dari suatu populasi. Analisis ANOVA sering digunakan pada penelitian eksperimen di mana terdapat beberapa perlakuan. Peneliti ingin menguji, apakah ada perbedaan bermakna antara perlakuan tersebut.

Asumsi ANOVA ada 3 yaitu:

1. populasi yang akan diuji berdistribusi normal;
2. antar pengamatan saling bebas;
3. ragam dari populasi-populasi tersebut sama dengan kata lain ragam populasi tersebut homogen

9.2 Ciri dan Prinsip ANOVA

Ciri khas ANOVA adalah adanya satu atau lebih variabel bebas sebagai faktor penyebab dan satu atau lebih variabel respons sebagai akibat atau efek faktor. Contoh penelitian yang dapat menggambarkan penjelasan ini: “Apakah ada pengaruh biaya iklan dan mutu produksi terhadap harga penjualan.” Dari judul tersebut jelas sekali bahwa biaya iklan dan mutu produksi adalah faktor penyebab sedangkan harga penjualan adalah akibat atau efek dari adanya perlakuan faktor.

Ciri lainnya adalah variabel respons berskala data rasio atau interval (numerik atau kuantitatif). Anova merupakan salah satu dari berbagai jenis uji parametrik, karena mensyaratkan adanya distribusi normal pada variabel terikat dalam perlakuan atau distribusi normal pada residual. Syarat normalitas ini mengasumsikan bahwa sampel diambil secara acak dan dapat mewakili keseluruhan populasi agar hasil penelitian dapat digunakan sebagai generalisasi.

Prinsip ANOVA adalah melakukan analisis variabilitas atau data menjadi dua sumber ragam yaitu ragam di dalam kelompok (within) dan ragam antar kelompok (between). Anova merupakan salah satu dari berbagai jenis uji parametris, karena mensyaratkan adanya distribusi normal pada variabel terikat per perlakuan atau distribusi normal pada residual. Syarat normalitas ini mengasumsikan bahwa sampel diambil secara acak dan dapat mewakili keseluruhan populasi agar hasil penelitian dapat digunakan sebagai generalisasi.

Prinsip uji ANOVA adalah melakukan analisis variabilitas data menjadi dua sumber ragam yaitu ragam di dalam kelompok (within) dan ragam antar kelompok (between) Contoh ANOVA Contohnya adalah seorang peneliti ingin menilai adakah perbedaan model pembelajaran A, B dan C terhadap hasil pembelajaran mata kuliah statistik. Di mana dalam penelitian tersebut, di section 1 diberi perlakuan A, di section 2 diberi perlakuan B dan section 3 diberi perlakuan C.

Setelah adanya perlakuan selama satu semester, kemudian dibandingkan hasil belajar semua pengambil kelas statistik 1 (A, B dan C). Masing-masing kelas jumlahnya berkisar antara 30 sampai dengan 40 mahasiswa. Hasil akhir yang didapatkan adalah nilai f hitung. Nilai tersebut dibandingkan dengan nilai dalam tabel f pada derajat kebebasan tertentu (degree of freedom). Jika F hitung $>$ F Tabel, maka disimpulkan bahwa H_1 (Hipotesis Alternatif) diterima, artinya ada

perbedaan secara signifikan pada hasil ujian mahasiswa antar perlakuan model pembelajaran.

9.3 Jenis ANOVA

Jenis ANOVA ada 2 yaitu: One Way ANOVA (melibatkan satu faktor) dan Two Way ANOVA

9.3.1 One Way Analysis of Variance

Pengertian One-Way ANOVA (Analysis Of Variance) Analisis of variance atau ANOVA merupakan salah satu teknik analisis multivariate berfungsi untuk membedakan rerata lebih dari dua kelompok data dengan cara membandingkan variansinya. One-way ANOVA merupakan prosedur yang digunakan untuk menghasilkan analisis variansi satu arah untuk variabel dependen dengan tipe data kuantitatif dengan sebuah variabel independen sebagai variabel faktor. Analisis variansi pertama kali diperkenalkan oleh Sir Ronald Fisher (Bapak Statistik modern). Dalam praktik, analisis variansi dapat merupakan uji hipotesis (lebih sering dipakai) maupun prediksi (estimasi khususnya di bidang genetika terapan).

Tujuan dilakukannya One-Way ANOVA (Analysis Of Variance) adalah untuk membandingkan dua rata-rata atau lebih yang akan digunakan untuk menguji kemampuan generalisasi. Jika hasil pengujian ditemukan kedua sampel tersebut berbeda, maka bisa digeneralisasikan. Sampel dianggap bisa mewakili populasi.

Analisis pada uji ANOVA ditegakkan pada beberapa asumsi. Pertama, populasi harus independen dan data yang diamati juga independen pada kelompoknya. Artinya, setiap sampel tidak berhubungan dengan sampel yang lain. Kedua, populasi yang diteliti dalam uji ini haruslah berdistribusi normal. Normalitas ini harus terpenuhi dikarenakan pada prinsipnya uji ANOVA merupakan uji beda rerata. Sama halnya dengan uji t atau uji Z . Sampel tentu saja harus diambil dari populasi normal dan usahakan sampel berukuran besar.

Normalitas bisa diukur dengan uji-uji normalitas seperti *chi square* atau *lilliefors*. Ketiga, populasi harus memiliki standar deviasi atau variansi yang sama. Kesamaan variansi ini akan berhubungan dengan uji F yang akan dilakukan nantinya. Jika variansi pada populasi tidak sama, maka tidak dapat dilakukan uji F . Syarat yang terakhir adalah apabila sampel yang diambil dari sebuah populasi adalah bersifat bebas maka sampel diambil secara acak.

9.3.2 Two Way Analysis of Variance

Two Way Analysis of Variance melibatkan dua faktor sedangkan *Multi way Analysis of Variance*, melibatkan lebih dari dua faktor. Analisis varians banyak dipergunakan pada penelitian-penelitian yang banyak melibatkan pengujian komparatif yaitu variabel terikat dengan cara membandingkannya pada kelompok-kelompok sampel independen yang diamati.

Ada beberapa asumsi yang dipakai dalam variansi dua arah:

1. Setiap skor dalam sel harus berdistribusi normal. Asumsi ini dapat sedikit diabaikan jika sampel tiap sel cukup banyak.
2. Variasi skor pada setiap sel hendaknya homogen atau sama.

Skor yang ada bebas dari pengaruh variabel yang tidak diteliti. Hal ini bisa dicapai dengan mengambil sampel acak dari populasi yang sudah diklasifikasikan sesuai dengan sel yang ada. Di samping itu perlu dilakukan kontrol atas terjadinya perembesan pengaruh faktor lain maupun antar kelompok itu sendiri.

ANOVA dua arah ini digunakan bila sumber keberagaman yang terjadi tidak hanya karena satu faktor (perlakuan). Faktor lain yang mungkin menjadi sumber keberagaman respons juga harus diperhatikan. Faktor lain ini bisa berupa perlakuan lain atau faktor yang sudah terkondisi. Pertimbangan memasukkan faktor kedua sebagai sumber keberagaman ini perlu bila faktor itu dikelompokkan (blok), sehingga keragaman antar kelompok sangat besar, tetapi kecil dalam kelompoknya sendiri.

Bila disusun dalam bentuk tabel, maka tampilan tabel dua arah terlihat pada Tabel 9.1 berikut:

Tabel 9.1: Anova Dua Arah

Blok	Populasi				Total	Ukuran
	1	2	...	K		
1	X_{11}	X_{12}	...	X_{1k}	B_1	K_1
2	X_{21}	X_{22}	...	X_{2k}	B_2	K_2
...
R	X_{r1}	X_{r2}	...	X_{rk}	B_r	K_r
Total	T_1	T_2	...	T_k	T	
Ukuran	n_1	n_2	...	nk	N	

Tabel 9.2: Tabel ANOVA Dua Arah

Sumber Keragaman	Derajat Bebas	Jumlah Kuadrat	Varian (Ragam)	F _{hitung}	F _{tabel}
Blok	r-1	JKB	(S ₁) ²	(S ₁) ² / (S ₃) ²	F(v ₁ , v ₃)
Antar kolom	k-1	JKK	(S ₂) ²	(S ₂) ² / (S ₃) ²	F(v ₂ , v ₃)
Sisaan	(k-1)(r-1)	JKS	(S ₃) ²		
Total	rk-1	JKT			

k = jumlah populasi atau perlakuan

N = banyaknya pengamatan = n₁ + n₂ + ... + n_k

JKB = Jumlah kuadrat antar baris

$$= \sum_{K_i} \frac{B_i^2}{K_i} - \frac{T^2}{N}$$

JKK = Jumlah kuadrat antar kolom

$$= \sum \frac{T_i^2}{n_i} - \frac{T^2}{N}$$

JKS = Jumlah kuadrat sisaan

$$= JKT - JKK - JKB$$

JKT = Jumlah kuadrat total

$$= \left(\sum X_{ij}^2 \right) - \frac{T^2}{N}$$

$$(S_1)^2 = JKB / v_1$$

$$(S_2)^2 = JKK / v_2$$

$$(S_3)^2 = JKS / v_3$$

Contoh

Suatu studi mengenai preferensi konsumen melibatkan tiga rancangan kemasan yang dijual pada empat pasar swalayan. Hasil penjualan selama satu bulan pengamatan dalam unit penjualan adalah sebagai berikut:

Tabel 9.3: Pengamatan Unit Penjualan Dalam 1 Bulan

Pasar Swalayan	Bentuk Kemasan		
	A	B	C
I	12	34	23
II	15	26	21
III	1	23	8
IV	6	22	16

Dengan menggunakan $\alpha = 0,05$, Buatlah pengujian hipotesis untuk mengetahui:

1. Apakah ada pengaruh bentuk kemasan pada tingkat penjualan?
2. Apakah ada pengaruh perbedaan Pasar Swalayan pada penjualan?

Penyelesaian:

Tabel 9.4: Pengaruh Bentuk Kemasan Pada Tingkat Penjualan

Pasar Swalayan	Bentuk Kemasan			Total
	A	B	C	
I	12	34	23	74
II	15	26	21	62
III	1	23	8	32
IV	6	22	16	44
Total	39	105	68	212

$$JKB = \frac{74^2}{3} + \frac{62^2}{3} + \frac{32^2}{3} + \frac{44^2}{3} - \frac{212^2}{12} = 348$$

$$JKK = \frac{39^2}{4} + \frac{105^2}{4} + \frac{68^2}{4} - \frac{212^2}{12} = 547,17$$

$$JKT = 17^2 + 15^2 + \dots + 16^2 - \frac{212^2}{12} = 940,67$$

$$JKS = JKT - JKK - JKB = 940,67 - 547,17 - 348 = 45,5$$

Tabel 9.5: ANOVA

Sumber Keragaman	Derajat Bebas	Jumlah Kuadrat	Varian (Ragam)	F_{hitung}	F_{tabel}
Blok	3	348	116	15,30	5,14
Antar kolom	2	547,17	273,58	36,09	4,76
Sisaan	16	45,50	7,58		
Total	11	940,67			

Pengujian pengaruh kemasan pada penjualan

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

H_a : tidak semua sama

$F_{hitung} > F_{tabel}$, maka tolak H_0

Kesimpulan: Ada pengaruh dari perbedaan bentuk kemasan pada tingkat penjualan

Pengujian pengaruh pasar swalayan pada penjualan

$$H_0 \mu_1 = \mu_2 = \mu_3 = \mu_4$$

H_0 Tidak ada perbedaan pasar swalayan pada tingkat penjualan

$$H_a \mu_1 \neq \mu_2 \neq \mu_3 \neq \mu_4$$

H_a Semua berbeda

$F_{hitung} > F_{tabel}$, maka tolak H_0

Kesimpulan: Ada pengaruh dari perbedaan pasar swalayan pada tingkat penjualan.

Bab 10

Analisis Korelasi dan Regresi

10.1 Analisis Korelasi

Korelasi merujuk pada hubungan statistik antara dua variabel atau lebih, baik bersifat sebab-akibat maupun tidak. Meskipun dalam arti yang lebih luas, "korelasi" umumnya menunjukkan berbagai jenis asosiasi, tetapi dalam statistik biasanya mengacu pada sejauh mana sepasang variabel berhubungan secara linear. Korelasi merupakan salah satu metode analisis statistik yang paling umum digunakan oleh para peneliti.

Hal ini karena para peneliti biasanya tertarik untuk memahami hubungan antara berbagai peristiwa yang terjadi. Sebagai contoh, seorang peneliti mungkin ingin mengidentifikasi kaitan antara tingkat intensitas belajar dan hasil belajar, atau antara media pembelajaran yang digunakan dengan penguasaan konsep siswa. Begitu juga, hubungan antara latar belakang pendidikan dengan pendapatan juga dapat menjadi perhatian dalam analisis korelasi. Dalam korelasi, hubungan antara dua variabel disebut *bivariate correlation*, sementara hubungan antar lebih dari dua variabel disebut *multivariate correlation*.

Hubungan antara dua variabel seperti korelasi antara intensitas partisipasi dalam tugas (variabel x) dengan produktivitas belajar (variabel y). Hubungan antara lebih dari dua variabel misalnya korelasi antara intensitas komunikasi guru ke

siswa (variabel x_1), kemampuan guru dalam memotivasi siswa (variabel x_2), dengan jumlah siswa yang lulus (variabel y).

Dalam analisis korelasi, tidak ada perbedaan antara variabel independen (X) dan variabel dependen (Y). Sebabnya adalah hubungan antara variabel independen dengan variabel dependen memiliki arti yang sama dengan hubungan variabel dependen terhadap variabel independen. Namun, dalam praktiknya, sering kali peneliti memberi label pada hubungan antara variabel independen dan variabel dependen. Meskipun demikian, hal ini bukanlah permasalahan besar, karena label tersebut sebenarnya hanya digunakan sebagai alat bantu agar pembaca lebih mudah memahami arah hubungan yang ingin disampaikan oleh peneliti dalam penelitiannya.

Korelasi bermanfaat karena dapat menunjukkan hubungan prediktif yang dapat dimanfaatkan dalam dunia praktik. Sebagai contoh, perusahaan pembangkit listrik mungkin akan memproduksi listrik lebih sedikit pada hari yang sejuk berdasarkan korelasi antara permintaan listrik dan cuaca. Dalam contoh ini, terdapat hubungan sebab-akibat, karena cuaca ekstrem menyebabkan orang menggunakan lebih banyak listrik untuk pemanasan atau pendinginan. Perlu diingat bahwa dalam konteks teknik korelasi, hubungan antara dua variabel tidak mengindikasikan hubungan sebab-akibat (hubungan timbal balik).

Arah Korelasi

Ketika kita mengamati hubungan antara dua variabel atau lebih, kita dapat membaginya menjadi dua kategori berdasarkan arahnya, yaitu hubungan searah dan berlawanan arah. Hubungan searah disebut sebagai korelasi positif, sedangkan hubungan berlawanan arah disebut sebagai korelasi negatif.

Dalam korelasi positif, ketika nilai variabel x meningkat, nilai variabel y juga cenderung meningkat. Sebagai contoh, jika motivasi belajar siswa meningkat (variabel x), maka kemungkinan besar akan diikuti peningkatan jumlah siswa yang lulus (variabel y).

Di sisi lain, korelasi negatif terjadi ketika meningkatnya nilai variabel x berdampak pada penurunan nilai variabel y , dan sebaliknya. Sebagai ilustrasi, semakin tinggi intensitas guru marah terhadap siswa (variabel x), maka kemungkinan produktivitas belajar siswa akan menurun (variabel y).

Macam Teknik Korelasi

Berdasarkan tipe data, ada beberapa teknik korelasi yang dapat digunakan, dapat dilihat pada Tabel 10.1.

Tabel 10.1: Tipe Data dan Teknik Korelasinya

Tipe data	Teknik Analisis	Jenis Statistik
Nominal	Contingency Coefficient	Non-Parametrik
Ordinal	Spearman Rank Correlation Kendall Tau	
Interval Rasio	Pearson Product Moment Partial Correlation Multiple Correlation	Parametrik

Arah dari hubungan korelasi ini dipresentasikan oleh suatu nilai yang disebut sebagai koefisien korelasi (r). Koefisien korelasi ini berkisar dari $-1,0$ hingga $+1,0$. Korelasi tertinggi dinyatakan sebagai $-1,0$ atau $+1,0$, sedangkan korelasi terendah adalah 0 .

Jika dalam perhitungan ditemukan nilai korelasi yang melebihi angka 1 ($+/-$), maka dapat disimpulkan bahwa terdapat kesalahan dalam perhitungan. Tanda minus ($-$) menandakan adanya hubungan yang berlawanan arah (tidak searah), sedangkan tanda plus ($+$) mengindikasikan adanya hubungan yang searah. Meskipun jarang dijumpai, ada korelasi dengan koefisien korelasi yang benar-benar sempurna ($+1,0$ atau $-1,0$), atau bahkan benar-benar tidak ada korelasi (nihil 0). Interpretasi nilai koefisien korelasi (r) dapat dilihat pada Tabel 11.2.

Tabel 10.2: Nilai Koefisien Korelasi (r)

Nilai r	Makna
$0,9$ s.d. 1 / $-0,9$ s.d. -1	Sangat Kuat
$0,7$ s.d. $0,9$ / $-0,7$ s.d. $-0,9$	Kuat
$0,5$ s.d. $0,7$ / $-0,5$ s.d. $-0,7$	Sedang
$0,3$ s.d. $0,5$ / $-0,3$ s.d. $-0,5$	Lemah
0 s.d. $0,3$ / -0 s.d. $-0,3$	Sangat Lemah

Korelasi Product Moment

Metode korelasi ini digunakan untuk mengidentifikasi keterkaitan dan menguji hipotesis mengenai hubungan antara dua variabel jika data dari kedua variabel tersebut bersifat interval atau rasio. Karena metode koefisien korelasi Pearson termasuk dalam kategori parametrik, maka diperlukan uji asumsi bahwa kedua variabel memiliki distribusi yang berbentuk normal.

Terdapat beberapa prasyarat atau asumsi dasar yang harus dipenuhi:

1. Data yang digunakan untuk setiap variabel minimal memiliki skala rasio atau interval (yaitu data berupa angka aktual atau data kuantitatif). Meskipun demikian, analisis ini juga bisa diterapkan pada data dari kuesioner dengan skala Likert.
2. Data yang diterapkan pada setiap variabel yang saling terhubung harus memiliki distribusi yang mengikuti pola normal.
3. Harus terdapat hubungan linier antara variabel-variabel penelitian.

$$r_{XY} = \frac{n \sum XY - (\sum X)(\sum Y)}{\sqrt{(n \sum X^2 - (\sum X)^2)(n \sum Y^2 - (\sum Y)^2)}}$$

Contoh:

Sebuah penelitian tentang nilai rata-rata ulangan biologi siswa terhadap nilai ujian akhir semester siswa di sekolah menengah atas. Sampel penelitian melibatkan 25 siswa dengan data sebagai berikut:

Tabel 10.3: Rata-rata Nilai Ulangan dan UAS siswa

Siswa	Ulangan	UAS	Siswa	Ulangan	UAS
Siswa 1	77	81	Siswa 14	64	76
Siswa 2	67	89	Siswa 15	73	94
Siswa 3	73	79	Siswa 16	61	88
Siswa 4	71	81	Siswa 17	73	87
Siswa 5	69	83	Siswa 18	62	89
Siswa 6	68	75	Siswa 19	69	89
Siswa 7	72	81	Siswa 20	75	87
Siswa 8	79	86	Siswa 21	74	84
Siswa 9	62	80	Siswa 22	74	92
Siswa 10	80	89	Siswa 23	61	87
Siswa 11	80	91	Siswa 24	76	86
Siswa 12	80	81	Siswa 25	69	79
Siswa 13	67	81			

Hipotesis:

H_0 : Tidak terdapat hubungan yang signifikan antara rata-rata nilai ulangan dengan nilai UAS biologi siswa.

H_a : Terdapat hubungan yang signifikan antara rata-rata nilai ulangan dengan nilai UAS biologi siswa.

Guna memudahkan menginput perhitungan, mari buat tabel pembantu:

Tabel 10.4: Tabel Pembantu Nilai Ulangan dan UAS Siswa

Siswa	X	Y	X ²	Y ²	XY
Siswa 1	77	81	5929	6561	6237
Siswa 2	67	89	4489	7921	5963
Siswa 3	73	79	5329	6241	5767
Siswa 4	71	81	5041	6561	5751
Siswa 5	69	83	4761	6889	5727
Siswa 6	68	75	4624	5625	5100
Siswa 7	72	81	5184	6561	5832
Siswa 8	79	86	6241	7396	6794
Siswa 9	62	80	3844	6400	4960
Siswa 10	80	89	6400	7921	7120
Siswa 11	80	91	6400	8281	7280
Siswa 12	80	81	6400	6561	6480
Siswa 13	67	81	4489	6561	5427
Siswa 14	64	76	4096	5776	4864
Siswa 15	73	94	5329	8836	6862
Siswa 16	61	88	3721	7744	5368
Siswa 17	73	87	5329	7569	6351
Siswa 18	62	89	3844	7921	5518
Siswa 19	69	89	4761	7921	6141
Siswa 20	75	87	5625	7569	6525
Siswa 21	74	84	5476	7056	6216
Siswa 22	74	92	5476	8464	6808

Siswa 23	61	87	3721	7569	5307
Siswa 24	76	86	5776	7396	6536
Siswa 25	69	79	4761	6241	5451
Σ	1776	2115	127046	179541	150385

Memasukkan skor ΣX , ΣX^2 , ΣY , ΣY^2 , dan ΣXY ke dalam rumus korelasi *Product Moment*.

$$r_{XY} = \frac{25.150385 - (1776)(2115)}{\sqrt{(25.127046 - 3154176)(25.179541 - 4473225)}}$$

$$r_{XY} = \frac{3385}{18335.81741}$$

$$r_{XY} = 0,185$$

Dengan demikian nilai koefisien korelasi (r) adalah sebesar 0,18. Sekarang mari bandingkan nilai r tersebut dengan nilai r_{tabel} untuk menguji hipotesis yang dirumuskan.

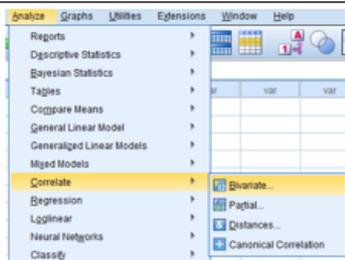
1. Tentukan apakah menggunakan taraf signifikansi 5% (taraf kepercayaan 95%) atau 1%. Dalam analisis ini kita akan gunakan 5%.
2. Silakan baca tabel r , tentukan angka yang berada di ukuran sampel (dalam analisis ini 25) dengan taraf signifikansi (5%). Contoh pada tabel r yang ditampilkan di bawah ini didapatkan nilai 0,396.
3. Bila $r_{xy} \geq r_{\text{tabel}}$ maka H_0 ditolak (dan H_a diterima), begitu juga jika $r_{xy} < r_{\text{tabel}}$ maka H_a ditolak (dan H_0 diterima). Pada analisis ini, statusnya adalah $r_{xy} < r_{\text{tabel}}$ ($0,185 < 0,396$), dengan demikian H_0 ditolak (dan H_a diterima).
4. Dapat disimpulkan, tidak terdapat hubungan yang signifikan antara rata-rata nilai ulangan dengan nilai UAS biologi siswa.

Tabel 10.5: Tabel r

N	The Level of Significance		N	The Level of Significance	
	5%	1%		5%	1%
3	0.997	0.999	38	0.320	0.413
4	0.950	0.990	39	0.316	0.408
5	0.878	0.959	40	0.312	0.403
6	0.811	0.917	41	0.308	0.398
7	0.754	0.874	42	0.304	0.393
8	0.707	0.834	43	0.301	0.389
9	0.666	0.798	44	0.297	0.384
10	0.632	0.765	45	0.294	0.380
11	0.602	0.735	46	0.291	0.376
12	0.576	0.708	47	0.288	0.372
13	0.553	0.684	48	0.284	0.368
14	0.532	0.661	49	0.281	0.364
15	0.514	0.641	50	0.279	0.361
16	0.497	0.623	55	0.266	0.345
17	0.482	0.606	60	0.254	0.330
18	0.468	0.590	65	0.244	0.317
19	0.456	0.575	70	0.235	0.306
20	0.444	0.561	75	0.227	0.296
21	0.433	0.549	80	0.220	0.286
22	0.432	0.537	85	0.213	0.278
23	0.413	0.526	90	0.207	0.267
24	0.404	0.515	95	0.202	0.263
25	0.396	0.505	100	0.195	0.256
26	0.388	0.496	125	0.176	0.230

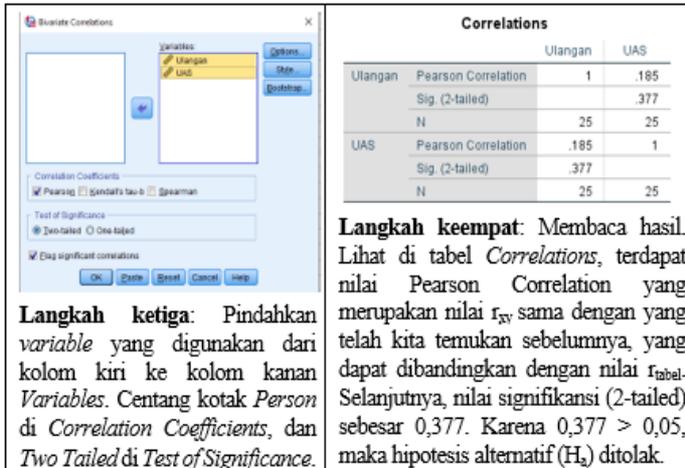
Penyelesaian menggunakan SPSS

	Ulangan	UAS	var
1	77	81	
2	67	89	
3	73	79	
4	71	81	
5	69	83	
6	68	75	
7	72	81	
8	79	86	
9	62	80	
10	80	89	
11	80	91	
12	80	81	
13	67	81	
14	64	76	
15	73	94	
16	61	88	
17	73	87	
18	62	89	
19	69	89	
20	75	87	
21	74	84	
22	74	92	
23	61	87	
24	76	86	
25	69	79	



Langkah pertama: Masukkan semua data pada sheet data secara vertikal. Pada kolom pertama adalah nilai ulangan dan kolom kedua adalah nilai UAS. Tipe data *numeric* dan bentuknya skala.

Langkah kedua: Pilih menu pada bagian atas `analyze > Correlate > Bivariate`.



Langkah ketiga: Pindahkan *variable* yang digunakan dari kolom kiri ke kolom kanan *Variables*. Centang kotak *Person* di *Correlation Coefficients*, dan *Two Tailed* di *Test of Significance*.

Langkah keempat: Membaca hasil. Lihat di tabel *Correlations*, terdapat nilai *Pearson Correlation* yang merupakan nilai r_{xy} sama dengan yang telah kita temukan sebelumnya, yang dapat dibandingkan dengan nilai r_{tabel} . Selanjutnya, nilai signifikansi (*2-tailed*) sebesar 0,377. Karena $0,377 > 0,05$, maka hipotesis alternatif (H_2) ditolak.

		Ulangan	UAS
Ulangan	Pearson Correlation	1	.185
	Sig. (2-tailed)		.377
	N	25	25
UAS	Pearson Correlation	.185	1
	Sig. (2-tailed)	.377	
	N	25	25

Korelasi Rank Spearman

Korelasi Rank Spearman digunakan untuk mengukur hubungan antara dua variabel ordinal atau interval yang tidak berdistribusi normal. Metode ini dinamakan setelah penemu pertama metode ini, yakni Charles Spearman. Korelasi Rank Spearman menghitung tingkat kekuatan dan arah hubungan antara kedua variabel, tetapi tidak mengasumsikan bahwa hubungan tersebut harus bersifat linear.

Korelasi Spearman bekerja dengan cara mengubah data menjadi rangking atau peringkat berdasarkan urutan dari yang terkecil hingga terbesar. Setelah peringkat ini dihitung, korelasi dihitung dengan menggunakan rumus yang mirip dengan korelasi Pearson. Namun, perbedaannya adalah data dihitung berdasarkan peringkat.

Keuntungan dari Korelasi Spearman adalah bahwa metode ini tidak sensitif terhadap data yang terlalu jauh dari distribusi normal, sehingga cocok digunakan pada data yang tidak memenuhi asumsi normalitas. Metode ini juga cocok digunakan ketika hubungan antara variabel tidak linear, karena tidak ada asumsi linearitas dalam korelasi Spearman.

Secara umum, Korelasi Spearman menghasilkan koefisien korelasi yang berkisar antara -1 dan 1, di mana nilai positif menunjukkan hubungan yang searah, dan nilai negatif menunjukkan hubungan yang berlawanan arah antara variabel. Koefisien 0 menunjukkan tidak adanya korelasi.

$$\rho = 1 - \frac{6D^2}{N(N^2 - 1)}$$

Deskripsi:

ρ : Koefisien Korelasi Spearman

D : *Difference* (perbedaan peringkat antara kedua kelompok pasangan)

N : Jumlah Kasus

1 dan 6 : Bilangan Konstan

Contoh:

Sebuah penelitian tentang versi/kualitas smartphone siswa terhadap tingkat motivasi belajar siswa di sekolah menengah atas. Sampel penelitian melibatkan 10 siswa. Peneliti mengukur versi/kualitas smartphone siswa dan tingkat motivasi belajar siswa ke dalam kategori sebagai berikut.

Tabel 10.6: Kategori Versi/Kualitas Smartphone dan Motivasi Belajar

Versi/Kualitas Smartphone Siswa		Tingkat Motivasi Belajar Siswa	
Kategori	Nilai	Kategori	Nilai
Sangat Baru (SB)	5	Sangat Tinggi (ST)	5
Baru (B)	4	Tinggi (T)	4
Cukup Baru (CB)	3	Cukup Tinggi (CT)	3
Tidak Baru (TB)	2	Rendah (R)	2
Sangat Tidak Baru (STB)	1	Sangat Rendah (SR)	1

Data yang didapatkan ditampilkan pada Tabel 10.7

Tabel 10.7: Data Versi/Kualitas Smartphone dan Motivasi Belajar Siswa

Siswa		1	2	3	4	5	6	7	8	9	10
Versi/ Kualitas Smartphone	Kategori	TB	TB	CB	B	SB	B	B	B	STB	STB
	Nilai	2	2	3	4	5	4	4	4	1	1
Motivasi Belajar	Kategori	R	R	R	CT	CT	R	T	T	T	CT
	Nilai	2	2	2	3	3	2	4	4	4	3

Hipotesis:

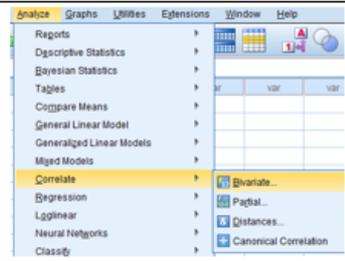
H_1 : Terdapat hubungan yang signifikan antara versi/kualitas smartphone siswa dengan tingkat motivasi belajar siswa.

H_0 : Tidak terdapat hubungan yang signifikan antara versi/kualitas smartphone siswa dengan tingkat motivasi belajar siswa.

Penyelesaian menggunakan SPSS

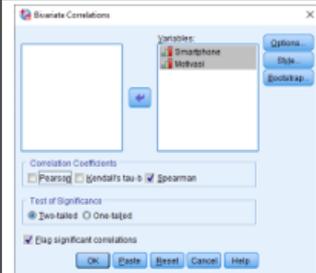
	Smartphone	Motivasi	vs
1	2	2	
2	2	2	
3	3	2	
4	4	3	
5	5	3	
6	4	2	
7	4	4	
8	4	4	
9	1	4	
10	1	3	
11			

Langkah pertama: Masukkan semua data pada sheet data secara vertikal. Pada kolom pertama adalah nilai dari versi/kualitas smartphone dan kolom kedua adalah nilai dari motivasi belajar. Tipe data *numeric* dan bentuknya ordinal.



Langkah kedua: Pilih menu pada bagian atas analyze > Correlate > *Bivariate*.

Langkah ketiga: Pindahkan *variable* yang digunakan dari kolom kiri ke kolom kanan *Variables*. Centang kotak



Langkah keempat: Membaca hasil. Lihat di tabel *Correlations*, terdapat nilai Spearman's rho. Selanjutnya, nilai signifikansi (2-tailed) sebesar 0,748. Karena $0,748 > 0,05$, maka hipotesis alternatif (H_1) ditolak.

Correlations				
Spearman's rho	Smartphone	Smartphone	Motivasi	
	Correlation Coefficient	1,000	,117	
	Sig. (2-tailed)		,748	
	N	10	10	
	Motivasi	Correlation Coefficient	,117	1,000
		Sig. (2-tailed)		,748
		N	10	10

10.2 Analisis Regresi

Pengertian

Analisis regresi adalah metode statistik yang digunakan untuk memahami hubungan fungsional antara satu variabel (variabel dependen) dan satu atau lebih variabel lainnya (variabel independen). Tujuan dari analisis regresi adalah untuk memprediksi nilai variabel dependen berdasarkan nilai variabel independen. Regresi dapat membantu mengidentifikasi pengaruh dan kontribusi relatif dari setiap variabel independen terhadap variabel dependen. Hasil analisis regresi sering dinyatakan dalam bentuk persamaan regresi yang menggambarkan hubungan matematis antara variabel-variabel tersebut.

Meskipun analisis regresi dan analisis korelasi sama-sama ingin memahami hubungan antar variabel, tetapi keduanya merupakan analisis yang berbeda. Setiap regresi memiliki korelasinya, tetapi tidak semua korelasi diikuti oleh regresi. Korelasi yang tidak diikuti oleh regresi terjadi ketika dua variabel tidak memiliki hubungan sebab akibat atau fungsional. Untuk menentukan apakah hubungan antara dua variabel bersifat kausal atau tidak, diperlukan dasar teori atau konsep tentang variabel tersebut.

Contohnya, hubungan antara motivasi belajar dan hasil belajar bisa dianggap kausal, sedangkan hubungan antara jumlah gaji guru dengan hasil belajar siswa bisa dianggap fungsional. Namun, hubungan antara berapa kali siswa fotokopi hasil ulangannya dengan nilai hasil belajar siswa tidak bersifat kausal atau fungsional, sehingga tidak ada landasan untuk dianalisis regresi.

Analisis regresi digunakan untuk mengukur pengaruh variabel bebas terhadap variabel tak bebas. Analisis ini hanya sesuai untuk data kuantitatif saja. Dari sisi tujuan, analisis regresi memahami hubungan fungsional antara variabel dependen dan variabel independen serta memprediksi nilai variabel dependen berdasarkan nilai variabel independen, sedangkan analisis korelasi adalah untuk mengukur sejauh mana hubungan linier antara dua variabel tanpa memprediksi nilai variabel lain. Dengan demikian, pada analisis regresi, ada variabel dependen dan independen, sedangkan pada analisis korelasi tidak ada.

Analisis regresi adalah alat statistik parametrik yang digunakan untuk memahami hubungan antara variabel independen dan variabel dependen. Sebelum menggunakan analisis regresi, peneliti perlu menguji sejumlah asumsi yang relevan untuk memastikan hasil analisis yang valid.

Asumsi-asumsi yang diuji meliputi:

1. Normalitas distribusi data: data mengikuti distribusi normal.
2. Linieritas: jika menggunakan regresi linear, hubungan antara variabel harus bersifat linier.
3. Ketidak-heteroskedastisitas-an (homoskedastisitas): varians residual antar data harus seragam.
4. Tidak ada multikolinearitas: korelasi tinggi antara variabel independen harus dihindari dalam regresi berganda.
5. Tidak ada autokorelasi: korelasi antara kesalahan pada periode t dengan kesalahan periode sebelumnya harus minim.

Pengujian asumsi ini penting karena hasil regresi dapat menjadi tidak valid jika salah satu asumsi tidak terpenuhi. Peneliti harus memahami implikasi dari setiap asumsi yang tidak terpenuhi dan mungkin perlu mengambil tindakan korektif sebelum melanjutkan analisis. Uji statistik dan teknik lainnya digunakan untuk menguji asumsi-asumsi tersebut, seperti uji normalitas, *scatter plot*, VIF untuk multikolinearitas, dan uji Durbin-Watson untuk autokorelasi. Kualitas dan validitas hasil analisis regresi sangat bergantung pada pemenuhan asumsi-asumsi ini.

Analisis Regresi Linier Sederhana

Analisis regresi linear sederhana, juga dikenal sebagai *simple linear regression*, digunakan untuk mengukur sejauh mana pengaruh satu variabel bebas (variabel independen, predictor, atau X) terhadap satu variabel tergantung (variabel dependen atau Y). Rumus umum regresi linier sederhana adalah:

$$\hat{Y} = a + bX$$

Deskripsi:

\hat{Y} : Variabel Terikat (*Dependent Variable*).

a: Konstanta (titik potong Y)

b: Koefisien dari variabel X

X : Variabel Bebas (*Independent Variable*)

Contoh:

Sebuah penelitian tentang pengaruh motivasi belajar siswa terhadap hasil belajar siswa di mata pelajaran biologi di sekolah menengah atas. Sampel penelitian melibatkan 12 siswa. Data yang didapatkan ditampilkan pada Tabel 10.8

Tabel 10.8: Data Motivasi Belajar Siswa Terhadap Hasil Belajar Siswa

Siswa	1	2	3	4	5	6	7	8	9	10	11	12
Motivasi Belajar (X)	8	7	8	7	6	6	8	8	8	9	8	7
Hasil Belajar (Y)	90	80	65	60	50	40	85	80	85	90	80	80

Hipotesis:

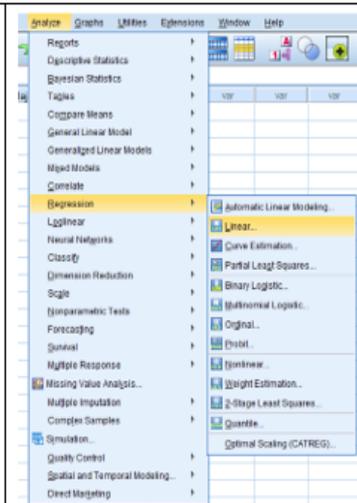
H_0 : Terdapat pengaruh motivasi belajar siswa (X) terhadap motivasi belajar (Y) pada mata pelajaran biologi.

H_a : Tidak terdapat pengaruh motivasi belajar siswa (X) terhadap motivasi belajar (Y) pada mata pelajaran biologi.

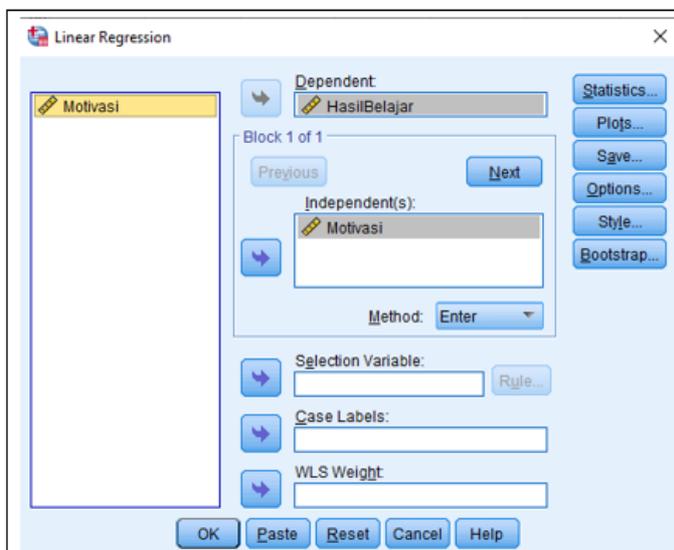
Penyelesaian menggunakan SPSS

	Motivasi	HasilBelajar
1	8	90
2	7	80
3	8	65
4	7	60
5	6	50
6	6	40
7	8	85
8	8	80
9	8	85
10	9	90
11	8	80
12	7	80

Langkah pertama: Masukkan semua data pada sheet data secara vertikal. Pada kolom pertama adalah nilai dari motivasi belajar (X) dan kolom kedua adalah nilai dari hasil belajar (Y). Tipe data *numeric* dan bentuknya *scale*.



Langkah kedua: Pilih menu pada bagian atas `analyze > Regression > Linear`.



Langkah ketiga: Pindahkan *variable* yang digunakan dari kolom kiri ke kolom kanan. Variabel bebas (X), dalam hal ini motivasi belajar, pindahkan ke kolom *Independent*. Variabel terikat (Y), dalam hal ini Hasil Belajar, pindahkan ke kolom *Dependent*. Pilih method *Enter* kemudian klik *OK*.

Langkah keempat: Membaca hasil. Berikut Tabel yang akan muncul di hasil SPSS.

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.819 ^a	.671	.638	9.775

a. Predictors: (Constant), Motivasi

ANOVA^a

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	1950.694	1	1950.694	20.414	.001 ^b
	Residual	955.556	10	95.556		
	Total	2906.250	11			

a. Dependent Variable: HasilBelajar

b. Predictors: (Constant), Motivasi

Coefficients^a

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	-36.667	24.601		-1.490	.167
	Motivasi	14.722	3.258	.819	4.518	.001

a. Dependent Variable: HasilBelajar

Pertama, mari kita uji hipotesisnya. Dengan melihat nilai signifikansi di output tabel Coefficients. Nilai Sig untuk model motivasi terhadap Hasil Belajar adalah 0,001. $0,001 < 0,05$, dengan demikian dapat disimpulkan H_a diterima, motivasi belajar siswa berpengaruh terhadap hasil belajar di mata pelajaran biologi.

Kedua, mari susun persamaan regresinya. Sebagaimana yang telah disebutkan sebelumnya, persamaannya adalah $\hat{Y} = a + bX$. Silakan lihat kembali output tabel Coefficients, a adalah angka konstan dari unstandardized coefficients, yang dalam hal ini adalah sebesar -36,667. Angka ini adalah angka konstan yang berarti jika tidak ada motivasi belajar siswa (X), maka nilai hasil belajar siswa (Y) adalah sebesar -36,667. Selanjutnya, b di persamaan adalah angka koefisien regresi, dalam hal ini bernilai 14,722. Angka ini dapat dimaknai bahwa setiap penambahan 1% motivasi belajar siswa (X) maka nilai hasil belajar siswa (Y) akan meningkat sebesar 14,722. Dengan demikian, persamaan regresi pada kasus ini adalah $\hat{Y} = -36,667 + 14,722X$.

Ketiga, mari lihat besarnya pengaruh variabel X (Motivasi belajar siswa) terhadap variabel Y (Hasil belajar siswa). Silakan lihat pada output tabel Model Summary, terutama di kolom R Square, didapatkan nilai sebesar 0,671. Angka ini berarti, motivasi belajar siswa mempengaruhi hasil belajar siswa sebesar 67,1%. Sedangkan sisanya, yaitu 32,9% dipengaruhi oleh berbagai faktor lain. Hal ini dapat diinterpretasikan bahwa jika siswa memiliki motivasi belajar 100%, maka paling tidak dia sudah memiliki 67,1% hasil belajar yang positif.

Analisis Regresi Linier berganda

Analisis regresi linear berganda, juga dikenal sebagai *multiple linear regression*, digunakan untuk mengukur sejauh mana pengaruh dua variabel bebas (variabel X_1 dan X_2) atau lebih terhadap satu variabel terikat (variabel dependen atau Y). Rumus umum regresi linier sederhana adalah:

$$\hat{Y} = a + b_1X_1 + b_2X_2 \dots b_nX_n$$

Deskripsi:

\hat{Y} : Variabel Terikat (*Dependent Variable*).

a: Konstanta (titik potong Y)

b: Koefisien dari variabel X(n)

X: Variabel Bebas (*Independent Variable*)

Contoh:

Sebuah penelitian tentang pengaruh motivasi belajar (X1) dan tingkat stress (X2) terhadap hasil belajar (Y) siswa di mata pelajaran biologi di sekolah menengah atas. Sampel penelitian melibatkan 12 siswa. Data yang didapatkan ditampilkan pada Tabel 10.9

Tabel 10.9: Data Motivasi Belajar Siswa Terhadap Hasil Belajar Siswa

Siswa	1	2	3	4	5	6	7	8	9	10	11	12
Motivasi Belajar (X1)	8	7	8	7	6	6	8	8	8	9	8	7
Tingkat Stres (X2)	5	5	5	5	4	4	3	3	3	3	3	5
Hasil Belajar (Y)	90	80	65	60	50	40	85	80	85	90	80	80

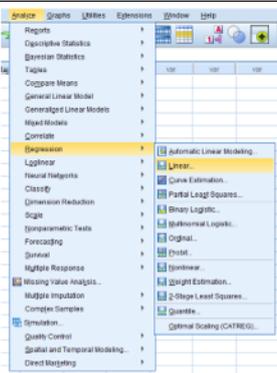
Hipotesis yang ingin diuji:

H_a : Terdapat pengaruh motivasi belajar siswa (X1) dan tingkat stress (X2) secara bersama-sama terhadap motivasi belajar (Y) pada mata pelajaran biologi.

H_0 : Tidak terdapat pengaruh motivasi belajar siswa (X1) dan tingkat stress (X2) secara bersama-sama terhadap motivasi belajar (Y) pada mata pelajaran biologi.

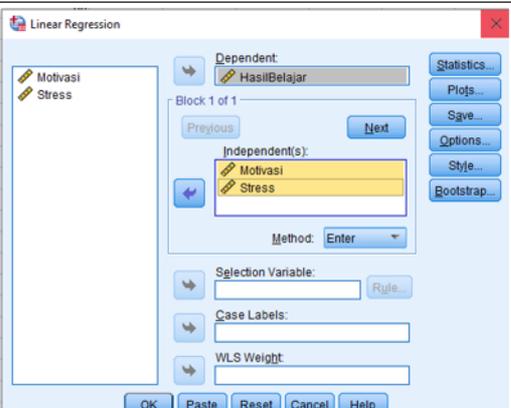
Penyelesaian menggunakan SPSS

	Motivasi	Stress	HasilBelajar
1	8	5	90
2	7	5	80
3	8	5	65
4	7	5	60
5	6	4	50
6	6	4	40
7	8	3	85
8	8	3	80
9	8	3	85
10	9	3	90
11	8	3	80
12	7	5	80



Langkah pertama: Masukkan semua data pada sheet data secara vertikal. Pada kolom pertama adalah nilai dari motivasi belajar (X1), pada kolom kedua adalah nilai dari tingkat stress siswa (X2), dan kolom ketiga adalah nilai dari hasil belajar (Y). Tipe data *numeric* dan bentuknya *scale*.

Langkah kedua: Pilih menu pada bagian atas analyze > Regression > Linear.



Langkah ketiga: Pindahkan *variable* yang digunakan dari kolom kiri ke kolom kanan. Variabel bebas (X1 dan X2), dalam hal ini motivasi belajar dan tingkat stress, pindahkan ke kolom *Independent*. Variabel terikat (Y), dalam hal ini Hasil Belajar, pindahkan ke kolom *Dependent*. Pilih method Enter kemudian klik OK.

Langkah keempat: Membaca hasil. Berikut Tabel yang akan muncul di hasil SPSS.

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.824 ^a	.679	.608	10.177

a. Predictors: (Constant), Stress, Motivasi

ANOVA^a

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	1974.155	2	987.078	9.531	.006 ^b
	Residual	932.095	9	103.566		
	Total	2906.250	11			

a. Dependent Variable: HasilBelajar

b. Predictors: (Constant), Stress, Motivasi

Coefficients^a

Model		Unstandardized Coefficients		Standardized Coefficients		Sig.
		B	Std. Error	Beta	t	
1	(Constant)	-49.054	36.514		-1.343	.212
	Motivasi	15.473	3.741	.861	4.136	.003
	Stress	1.689	3.549	.099	.476	.645

a. Dependent Variable: HasilBelajar

Pertama, mari kita uji hipotesisnya. Dengan melihat nilai signifikansi di output tabel *Coefficients*. Nilai Sig untuk model motivasi terhadap hasil belajar adalah 0,003 dan nilai Sig untuk model stres terhadap hasil belajar adalah 0,645. Untuk $0,003 < 0,05$ maka dapat disimpulkan motivasi belajar berpengaruh terhadap hasil belajar, sedangkan $0,645 > 0,05$ maka dapat disimpulkan tingkat stres tidak berpengaruh terhadap hasil belajar siswa.

Tetapi karena secara terpisah kita tidak hipotesiskan, yang kita hipotesiskan hanya secara bersama-sama, maka kita bisa menggunakan nilai Signifikansi di tabel ANOVA. Nilai signifikansi di tabel Anova adalah sebesar 0,006. Dengan demikian, H_0 diterima, motivasi belajar siswa dan tingkat stres siswa secara bersama-sama berpengaruh terhadap hasil belajar di mata pelajaran biologi.

Kedua, mari susun persamaan regresinya. Sebagaimana yang telah disebutkan sebelumnya, persamaannya adalah $\hat{Y} = a + b_1X_1 + b_2X_2$. Silahkan lihat kembali output tabel *Coefficients*, a adalah angka konstan dari *unstandardized coefficients*, yang dalam hal ini adalah sebesar -49,054. Angka ini adalah angka konstan yang berarti jika tidak ada motivasi belajar siswa (X_1) dan tingkat stress (X_2), maka nilai hasil belajar siswa (Y) adalah sebesar -49,054.

Selanjutnya, b di persamaan adalah angka koefisien regresi, pada variabel motivasi bernilai 15,473 dan variabel tingkat stres adalah 1,689. Angka ini dapat dimaknai bahwa setiap penambahan 1% motivasi belajar siswa (X_1) maka nilai hasil belajar siswa (Y) akan meningkat sebesar 15,473 dan setiap penambahan

1% tingkat stress (X_2) maka nilai hasil belajar siswa (Y) akan meningkat sebesar 1,689. Dengan demikian, persamaan regresi pada kasus ini adalah $\hat{Y} = -49,054 + 15,473X_1 + 1,689X_2$.

Ketiga, mari lihat besarnya pengaruh variabel X_1 (Motivasi belajar siswa) dan X_2 (tingkat stres siswa) terhadap variabel Y (Hasil belajar siswa). Silakan lihat pada output tabel *Model Summary*, terutama di kolom *R Square*, didapatkan nilai sebesar 0,679. Angka ini berarti, motivasi belajar dan tingkat stres secara bersama-sama mempengaruhi hasil belajar siswa pada mata pelajaran biologi sebesar 67,9%. Sedangkan sisanya, yaitu 32,1% dipengaruhi oleh berbagai faktor lain.

Daftar Pustaka

- Abigail Soesana, Hani Subakti, D. (2023) *Metodologi Penelitian Kuantitatif*. 1st edn. Yayasan Kita Menulis.
- Amir, F. (2014) 'Peranan Matematika Dan Statistika Dalam Pertanian Industrial Untuk Mewujudkan Ketahanan Pangan Nasional', *Prosiding Seminar Nasional Matematika*, 1(November), pp. 356–364.
- Anwar, A. (2009). *Statistika untuk penelitian pendidikan dan aplikasinya dengan SPSS dan excel*. Kediri: IAIT Press.
- Anwar, K. & Marzuki, I., (2020). *Statistik Terapan*. Yogyakarta: CV. Tangan Emas.
- Aprilina, K., Nuraini, T.A., dan Sopaheluwakan, A. (2018). "Kajian Awal uji Statistik Perbandingan Suhu Udara dari Peralatan Otomatis dan Manual," *Jurnal Meteorologi dan Geofisika*, 18(1).
- Balanda, K. P. and Macgillivray, H. L. (1988) 'Kurtosis: A critical review', *American Statistician*, 42(2). doi: 10.1080/00031305.1988.10475539.
- Birhan, W. et al. (2021) 'Exploring the context of teaching character education to children in preprimary and primary schools', *Social Sciences & Humanities Open*, 4(1), p. 100171. Available at: <https://doi.org/10.1016/j.ssaho.2021.100171>.
- Bowman, A. W., D'Agostino, R. B. and Stephens, M. A. (1988) 'Goodness-of-Fit Techniques.', *Journal of the Royal Statistical Society. Series A (Statistics in Society)*, 151(1). doi: 10.2307/2982198.
- Bulmer, M. (1979) 'Concepts in the analysis of qualitative data', *The Sociological Review*, 27(4). doi: 10.1111/j.1467-954X.1979.tb00354.x.

- Cohen, B.H. (2013) *Explaining Psychological Statistics*. 4th edn. New Jersey: Wiley.
- Coolican, H. (2019) *Research Methods and Statistics in Psychology*. 7th edn, *Research Methods and Statistics in Psychology*. 7th edn. London: Routledge.
- Cramer, D. (2002) *Basic Statistics for Social Research, Basic Statistics for Social Research*. doi: 10.4324/9780203430897.
- Furqon (2004) *Statistika Terapan untuk Penelitian*. Bandung: Alfabeta.
- Hamdi, A. S. (2004). "Metode Penelitian Kuantitatif Aplikasi dalam Pendidikan," Yogyakarta: CV. Budi Utama.
- Hanief, Y. N., & Himawanto, W. (2017). *Statistik pendidikan*. Yogyakarta: Deepublish.
- Hasan, M. Iqbal. (2011). *Pokok- Pokok Materi Statistika 1 (Statistik Deskriptif)*. Jakarta: PT Bumi Aksara.
- J.Supranto (1993) *Statistik dan Aplikasi Ekonomi dan Bisnis*.
- Jaya, I., (2010). *Statistika penelitian untuk pendidikan*. Bandung: Citapustaka media perintis.
- Joanes, D. N. and Gill, C. A. (1998) 'Comparing measures of sample skewness and kurtosis', *Journal of the Royal Statistical Society Series D: The Statistician*, 47(1). doi: 10.1111/1467-9884.00122.
- Johnson, R. A. (2010). "Statistics: Principles and Methods (6rd ed)," Hoboken, NJ: John Wiley & Sons.
- Kerlinger (2006) *Asas–Asas Penelitian Behaviour*. 3rd edn. Yogyakarta: Gadjah Mada University Press.
- Kurniadi, G. and Purwaningrum, J.P. (2018) 'Kemampuan Pemahaman Matematis Siswa melalui Discovery Learning Berbantuan Asesmen Hands on Activities', *ANARGYA: Jurnal Ilmiah Pendidikan Matematika*, 1(1), pp. 8–13. Available at: <https://doi.org/10.24176/anargya.v1i1.2463>.
- Malik, A. and Chusni, M. (2018) *Pengantar Statistika Pendidikan: Teori dan Aplikasi*, Deepublish. Yogyakarta: CV Budi Utama.

- Mundir (2012) Statistik Pendidikan: Pengantar Analisis Data untuk Penulisan Skripsi dan Tesis. Jember: STAIN Jember Press.
- Munfarikhatin, A. (2019) 'Musamus Jurnal of Mathematics Education', 2.
- Munfarikhatin, A. et al. (2021) 'Quizizz and Hard work Character in Geometry Online Lecture : How It Influence ?', 5(1), pp. 33–40.
- Murti; Bhisma (1996) Penerapan Metode Statistik Non Parametrik Dalam Ilmu-ilmu Kesehatan. Jakarta: PT.Gramedia Pustaka Utama.
- Murwani, S. (2001). "Statistika Terapan (Teknik Analisis Data)," Malang: Universitas Negeri Malang.
- Nazir (1988) Metode Penelitian. Jakarta: Ghalia Indonesia.
- Nursalam, (2015). Statistika Pendidikan. Makassar: Alauddin University Press
- Nuryadi (2017) Dasar-Dasar Statistik Penelitian. Yogyakarta: Sibuku media.
- Nuryadi, dkk. (2017). Dasar-Dasar Statistik Penelitian. Yogyakarta : Gramasurya.
- Priadana, Sidik., Sunarsi, D. (2021) Metode Penelitian Kuantitatif. 1st edn. Tangerang: Pascal Books.
- Purwanto. (2011). "Statistik untuk Penelitian," Yogyakarta: Pustaka Pelajar.
- Quraisy, A. (2020). "Normalitas Data Menggunakan Uji Kolmogorov-Smirnov dan Saphiro-Wilk," J-HEST: Journal of Healt, Education, Economics, Science, and Technology, 3(1), hal.7-11.
- Razali, N.M dan Wah, Y.B. (2011). "Power Comparisons Saphiro Wilk, Kolmogorov-Smirnov, Lilliefors and Anderson Darling Test," Jurnal of Statistical Modeling an Analytics, 2(1), hal.21-33.
- Real-Statistic. (2019). "Kolmogorov-Smirnov Test For Normality," Diakses 2023 <http://www.real-statistics.com/testsnormality-and-symmetry/statistical-testsnormality-symmetry/kolmogorov-smirnovtest/>.
- Riduwan, S.H. (2009) Pengantar Statistika untuk Penelitian: Pendidikan, Sosial, Komunikasi, Ekonomi, dan Bisnis, Buku. Bandung: Alfabeta.
- Riduwan, S.H. (2009) Pengantar Statistika untuk Penelitian: Pendidikan, Sosial, Komunikasi, Ekonomi, dan Bisnis, Buku. Bandung: Alfabeta.
- Riduwan. (2008). "Dasar-Dasar Statistika," Bandung: Alfabeta.

- Ruslau, M. F. V. & Suryani, D. R., (2017). *Statistika Dasar*. Surabaya: Mavendra Press.
- Sabri, L.H. (2008) *Statistik Kesehatan*. Edisi Revisi. Jakarta: Rajawali Pers.
- Santoso, S. (2002). “Buku Latihan SPSS Statistik Parametrik,” Jakarta: PT. Elex Media Komputindo.
- Sasmoko (2004) *Metode Penelitian, Pengukuran dan analisa Data*. Jakarta: Press Uki.
- Siegel, S. (1994). “Statistika Nonparametrik Untuk Ilmu-Ilmu Sosial,” Jakarta: PT.Garamedia.
- Siregar, Syofian. (2015). “Statistika Terapan untuk Perguruan Tinggi,” Jakarta: Prenadamedia Group.
- Soekadijo, R.G. (1983). *Logika Dasar: tradisional, simbolik, dan induktif*. Jakarta: Gramedia Pustaka Utama.
- Somantri, Ating dan Sambas Ali Muhidin. (2006). *Aplikasi statistika dalam Penelitian*. Bandung : Pustaka Ceria.
- Spiegel, M.. R. (1961) *Theory and Problem of Statistics*. New York: Mc Graw-Hill.
- Spiegel, M.R., Schiller, J. and Srinivasan, R.A. (2013) *Probability and Statistics*. 4th edn. New York: McGraw-Hill.
- Sudijono, A. (2015). “Pengantar Statistik Pendidikan,” Jakarta: PT. Raja Grafindo Persada.
- Sudijono, Anas. (2009). *Pengantar Statistik Pendidikan*. Jakarta: PT RajaGrafindo Persada.
- Sudjana. (2005). “Metode Statistika,” Bandung: Tarsito.
- Sugiono (1988) *Metode Penelitian Kuantitatif dan Kualitatif*. Bandung: Alfa Beta.
- Sugiyono (2009) *Metode Penelitian Pendidikan Pendekatan Kuantitatif, Kualitatif, dan R&D*. Bandung: Alfabeta.
- Sugiyono, (2004). “Metode Penelitian Bisnis,” Bandung: Alfabeta.
- Sugiyono, P.D. (2010) *Metod. Penelit. Kuantitatif, Kualitatif, dan R&D*. Bandung: Alfabeta.

- Sukardi (2018) *Metodologi Penelitian Pendidikan; Kompetensi dan Praktiknya*. REVISI. Jakarta: Bumi Aksara.
- Sukestiyarno, Y.L. (2013) *Olah Data Penelitian Berbantuan SPSS*. Semarang: Universitas Negeri Semarang.
- Sunaryo, S. et al. (2003) 'Sejarah Perkembangan Statistika Dan Aplikasinya', *Forum Statistika Dan Komputasi*, 8(1), pp. 1–7. Available at: <https://journal.ipb.ac.id/index.php/statistika/article/view/5506>.
- Supriadi, G. (2021). *Statistik Penelitian Pendidikan*. Yogyakarta: UNY Press.
- Utami, F. (2020). "Pengaruh Metode Pembelajaran Outing Class terhadap Kecerdasan Naturalis Anak Usia 5-6 Tahun," *Jurnal Obsesi: Jurnal Pendidikan Anak Usia Dini Sekolah Tinggi Theologia Jaffray*.
- Walpole, R.E. et al. (2016) *Probability & Statistics for Engineers & Scientists*. 9th edn. Pearson.
- Wasserman, L. (2004). *All of statistics: a concise course in statistical inference* (Vol. 26, p. 86). New York: Springer.
- Westfall, P. H. (2014) 'Kurtosis as Peakedness, 1905–2014. R.I.P.', *American Statistician*, 68(3). doi: 10.1080/00031305.2014.917055.

Biodata Penulis



Lahir di Bojonegoro 19 April 1992. Memperoleh gelar Sarjana di IKIP PGRI Bojonegoro pada tahun 2014 dan gelar Magister di Universitas Sebelas Maret Surakarta pada tahun 2016. Sebelumnya mengawali guru mata pelajaran Matematika di SMA Negeri Model Terpadu Bojonegoro pada tahun 2016. Sebelumnya juga pernah menjadi tentor di Lembaga bimbingan belajar dan saat ini Aktif sebagai Dosen di Program Studi Pendidikan Matematika Universitas Nahdlatul ‘Ulama Sunan Giri Bojonegoro.



Rani Rahim, S.Pd., M.Pd, Lahir di Palembang pada tanggal 22 Oktober 1990, merupakan putri kedua dari pasangan Purn. Serma Hamzani dan Syafrida. Menyelesaikan pendidikan S1 di Universitas Muhammadiyah Sumatera Utara (UMSU) pada program studi Pendidikan Matematika dan lulus pada tahun 2012. Selanjutnya pada tahun 2013, penulis melanjutkan pendidikan ke jenjang Magister Pendidikan (S2) di Program Pascasarjana Universitas Negeri Medan (UNIMED) pada program studi Pendidikan Matematika dan lulus pada tahun 2015.

Saat ini penulis sedang melanjutkan studi S3 di Program Pascasarjana Universitas Negeri Medan (UNIMED). Penulis merupakan dosen di Universitas Dharmawangsa Medan, Sumatera Utara sejak tahun 2015 s/d sekarang. Selain sebagai dosen, penulis juga bertugas sebagai guru di SMK Negeri 5 Medan sejak tahun 2013 s/d sekarang. Penulis juga aktif melakukan penelitian dalam bidang pendidikan matematika, pengembangan perangkat pembelajaran, serta pendekatan kontekstual. Penulis pernah mendapatkan Hibah Penelitian Dosen Pemula (PDP) pada tahun 2019, 2021 dan 2022 sebagai ketua atau anggota. Selain aktif dalam bidang penelitian, penulis juga aktif dalam melakukan kegiatan pengabdian kepada masyarakat. Pada tahun 2020, penulis

mendapatkan Hibah Program Kemitraan Masyarakat Stimulus (PKMS) yang diperoleh dari Hibah Kemenristek/BRIN sebagai anggota. Pada tahun ini, penulis mendapatkan kesempatan kembali mendapatkan hibah Pengabdian kepada Masyarakat yang diperoleh dari Kementerian Pendidikan, Kebudayaan, Riset dan Teknologi sebagai anggota. Buku ini merupakan buku kolaborasi keenam bagi penulis di Yayasan Kita Menulis. Semoga buku ini banyak memberikan manfaat bagi para pembaca dan dapat menjadi inspirasi bagi penulis lainnya.

E-mail: ranirahim@dharmawangsa.ac.id

Phone/WA: 0852 9644 0383



Nurhayati. Telah menempuh pendidikan S1 pada Jurusan Teknik Informatika di Universitas Musamus dan pendidikan S2 pada jurusan Pengajaran Matematika di Institut Teknologi Bandung. Saat ini adalah dosen tetap Program Studi Pendidikan Komputer, Fakultas Keguruan dan Ilmu Pendidikan, Universitas Musamus. Mengampu mata kuliah statistika pendidikan, matematika dasar dan profesi keguruan. Telah menulis buku ajar dengan judul aljabar linear dan matematika dasar.

E-mail: nurhayati_fkip@unmus.ac.id



Azis lahir di Kadolo Katapi tanggal 29 September 1987. Penulis adalah dosen tetap pada Program Studi Pendidikan Matematika Fakultas Keguruan dan Ilmu Pendidikan, Universitas Dayanu Ikhsanuddin. Menyelesaikan pendidikan S1 pada Program Studi Pendidikan Matematika di Universitas Dayanu Ikhsanuddin dan menyelesaikan pendidikan S2 pada Program Studi Pendidikan Matematika di Universitas Negeri Yogyakarta.



Sadrack Luden Pagiling lahir di Tana Toraja, pada 15 Maret 1989. Ia tercatat sebagai lulusan Magister Pendidikan Matematika, Universitas Negeri Surabaya, tahun 2017. Saat ini, Ia bekerja sebagai dosen tetap di Jurusan Pendidikan Matematika, Fakultas Keguruan dan Ilmu Pendidikan, Universitas Musamus Merauke sejak tahun 2019. Berbagai karya ilmiah berupa artikel telah dipublikasikan penulis di Jurnal Nasional, prosiding seminar internasional, dan jurnal internasional. Selain itu, Penulis juga secara rutin mempresentasikan artikel ilmiah di Seminar

Internasional Pendidikan Matematika. Penulis juga sudah mempublikasikan 2 (dua) buku ajar yang berjudul Logika Elementer yang diterbitkan oleh Penerbit Deepublish dan Geometri Analitik Bidang dan Ruang dan Aplikasinya Menggunakan Geogebra yang diterbitkan oleh UNY Press dan 3 (Tiga) buku referensi yang berjudul Inovasi Pembelajaran Digital Abad 21, Teknologi dalam Pembelajaran, dan Riset Pendidikan penerbit Yayasan Kita Menulis.



Irmawaty Natsir, lahir di Ujung Pandang, Sulawesi Selatan, Indonesia pada tanggal 20 Juli 1988. Penulis menyelesaikan studi S-1 di Universitas Muhammadiyah Makassar (2011); S-2 Bidang Pendidikan Matematika Kekhususan Matematika Sekolah di Universitas Negeri Makassar (2014). Sejak tahun 2019 - sekarang penulis menjadi dosen di Jurusan Pendidikan Matematika, Universitas Musamus.



Anis Munfarikhatin lahir di Pati, pada 22 Oktober 1989. Beliau adalah dosen Pendidikan Matematika di Universitas Musamus Merauke sejak tahun 2019. Selain menjadi dosen juga aktif dalam menulis buku diantaranya buku ajar Geometri Analitik Bidang dan Ruang Menggunakan Geogebra dan buku referensi hasil penelitian "Potret Literasi Numerasi di Ujung Timur Indonesia".



Daniel Simanjuntak, lahir di Balikpapan pada tanggal 03 Mei 1992. Anak pertama dari tiga bersaudara dari pasangan bapak Manogang Simanjuntak dan ibu Kartini br Hutagaol. Daniel Simanjuntak menempuh pendidikan formal di SD Advet Lab School UNAI, Kecamatan Parongpong, Bandung Barat, selesai pada tahun 2004, pada tahun yang sama melanjutkan pendidikan di SMP Advet Lab School UNAI, Kecamatan Parongpong, Bandung Barat selesai tahun 2007, kemudian pada tahun yang sama melanjutkan pendidikan Sekolah Menengah Atas di Perguruan Advent Bandung , selesai tahun 2010. Kemudian pada tahun 2014 menyelesaikan pendidikan S1 dengan Jurusan Akuntansi dari kampus Universitas Advent Bandung. Pada tahun 2017, menyelesaikan pendidikan program Master (S2) dengan jurusan Akuntansi dari Universitas Maranata Bandung.



Kartini Hutagaol, lahir di Balige pada tanggal 21 April 1966. Anak ke-empat dari pasangan bapak Israel Hutagaol dan ibu Keberlina br Tambunan. Kartini Hutagaol menempuh pendidikan formal di SDN Hutagaol Kecamatan Balige, selesai pada tahun 1979, pada tahun yang sama melanjutkan pendidikan di SMPN III Balige, selesai tahun 1982, kemudian pada tahun yang sama melanjutkan pendidikan Kampus Menengah Atas di SMAN I Balige, selesai tahun 1985. Kemudian pada tahun 2001 menyelesaikan pendidikan S1 dengan Jurusan Pendidikan Matematika di STKIP Siliwangi Bandung. Pada tahun 2007, menyelesaikan pendidikan program Master (S2) dengan jurusan Pendidikan Matematika dari UPI Bandung, dan menyelesaikan program Doktor (S3) dengan jurusan yang sama dan dari kampus yang sama di awal tahun 2010.



Dr. Nanda Eska Anugrah Nasution, M.Pd. adalah dosen pendidikan biologi dan biologi lingkungan di UIN Kiai Haji Achmad Siddiq Jember. Menyelesaikan pendidikan S1 Bilingual Pendidikan Biologi di Universitas Negeri Medan, S2 Pendidikan Biologi di Universitas Negeri Medan, dan S3 Pendidikan Biologi di Universitas Negeri Malang.

Mengampu mata kuliah metode penelitian pendidikan biologi, media pembelajaran biologi, pengembangan bahan ajar biologi, statistika pendidikan, biokonservasi, genetika, pengetahuan lingkungan, pendidikan lingkungan hidup, biologi sel dan molekular, dan beberapa mata kuliah lain. Selama ini terlibat aktif melaksanakan Tri Dharma Perguruan Tinggi dan telah mempublikasikan sejumlah artikel jurnal dan buku referensi.

E-mail: eska@uinkhas.ac.id

DASAR-DASAR STATISTIKA UNTUK PENELITIAN

Buku Dasar-Dasar Statistika Untuk Penelitian membahas topik-topik penting yang perlu diketahui oleh periset sebagai dasar ketika melakukan penelitian. Semoga dengan dibuatkan buku ini para pendidik, akademisi, periset, serta setiap pihak yang melakukan studi dan riset dapat memahami tentang analisis-analisis dalam Statistika yang lebih berorientasi pada Penelitian.

Buku ini berisi penjabaran mengenai Dasar-Dasar Statistika Untuk Penelitian yang terdiri dari 11 (sebelas) bab, yaitu:

Bab 1 Statistik dan Statistika

Bab 2 Penyajian Data Penelitian

Bab 3 Distribusi Frekuensi

Bab 4 Kemencengan dan Kurtosis

Bab 5 Statistik Inferensi

Bab 6 Uji Normalitas dan Uji Homogenitas Data

Bab 7 Uji T-Test

Bab 8 Uji Chi Square

Bab 9 Analisis Varians

Bab 10 Analisis Korelasi dan Regresi



YAYASAN KITA MENULIS
press@kitamenulis.id
www.kitamenulis.id

