

Statistik untuk Psikologi dan Ilmu Sosial

Buku Panduan Praktis untuk Mahasiswa dan Peneliti Muda

Sunu Bagaskara Entin Nurhayati Sari Zakiah Akmal
Titi Sahidah Fitriana Aswin Januarsjaf

30 Agustus 2025

Daftar Isi

Prakata	1
Tentang Penulis	2
I PENGANTAR STATISTIK	4
1 Pengertian & Sejarah Ringkas Statistik	6
Linimasa Sejarah Statistik	6
1.1 Definisi Statistik	7
1.2 Sampel dan Populasi	7
1.3 Parameter dan Statistik	7
2 Statistik dalam Penelitian Psikologi	8
3 Jenis-jenis Statistik: Deskriptif & Inferensial	9
Soal Latihan	10
II PERSIAPAN DATA	11
4 Jenis Data	13
4.1 Mengidentifikasi Jenis Data	13
4.2 Penamaan dan Pengkodean Data/Variabel	14
5 Pembersihan Data	16
5.1 Melakukan Analisis Deskriptif	16
5.2 Data Hilang	17
5.3 Data Pencilan	18
6 Akuntabilitas Data	20
Soal Latihan	21
III STATISTIK DESKRIPTIF	22
7 Ukuran Pemusatan Data	24
7.1 <i>Mean</i>	24
7.2 Median	24
7.3 Modus	25

8 Ukuran Penyebaran Data	26
8.1 <i>Range</i>	26
8.2 Varians	27
8.3 Standar Deviasi	28
9 Bentuk Distribusi Data	30
9.1 Kurva Normal	30
9.2 <i>Skewness</i>	32
9.3 Kurtosis	33
9.4 <i>Z-Score</i>	34
Soal Latihan	35
IV STATISTIK INFERENSIAL	36
10 Statistik Parametrik dan Non-Parametrik	38
10.1 Asumsi Parametrik	38
10.2 Metode Parametrik	39
10.3 Metode Non-Parametrik	39
11 <i>Statistical Power</i>	41
12 Pengujian Hipotesis	42
13 Signifikansi Statistik & <i>p-value</i>	44
13.1 Signifikansi	44
13.2 <i>p-value</i>	45
13.3 Interpretasi Signifikansi & <i>p-value</i>	45
14 <i>Confidence Interval</i>	47
14.1 Prinsip Dasar <i>Confidence Interval</i>	47
14.2 <i>Margin of Error</i>	47
14.3 <i>Level of Confidence</i>	48
14.4 Menginterpretasi <i>CI</i> , <i>MoE</i> , & <i>LoC</i>	48
15 <i>Effect Size</i>	50
15.1 Prinsip Dasar	50
15.2 Klasifikasi <i>Effect Size</i>	50
15.3 Ukuran <i>Effect Size</i> Lainnya	52
Soal Latihan	53
V PERBEDAAN RATA-RATA DUA KELOMPOK	54
16 Memahami Perbedaan Rata-rata Antar Kelompok	56
17 Perbedaan Rata-rata Antar Kelompok vs. Dalam Kelompok	57
18 <i>Independent Samples t-test</i>: Parametrik dan Non-Parametrik	58

19 Paired Samples t-test: Parametrik dan Non-Parametrik	61
Soal Latihan	64
VI MEMBANDINGKAN RATA-RATA TIGA ATAU LEBIH KELOMPOK	65
20 Tujuan Perbandingan Tiga Kelompok atau Lebih	67
21 Prinsip-prinsip Dasar dan Asumsi-asumsi ANOVA	68
21.1 Statistik F: Rasio Variabilitas	68
21.2 Asumsi-asumsi Dasar ANOVA	69
22 Jenis-jenis ANOVA dan Interpretasi Hasil Analisis	71
22.1 ANOVA Satu Arah	71
22.2 ANOVA Dua Arah	72
22.3 ANOVA Pengukuran Berulang	73
22.4 ANOVA Desain Campuran	74
22.5 Uji Post-hoc vs. Planned Contrasts	75
22.6 Effect Size dan Visualisasi Hasil Analisis	76
23 Alternatif Non-Parametrik untuk ANOVA	77
23.1 Uji Kruskal-Wallis: Alternatif Non-Parametrik untuk ANOVA Satu Arah	77
23.2 Uji Friedman: Alternatif Non-Parametrik untuk ANOVA Pengukuran Berulang	78
Soal Latihan	79
VII ANALISIS KORELASI	80
24 Karakteristik Hubungan Antara Dua Variabel	82
24.1 Arah Hubungan Korelasi	82
24.2 Bentuk Hubungan Korelasi	83
24.3 Besaran Hubungan Korelasi	83
25 Asumsi Korelasi, Interpretasi, dan Penulisan Hasil Analisis	85
25.1 Asumsi dalam Hubungan Antara Dua Variabel	85
25.2 Interpretasi dan Penulisan Hasil Uji Korelasi	85
Soal Latihan	88
VIII MODEL REGRESI LINEAR	89
26 Prinsip Dasar Model Regresi Linear	91
26.1 Regresi vs. Korelasi	91
26.2 Persamaan dan Garis Regresi Linear	91
27 Regresi Linear Sederhana	93
27.1 Asumsi-asumsi dalam Regresi Sederhana	93

27.2 Interpretasi Hasil Regresi Sederhana	93
27.3 Penulisan Laporan Hasil Regresi Sederhana	95
28 Regresi Linear Majemuk	96
28.1 Konsep Dasar	96
28.2 Asumsi-asumsi dalam Regresi Majemuk	96
28.3 Interpretasi Hasil Analisis Regresi Majemuk	97
28.4 Penulisan Laporan Hasil Regresi Majemuk	97
29 Regresi Linear Hierarkis	99
29.1 Konsep Dasar	99
29.2 Asumsi-asumsi dalam Regresi Hierarkis	99
29.3 Interpretasi Hasil Analisis Regresi Hierarkis: <i>Delta (Δ) R²</i>	100
29.4 Penulisan Laporan Hasil Regresi Hierarkis	101
Soal Latihan	102
Glosarium	103
Daftar Pustaka	107

Daftar Gambar

5.1 Outlier dalam boxplot	18
8.1 Visualisasi data dengan: a) varians kecil & b) varians besar	29
9.1 Bentuk kurva normal	30
9.2 Proporsi sebaran data pada kurva normal	31
9.3 Distribusi skewed	32
9.4 Skewness data berdasarkan lokasi tendensi sentral	33
9.5 Visualisasi kurtosis dalam distribusi data	34
10.1 Kurva distribusi data	39
21.1 Asumsi-asumsi ANOVA	69
22.1 Contoh penggunaan visualisasi hasil analisis	76
24.1 Visualisasi arah korelasi	82
24.2 Visualisasi korelasi non-linear	83
24.3 Visualisasi kekuatan korelasi	83
25.1 Output hasil uji korelasi	86
26.1 Garis persamaan regresi	92
27.1 Contoh output hasil analisis regresi sederhana	94
28.1 Contoh <i>output</i> hasil analisis regresi majemuk	98
29.1 Contoh output hasil analisis regresi hierarkis	100

Daftar Tabel

8.3 Perhitungan simpangan dan varians dua kelompok	28
10.1 Perbandingan metode statistik parametrik dan non-parametrik	40
15.1 Tabel interpretasi <i>effect size</i> (Cumming, 2011)	52
18.1 Contoh hasil uji normalitas (<i>Shapiro-Wilk</i>)	58
18.2 Contoh hasil uji homogenitas <i>Levene</i>	59
18.3 Contoh hasil analisis <i>independent samples t-test</i>	59
18.4 Contoh statistik deskriptif <i>independent samples t-test</i>	59
18.5 Contoh hasil analisis Mann-Whitney	60
18.6 Contoh statistik deskriptif <i>Mann-Whitney</i>	60
19.1 Contoh hasil uji normalitas normalitas sampel berpasangan	61
19.2 Contoh hasil analisis <i>paired samples t-test</i>	62
19.3 Contoh statistik deskriptif <i>paired samples t-test</i>	62
19.4 Contoh hasil analisis <i>Wilcoxon signed-rank</i>	63
22.1 Contoh tabel <i>output</i> hasil ANOVA satu-arah	72
22.2 Contoh tabel <i>output</i> hasil ANOVA dua-arah	73
22.3 Contoh tabel <i>output</i> hasil ANOVA pengukuran berulang	74
22.4 Contoh tabel output hasil ANOVA campuran	74
22.6 Contoh hasil uji <i>post hoc</i>	75
24.1 Kekuatan koefisien korelasi	83
25.1 Koefisien korelasi variabel penelitian ($N = 267$)	86

Prakata

Buku panduan ini disusun sebagai sumber belajar praktis bagi mahasiswa Psikologi dan ilmu sosial, serta para peneliti pemula yang sedang mempelajari statistik. Tujuannya adalah membantu mereka memahami prinsip-prinsip dasar analisis statistik dengan cara yang sederhana, mudah dipahami, dan relevan dengan kebutuhan akademik mereka.

Buku ini lahir dari pengalaman sehari-hari di kelas, saat mendampingi mahasiswa yang sering kali merasa “alergi” begitu mendengar kata *statistik*. Padahal, kalau dipahami dengan cara yang tepat, statistik bukan sesuatu yang menakutkan. Justru ia adalah alat yang sangat membantu kita untuk membaca, memahami, dan menjelaskan fenomena psikologi secara lebih objektif.

Berbeda dengan banyak buku statistik yang sarat dengan rumus-rumus matematika yang sering kali terasa mengintimidasi, buku ini disajikan dengan pendekatan yang lebih ramah bagi mahasiswa tahun-tahun awal. Penjelasan konsep lebih ditekankan pada pemahaman makna dan penerapannya, bukan sekadar pada perhitungan matematis.

Untuk memperkaya proses belajar, materi dilengkapi dengan soal-soal latihan serta simulasi interaktif yang memungkinkan mahasiswa berlatih secara mandiri dan lebih aktif terlibat dalam pembelajaran. Dengan demikian, buku ini tidak hanya berfungsi sebagai referensi teoretis, tetapi juga sebagai panduan praktis yang membantu mahasiswa membangun keterampilan analisis statistik yang aplikatif.

Harapannya, buku ini dapat menjadi bekal awal yang menyenangkan dan bermanfaat bagi mahasiswa dalam memahami statistik, sekaligus menumbuhkan kepercayaan diri mereka untuk menggunakan sebagai alat penting dalam riset psikologi dan ilmu sosial.

Kami menyadari buku ini masih jauh dari sempurna. Kritik dan saran dari para pembaca tentu sangat kami nantikan, agar pada edisi-edisi berikutnya buku ini bisa semakin baik.

Akhir kata, semoga buku ini tidak hanya menjadi pegangan dalam belajar, tetapi juga menjadi sahabat yang menemani perjalanan Anda meneliti, menulis, dan memahami psikologi dengan lebih ilmiah.

Jakarta, Agustus 2025

Tim Penulis

Tentang Penulis

Dr. Sunu Bagaskara

adalah akademisi dan peneliti di bidang Psikologi dengan latar belakang pendidikan S1 hingga S3 di Fakultas Psikologi Universitas Indonesia. Ia berpengalaman mengajar Psikologi Sosial, Metode Penelitian Kuantitatif dan Statistik, serta Konstruksi Alat Ukur Psikologi. Minat penelitiannya mencakup kognisi sosial, psikologi lalu lintas, dan penerapan statistik dalam psikologi, dengan berbagai publikasi di jurnal ilmiah nasional dan internasional yang mencerminkan dedikasinya pada pengembangan ilmu psikologi dan keselamatan masyarakat.

Profil akademik:  Google Scholar  ORCiD

Dr. Entin Nurhayati

berpengalaman mengajar kurang lebih selama 20 tahun. Pernah mengajar berbagai mata kuliah, seperti Faal, Psi Perkembangan, Psikodiagnosis (TAT/CAT, Rorschach) hingga beberapa tahun terakhir fokus di rumpun penelitian, khususnya mata kuliah metodologi penelitian dan statistik. Berdasar pengalaman mengajar ini serta membimbing penelitian mahasiswa, penulis memahami berbagai kesulitan yang dihadapi mahasiswa dalam memahami metopen dan statistik. Bersama sejawat, penulis berusaha melahirkan rangkaian buku praktis, agar mahasiswa dapat dengan mudah dan cepat memahami ilmu alat mengembangkan pengetahuan ini. Buku ini merupakan buku pertama yang dirancang untuk keperluan tersebut.

Profil akademik:  Google Scholar  ORCiD

Sari Zakiah Akmal, Ph.D., Psikolog

adalah akademisi dan peneliti di bidang Psikologi Pendidikan. Ia lulusan dari program Sarjana dan Magister Profesi Psikologi Pendidikan, Fakultas Psikologi Universitas Indonesia dan program doktoral dari School of Applied Psychology, Griffith University, Australia. Sebagai dosen psikologi, ia pernah mengampu matakuliah terkait Metode Penelitian Kuantitatif, Statistik, Logika – Penulisan Ilmiah, Konstruksi Alat Ukur Psikologi, Asesmen Inteligensi, dan Pengembangan Diri dan Karier. Ia juga aktif dalam melakukan penelitian dan pengabdian masyarakat dengan fokus utama pada pengambilan keputusan karier, adaptabilitas karier, ketahanan akademik di kalangan mahasiswa, dan berbagai permasalahan terkait pendidikan siswa SMA/SMK dan mahasiswa di perguruan tinggi. Karya-karyanya telah dipublikasikan dalam berbagai jurnal akademik dan konferensi, mencerminkan komitmennya dalam mengembangkan penelitian dan pendidikan psikologi.

Profil akademik:  Google Scholar  ORCiD

Titi Sahidah, M.Psi, M.Sc., Ph.D., Psikolog

adalah pengajar di Fakultas Psikologi Universitas YARSI, Jakarta. Sejak memulai karier akademiknya pada tahun 2012, ia secara konsisten mengampu mata kuliah Psikometri dan Konstruksi Alat Ukur, yang menjadi bidang minat dan keahliannya. Ia melanjutkan studi doktoralnya di bidang Psikologi Medis di Erasmus MC, Rotterdam, Belanda. Di bawah bimbingan ahli dalam bidang psikometri, Titi menulis disertasi mengenai validitas, reliabilitas dan norma alat ukur EQ-5D-Y di Indonesia. Karya-karya publikasinya, khususnya dalam pengembangan alat ukur psikologi kesehatan yang dipresentasikan di berbagai konferensi dan diterbitkan di jurnal internasional, menjadikannya salah satu pakar dalam bidang pengukuran psikologi kesehatan di Indonesia.

Profil akademik:  [Google Scholar](#)  [ORCiD](#)

Aswin Januarsjaf, S.Psi., M.M

adalah akademisi dan praktisi di bidang Psikologi serta Manajemen Sumber Daya Manusia. Lulusan Psikologi Universitas Padjadjaran dan Magister Manajemen PPM School of Management ini memiliki pengalaman lebih dari dua dekade di berbagai industri, khususnya dalam bidang rekrutmen, asesmen, dan pengembangan SDM. Ia mendirikan Talentlytica, sebuah platform asesmen berbasis data, serta aktif mengajar statistik psikologi dan melatih topik manajemen SDM di berbagai institusi. Karya-karya dan inovasi perangkat asesmennya mencerminkan komitmennya untuk menghubungkan ilmu psikologi dengan praktik manajemen modern di Indonesia.

Profil profesional:  [LinkedIn](#)

Bagian I

PENGANTAR STATISTIK

Statistik memainkan peran yang sangat penting dalam dunia psikologi, khususnya dalam konteks penelitian dan pengambilan keputusan berbasis data. Bab-bab awal buku ini bertujuan memberikan pemahaman dasar tentang konsep-konsep statistik yang relevan bagi mahasiswa psikologi. Mulai dari sejarah dan definisi statistik, pembahasan ini menjelaskan bagaimana perkembangan ilmu ini dipengaruhi oleh berbagai tokoh dan kebutuhan untuk memahami variasi dalam kehidupan manusia. Selain itu, bab-bab ini juga menguraikan prinsip-prinsip dasar seperti hubungan antara sampel dan populasi, serta peran parameter dan statistik dalam menginterpretasikan data penelitian.

Lebih jauh, buku ini membahas pentingnya statistik dalam penelitian psikologi, terutama untuk mengakomodasi kompleksitas perilaku manusia. Dengan menggunakan pendekatan deskriptif dan inferensial, statistik membantu peneliti menggambarkan data secara efektif sekaligus membuat generalisasi yang valid dari sampel ke populasi. Bab-bab ini dirancang untuk menjadi landasan kokoh bagi mahasiswa dalam memahami bagaimana statistik dapat digunakan untuk menjawab berbagai pertanyaan penelitian, mengatasi tantangan analisis data, dan mendukung validitas ilmiah psikologi sebagai disiplin ilmu.

i Tujuan Pembelajaran

Bab 1. Pengertian & Sejarah Ringkas Statistik:

- Memahami definisi statistik dan konsep dasar yang meliputi parameter dan statistik.
- Mengidentifikasi hubungan antara sampel dan populasi dalam konteks penelitian.
- Memahami peran statistik sebagai alat untuk mengorganisasi, merangkum, dan menginterpretasi informasi.

Bab 2. Statistik dalam Penelitian Psikologi:

- Menjelaskan peran statistik dalam mendukung psikologi sebagai ilmu ilmiah.
- Mengevaluasi signifikansi statistik dalam memahami variasi perilaku dan membuat inferensi dari data penelitian.

Bab 3. Jenis-jenis Statistik: Deskriptif & Inferensial:

- Mengidentifikasi perbedaan utama antara statistik deskriptif dan statistik inferensial, termasuk ruang lingkup dan tujuan keduanya.
- Memahami bagaimana kombinasi statistik deskriptif dan inferensial mendukung analisis data yang kuat, khususnya dalam ilmu perilaku dan sosial.

Bab 1

Pengertian & Sejarah Ringkas Statistik

Sejarah statistik bermula dari kebutuhan memahami **variasi dalam ilmu hayati** dan sosial (Gravetter & Wallnau, 2017). Variasi ini menjadi pusat perhatian dengan diperkenalkannya teori evolusi oleh Charles Darwin pada abad ke-19. Pionir seperti Francis Galton, sepupu Darwin, berusaha menguantifikasi pewarisan sifat biologis dan memunculkan **konsep regresi terhadap rata-rata**. Galton juga mengembangkan metode **korelasi**, yang menjadi dasar statistik modern. Selanjutnya, Karl Pearson memperkuat landasan matematis statistik dengan memperkenalkan koefisien korelasi dan uji chi-kuadrat. Perkembangan ini melibatkan kajian biometrika dan genetik, dengan fokus pada pewarisan sifat yang dipengaruhi oleh gerakan eugenika.

Pada abad ke-20, Ronald Fisher memperkenalkan **analisis varians** dan **metode eksperimental** untuk memperluas penerapan statistik, terutama dalam ilmu perilaku (Cowles, 2000). Fisher, bersama Pearson, menekankan pentingnya inferensi statistik untuk mengatasi variasi alami. Pada saat yang sama, probabilitas berkembang sebagai dasar teoretis statistik, berakar dari upaya menjawab masalah dunia nyata seperti perjudian dan prediksi aktuarial. Normalisasi distribusi data, yang diperkenalkan oleh Lambert Quetelet, menjadi inti statistik inferensial. Hingga kini, statistik terus berkembang dengan mengintegrasikan metode baru untuk menjawab tantangan dalam berbagai disiplin ilmu.

 Simulasi Interaktif

Linimasa Sejarah Statistik



Scan QR atau klik [link ini](#) untuk membuka simulasi interaktif.

1.1 Definisi Statistik

Statistik adalah **kumpulan prosedur matematis** yang digunakan untuk **mengorganisasikan, merangkum, dan menginterpretasi** informasi (Gravetter & Wallnau, 2017; Sugiyono, 2015). Dalam konteks penelitian, statistik membantu para peneliti untuk memahami hasil studi dan menyampaikan informasi tersebut secara akurat dan informatif. Statistik juga menyediakan teknik-teknik standar yang diakui dalam komunitas ilmiah sehingga hasil analisis dapat diinterpretasikan dengan konsistensi yang tinggi oleh peneliti lainnya.

1.2 Sampel dan Populasi

Populasi adalah **keseluruhan individu** atau elemen yang menjadi fokus penelitian. Karena populasi biasanya terlalu besar untuk dianalisis secara langsung, peneliti memilih **sampel**, yaitu **sekelompok individu yang mewakili populasi**. Sampel bertujuan untuk menyederhanakan penelitian sambil tetap memungkinkan penarikan kesimpulan yang dapat digeneralisasi ke populasi. Hubungan antara sampel dan populasi bersifat dua arah: sampel diambil dari populasi, dan hasil penelitian pada sampel digeneralisasikan kembali ke populasi (Azwar, 2019; Moore dkk., 2018).

1.3 Parameter dan Statistik

Parameter adalah nilai yang menggambarkan suatu **karakteristik dari populasi**, seperti rata-rata populasi, sedangkan **statistik** adalah nilai yang menggambarkan karakteristik dari sampel, seperti rata-rata sampel. Dalam penelitian, data yang diperoleh berasal dari sampel, sehingga **statistik** digunakan untuk **memperkirakan parameter populasi**. Proses ini menciptakan hubungan erat antara statistik sampel dan parameter populasi, yang menjadi dasar dari banyak prosedur statistik.

Bab 2

Statistik dalam Penelitian Psikologi

Statistik memegang peran penting dalam perkembangan ilmu psikologi, khususnya dalam upaya menjadikannya **ilmu yang lebih ilmiah** (Dancey & Reidy, 2017). Dimulai dengan prinsip **determinisme** yang mengasumsikan adanya **keteraturan dalam alam**, psikologi mencoba meniru pendekatan ilmu alam untuk memprediksi dan mengontrol perilaku. Hal ini dipopulerkan oleh tokoh seperti John B. Watson melalui **behaviorisme** yang mengedepankan studi eksperimental untuk mengendalikan variabel-variabel perilaku. Namun, pendekatan **deterministik** ini dipadukan dengan model **probabilistik** untuk mengakomodasi kompleksitas dan variasi yang melekat pada manusia.

Statistik memungkinkan psikologi menangani variasi data dan membuat **inferensi** dari sampel ke populasi. Misalnya, Ronald Fisher mengembangkan desain eksperimen dan analisis varians (*ANOVA*) untuk membantu peneliti memahami **hubungan sebab-akibat** meskipun terdapat variabilitas. Selain itu, pengembangan metode korelasi oleh Francis Galton dan Karl Pearson membantu mengukur **hubungan antar variabel**, sementara analisis faktor yang diperkenalkan Charles Spearman mengidentifikasi faktor-faktor dalam kecerdasan (Coolican, 2014).

Dalam sejarahnya, statistik tidak hanya digunakan untuk menganalisis data, tetapi juga sebagai **alat eksplorasi**. Contohnya, metode **analisis faktor** sering digunakan untuk memetakan domain baru dalam psikologi sebelum eksperimen langsung dilakukan. Statistik juga membantu memperjelas perbedaan antara pendekatan eksperimental yang berfokus pada manipulasi variabel independen dan pendekatan korelasional yang mempelajari hubungan antar variabel dalam kondisi alami.

Signifikansi statistik dalam psikologi tidak lepas dari pengaruh pionir seperti Gustav Fechner, yang memulai pendekatan kuantitatif untuk mengukur hubungan antara stimulus dan sensasi. Meskipun beberapa konsep awal seperti hukum psikofisik Fechner kini telah direvisi, kontribusinya mendasari metode eksperimental modern. Statistik memungkinkan peneliti psikologi untuk mengatasi tantangan dalam memahami kondisi manusia yang penuh variasi, sekaligus menyediakan dasar untuk menguji hipotesis secara objektif (Salsburg, 2002).

Statistik juga memperkuat nilai-nilai keilmianan psikologi, seperti keterbukaan terhadap koreksi, validasi empiris, dan kemampuan untuk memprediksi. Dengan demikian, statistik tidak hanya menjadi alat teknis tetapi juga **kerangka filosofis** yang mendukung psikologi sebagai ilmu yang dapat menjembatani antara determinisme dan kebebasan manusia, serta antara pandangan **positivistik** dan **humanisme**.

Bab 3

Jenis-jenis Statistik: Deskriptif & Inferensial

Statistik dapat dibagi menjadi dua kategori utama, yaitu **statistik deskriptif** dan **statistik inferensial**, yang masing-masing memiliki peran penting dalam analisis data. Statistik deskriptif (**Bagian 3**) digunakan untuk **merangkum, mengorganisasi, dan menyederhanakan data** mentah menjadi informasi yang lebih mudah dipahami. Teknik-teknik ini termasuk distribusi frekuensi, tabel, grafik, rata-rata (*mean*), median, mode, serta ukuran dispersi seperti variansi dan standar deviasi. Statistik deskriptif membantu memberikan gambaran umum tentang data tanpa membuat kesimpulan lebih luas.

Sebaliknya, statistik inferensial (**Bagian 4**) digunakan untuk membuat **generalisasi tentang populasi** berdasarkan data sampel. Teknik ini melibatkan analisis data dari sampel untuk menjawab pertanyaan penelitian dan menyimpulkan parameter populasi. Statistik inferensial sering menggunakan konsep probabilitas untuk memperkirakan sejauh mana kesimpulan dari sampel dapat mewakili populasi secara keseluruhan. Contoh aplikasinya termasuk uji hipotesis, analisis varians (*ANOVA*), dan regresi.

Perbedaan utama antara keduanya terletak pada tujuan dan ruang lingkupnya. Statistik deskriptif fokus pada **penggambaran data** yang ada, sedangkan statistik inferensial fokus pada **pengambilan keputusan atau prediksi** berdasarkan data tersebut. Misalnya, statistik deskriptif digunakan untuk menghitung rata-rata usia partisipan dalam sebuah survei, sementara statistik inferensial menggunakan data tersebut untuk memprediksi rata-rata usia seluruh populasi.

Dalam praktiknya, kedua jenis statistik ini **saling melengkapi**. Statistik deskriptif menyediakan dasar untuk memahami data, sedangkan statistik inferensial memberikan kerangka untuk mengambil keputusan yang lebih luas. Sebagai contoh, analisis distribusi frekuensi dalam statistik deskriptif dapat menjadi langkah awal sebelum melakukan uji-t atau *ANOVA* dalam statistik inferensial.

Kombinasi keduanya pendekatan ini mendukung analisis data yang kuat, terutama dalam ilmu perilaku dan sosial. Statistik deskriptif membantu menggambarkan data dengan jelas, sementara statistik inferensial memberikan justifikasi yang didukung oleh probabilitas untuk menjawab pertanyaan-pertanyaan penelitian yang kompleks. Hal ini menjadikan statistik alat yang tak tergantikan dalam proses penelitian ilmiah.

Soal Latihan

Penilaian mengenai hasil belajar merupakan bagian penting dalam sebuah proses pembelajaran. Ujilah pengetahuan dan pemahaman Anda mengenai materi Bagian 1 ini dengan menjawab sejumlah pertanyaan pada link di bawah ini:



*Scan QR untuk membuka versi interaktif: [https://www.yarsi.ac.id/fakultas-psikologi/
bukustatistik/latihan1/](https://www.yarsi.ac.id/fakultas-psikologi/bukustatistik/latihan1/)*

Bagian II

PERSIAPAN DATA

Bagian ini akan membahas mengenai mengenali dan mempersiapkan data penelitian agar dapat diolah lebih lanjut secara statistik. Untuk menghasilkan analisis yang valid dan bermakna, peneliti perlu memahami jenis data yang dimiliki, melakukan proses pembersihan data secara sistematis, serta menjaga akuntabilitas dalam setiap tahap pengelolaan data. Tahapan ini bukan sekadar teknis, melainkan fondasi penting untuk memastikan bahwa hasil analisis benar-benar mencerminkan realitas yang diteliti.

Bab ini diawali dengan penjelasan tentang jenis data, termasuk cara mengidentifikasi tipe data serta prinsip dalam penamaan dan pengkodean variabel agar konsisten dan mudah diinterpretasikan. Selanjutnya, pembahasan berlanjut pada proses pembersihan data, yang mencakup analisis deskriptif awal, penanganan missing value, serta deteksi penciran (*outlier*) yang dapat memengaruhi hasil analisis. Terakhir, bab ini menekankan pentingnya akuntabilitas data, yaitu tanggung jawab peneliti untuk mendokumentasikan, menyimpan, dan mengelola data secara transparan agar dapat dipertanggungjawabkan secara ilmiah. Dengan memahami dan menerapkan ketiga aspek ini, mahasiswa akan memiliki dasar yang kuat dalam mempersiapkan data untuk analisis statistik yang sah dan terpercaya.

i Tujuan Pembelajaran

Bab 4. Jenis Data:

- Mengidentifikasi berbagai jenis data dalam konteks penelitian psikologi (nominal, ordinal, interval, rasio).
- Menerapkan prinsip penamaan dan pengkodean variabel secara konsisten dan sistematis.
- Memahami pentingnya kesesuaian antara jenis data dan teknik analisis statistik yang akan digunakan.

Bab 5. Pembersihan Data:

- Melakukan analisis deskriptif awal untuk mengevaluasi distribusi dan karakteristik dasar data.
- Mengidentifikasi dan menangani data hilang (*missing value*) serta *outlier* secara tepat untuk memastikan keakuratan analisis lanjutan.

Bab 6. Akuntabilitas Data:

- Menjelaskan pentingnya pencatatan proses analisis statistik sebagai bentuk akuntabilitas ilmiah.
- Memahami pentingnya transparansi dalam meningkatkan kualitas dan integritas penelitian psikologi.

Bab 4

Jenis Data

Sebelum data dianalisis, peneliti perlu memahami terlebih dahulu **jenis data** yang dimiliki. Pengklasifikasian data menjadi **nominal**, **ordinal**, **interval**, atau **rasio** akan menentukan teknik statistik apa yang tepat untuk digunakan. Selain itu, proses penamaan dan pengkodean variabel juga merupakan bagian penting dalam pengelolaan data yang baik. **Pengkodean yang konsisten** akan memudahkan dalam proses input data, analisis, maupun interpretasi hasil. Oleh karena itu, pengenalan jenis data dan cara mengelola variabel adalah fondasi penting sebelum melangkah ke tahap analisis statistik.

4.1 Mengidentifikasi Jenis Data

Data adalah kumpulan informasi berupa angka. Informasi ini biasanya dicatat sesuai dengan definisi yang diinginkan oleh peneliti. Sebagai contoh, data tinggi badan dicatat dalam centimeter per individu untuk anak-anak usia 0-10 tahun di Indonesia. Pada data tersebut terdapat data tinggi badan anak yang dapat diolah per individu. Di saat yang sama, terdapat pula kelompok data yang lebih besar, misalnya data tinggi badan untuk wilayah DKI Jakarta, Jawa Barat, dan wilayah lainnya yang kemudian dapat diolah dan dibandingkan per wilayah.

Data disebut dengan *raw score (data mentah)* apabila data tersebut adalah data langsung dari partisipan, tidak dimodifikasi, dan tidak dikonversi ke dalam nilai tertentu. Contoh data mentah:

ID Partisipan	Skor Mentah Matematika (jumlah menjawab benar dari 50 soal)	Skor Konversi dalam rentang 0-100 ($raw/50 \times 100$)
1	25	50
2	35	70
3	45	90
4	15	30
5	37	74

Data dapat memberikan berbagai jenis informasi yang berbeda. Misalnya, informasi dasar seperti jenis kelamin, atau informasi lainnya yang lebih rumit seperti sikap atau perasaan seseorang. Terdapat berbagai jenis data di dalam penelitian psikologi, yaitu:

1. Data kategorikal atau nominal

Data kategorikal atau nominal adalah penunjuk **karakteristik**, **label**, atau **go-longan** pada subjek. Sebagai contoh: jenis kelamin (laki-laki dan perempuan),

daerah tempat tinggal (Jakarta, Bekasi, Bogor) dan lain sebagainya. Pemberian kode pada data-data ini hanya bersifat **dummy** dan tidak memiliki makna apa-apa secara statistik (Field, 2018; Stevens, 1946). Sebagai contoh, pada sebuah data penelitian, laki-laki diberikan kode 0 dan perempuan kode 1. Hal ini tidak berarti perempuan memiliki nilai yang lebih tinggi daripada laki-laki, begitu pula sebaliknya. Pemberian kode ini hanya untuk kemudahan pengolahan data dan tidak berarti secara statistik. Misalnya pada jenis kelamin, nilai rata-rata (**mean**) tidak dapat dihitung meski laki-laki disematkan angka 0 dan perempuan disematkan angka 1.

2. Data ordinal

Data ordinal adalah data yang bersifat kategorik dan diketahui urutannya dari yang paling atas hingga paling bawah (Stevens, 1946). Namun demikian, tidak diketahui jarak skor antara satu kategori dengan kategori lainnya dan seberapa besar jarak tersebut. Sebagai contoh, penggolongan status sosial ekonomi (SES) subjek berdasarkan kombinasi sejumlah indikator. SES dapat digolongkan menjadi rendah, sedang, dan tinggi. Kategori tinggi adalah mereka dengan SES yang lebih baik dari sedang dan rendah. Kategori rendah adalah mereka dengan SES yang lebih buruk daripada mereka dengan SES sedang. Tidak ada informasi lebih lanjut bagaimana perbedaan jarak situasi ekonomi antara satu level dengan level yang lain.

3. Data interval

Data interval adalah data yang memiliki makna, terdapat **informasi mengenai jarak** antara satu skor dengan skor lainnya namun tidak ada informasi mengenai seberapa mutlak nilai tersebut karena **titik 0 tidak definitif** (Field, 2018). Misalnya suhu udara 0° Celsius tidak berarti ‘tidak ada temperatur’. Titik tersebut adalah definisi dari kondisi di mana air membeku. Perbedaan suhu antara 30°C dan 20°C adalah 10°C . Namun, tidak dapat dikatakan bahwa 20°C adalah dua kali lebih panas daripada 10°C karena suhu tidak memiliki titik 0°C yang definitif.

4. Data rasio

Data rasio adalah data yang memiliki **titik 0 yang mutlak** sehingga jarak antara satu skor dengan skor lainnya dapat dilakukan secara definit (Field, 2018; Stevens, 1946). Sebagai contoh, berat badan 100 kilogram dapat dikatakan lebih berat dua kali lipat dibandingkan seseorang dengan berat badan 50 kg.

4.2 Penamaan dan Pengkodean Data/Variabel

Langkah awal yang penting dalam mempersiapkan data adalah **memberikan nama atau kode variabel** dengan jelas, mudah dipahami, dan mengikuti pola penamaan yang konsisten. Penamaan variabel yang baik tidak hanya memudahkan peneliti dalam mengolah data, tetapi juga memastikan bahwa **data dapat dipahami dan digunakan** kembali oleh peneliti lain, terutama ketika bekerja dengan data set yang besar atau di dalam tim.

Sebagai contoh, dalam sebuah penelitian yang melibatkan tiga kali pengambilan data menggunakan dua kuesioner – EQ-5D dan PedsQL – penamaan variabel dapat menggu-

nakan kode yang menggambarkan waktu pengambilan data dan jenis instrumen secara ringkas dan jelas. Misalnya:

- Data pengambilan pertama diberi kode B (*Baseline*)
- Pengambilan kedua diberi kode R (*Retest*),
- Dan pengambilan ketiga diberi kode F (*Follow-up*).

Selanjutnya, nama variabel dibentuk dengan menggabungkan kode waktu dan nama instrumen. Dengan demikian:

- Untuk pengambilan pertama dengan EQ-5D, variabel dinamai BEQ5D, dan untuk PedsQL dinamai BPedsQL.
- Untuk pengambilan kedua: REQ5D dan RPedsQL.
- Dan seterusnya untuk pengambilan ketiga.

Pola penamaan yang konsisten seperti ini akan mempermudah proses analisis, mengurangi risiko kesalahan, serta meningkatkan transparansi dan akuntabilitas, khususnya jika pengolahan data dilakukan oleh pihak lain atau pada proyek penelitian berskala besar.

Bab 5

Pembersihan Data

Setelah data dikumpulkan dan diklasifikasikan, langkah selanjutnya adalah memastikan data tersebut layak untuk dianalisis. Proses ini dikenal sebagai ***data cleaning*** atau pembersihan data. Dalam tahap ini, peneliti melakukan **analisis deskriptif awal** untuk mendapatkan gambaran umum data, serta mengecek kemungkinan adanya ***missing value*** atau ***outlier*** yang dapat memengaruhi hasil analisis. Pembersihan data bukan hanya soal menghapus skor-skor yang tidak wajar, tetapi juga tentang membuat keputusan yang tepat berdasarkan konteks penelitian. Dengan data yang bersih, hasil analisis akan menjadi lebih **akurat dan dapat dipercaya**.

5.1 Melakukan Analisis Deskriptif

Setelah peneliti memberi nama variabel dan mengidentifikasi jenis data sesuai dengan sifat masing-masing variabel, langkah berikutnya adalah melakukan analisis deskriptif awal terhadap data. Analisis deskriptif ini bertujuan untuk memastikan kualitas data sebelum masuk ke tahap analisis lebih lanjut.

Beberapa hal penting yang perlu diperhatikan dalam tahap ini antara lain:

- **Memeriksa sebaran skor**

Peneliti perlu melihat apakah nilai total skor, minimum, maksimum, rata-rata (*mean*), dan simpangan baku (standar deviasi) dari setiap variabel berada dalam rentang yang sesuai dengan instrumen alat ukur dan sampel yang diteliti. Hal ini membantu mendeteksi adanya kesalahan input data atau anomali yang tidak sesuai dengan ekspektasi.

- **Memeriksa data hilang (*missing values*)**

Identifikasi jumlah dan pola *missing values* pada setiap variabel sangat penting, karena dapat mempengaruhi validitas hasil analisis.

- **Mendeteksi nilai pencilan (*outliers*)**

Nilai-nilai pencilan dapat mengindikasikan kesalahan pencatatan data atau variasi yang ekstrem di dalam sampel. Oleh karena itu, penting untuk mengidentifikasi *outlier* pada tahap awal, kemudian menentukan apakah nilai tersebut merupakan bagian dari variasi alami atau justru kesalahan yang perlu diperbaiki atau dikeluarkan dari analisis.

Dengan melakukan analisis deskriptif ini secara teliti, peneliti dapat memastikan bahwa data yang akan dianalisis lebih lanjut sudah bersih, valid, dan siap untuk digunakan, sehingga hasil penelitian menjadi lebih akurat dan dapat dipertanggungjawabkan.

5.2 Data Hilang

Terdapat sejumlah hal yang perlu Anda lakukan sebelum memulai analisis utama terhadap data Anda. Hal pertama adalah memeriksa apakah terdapat *missing data* (data hilang). Data hilang terjadi ketika terdapat pernyataan yang tidak diisi oleh seseorang yang menjawab survei atau kuesioner Anda. Partisipan mungkin tidak menjawab salah satu item pada survei karena berbagai alasan. Mereka mungkin melewatkannya sebuah pertanyaan, enggan menjawab pertanyaan tertentu, atau merasa bosan dan berhenti mengisi survei. Ketidaklengkapan jawaban pada satu atau beberapa item ini menimbulkan masalah dalam analisis.

Data hilang dapat dikategorikan menjadi (1) *missing completely at random* (MCAR), (2) *missing at random* (MAR), atau (3) *missing not at random* (MNAR) (Buuren, 2018).

1. **Missing completely at random (MCAR)** berarti probabilitas sebuah data hilang sama besar untuk semua responden. Misalnya, seseorang secara acak melewatkannya pada pertanyaan di survei Anda, sehingga data hilang itu benar-benar acak.
2. Jika probabilitas sebuah nilai hilang hanya sama dalam kelompok-kelompok tertentu yang didefinisikan oleh data yang teramat, maka data tersebut disebut **missing at random (MAR)**. Karena itu, MCAR dan MAR adalah konsep yang saling terkait. Sebagai contoh, ditemukan bahwa responden yang tingkat stresnya tinggi cenderung melewatkannya tentang tidur, mungkin karena merasa pertanyaan tersebut terlalu sensitif atau tidak relevan bagi mereka dalam kondisi stres.
3. Jika data hilang tidak memenuhi asumsi MCAR maupun MAR, maka disebut **missing not at random (MNAR)**. Sebagai contoh, jika sejumlah besar peserta sengaja melewati satu pertanyaan tertentu pada survei, maka hal itu tidak terjadi secara acak. Kemungkinan ada alasan tertentu mengapa pertanyaan itu dilewati, misalnya karena pertanyaan kurang jelas, terlalu pribadi, atau tersembunyi di bagian bawah halaman.

Terdapat sebuah uji untuk mengetahui apakah data hilang secara acak atau tidak, yaitu *Little's MCAR test*. Secara sederhana, jika hasil uji tidak signifikan, maka data hilang kemungkinan terjadi secara acak. Sebaliknya, jika hasil uji signifikan, ada kemungkinan data hilang karena alasan sistematis atau tidak acak.

Lalu, apa yang bisa dilakukan peneliti untuk menangani data hilang? Secara umum, terdapat dua cara utama untuk menangani data hilang, yaitu:

1. **Penggantian dengan rata-rata (mean replacement)**. Dalam prosedur ini, Anda mengganti titik data yang hilang dengan nilai rata-rata variabel tersebut. Namun, teknik ini hanya disarankan jika data hilang secara acak (*at random*) dan proporsinya kurang dari 5% pada variabel yang bersangkutan.
2. **Imputasi majemuk (multiple imputation)**. Teknik ini dilakukan dengan bantuan program statistik yang Anda gunakan, yang menganalisis pola data dan menetapkan nilai untuk variabel yang hilang pada kasus tertentu. Nilai ini didasarkan pada jawaban sebelumnya pada variabel terkait, serta jawaban tidak hilang pada variabel tersebut di kasus lain. Disarankan menggunakan metode ini jika data hilang secara acak dengan proporsi antara 5–10% dari total respons pada variabel tersebut.

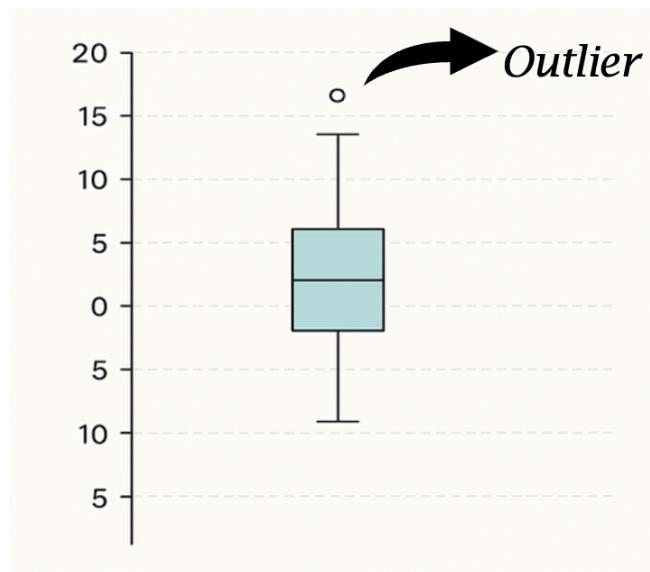
5.3 Data Penculan

Salah satu hal yang perlu diperhatikan saat menganalisis data adalah pengaruh *outlier* — yaitu data yang nilainya sangat tinggi atau sangat rendah dibandingkan data lainnya — terhadap hasil keseluruhan. Contohnya, jika Anda bertanya kepada orang-orang tentang berapa cangkir kopi yang mereka minum dalam sehari, sebagian besar jawabannya mungkin berkisar antara nol hingga empat cangkir. Sebaran ini disebut dengan sebaran jawaban yang normal. Namun, jika ada satu orang yang menjawab bahwa ia minum **17 cangkir kopi sehari**, maka jawaban ini bisa dianggap sebagai *outlier* bila dibandingkan dengan partisipan lainnya.

Kita bisa mengasumsikan bahwa orang tersebut memang memiliki masalah dengan kafein atau mungkin salah memasukkan jawaban. Apa pun alasannya, jawaban ini bisa meningkatkan rata-rata (*mean*) konsumsi kopi dalam sampel, padahal sebetulnya jawaban ini tidak mewakili kebiasaan rata-rata peminum kopi.

Ada beberapa cara secara statistik untuk mengidentifikasi *outlier* dalam kumpulan data Anda (Tabachnick & Fidell, 2019a):

1. Mengubah jawaban peserta untuk setiap variabel menjadi skor-z (*z-score*).
Skor-z adalah transformasi dasar untuk membandingkan jawaban peserta dengan distribusi standar yang memiliki rata-rata 0 dan standar deviasi 1.
Jika suatu jawaban memiliki skor-z lebih besar dari +3.3 atau kurang dari -3.3, maka jawaban ini dianggap sebagai *outlier*.
2. Memvisualisasikan data menggunakan *box plot* atau *bar graph*, lalu melihat secara langsung apakah ada data yang tampak menyimpang jauh. Lihat Gambar 5.1 untuk visualisasi dari *outlier* menggunakan *box plot*.



Gambar 5.1: Outlier dalam boxplot

Setelah menemukan *outlier*, Anda punya dua pilihan (Dancey & Reidy, 2017):

1. **Tetap menyertakan** dalam analisis jika Anda menganggap *outlier* ini bukan kebetulan.
2. **Menghapus** dari analisis jika *outlier* ini muncul karena alasan yang tidak valid, misalnya kesalahan input atau jawaban yang tidak realistik.

Jika jawaban *outlier* tersebut tidak masuk akal dalam konteks pertanyaan, atau terlalu ekstrem tanpa alasan yang jelas, maka sebaiknya dianggap sebagai jawaban tidak valid (*spurious response*) dan dihapus dari analisis. Namun, jika jawaban tersebut masih masuk akal atau hanya sedikit melampaui batas skor-z (± 3.3), Anda bisa mempertimbangkannya sebagai *outlier* yang valid dan tetap menyertakannya dalam analisis.

Bab 6

Akuntabilitas Data

Salah satu langkah penting yang sering kali terabaikan dalam pengolahan data statistik adalah **mencatat secara sistematis** setiap proses analisis yang dilakukan. Pencatatan ini merupakan inti dari prinsip **akuntabilitas**, yaitu kemampuan untuk mempertanggungjawabkan seluruh proses analisis kepada diri sendiri maupun komunitas ilmiah (Gelfond dkk., 2014). Dengan memiliki catatan lengkap, peneliti lain dapat menelusuri, memverifikasi, atau mereplikasi prosedur yang telah dilakukan, sehingga penelitian menjadi lebih transparan dan dapat dipercaya.

Pentingnya akuntabilitas semakin terasa ketika terjadi hasil yang tidak sesuai harapan atau tampak tidak wajar. Dalam situasi ini, dokumentasi proses analisis akan sangat membantu peneliti untuk menelusuri kembali langkah-langkah yang telah diambil, **mengidentifikasi potensi kesalahan** (seperti pengkodean yang keliru, pemilihan teknik analisis yang tidak tepat, atau salah input data), dan memperbaikinya tanpa harus mengulang seluruh proses dari awal.

Pencatatan proses analisis dapat dilakukan melalui *script* atau *syntax* di perangkat lunak statistik. **Tools** ini memungkinkan setiap langkah (misalnya transformasi variabel dan *filtering* data) untuk terekam secara otomatis dan dapat dijalankan ulang. Penggunaan *syntax* bukan hanya efisien, tetapi juga menciptakan **audit trail** yang sangat berguna dalam kerja tim atau revisi laporan.

Jika perangkat lunak yang digunakan tidak menyimpan jejak analisis secara otomatis, maka **pencatatan manual** menjadi penting. Hal ini bisa dilakukan dengan menuliskan setiap prosedur analisis dalam dokumen pendamping, seperti di *Microsoft Word*, *Google Docs*, atau aplikasi pencatat lain. Format catatan sebaiknya mencakup: nama *file* data, langkah-langkah analisis yang dilakukan, alasan memilih teknik tertentu, serta ringkasan hasil yang diperoleh.

Menjaga akuntabilitas bukan hanya menunjukkan **profesionalisme** peneliti, tetapi juga mendukung prinsip replikasi, transparansi, dan validitas ilmiah. Dokumentasi proses analisis dianggap sebagai bagian dari etika penelitian, terutama di era *open science*. Oleh karena itu, mahasiswa perlu membiasakan diri untuk tidak hanya fokus pada hasil akhir, tetapi juga menaruh perhatian serius pada proses yang ditempuh untuk sampai ke hasil tersebut.

Soal Latihan

Penilaian mengenai hasil belajar merupakan bagian penting dalam sebuah proses pembelajaran. Ujilah pengetahuan dan pemahaman Anda mengenai materi Bagian 2 ini dengan menjawab sejumlah pertanyaan pada link di bawah ini:



Scan QR untuk membuka versi interaktif: [https://www.yarsi.ac.id/fakultas-psikologi/
bukustatistik/latihan2/](https://www.yarsi.ac.id/fakultas-psikologi/bukustatistik/latihan2/)

Bagian III

STATISTIK DESKRIPTIF

Setelah data dikumpulkan dan dipersiapkan, langkah selanjutnya dalam proses analisis adalah memahami karakteristik umum dari data tersebut. Bab ini membahas tentang statistik deskriptif, yaitu serangkaian teknik yang digunakan untuk merangkum, menyederhanakan, dan menyajikan data mentah menjadi informasi yang lebih mudah dipahami. Statistik deskriptif tidak berfokus pada pengambilan keputusan atau generalisasi, melainkan pada penggambaran pola, kecenderungan, dan sebaran data yang ada.

Bagian ini terdiri dari tiga bagian utama. Pertama, pembahasan mengenai ukuran pemusatan data (*central tendency*), mencakup *mean* (rata-rata), median, dan modus yang menggambarkan nilai pusat dari sekumpulan data. Kedua, penjelasan tentang sebaran data (dispersi) yang mencakup rentang, varians, dan standar deviasi untuk memahami seberapa jauh data menyebar dari pusatnya. Terakhir, akan dibahas mengenai bentuk distribusi data, termasuk konsep simetri, *skewness*, dan *kurtosis*, yang penting untuk mengetahui apakah data berdistribusi normal atau memiliki penyimpangan tertentu.

Tujuan Pembelajaran

Bab 7. Ukuran Pemusatan Data:

- Menjelaskan konsep ukuran pemusatan data dan peranannya dalam menggambarkan nilai pusat data.
- Menghitung dan menginterpretasikan *mean*, median, dan modus dari data.
- Memilih ukuran pemusatan data yang paling tepat berdasarkan karakteristik distribusi data.

Bab 8. Pembersihan Data:

- Melakukan analisis deskriptif awal untuk mengevaluasi distribusi dan karakteristik dasar data.
- Mengidentifikasi dan menangani data hilang (*missing value*) serta *outlier* secara tepat untuk memastikan keakuratan analisis lanjutan.

Bab 9. Akuntabilitas Data:

- Menjelaskan pentingnya pencatatan proses analisis statistik sebagai bentuk akuntabilitas ilmiah.
- Memahami pentingnya transparansi dalam meningkatkan kualitas dan integritas penelitian psikologi.

Bab 7

Ukuran Pemusatan Data

Bayangkan, Anda sedang melakukan pengamatan terhadap hasil belajar di satu sekolah, misalnya pelajaran matematika dari 150 anak kelas X. Dari daftar nilai itu, Anda mempunyai informasi mengenai nilai dari setiap anak. Namun, bagaimana Anda akan mengatakan kepada khalayak, mengenai capaian nilai matematika dari ke-150 anak ini? Seberapa baik capaian tersebut? Untuk menjawabnya, Anda memerlukan satu **perwakilan angka** yang dapat dinilai sebagai baik, cukup atau buruk. Inilah yang disebut dengan konsep **central tendency** atau ukuran pemusatan data, yaitu satu angka yang menjadi ukuran sentral dari semua kumpulan angka yang diwakilinya. Terdapat tiga macam ukuran sentral, yaitu *mean*, median dan modus.

7.1 *Mean*

Mean (rerata) merupakan ukuran tendensi sentral yang paling umum digunakan karena mempertimbangkan semua nilai dalam distribusi. **Mean** diperoleh dengan cara menjumlahkan seluruh nilai data dibagi dengan banyaknya data. **Mean** sangat sensitif terhadap nilai ekstrem (*outlier*) yang dapat memengaruhi hasilnya secara signifikan.

Misalnya kita memperoleh data dari 10 orang seperti di bawah ini:

$$\begin{array}{cccccccccc} \hline 4 & 4 & 6 & 6 & 7 & 4 & 7 & 3 & 5 & 4 \\ \hline \end{array}$$

Setiap angka menggambarkan skor dari setiap orang. Berikut adalah cara menghitung *mean* dari data tersebut:

$$\text{Mean} = \frac{\sum x}{n} = \frac{4 + 4 + 6 + 6 + 7 + 4 + 7 + 3 + 5 + 4}{10} = \frac{50}{10} = 5$$

7.2 *Median*

Median adalah **nilai tengah** dari data yang telah diurutkan dari yang terkecil ke terbesar. Jika jumlah data ganjil, median adalah nilai di posisi tengah; jika genap, median adalah rata-rata dari dua nilai tengah. Median berguna ketika data mengandung pencilan atau distribusinya tidak simetris, karena tidak terpengaruh oleh nilai-nilai ekstrem.

Menggunakan data yang sama dengan yang di atas, maka untuk memperoleh nilai median kita perlu mengurutkan data dari yang paling kecil hingga yang paling besar.

Nilai median merupakan nilai yang ada di tengah dari urutan data tersebut. Sehingga, dari data tersebut, dapat ditentukan bahwa nilai median adalah sebagai berikut:

3	4	4	4	4	5	6	6	7	7
---	---	---	---	---	---	---	---	---	---

Mengingat jumlah data yang diperoleh genap (10), maka nilai median dihitung dengan menghitung rata-rata dari dua nilai yang berada di tengah. Dalam hal ini, maka median dari data adalah:

$$\text{Median} = \frac{4+5}{2} = 4.5$$

7.3 Modus

Modus adalah nilai yang **paling sering** muncul dalam suatu kumpulan data. Tidak seperti *mean* dan median, modus bisa digunakan pada data kategorikal dan dapat memiliki lebih dari satu nilai (bimodal atau multimodal). Modus cocok digunakan untuk mengidentifikasi nilai yang paling umum dalam suatu populasi atau kelompok.

Langkah penting untuk memperoleh nilai modus dalam kelompok data adalah dengan membuat tabel **distribusi frekuensi**. Tabel ini berfungsi untuk mengidentifikasi nilai, skor, atau data apa yang paling sering muncul (frekuensi tertinggi). Masih menggunakan data yang sama dengan yang di atas, perlu dibuat tabel distribusi frekuensi sebagai berikut:

Nilai	Frekuensi
3	1
4	4
5	1
6	2
7	2

Berdasarkan pengelompokan data berdasarkan frekuensi kemunculannya yang ditampilkan pada tabel Distribusi frekuensi tersebut, maka kita dapat menentukan bahwa nilai Modus dari data adalah 4.

Bab 8

Ukuran Penyebaran Data

Setelah mengetahui nilai pusat dari suatu data melalui ukuran pemusatan data, langkah berikutnya adalah memahami seberapa jauh data menyebar dari nilai pusat tersebut.

Ukuran penyebaran data, atau **dispersi**, memberikan informasi penting tentang variasi atau keragaman dalam kumpulan data. Dua set data bisa memiliki *mean* yang sama, tetapi penyebarannya bisa sangat berbeda, dan perbedaan ini dapat memengaruhi cara interpretasi data secara keseluruhan. Dalam sub-bab ini, akan dibahas beberapa ukuran utama penyebaran, yaitu rentang (*range*), standar deviasi, dan varians, yang membantu menggambarkan tingkat homogenitas atau heterogenitas suatu data secara lebih mendalam.

8.1 Range

Range, atau rentang data, adalah selisih antara nilai tertinggi dengan nilai terendah dalam sebuah set data. **Range** menunjukkan seberapa **lebar sebaran nilai** dalam data, namun karena hanya mempertimbangkan dua nilai ekstrem, ukuran ini sangat sensitif terhadap *outlier* dan tidak menggambarkan variasi data secara keseluruhan. Meskipun demikian, *range* tetap berguna sebagai gambaran awal tentang sebaran data.

Untuk dapat memahami fungsi *range* dalam memahami set data, mari kita pelajari kasus berikut ini. Misalnya, seorang peneliti memiliki data nilai tugas mata pelajaran Matematika dari 3 kelas yang berbeda:

Kelas A	4	4	6	6	7	4	7	3	5	4
Kelas B	2	6	6	7	7	8	4	4	3	3
Kelas C	6	5	6	4	5	6	4	4	5	5

Pertanyaan penting mengenai data tersebut adalah apakah ketiga kelas tersebut memperoleh capaian belajar yang sama. Jika hanya mengandalkan nilai rata-rata, maka ketiga kelas tersebut akan tampak sama karena memiliki nilai rata-rata yang sama, yaitu 5. Padahal jika dilihat ke perolehan nilai individual, terlihat berbeda. Nilai *range* pada data dapat membantu kita untuk menggambarkan seberapa berbeda dan bervariasi data yang dimiliki dan makna variasi tersebut dalam memahami setiap poin data. Untuk itu, kita perlu menghitung selisih skor terbesar dengan skor terkecil untuk setiap data kelas.

$$\text{Range kelas A} = \text{nilai tertinggi} - \text{nilai terendah} = 7 - 3 = 4$$

$$\text{Range kelas B} = \text{nilai tertinggi} - \text{nilai terendah} = 8 - 2 = 6$$

$$\text{Range kelas C} = \text{nilai tertinggi} - \text{nilai terendah} = 6 - 2 = 4$$

Dari hasil penghitungan, dapat dilihat bahwa kelas B memiliki lebar data yang paling besar, artinya terdapat perbedaan atau variasi yang lebih besar di dalam kelas tersebut dibandingkan dengan kelas-kelas lainnya. Sebaliknya, kelas C memiliki rentang skor paling kecil di antara ketiganya. Dalam hal ini, satu skor yang sama (misalnya, 4) dapat dimaknai secara berbeda berdasarkan nilai *range* tiap kelas.

Satu hal yang perlu diingat adalah bahwa variasi skor yang digambarkan dalam *range* hanya menunjukkan perbedaan antara yang mendapatkan nilai paling tinggi dan yang mendapatkan nilai paling rendah. Dari nilai *range* tersebut kita tidak memiliki informasi mengenai seberapa besar ukuran variabilitas (atau seberapa bervariasi) sebuah set data. Untuk mendapatkan gambaran ini, kita menggunakan ukuran penyebaran data berikutnya, yaitu varians.

8.2 Varians

Varians (s^2) adalah ukuran penyebaran data yang menunjukkan **seberapa jauh nilai-nilai dalam suatu kumpulan data menyimpang dari nilai mean**. Varians dihitung dengan menjumlahkan kuadrat selisih setiap nilai (X) terhadap *mean* (\bar{x}), kemudian dibagi dengan jumlah data (N) (untuk populasi) atau jumlah data dikurangi satu ($n-1$) (untuk sampel). Dapat juga dituliskan sebagai:

$$s^2 = \frac{\sum(X - \bar{x})^2}{N - 1}$$

Hasil varians menunjukkan “rata-rata kuadrat penyimpangan” dari *mean*, sehingga semakin besar nilai varians, semakin besar pula variasi data di sekitar *mean*. Dengan kata lain, varians juga menggambarkan **homogenitas** data, di mana semakin kecil varians maka semakin kecil perbedaan antar poin data, dan sebaliknya semakin besar varians semakin besar pula perbedaan antar poin data.

Sebagai ilustrasi, misalnya terdapat data set skor skala sikap dari dua kelompok partisipan yang berbeda (a & b), dengan masing-masing sebaran data sebagai berikut:

Kelompok a	6	6	6	6	7	7	7	7	8	8	8	$\bar{x}=6.9$
Kelompok b	2	3	5	5	6	6	7	9	9	10	12	$\bar{x}=6.7$

Untuk dapat memperoleh nilai varians, kita perlu menghitung simpangan atau selisih dari setiap data poin terhadap mean kemudian dikuadratkan, seperti pada **Tabel 8.3**.

Dari hasil penghitungan tersebut, kita menemukan perbedaan varians yang cukup besar antara kedua kelompok, meskipun *mean* keduanya tidak jauh berbeda. Hal ini menandakan bahwa kelompok b memiliki sebaran data yang jauh lebih beragam (heterogen) dibandingkan kelompok a (lihat Gambar 8.1). Oleh karena itu, pemaknaan terhadap nilai *mean* tiap kelompok juga berbeda, relatif terhadap variansnya masing-masing.

Tabel 8.3: Perhitungan simpangan dan varians dua kelompok

Kelompok a			Kelompok b		
Skor (X)	Simpangan $(X - \bar{x})$	Kuadrat simpangan $(X - \bar{x})^2$	Skor (X)	Simpangan $(X - \bar{x})$	Kuadrat simpangan $(X - \bar{x})^2$
6	-0.9	0.81	2	-4.7	22.09
6	-0.9	0.81	3	-3.7	13.69
6	-0.9	0.81	5	-1.7	2.89
6	-0.9	0.81	5	-1.7	2.89
7	0.1	0.01	6	-0.7	0.49
7	0.1	0.01	6	-0.7	0.49
7	0.1	0.01	6	-0.7	0.49
7	0.1	0.01	7	0.3	0.09
7	0.1	0.01	9	2.3	5.29
8	1.1	1.21	9	2.3	5.29
8	1.1	1.21	10	3.3	10.89
8	1.1	1.21	12	5.3	28.09
$\Sigma(X - \bar{x})^2$			$\Sigma(X - \bar{x})^2$		
$s^2 = \frac{\Sigma(X - \bar{x})^2}{N - 1}$			$s^2 = \frac{\Sigma(X - \bar{x})^2}{N - 1}$		
0.6			8.4		

Varians memiliki fungsi penting dalam memahami keragaman atau variasi data di sekitar nilai rata-rata. Informasi ini sangat berguna untuk a) menilai konsistensi data (misalnya, dua kelompok dengan rata-rata yang sama bisa memiliki varians yang berbeda; kelompok dengan varians kecil berarti anggotanya lebih homogen.) dan b) dasar dari analisis statistik lanjutan, di mana varians menjadi komponen penting dalam berbagai teknik analisis, seperti standar deviasi, analisis varians (*ANOVA*), uji-t, dan regresi. Dengan demikian, varians bukan hanya ukuran penyebaran, tetapi juga alat untuk mengevaluasi struktur dan kualitas data sebelum melakukan interpretasi atau pengambilan keputusan lebih lanjut.

8.3 Standar Deviasi

Standar deviasi (s) adalah ukuran penyebaran data yang menunjukkan **rata-rata penyimpangan (deviasi) nilai data dari mean**. Ukuran ini memberikan gambaran yang lebih intuitif tentang seberapa besar variasi dalam data. Artinya, semakin besar standar deviasi, semakin lebar sebaran data dari nilai tengahnya. Standar deviasi diperoleh dengan mengambil akar kuadrat dari varians:

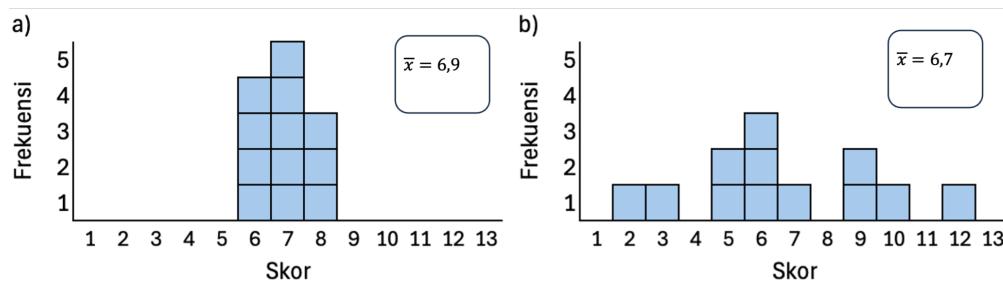
$$s = \sqrt{s^2}$$

Dengan menggunakan data pada pembahasan varians, kita dapat menghitung standar deviasi dari data tiap kelompok, yaitu:

$$\text{Kelompok a: } s = \sqrt{s^2} = \sqrt{0.6} = 0.77$$

$$\text{Kelompok a: } s = \sqrt{s^2} = \sqrt{8.4} = 2.90$$

Hasil penghitungan standar deviasi di atas memperkuat pemahaman bahwa Kelompok a lebih homogen ($s = 0,77$) dibandingkan Kelompok b yang lebih bervariasi ($s = 2,90$). Meskipun informasi ini serupa dengan yang diperoleh dari varians (0,6 vs 8,4), standar deviasi lebih mudah diinterpretasikan karena satuannya kembali ke satuan asli data, tidak dalam bentuk kuadrat seperti varians. Dengan demikian, standar deviasi memberikan gambaran yang lebih intuitif tentang jarak rata-rata setiap data dari *mean*, sehingga lebih sering digunakan dalam pelaporan dan interpretasi hasil statistik.



Gambar 8.1: Visualisasi data dengan: a) varians kecil & b) varians besar

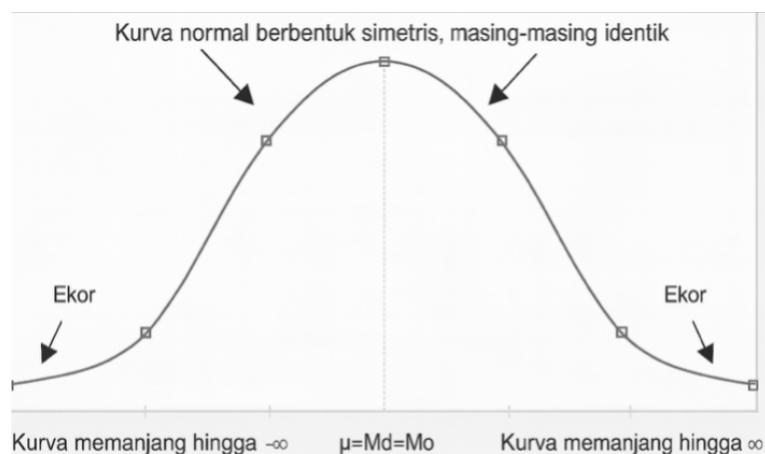
Bab 9

Bentuk Distribusi Data

Setiap nilai dalam sekumpulan data memiliki frekuensi kemunculan yang berbeda-beda; nilai ekstrem biasanya muncul lebih jarang, sedangkan nilai yang mendekati **mean** lebih sering muncul. Pola ini membentuk **distribusi data**, yang jika digambarkan dalam grafik seperti poligon frekuensi, akan membentuk kurva. Pada distribusi yang ideal, yaitu **distribusi normal**, kurva berbentuk lonceng (**bell-shaped**) dan **simetris** di sekitar nilai ukuran pemusatan data seperti **mean** atau median [(Field, 2018; Gravetter & Wallnau, 2017)]. Namun, pada kenyataannya, data bisa saja miring (**skewed**) jika nilai-nilai terkonsentrasi di satu sisi, atau memiliki kurtosis jika kurvanya lebih runcing atau lebih datar dari distribusi normal. Pemahaman tentang bentuk distribusi ini penting karena banyak analisis statistik, terutama yang bersifat inferensial, mengasumsikan bahwa data terdistribusi normal. Untuk itu, bagian-bagian berikut akan membahas lebih lanjut mengenai **kurva normal**, **skewness**, **kurtosis**, dan **z-score** sebagai alat untuk menstandarkan data.

9.1 Kurva Normal

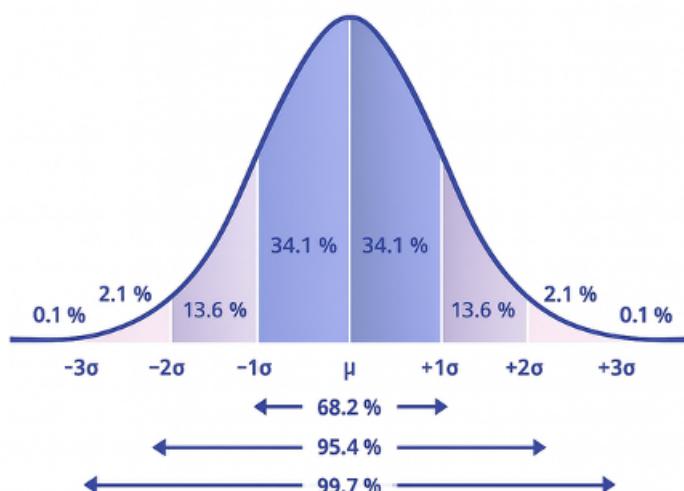
Seperti telah disampaikan sebelumnya, distribusi data dapat dikatakan normal jika kurvanya membentuk lonceng dan simetris. Pada distribusi data yang demikian, frekuensi dari setiap nilai **data tersebar secara simetris** antara yang terletak di atas tendensi sentral dan yang di bawah tendensi sentral. Data tersebar 50% di atas dan 50% di bawah tendensi sentral, dengan frekuensi yang semakin mengecil pada nilai data yang semakin ekstrem, baik yang kecil maupun yang besar (lihat Gambar 9.1).



Gambar 9.1: Bentuk kurva normal

Pada kurva normal, tiga pengukuran data terpusat (*mean*, median, dan modus) berada pada satu titik atau nilai yang relatif sama. Dari Gambar 9.1 dapat lebih terlihat jelas bahwa 50% data tersebar di atas tendensi sentral dan 50% data tersebar di bawah tendensi sentral, di mana semakin jauh dari tendensi sentral maka semakin sedikit frekuensinya.

Untuk menjawab masalah-masalah tertentu, kita perlu mencari besaran frekuensi sebuah nilai atau cakupan nilai berdasarkan sebaran data yang ada. Misalnya, kita ingin memperkirakan ada berapa jumlah mahasiswa yang nilai ujiannya antara 70 dan 85, di mana nilai $\text{mean} = 80$, $SD = 5$, dan $N = 250$. Dengan memanfaatkan informasi standar deviasi (s), proporsi frekuensi data pada distribusi normal dapat kita hitung berdasarkan luas area dalam kurva. Gambar 9.2 menunjukkan besaran proporsi frekuensi pada data dilihat dari luas area berdasarkan nilai standar deviasi.



Gambar 9.2: Proporsi sebaran data pada kurva normal

Dari kurva normal yang ditampilkan pada Gambar 3.3, kita dapat mengetahui bahwa setiap belahan kurva dibagi menjadi 4 area berdasarkan standar deviasi, yaitu (1) $\text{mean} \pm 1 SD$ (34,1%), $1 SD - 2 SD$ (13,6%), $2 SD - 3 SD$ (2,1%), dan $> 3 SD$ (0,1%). Artinya, proporsi terbanyak dari data adalah nilai-nilai yang mendekati *mean*, dan sebaliknya, semakin jauh dari *mean* maka proporsinya semakin kecil, terutama untuk nilai-nilai ekstrem ($> \pm 3 SD$).

Proporsi frekuensi data dengan nilai antara *mean* hingga $1 SD$ di bawah *mean* adalah sebesar 34,1% dari keseluruhan data, begitu juga dengan yang $1 SD$ di atas *mean*. Selanjutnya, jika kita ingin mengetahui proporsi frekuensi data antara *mean* hingga $2 SD$ di atas *mean*, maka kita menjumlahkan proporsi $(\text{mean} - 1 SD) + (1 SD - 2 SD) = 34,1\% + 13,6\% = 47,7\%$.

Dengan demikian, untuk menjawab pertanyaan sebelumnya (proporsi mahasiswa yang memperoleh nilai ujian $70 (X_1) - 85 (X_2)$ dengan $\text{mean} = 80$, $SD = 5$, dan $N = 250$), maka kita menghitungnya dengan mencari besaran SD dari batas bawah dan atas rentang skor:

$$X_1 - \text{mean} = 70 - 80 = -10; \text{ karena } SD = 5, \text{ maka skor } 70 = -2SD$$

$$X_2 - \text{mean} = 85 - 80 = 5; \text{ karena } SD = 5, \text{ maka skor } 85 = 1SD$$

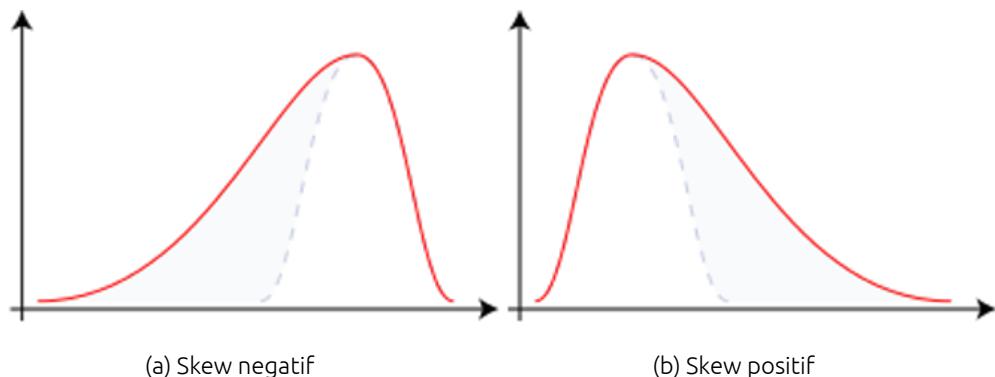
$$\begin{aligned} \text{Maka, proporsi } X_1 - X_2 &= (\text{proporsi } -2SD - \text{mean}) + (\text{proporsi mean} - 1SD) \\ &= (13,6\% + 34,1\%) + (34,1\%) \\ &= 81,8\% \end{aligned}$$

Sehingga, jumlah mahasiswa dengan skor antara $70 - 85 = 81,8\% \times 250$ orang = 204 orang (pembulatan).

9.2 Skewness

Data pada sampel yang diambil dari populasi penelitian tidak selalu terdistribusi secara normal. Sebaran yang tidak normal ini seringnya terjadi pada sekumpulan data yang memiliki nilai ekstrem (*outlier*) yang proporsinya cukup jauh melebihi data yang tersebar secara normal. Besarnya *outlier* ini membuat sebaran data seolah-olah terpusat di bawah nilai tengah, atau sebaliknya terpusat di atas nilai tengah.

Jika data yang tersebar secara normal membentuk kurva berbentuk lonceng yang simetris, maka **data yang tidak normal** tersebut bisa membentuk **kurva yang mirroring** atau juling (*skewed*). Arah kemiringan (*skewness*) kurva bergantung pada besaran frekuensi data yang menyimpang dari nilai tengah (Cumming, 2011). Jika frekuensi nilai yang berada di bawah *mean* jauh lebih kecil daripada frekuensi nilai di atas *mean*, maka puncak kurva akan condong ke kanan, yang disebut sebagai *skew negatif* (Gambar 9.3a). Sebaliknya, jika frekuensi nilai yang berada di bawah *mean* jauh lebih besar daripada frekuensi nilai di atas *mean*, maka puncak kurva akan condong ke kiri, yang disebut sebagai *skew positif* (Gambar 9.3b).



Gambar 9.3: Distribusi skewed

Terdapat dua cara yang umum digunakan untuk mengetahui arah kemiringan (*skewness*) dari sebuah distribusi data, yaitu dengan (a) menggunakan rumus Pearson, dan (b) membandingkan nilai-nilai ukuran pemusatan data.

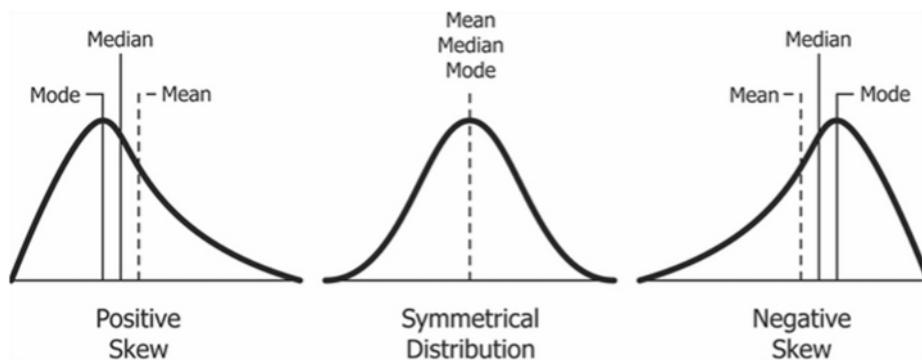
1. Rumus Pearson:

$$\text{Skewness} = 3 \times \frac{\text{Mean} - \text{Median}}{\text{SD}}$$

Jika hasil penghitungan nilai *skewness* adalah positif, maka disebut dengan kurva *skewed* positif. Jika hasil hitungnya negatif, maka disebut dengan kurva *skewed* negatif. Apabila hasil hitungnya (mendekati) nol, maka disebut dengan kurva normal. Tidak ada angka pasti, seberapa dekat dengan nol dapat disebut sebagai kurva normal, tetapi kesepakatan umum adalah $\pm 0,5$.

2. Membandingkan nilai ukuran pemusatan data

Jika *mean* lebih kecil dari median, maka bentuk distribusinya adalah *skewed* negatif. Sebaliknya, jika *mean* lebih besar dari median, maka bentuk distribusinya adalah *skewed* positif. Namun, jika *mean*, median, dan modus berada pada satu titik yang relatif sama atau berdekatan, maka bentuk distribusi datanya adalah normal (Gambar 9.4).

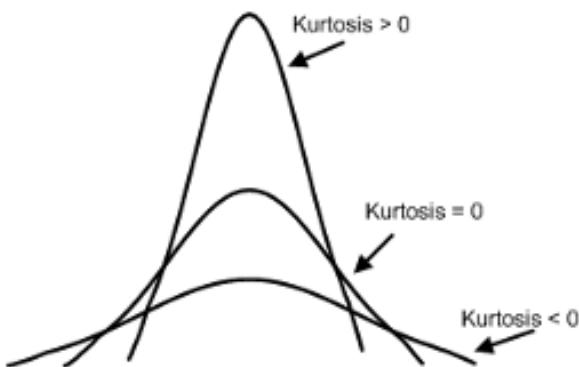


Gambar 9.4: Skewness data berdasarkan lokasi tendensi sentral

9.3 Kurtosis

Kurtosis adalah seberapa **berbeda bentuk kurva dari kurva normal**, dilihat dari seberapa tebal-tipis bagian ekor dari kurva tersebut [(Field, 2018; Howell, 2013)]. Selain dilihat dari ekor kurva, kurtosis sebenarnya juga menunjukkan seberapa tajam bentuk puncak dari kurva. Terdapat tiga bentuk kurtosis, yaitu normal, ekor panjang (***heavy-tailed***) dan ekor pendek (***light-tailed***) (Gambar 9.5).

Kurtosis pada distribusi data yang normal adalah ketika ekor kurva tersebar secara proporsional dan puncak kurva tidak terlalu curam maupun terlalu landai. Jika nilai kurtosis > 0 , maka kurva akan memiliki ekor yang pendek dan puncak yang curam, menandakan data menumpuk di nilai-nilai sekitar nilai tengah. Sebaliknya, jika nilai kurtosis < 0 maka ekor kurva akan memanjang dan puncaknya melandai, yang artinya hampir seluruh nilai memiliki frekuensi atau jumlah kejadian yang serupa. Kurva yang menunjukkan baik ekor panjang maupun ekor pendek menandakan bahwa data tidak terdistribusi secara normal.



Gambar 9.5: Visualisasi kurtosis dalam distribusi data

9.4 Z-Score

z-score, yang dikenal juga dengan nama **skor baku**, menunjukkan **posisi sebuah skor** dibandingkan dengan rata-rata dalam satuan standar deviasi. Dengan kata lain, **z-score** menyatakan posisi relatif suatu nilai (X) dalam distribusi data. $z = 0$ berarti nilai tersebut tepat di **mean**, **z-score** negatif menandakan bahwa nilainya lebih rendah dari **mean**, dan **z-score** positif berarti nilainya di atas **mean**. **z-score** dapat dihitung dengan cara:

$$z = \frac{X - \bar{x}}{\text{SD}}$$

Tujuan utama penggunaan **z-score** adalah untuk menstandarkan data, sehingga memungkinkan perbandingan antar nilai dari distribusi yang berbeda. Sebagai analogi, kita tidak bisa menilai mana yang memiliki nilai yang lebih tinggi antara nilai dua mata uang yang berbeda (misalnya, 10 Rupee India dengan 10 Baht Thailand) tanpa mengkonversinya ke dalam satuan mata uang yang sama atau standar.

Dalam hal data, dua nilai dari distribusi yang berbeda (misalnya, nilai 8 pada subtes aritmatika dan nilai 8 pada subtest logika numerikal) bisa jadi memiliki posisi yang berbeda dalam distribusi data masing-masing, sehingga tidak dapat dibandingkan secara langsung. Oleh karena itu, kedua nilai tersebut perlu ditransformasi ke **z-score** agar memiliki satuan yang sama untuk dapat dibandingkan.

z-score juga bermanfaat dalam mendeteksi **outlier**, serta dalam berbagai analisis statistik lanjutan seperti uji hipotesis dan analisis distribusi normal. Karena **z-score** mengubah data ke skala yang seragam, ia menjadi alat penting dalam interpretasi dan generalisasi hasil penelitian.

Soal Latihan

Penilaian mengenai hasil belajar merupakan bagian penting dalam sebuah proses pembelajaran. Ujilah pengetahuan dan pemahaman Anda mengenai materi Bagian 3 ini dengan menjawab sejumlah pertanyaan pada link di bawah ini:



Scan QR untuk membuka versi interaktif: [https://www.yarsi.ac.id/fakultas-psikologi/
bukustatistik/latihan3/](https://www.yarsi.ac.id/fakultas-psikologi/bukustatistik/latihan3/)

Bagian IV

STATISTIK INFERENSIAL

Statistik inferensial merupakan salah satu cabang dalam statistik yang digunakan untuk mengambil kesimpulan. Berbeda dengan statistik deskriptif yang lebih bertujuan untuk menyajikan data secara ringkas dan mudah dipahami pembaca, statistik inferensial merupakan teknik perhitungan yang ditujukan untuk dapat lebih melihat makna dari pola-pola yang ditunjukkan oleh data. Misalnya, dari dua set ada yang ada, dengan statistik inferensial ini dapat kita lihat apakah menunjukkan pola teratur yang konsisten sehingga dapat dikatakan terdapat keterkaitan antara dua set data itu. Contoh lainnya, dengan statistik inferensial, kita dapat membandingkan dua set data sehingga kita tahu apakah dua set data itu benar-benar berbeda atau hanya tampak berbeda.

i Tujuan Pembelajaran

Bab 10. Statistik Parametrik dan Non-Parametrik:

- Memahami perbedaan antara statistik parametrik dan non-parametrik

Bab 11. Statistical Power:

- Memahami konsep *statistical power* sebagai probabilitas mendekripsi efek yang benar-benar ada.
- Menjelaskan Eror Tipe I & Eror Tipe II dalam statistik

Bab 12. Pengujian Hipotesis:

- Memahami konsep dasar jenis-jenis hipotesis penelitian dan hipotesis statistik.

Bab 13. Signifikansi Statistik dan p-value:

- Menjelaskan arti & interpretasi *p-value*, serta hubungannya dengan tingkat signifikansi.

Bab 14. Confidence Interval:

- Memahami konsep *confidence interval* dan *level of confidence* dalam estimasi parameter populasi.

Bab 15. Effect Size:

- Menjelaskan konsep dan interpretasi nilai *effect size* sebagai ukuran besarnya pengaruh atau hubungan dalam data.

Bab 10

Statistik Parametrik dan Non-Parametrik

Dalam statistik inferensial, metode analisis dapat dibedakan menjadi dua kelompok besar, yaitu statistik **parametrik** dan statistik **non-parametrik**. Pemilihan metode ini berpengaruh pada validitas hasil analisis karena masing-masing memiliki asumsi dan prinsip kerja yang berbeda. Meskipun keduanya bertujuan untuk menarik kesimpulan dari data sampel ke populasi, perbedaan pendekatan dapat memengaruhi akurasi dan kekuatan analisis.

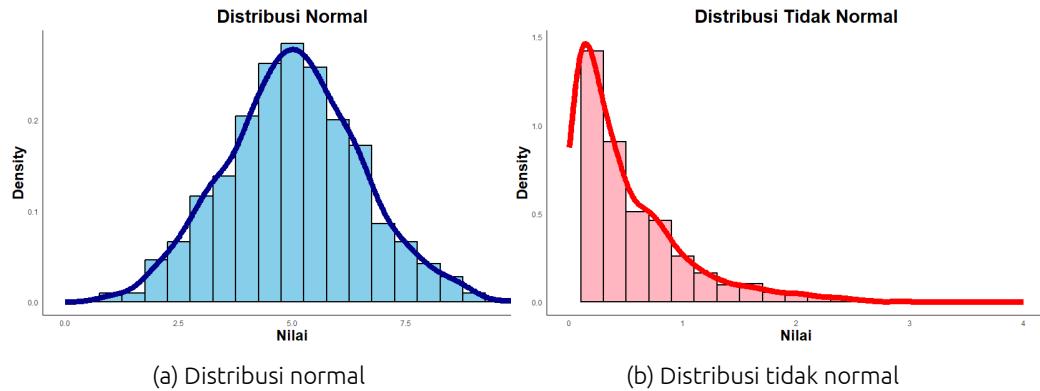
10.1 Asumsi Parametrik

Metode analisis parametrik bekerja dengan optimal jika data memiliki distribusi normal (lihat Bab 9). Karena berbagai prosedur parametrik—seperti estimasi parameter dan pengujian hipotesis—mengandalkan sifat distribusi normal, penyimpangan yang besar dari normalitas dapat membuat hasil analisis menjadi bias atau kurang akurat. Oleh karena itu, sebelum menggunakan metode parametrik, peneliti perlu memeriksa apakah data mendekati distribusi normal melalui uji statistik maupun pemeriksaan visual.

Pemeriksaan melalui **metode visual** bisa dilihat berdasarkan kurva sebaran data. Salah satu yang paling umum digunakan adalah grafik histogram. Jika histogram tampak seperti lonceng, maka distribusi data tersebut dikatakan normal. Sebaliknya, jika terdapat kemiringan sehingga kurva terlihat jelas tidak simetris, maka data tidak terdistribusi dengan normal (lihat Gambar 10.1).

Metode **pengujian statsitik** dapat menjadi cara yang paling akurat untuk menentukan normalitas distribusi data (Field, 2018; Tabachnick & Fidell, 2019b). Terdapat dua uji statistik yang cukup umum digunakan dalam konteks riset ilmu sosial, yaitu **Kolmogorov-Smirnov** (KS) dan **Shapiro-Wilk** (SW). Sejumlah studi mengungkapkan bahwa uji SW lebih andal dibanding KS karena memiliki daya uji (*power*) yang lebih tinggi dalam mendekripsi berbagai bentuk penyimpangan dari normalitas, baik akibat *skewness* maupun *kurtosis*, pada hampir semua ukuran sampel dan jenis distribusi (Oktaviani & Notobroto, 2014; Razali & Wah, 2011; Yap & Sim, 2011). Selain itu, KS cenderung kurang *powerful* untuk sampel kecil (Ghasemi & Zahediasl, 2012). Oleh karena itu, tes SW lebih disarankan untuk digunakan dalam pengujian normalitas distribusi data.

Data yang memenuhi asumsi parametrik dapat dianalisis lebih lanjut dengan statistik parametrik, sedangkan data yang tidak memenuhi asumsi dianalisis dengan menggunakan teknik statistik non-parametrik. Konsekuensi hasil analisisnya adalah pada generalisasi hasil analisis data. Pada statistik parameterik — karena data sampel terdistibusi normal — maka data tersebut dianggap mewakili populasi, sehingga setiap hasil analisis



Gambar 10.1: Kurva distribusi data

data tersebut dapat digeneralisasikan kepada populasinya. Dengan kata lain, hasil analisis data dianggap mencerminkan populasinya. Sedangkan jika data tidak memenuhi asumsi dan kemudian dianalisis dengan statistik non-parametrik, maka hasilnya tidak dapat disebut mencerminkan populasi, melainkan hanya menggambarkan sekelompok partisipan penelitian tersebut.

10.2 Metode Parametrik

Metode **parametrik** adalah pendekatan analisis statistik yang mendasarkan perhitungannya pada **parameter populasi** yang diestimasi dari data sampel, seperti *mean*, varians, dan standar deviasi. Metode ini beroperasi di bawah asumsi bahwa data berasal dari populasi dengan distribusi tertentu, umumnya distribusi normal, sehingga sifat-sifat distribusi tersebut dapat digunakan untuk membangun **model matematis yang akurat**. Proses analisis dalam metode parametrik biasanya melibatkan penggunaan rumus yang memanfaatkan parameter-parameter ini untuk menghitung ukuran efek, menguji hipotesis, atau membuat estimasi terhadap populasi.

Kelebihan utama metode parametrik adalah **presisi** dan **kekuanan statistik yang tinggi**, artinya metode ini lebih mampu mendekripsi perbedaan atau hubungan yang benar-benar ada dalam data, asalkan asumsi normalitas terpenuhi. Hal ini karena informasi yang digunakan berasal dari seluruh nilai data, bukan hanya peringkat atau kategori. Akan tetapi, jika asumsi distribusi dilanggar—misalnya data tidak normal atau mengandung outlier ekstrem—maka metode parametrik bisa menghasilkan estimasi yang bias dan kesimpulan yang menyesatkan. Oleh sebab itu, pengujian asumsi, khususnya normalitas, menjadi langkah krusial sebelum memilih metode ini.

10.3 Metode Non-Parametrik

Metode **non-parametrik** adalah pendekatan analisis statistik yang **tidak bergantung pada asumsi** bentuk distribusi tertentu dan tidak secara langsung menggunakan parameter populasi seperti *mean* atau varians dalam perhitungannya (Conover, 1999). Metode ini sering digunakan ketika data tidak memenuhi asumsi normalitas atau ketika data yang

dimiliki berskala ordinal dan nominal, di mana informasi yang tersedia hanya berupa urutan atau kategori. Dalam praktiknya, metode non-parametrik banyak mengandalkan **peringkat (ranking)** untuk melakukan perbandingan antar kelompok atau mengukur hubungan antar variabel.

Kelebihan metode non-parametrik terletak pada **fleksibilitasnya** terhadap bentuk data —metode ini tetap dapat digunakan meskipun data terdistribusi miring, memiliki *outlier*, atau jumlah sampel relatif kecil. Selain itu, metode ini lebih “aman” digunakan jika peneliti tidak yakin bahwa asumsi parametrik terpenuhi. Namun, konsekuensi dari fleksibilitas ini adalah **kekuatan statistik yang umumnya lebih rendah** dibandingkan metode parametrik ketika asumsi normalitas sebenarnya terpenuhi. Artinya, metode non-parametrik mungkin **kurang sensitif** dalam mendekripsi perbedaan atau hubungan yang ada, sehingga hasil yang signifikan memerlukan efek yang lebih besar atau jumlah sampel yang lebih banyak.

Rangkuman perbandingan secara umum antara kedua metode analisis (parametrik vs. non-parametrik) dapat dilihat di Tabel 10.1.

Tabel 10.1: Perbandingan metode statistik parametrik dan non-parametrik

Aspek	Statistik Parametrik	Statistik Non-parametrik
Asumsi utama	Data berasal dari populasi yang berdistribusi normal.	Tidak memerlukan asumsi distribusi normal
Jenis data yang cocok	Data berskala interval atau rasio	Data berskala ordinal atau nominal, atau data interval/rasio yang tidak normal
Kelebihan	Memiliki daya statistik lebih tinggi jika asumsi normalitas terpenuhi	Lebih fleksibel terhadap pelanggaran asumsi dan dapat digunakan pada data yang tidak normal
Kekurangan	Data harus memenuhi asumsi-asumsi yang ketat	Umumnya memiliki kekuatan statistik lebih rendah
Konsekuensi praktis	Memberikan hasil yang lebih presisi pada data yang memenuhi asumsi normalitas	Memberikan hasil yang kurang sensitif dalam mendekripsi perbedaan atau hubungan yang ada

Bab 11

Statistical Power

Statistical power, atau **tingkat sensitivitas**, adalah kemampuan suatu pengujian statistik untuk mendeteksi adanya efek suatu variabel terhadap variabel lain atau perbedaan antar kelompok jika efek atau perbedaan tersebut memang benar-benar ada. Studi dengan **power** yang tinggi memiliki peluang besar untuk menemukan efek, bahkan jika efek tersebut kecil, selama efek itu memang nyata. Sebaliknya, studi dengan **power** yang rendah cenderung gagal mendeteksi perbedaan atau hubungan yang sebenarnya ada, dan hasil signifikansinya lebih rentan dipengaruhi oleh **error**, baik yang bersifat sistematis maupun acak.

Konsep ini sangat penting karena berkaitan langsung dengan risiko melakukan **kesalahan dalam pengujian hipotesis**, khususnya **kesalahan tipe II (β)**. Dalam statistik, terdapat dua jenis kesalahan:

1. **Kesalahan tipe I (α)**: Menolak hipotesis nol padahal sebenarnya tidak ada perbedaan atau hubungan.
2. **Kesalahan tipe II (β)**: Gagal menolak hipotesis nol padahal sebenarnya ada perbedaan atau hubungan.

Power penelitian dihitung sebagai $1 - \beta$, sehingga semakin kecil β , semakin besar **power**. Misalnya, **power** sebesar 80% berarti penelitian memiliki peluang 80% untuk mendeteksi perbedaan atau hubungan yang sebenarnya ada di populasi.

Menentukan tingkat **power** biasanya dilakukan sebelum pengambilan data agar peneliti dapat memperkirakan ukuran sampel yang memadai. Tingkat **power** yang umum digunakan adalah 80%, 90%, atau 95%. Penentuan ini mempertimbangkan beberapa faktor utama: ukuran sampel, besaran efek (**effect size**), tingkat signifikansi (α), dan variabilitas data. Dengan perencanaan yang baik, peneliti dapat meminimalkan risiko kesalahan tipe II dan meningkatkan peluang memperoleh hasil yang akurat dan dapat diandalkan.

Bab 12

Pengujian Hipotesis

Secara etimologi, hipotesis berasal dari dua kata, yaitu *hipo* yang berarti "di bawah" dan *thesa* yang berarti "kebenaran". Secara harfiah, hipotesis dapat diartikan "di bawah kebenaran" atau **kebenaran sementara** yang masih harus diuji kesahihannya. Dalam penelitian, hipotesis adalah jawaban sementara terhadap pertanyaan penelitian yang disusun berdasarkan teori, temuan penelitian sebelumnya, maupun logika dan akal sehat peneliti.

Hipotesis dibuat untuk dikonfirmasikan kepada data lapangan; jika terkonfirmasi, maka hipotesis diterima sebagai kebenaran berdasarkan data penelitian. Dalam analisis statistik, terdapat dua macam hipotesis yang wajib dipahami:

1. **Hipotesis nol (H_0)**, yaitu hipotesis yang menyatakan ketiadaan perbedaan antara dua keadaan atau ketiadaan hubungan antar variabel yang diamati. Sebagai contoh:
 - Tidak ada perbedaan tingkat konsentrasi belajar antara kelas A yang diberi sarapan dan kelas B yang tidak diberi sarapan.
 - Tidak ada hubungan antara tingkat pemahaman bahaya merokok dengan kecenderungan merokok di kalangan remaja perokok.
2. **Hipotesis alternatif (H_1)**, yaitu hipotesis yang menyatakan adanya perbedaan atau hubungan antar variabel yang diamati. Contohnya:
 - Terdapat perbedaan gaya *parenting* berdasarkan tingkat pendidikan orang tua.
 - Terdapat hubungan antara tingkat keyakinan pada akhirat dengan ketaatan menjalankan sholat lima waktu.

Hipotesis juga dapat diklasifikasikan berdasarkan tujuan dan teknik analisis data yang dibutuhkan:

1. **Hipotesis deskriptif**: digunakan pada penelitian yang hanya melibatkan satu variabel. Misalnya: "*Terdapat kesadaran yang tinggi pada remaja usia belasan terhadap merek sabun 'X'*". Analisisnya meliputi statistik deskriptif, tendensi sentral, dispersi, dan uji normalitas.
2. **Hipotesis komparatif**: digunakan untuk membandingkan dua atau lebih keadaan. Misalnya: "*Terdapat perbedaan konsentrasi belajar antara kelas A yang diberi sarapan dan kelas B yang tidak*". Analisisnya dapat menggunakan uji-t atau ANOVA untuk data parametrik, dan uji *Wilcoxon* atau *Kruskal-Wallis* untuk data non-parametrik.
3. **Hipotesis asosiatif**: digunakan untuk menilai keterhubungan antara dua peristiwa atau variabel. Misalnya: "*Terdapat hubungan antara tingkat eksposur isu di media sosial dengan literasi masyarakat terhadap isu tersebut*". Analisisnya dapat

menggunakan korelasi *Pearson* untuk data parametrik, dan korelasi *Spearman* untuk data non-parametrik.

Dalam pengujian hipotesis, **arah pengujian** menentukan di sisi mana peneliti mencari bukti untuk menolak hipotesis nol. Uji satu arah (**one-tailed test**) digunakan jika hipotesis alternatif memprediksi arah perbedaan atau hubungan yang diharapkan, misalnya rata-rata nilai kelas A lebih tinggi daripada kelas B. Seluruh tingkat signifikansi ditempatkan pada satu sisi distribusi sehingga lebih sensitif untuk mendekripsi perbedaan ke arah tersebut, namun tidak dapat menangkap perbedaan ke arah sebaliknya.

Sebaliknya, uji dua arah (**two-tailed test**) digunakan jika hipotesis alternatif tidak menentukan arah perbedaan, misalnya hanya ingin mengetahui apakah dua rata-rata berbeda tanpa memprediksi lebih tinggi atau lebih rendah. Dalam uji ini, tingkat signifikansi dibagi pada kedua sisi distribusi sehingga dapat mendekripsi perbedaan di kedua arah. Pemilihan jenis uji sebaiknya ditentukan sejak awal penelitian untuk menjaga validitas hasil.

Proses pengujian hipotesis dilakukan secara sistematis. Langkah pertama adalah merumuskan H_0 dan H_1 , kemudian menetapkan tingkat signifikansi (α), umumnya 0,05, untuk membatasi risiko kesalahan tipe I. Selanjutnya, peneliti memilih uji statistik yang sesuai dengan jenis data dan desain penelitian, lalu mengumpulkan dan menganalisis data untuk mendapatkan nilai statistik uji. Nilai ini dibandingkan dengan nilai kritis atau diinterpretasikan menggunakan **p-value**. Jika $p\text{-value} \leq \alpha$, maka H_0 ditolak; jika $p\text{-value} > \alpha$, maka H_0 tidak dapat ditolak. Penjelasan lebih lanjut mengenai signifikansi dan **p-value** akan diuraikan pada Bab 13.

Penting dipahami bahwa menolak H_0 tidak berarti H_0 salah secara mutlak, melainkan data memberikan bukti cukup untuk mendukung H_1 dalam batas risiko yang ditetapkan. Sebaliknya, gagal menolak H_0 tidak berarti H_0 benar, tetapi menunjukkan bukti yang ada belum cukup untuk mendukung H_1 . Dengan pemahaman ini, pengujian hipotesis menjadi alat penting untuk menarik kesimpulan ilmiah yang terukur, transparan, dan dapat dipertanggungjawabkan.

Bab 13

Signifikansi Statistik & *p-value*

Signifikansi statistik dan ***p-value*** merupakan dua konsep yang saling terkait dan sangat penting dalam pengujian hipotesis. Keduanya membantu peneliti menentukan apakah hasil penelitian cukup kuat untuk menolak hipotesis nol atau tidak. Pemahaman yang tepat mengenai kedua konsep ini tidak hanya mencegah kesalahan interpretasi, melainkan juga memastikan bahwa keputusan yang diambil mempertimbangkan konteks penelitian, ukuran efek, dan ***statistical power***, sehingga hasil yang diperoleh lebih akurat dan bermakna secara praktis.

13.1 Signifikansi

Dalam statistik, **signifikansi** merujuk pada **tingkat keyakinan** bahwa suatu hasil penelitian bukan semata-mata disebabkan oleh kebetulan. Tingkat signifikansi biasanya dinyatakan dengan simbol α (alpha), yang mewakili probabilitas melakukan kesalahan tipe I — yaitu menolak hipotesis nol (H_0) padahal H_0 benar. Nilai α yang umum digunakan adalah 0,05 (5%) atau 0,01 (1%). Pemilihan nilai ini merupakan batas yang ditetapkan peneliti sebelum analisis, untuk mengontrol risiko salah mendtolak H_0 .

Tingkat signifikansi berfungsi sebagai **ambang batas** untuk memutuskan apakah bukti yang diperoleh dari data cukup kuat untuk menolak H_0 . Semakin kecil nilai α , semakin ketat kriteria yang digunakan, sehingga peluang membuat kesalahan tipe I menjadi lebih kecil, tetapi sekaligus meningkatkan risiko kesalahan tipe II (gagal menolak H_0 padahal H_0 salah). Oleh karena itu, penentuan nilai α harus mempertimbangkan keseimbangan antara ketelitian statistik dan kebutuhan praktis penelitian, termasuk konteks bidang ilmu, risiko kesalahan yang dapat ditoleransi, dan implikasi keputusan yang akan diambil berdasarkan hasil analisis (Moore dkk., 2018).

Ilmu sosial umumnya menggunakan tingkat signifikansi 0,05 karena fenomena yang diteliti sering kali melibatkan variabilitas tinggi dan faktor-faktor yang sulit dikendalikan sepenuhnya, seperti perilaku, persepsi, atau interaksi sosial (Gravetter & Wallnau, 2017). Variabilitas ini membuat data cenderung mengandung banyak ***noise***, sehingga menetapkan ambang yang terlalu ketat (misalnya 0,01) dapat meningkatkan risiko kesalahan tipe II — gagal mendeteksi efek yang sebenarnya ada. Dengan $\alpha = 0,05$, peneliti di ilmu sosial masih memiliki keseimbangan yang cukup baik antara menghindari kesalahan tipe I dan tetap sensitif terhadap efek yang relevan.

Sebaliknya, di ilmu eksak seperti fisika, kimia, atau teknik, eksperimen biasanya dilakukan dalam kondisi yang lebih terkontrol, dengan variabel-variabel yang dapat diukur secara presisi dan replikasi yang konsisten. Karena kontrol yang ketat ini, tingkat kesalahan acak (***random***) lebih rendah, sehingga dimungkinkan untuk menetapkan α lebih kecil (misalnya 0,01 atau 0,001) tanpa mengorbankan terlalu banyak sensitivitas. Di bidang

ini, kesalahan tipe I sering kali memiliki konsekuensi besar — misalnya, klaim penemuan efektivitas obat baru — sehingga diperlukan standar bukti yang lebih kuat sebelum menolak H_0 .

13.2 *p-value*

p-value atau ***probability value*** adalah **probabilitas** mendapatkan hasil pengamatan, atau hasil yang lebih ekstrem, jika H_0 benar. Nilai ini digunakan untuk menilai apakah hasil yang diperoleh cukup kuat untuk menolak H_0 berdasarkan tingkat signifikansi yang ditetapkan.

Aturan pengambilan keputusan:

- $p\text{-value} \leq \alpha$ \square **Hasil signifikan secara statistik**, H_0 ditolak, artinya data memberikan bukti yang cukup untuk mendukung H_1 .
- $p\text{-value} > \alpha$ \square **Hasil tidak signifikan secara statistik**, H_0 tidak ditolak, artinya data tidak memberikan bukti yang cukup untuk menolak H_0 .

Mari kita gunakan ilustrasi sederhana untuk dapat lebih memahami mengenai konsep ***p-value*** ini. Bayangkan Anda sedang bermain permainan melempar koin yang menurut klaim teman Anda adalah koin normal (punya peluang 50% muncul gambar dan 50% muncul angka). Jika diterjemahkan ke dalam istilah hipotesis, maka:

- (a) Hipotesis nol (H_0): Koin itu seimbang (tidak curang)
- (b) Hipotesis alternatif (H_1): Koin itu tidak seimbang (curang)

Lalu Anda melempar koin 10 kali dan hasilnya 9 kali gambar, 1 kali angka. Sekarang Anda bertanya: “*Kalau koin ini benar-benar seimbang (H_0 benar), seberapa besar kemungkinan saya mendapatkan hasil yang ekstrem seperti ini atau lebih ekstrem?*”

Di sinilah ***p-value*** berperan, yaitu menghitung peluang mendapatkan hasil se-ekstrem ini jika H_0 benar:

- Jika peluangnya sangat kecil (misalnya < 5%), maka hasil yang Anda dapatkan tidak cocok dengan asumsi bahwa koin seimbang, sehingga Anda punya alasan kuat untuk menolak H_0 dan curiga koinnya curang.
- Jika peluangnya masih cukup besar (misalnya 30%), maka hasil yang Anda dapatkan masih wajar untuk koin seimbang, sehingga Anda tidak punya cukup alasan untuk menolak H_0 .

Jadi, ***p-value*** bukanlah peluang koin curang atau tidak curang, melainkan peluang mendapatkan data yang Anda lihat jika koin benar-benar seimbang.

13.3 Interpretasi Signifikansi & *p-value*

Terdapat dua hal penting yang perlu diingat diingat dalam menginterpretasi signifikansi statistik dan ***p-value***. Pertama, signifikan secara statistik tidak selalu berarti signifikan secara praktis; ukuran efek (***effect size***) perlu dipertimbangkan (lihat Bab 15). Misalnya, ditemukan perbedaan skor ujian antara kelas yang menggunakan metode pembelajaran daring (***Mean*** = 78,2) dan metode pembelajaran hibrid (***Mean*** = 79,2). Meskipun secara

statistik perbedaan kedua skor tersebut ditemukan signifikan (kemungkinan karena jumlah sampel yang besar), belum tentu memiliki nilai praktis yang bermakna. Dalam hal ini, perbedaan nilai 1 poin nampaknya terlalu “mahal” untuk dibayar dengan kerumitan dalam pelaksanaan dan pengelolaan metode belajar hibrid.

Kedua, *p-value* tidak menunjukkan peluang hipotesis benar atau salah, dan tidak membuktikan H_0 atau H_1 secara mutlak. *p-value* kecil menunjukkan bahwa jika H_0 benar, peluang memperoleh data seperti yang diamati sangat kecil; *p-value* besar menunjukkan data konsisten dengan H_0 , tetapi bukan bukti bahwa H_0 benar.

Dengan memahami hubungan antara signifikansi dan *p-value*, peneliti dapat membuat keputusan yang lebih tepat dalam pengujian hipotesis, sekaligus menghindari kesalahan interpretasi yang umum terjadi.

Bab 14

Confidence Interval

Sebelum memahami konsep *Confidence Interval (CI)*, kita perlu mengingat kembali prinsip dasar statistik, yaitu bahwa statistik adalah estimasi parameter-parameter populasi (misalnya, nilai rata-rata, standar deviasi) berdasarkan data yang diambil dari sampel. Dengan kata lain, nilai yang kita peroleh dalam statistik kemungkinan besar (atau bahkan hampir pasti) tidak sama dengan nilai yang sesungguhnya pada populasi. Akan selalu ada selisih (error) antara nilai pada sampel dan pada populasi.

14.1 Prinsip Dasar *Confidence Interval*

Confidence interval adalah rentang nilai yang digunakan untuk memperkirakan posisi parameter populasi berdasarkan data sampel. Interval ini dibentuk dari estimasi titik, seperti mean atau proporsi, kemudian ditambahkan batas bawah dan batas atas yang dihitung dari data.

Prinsip dasarnya adalah memberikan gambaran bahwa parameter populasi tidak dinyatakan sebagai satu angka pasti, tetapi berada dalam suatu rentang yang masuk akal menurut data yang dikumpulkan. Dengan menggunakan *CI*, peneliti dapat menyampaikan hasil estimasi secara lebih informatif dan transparan, karena selain nilai estimasi, juga disertakan rentang ketidakpastian yang menyertainya.

Tingkat kesahihan dan keterpercayaan sebuah hasil riset bergantung pada seberapa representatif data yang diperoleh dari sampel dan seberapa yakin bahwa hasil analisis data yang dilakukan menggambarkan kondisi populasi yang sebenarnya. Oleh karena itu, pembahasan mengenai *CI* tidak dapat dilepaskan dari dua konsep yang terkait, yaitu *Margin of Error (MoE)* dan derajat keyakinan (*Level of Confidence*).

14.2 *Margin of Error*

Margin of error adalah batas toleransi kesalahan yang menunjukkan seberapa jauh nilai estimasi sampel dapat berbeda dari nilai sebenarnya di populasi. Dalam konteks *CI*, *MoE* menentukan jarak antara estimasi titik (misalnya *mean* sampel) dan batas atas atau batas bawah interval kepercayaan. Semakin kecil *MoE*, semakin sempit intervalnya, yang berarti estimasi lebih presisi.

Besarnya *MoE* dipengaruhi oleh *level of confidence (LoC)* yang dipilih (misalnya 95%), ukuran sampel, dan variasi data. Pada *LoC* yang sama, ukuran sampel yang lebih besar atau variasi data yang lebih kecil akan menghasilkan *MoE* yang lebih kecil, sehingga interval kepercayaan menjadi lebih akurat.

14.3 Level of Confidence

LoC adalah tingkat keyakinan yang digunakan untuk menyatakan seberapa besar kemungkinan *CI* mencakup nilai parameter sebenarnya di populasi. Misalnya, *LoC* 95% berarti jika penelitian yang sama diulang berkali-kali dengan metode yang sama, sekitar 95% dari *CI* yang dihasilkan akan mengandung nilai parameter populasi. Semakin tinggi *LoC*, semakin besar jaminan bahwa interval mencakup parameter yang benar, namun konsekuensinya interval akan menjadi lebih lebar. Sebaliknya, level yang lebih rendah menghasilkan interval yang lebih sempit tetapi meningkatkan risiko parameter sebenarnya berada di luar interval tersebut.

Mari kita gunakan analogi untuk dapat lebih mudah memahaminya. Misalnya, Anda diminta untuk mengestimasi jarak antara Jakarta (dihitung dari lokasi Monumen Nasional) dan Yogyakarta (dihitung dari lokasi Monumen Yogyakarta Kembali). Jika Anda bukan orang yang terbiasa bepergian Jakarta-Yogyakarta, Anda bisa menjawab bahwa jarak keduanya antara 550-600 KM (rentang sempit) dengan derajat keyakinan 45% (level rendah) karena Anda tidak terlalu yakin bahwa jarak sebenarnya ada di antara rentang tersebut.

Atau, Anda bisa menjawab dengan tingkat keyakinan yang jauh lebih tinggi (misalnya 99%) bahwa jaraknya antara 100-800KM (rentang sangat lebar). Memang, hampir pasti jawabannya ada di dalam rentang tersebut, tetapi kemampuan estimasi Anda menjadi sangat diragukan karena rentangnya sangat lebar (700 KM), sehingga sulit untuk memperkirakan berapa jarak sebenarnya (menurut perhitungan *Google Maps* dengan mode kendaraan roda 4, jaraknya adalah 578 KM).

14.4 Menginterpretasi *CI*, *MoE*, & *LoC*

Rentang keyakinan atau *CI* diperoleh dengan cara sebagai berikut:

$$CI = \bar{x} \pm MoE$$

$$\text{di mana } MoE = z \times \frac{s}{\sqrt{n}}$$

Ket.: \bar{x} = rata-rata (*mean*) sampel; z = nilai *z-score* yang sesuai dengan tingkat kepercayaan yang dipilih (misalnya, 1,96 untuk *LoC* 95% dan 2,58 untuk *LoC* 99%); s = standar deviasi sampel; n = ukuran sampel

Jika *mean* dari data adalah 25 dan *MoE* (*LoC* 95%) = 4, maka $CI = 25 \pm 4 = 21-29$, yang artinya bahwa nilai populasi diestimasi berada pada kisaran 21 hingga 29. Dengan menggunakan *LoC* yang lebih tinggi (99%), maka *MoE* = 5,25, sehingga $CI = 25 \pm 5,25 = 19,75-30,25$. Dengan kata lain, dengan tingkat keyakinan 99%, nilai *mean* populasi berada di antara 19,75 hingga 30,25. Kita dapat melihat bahwa semakin tinggi *LoC* maka semakin lebar rentang keyakinannya, begitu pula sebaliknya.

Pemahaman mengenai *CI* dan *MoE* ini sangat penting bagi peneliti untuk mengambil kesimpulan dan keputusan dari data yang telah dianalisis. Kesalahan atau kurangnya pemahaman dapat membawa peneliti pada penyimpulan yang keliru. Sebagai ilustrasi,

di masa Pemilu sebuah lembaga survei politik melakukan riset tingkat keterpilihan (elektabilitas) para pasangan kandidat presiden dan wakil presiden. Dari dua pasang kandidat, berdasarkan risetnya lembaga survei tersebut menemukan bahwa elektabilitas pasangan A mencapai 48% dan pasangan B 52% dengan *MoE* 3% pada *LoC* 95%.

Jika tidak memahami bagaimana hasil ini seharusnya diinterpretasikan, maka mereka dapat berkesimpulan bahwa pasangan B unggul dan akan memenangi persaingan dengan pasangan A. Padahal, jika memperhitungkan *MoE* dalam memahami hasil tersebut, kita akan menemukan bahwa pada kondisi aktualnya, elektabilitas A berada di antara 45–51%, sedangkan B 49–55%. Artinya, masih ada kemungkinan bahwa pasangan A meraih skor lebih tinggi (misalnya 51%) dibandingkan B (misalnya 49%). Oleh karena itu, kemenangan di antara kedua pasangan tersebut masih belum dapat diestimasi secara meyakinkan karena ada area skor elektabilitas yang beririsan. Terlebih lagi jika penghitungannya menggunakan *LoC* 99%, maka irisan skornya akan makin besar, sehingga makin sulit menentukan siapa di antara mereka yang secara aktual lebih unggul.

Bab 15

Effect Size

Effect size adalah ukuran yang menunjukkan **seberapa besar perbedaan atau hubungan yang benar-benar berarti dalam penelitian**, melampaui sekadar signifikansi statistik. Jika nilai p hanya memberi tahu apakah suatu efek ada, maka **effect size** menjelaskan seberapa kuat atau penting efek tersebut secara praktis. Dengan demikian, memahami **effect size** menjadi penting agar peneliti tidak hanya tahu "ada atau tidaknya perbedaan", tetapi juga "seberapa besar perbedaan itu" dalam konteks nyata. Dari titik inilah pembahasan dapat diarahkan pada prinsip dasar **effect size** yang menjelaskan konsep, jenis, hingga cara menghitungnya.

15.1 Prinsip Dasar

Effect size berangkat dari gagasan bahwa **signifikansi statistik** saja tidak cukup untuk menjawab apakah suatu temuan penelitian benar-benar penting secara praktis. Sebuah hasil bisa saja signifikan karena ukuran sampelnya besar, padahal perbedaan yang ditemukan sebenarnya sangat kecil dan tidak relevan dalam praktik. Di sinilah **effect size** berperan, karena ia memberi informasi mengenai **kekuatan hubungan antar variabel** atau **besarnya perbedaan antar kelompok** dengan satuan yang lebih mudah dipahami secara kuantitatif.

Secara prinsip, **effect size** mengukur seberapa besar "efek nyata" yang terjadi dalam sebuah studi. Efek ini bisa berupa perbedaan rata-rata antar kelompok, kekuatan korelasi antara dua variabel, atau besarnya varians yang dapat dijelaskan oleh suatu model. Dengan demikian, **effect size** menjadi jembatan penting yang menghubungkan antara hasil analisis statistik dan implikasi praktis dari penelitian. Ia membantu peneliti untuk tidak hanya menjawab "*apakah ada efek?*" tetapi juga "*seberapa besar efek itu?*".

Pelaporan **effect size** penting karena melengkapi informasi yang tidak diberikan oleh **p -value**. Nilai signifikansi hanya menunjukkan apakah hasil mungkin terjadi karena kebetulan atau tidak, tetapi tidak menjelaskan seberapa besar pengaruh yang sebenarnya. Dengan menyertakan **effect size**, peneliti dapat menilai apakah temuan yang signifikan juga memiliki arti praktis, serta memungkinkan perbandingan lintas penelitian secara lebih adil, misalnya dalam **meta-analisis**. Hal ini mencegah kesalahpahaman bahwa hasil signifikan selalu berarti penting, padahal efeknya bisa saja sangat kecil dan kurang relevan secara praktis.

15.2 Klasifikasi *Effect Size*

Effect size dapat dinyatakan dalam beberapa bentuk tergantung pada teknik analisis statistik yang digunakan. Secara umum, ada dua kategori utama:

- **Perbedaan *mean* yang distandarisasi**

Ukuran ini digunakan untuk membandingkan rata-rata dua kelompok atau lebih. Rumus yang paling populer adalah *Cohen's d* (Cohen, 1988; Lakens, 2013):

$$d = \frac{M_1 - M_2}{SD_{\text{pooled}}}$$

dengan

$$SD_{\text{pooled}} = \sqrt{\frac{(n_1 - 1)SD_1^2 + (n_2 - 1)SD_2^2}{n_1 + n_2 - 2}}$$

Proporsi varians yang dapat dijelaskan

Ukuran ini menjawab pertanyaan: seberapa besar varians pada suatu variabel dapat dijelaskan oleh varians variabel lain. Bentuk yang sering digunakan antara lain:

- ***r*² untuk uji beda *mean***

Dari uji t, *effect size* dapat dihitung dengan:

$$r^2 = \frac{t^2}{t^2 + df}$$

- ***r*² untuk korelasi**

Jika analisis berupa korelasi, *effect size* dihitung dengan menguadratkan koefisien korelasi:

$$r^2 = (r_{xy})^2$$

Semakin besar nilai r^2 , semakin banyak varians satu variabel yang dapat dijelaskan oleh variabel lain.

- ***R*² dalam regresi**

Dalam analisis regresi, *effect size* dinyatakan dengan proporsi varians total yang dijelaskan oleh model:

$$R^2 = \frac{JK_{\text{reg}}}{JK_{\text{tot}}}$$

Nilai R^2 berkisar antara 0–1, dengan interpretasi bahwa semakin tinggi R^2 , semakin baik model regresi menjelaskan varians variabel dependen.

15.3 Ukuran *Effect Size* Lainnya

Selain *Cohen's d* dan korelasi (r), terdapat ukuran *effect size* lain yang digunakan sesuai jenis analisis. Setiap ukuran ini membantu peneliti membaca makna praktis hasil penelitian dalam konteks yang lebih spesifik (lihat Tabel 15.1).

1. *Cramer's V/Phi coefficient*

Digunakan pada uji *chi-square* atau tabel kontingensi untuk melihat kekuatan hubungan antar variabel kategorikal. Nilainya mirip dengan korelasi, semakin mendekati 1 berarti hubungan semakin kuat.

2. *Eta squared (η^2)/Partial eta squared (ηp^2)*

Sering digunakan pada *ANOVA*. Menunjukkan proporsi varians yang dijelaskan oleh faktor independen. *Partial eta squared* lebih umum dipakai dalam penelitian psikologi karena memperhitungkan pengaruh faktor lain.

3. *Odds ratio*

Digunakan dalam penelitian dengan data kategorikal (terutama di ilmu kesehatan). Menggambarkan seberapa besar peluang suatu kejadian terjadi pada satu kelompok dibandingkan kelompok lain.

Tabel 15.1: Tabel interpretasi *effect size* (Cumming, 2011)

Ukuran	Kecil	Sedang	Besar
<i>Cohen's d</i>	0.2	0.5	0.8
<i>r/r²</i>	$r = 0.1 (r^2 \approx 0.01)$	$r = 0.3 (r^2 \approx 0.09)$	$r = 0.5 (r^2 \approx 0.25)$
<i>Cramer's V/Phi</i>	0.1	0.3	0.5
<i>Eta squared (η^2)</i>	0.01	0.06	0.14
<i>Partial eta squared (ηp^2)</i>	0.01	0.06	0.14
<i>Odds ratio (OR)</i>	~1.2	~1.5	≥2.0

Soal Latihan

Penilaian mengenai hasil belajar merupakan bagian penting dalam sebuah proses pembelajaran. Ujilah pengetahuan dan pemahaman Anda mengenai materi Bagian 4 ini dengan menjawab sejumlah pertanyaan pada link di bawah ini:



Scan QR untuk membuka versi interaktif: [https://www.yarsi.ac.id/fakultas-psikologi/
bukustatistik/latihan4/](https://www.yarsi.ac.id/fakultas-psikologi/bukustatistik/latihan4/)

Bagian V

**PERBEDAAN RATA-RATA DUA
KELOMPOK**

Pada bagian ini akan dibahas mengenai perbedaan *mean* variabel yang sama antara dua kelompok berbeda. Pemahaman mengenai hal ini diperlukan untuk membandingkan antara dua karakteristik atau kondisi yang berbeda, misalnya perbedaan antara kelompok eksperimen/*treatment group* (KE) dan kelompok kontrol/*control group* (KK), atau perbedaan antara kelompok pertama dan kelompok kedua, atau perbedaan antara sebelum dan sesudah mendapatkan *treatment*.

i Tujuan Pembelajaran

Bab 16. Memahami Perbedaan Rata-rata Antar Kelompok:

- Memahami perbedaan antara kelompok eksperimen dan kelompok kontrol.

Bab 17. Perbedaan Rata-rata Antar Kelompok vs. Dalam Kelompok:

- Mengidentifikasi perbedaan penggunaan uji beda antar kelompok dan dalam kelompok.
- Memahami makna perbedaan rata-rata dalam konteks analisa statistik.

Bab 18. *Independent Samples t-test: Parametrik dan Non-Parametrik:*

- Menjelaskan asumsi yang perlu dipenuhi dalam uji *independent samples t-test* (parametrik dan non-parametrik), interpretasi dan cara penulisannya dalam laporan.

Bab 19. *Paired Samples t-test: Parametrik dan Non-Parametrik:*

- Menjelaskan asumsi yang perlu dipenuhi dalam uji *paired samples t-test test* (parametrik dan non-parametrik), interpretasi dan cara penulisannya dalam laporan.

Bab 16

Memahami Perbedaan Rata-rata Antar Kelompok

Meneliti dan memahami bagaimana suatu kelompok dapat berbeda jika dibandingkan dengan kelompok lainnya, merupakan salah satu hal penting dan sering menjadi tujuan penelitian pada ranah ilmu sosial, termasuk psikologi. Pemahaman ini sangat diperlukan ketika peneliti ingin mengetahui **efektivitas** dari intervensi atau manipulasi yang dilakukan. Misalnya, seorang dosen ingin mengetahui apakah penjelasan materi dengan pemberian video pembelajaran dapat meningkatkan pemahaman mahasiswa pada mata kuliah statistik. Untuk mengujinya, dosen tersebut melakukan penelitian eksperimental di mana satu kelompok/kelas mahasiswa mendapatkan video pembelajaran, sedangkan kelompok/kelas lainnya tidak.

Kelompok yang mendapatkan video pembelajaran disebut dengan kelompok eksperimen/*treatment group* (KE). KE adalah sekelompok partisipan penelitian yang mendapatkan perlakuan atau intervensi tertentu yang diharapkan dapat menyebabkan terjadinya perbedaan hasil (misalnya, penurunan atau peningkatan skor). Pada contoh sebelumnya, video pembelajaran adalah intervensi yang diberikan untuk meningkatkan skor pemahaman mahasiswa. Sementara itu, kelompok yang tidak mendapatkan perlakuan atau intervensi, disebut dengan kelompok kontrol/*control group* (KK).

Desain penelitian yang melibatkan KE dan KK ini umumnya ditemui pada *setting eksperimental*, di mana peneliti dengan sengaja membagi kelompok dan memberikan perlakuan yang berbeda, seperti pada uji coba klinis. Akan tetapi, dalam konteks ilmu sosial, sangat memungkinkan kelompok tertentu mendapatkan perlakuan atau mengalami hal yang berbeda, dan seolah menjadi "*treatment*" yang terjadi secara alami/natural. Misalnya, adanya pembentukan kelompok karena perbedaan kelas di sekolah, kelompok pekerja, perbedaan generasi, dan sebagainya.

Adanya perbedaan kelompok tersebut sangat memungkinkan untuk menimbulkan perbedaan karakteristik, perilaku atau performa antar kelompok. Contohnya, variabel "*risk-taking behavior*" pada karyawan. Apakah perilaku tersebut akan sama antara karyawan yang baru masuk perusahaan (awal karier) dengan mereka yang sudah mendekati masa pensiun (akhir karier)? Kedua kelompok tersebut akan memiliki perbedaan pemikiran, keterikatan terhadap organisasi dan perilaku kerja. Jika tidak ada perbedaan perlakuan yang dengan sengaja diciptakan atau dimanipulasi oleh peneliti, peneliti umumnya membedakan kedua kelompok dengan menyebutkan Kelompok 1 vs Kelompok 2 atau secara spesifik menyebutkan hal yang membedakan kedua kelompok (misal: kelompok awal karier vs kelompok akhir karier).

Bab 17

Perbedaan Rata-rata Antar Kelompok vs. Dalam Kelompok

Secara statistik, uji beda dilakukan untuk mengetahui apakah terdapat perbedaan **antar kelompok (between groups)** atau perbedaan **dalam kelompok (within groups)**. Perbedaan antar kelompok dilakukan untuk mengetahui apakah dua kelompok *independent* (kelompok yang tidak sama) memiliki perbedaan satu sama lain dalam hal yang sedang diteliti (*dependent variable*).

Uji beda ini umum digunakan untuk mengukur seberapa efektif sebuah *treatment* atau intervensi yang diberikan. Misalnya, peneliti ingin mengetahui apakah intervensi yang dirancang efektif dalam menurunkan tingkat kecemasan mahasiswa terhadap pelajaran statistik. Satu kelompok akan mendapatkan *treatment* penurunan kecemasan, sementara kelompok lain sebagai kelompok pembanding, yaitu mereka yang tidak mendapatkan *treatment*. Kedua kelompok kemudian diukur tingkat kecemasannya setelah mendapatkan *treatment*, untuk selanjutnya dibandingkan apakah ada perbedaan tingkat kecemasan antara kedua kelompok.

Perbedaan dalam kelompok dapat digunakan untuk mengetahui efektivitas dari sebuah program intervensi, hanya saja tidak membandingkan dua kelompok yang berbeda. Melainkan, perbedaan dilihat pada kelompok yang sama namun dalam dua kondisi atau waktu yang berbeda. Misalnya, peneliti ingin mengetahui apakah program olah raga yang dirancang efektif dalam menurunkan berat badan dengan membandingkan BMI partisipan antara sebelum (waktu 1) dengan sesudah (waktu 2) mengikuti program olah raga.

Pada kedua jenis uji beda yang telah dibahas, hal yang dibandingkan adalah perbedaan rata-rata (*mean differences*) variabel yang diteliti. Pada uji beda antar kelompok, hal yang dilihat adalah perbedaan rata-rata antara KE dan KK (perbedaan rata-rata skor kecemasan antara kelompok yang mendapatkan dan tidak mendapatkan intervensi kecemasan). Sementara itu, ketika melakukan uji beda dalam kelompok, hal yang dibandingkan adalah perbedaan rata-rata pada dua waktu pengukuran yang berbeda (perbedaan rata-rata BMI antara sebelum dan sesudah mengikuti program).

Selain perbedaan rata-rata, hal yang juga perlu diperhatikan adalah standar deviasi dari skor yang dibandingkan dan jumlah partisipan pada kedua kelompok. Jika ketiga hal tersebut diketahui, maka uji beda antar dan dalam kelompok dapat dilakukan dengan penghitungan *t-test*, untuk mengetahui apakah perbedaan rata-rata tersebut signifikan secara statistik. Terdapat dua jenis uji *t-test* yang akan dibahas lebih lanjut, yaitu **independent sample t-test** (uji beda antar kelompok berbeda) dan **paired sample t-test** (uji beda dalam kelompok).

Bab 18

Independent Samples t-test: Parametrik dan Non-Parametrik

Asumsi dasar yang perlu dipenuhi untuk melakukan uji beda menggunakan *independent t-test* adalah:

- **Dependent Variable** (DV – variabel yang menjadi fokus penelitian) berupa data kontinum (data interval atau rasio), misalnya skor tes, waktu respon.
- **Independent Variable** (IV – kelompok yang diuji) merupakan dua kelompok yang berbeda, misalnya perbedaan kemampuan bahasa antara kelompok laki-laki dan perempuan.
- Data yang diobservasi harus independen, tidak berasal dari kelompok yang sama.
- Data harus terdistribusi normal. Gunakan uji beda non-parametrik jika data tidak terdistribusi dengan normal.
- Tidak ada *outliers* (data yang sangat berbeda dari rata-rata).
- Data harus memiliki varians yang homogen (*homogeneity*). Asumsi ini dapat dilihat dengan menggunakan *Levene's test for homogeneity of variance*.

Interpretasi dari hasil uji beda menggunakan *independent t-test* perlu dilakukan secara bertahap. Misalnya, peneliti ingin mengetahui perbedaan *self-efficacy* berdasarkan jenis kelamin (laki-laki dan perempuan). Tahapan interpretasi hasil uji *independent t-test* dapat dipahami dengan mempelajari contoh *output* penghitungan *independent t-test* menggunakan *software* JASP berikut:

1. Uji asumsi

Hal pertama yang perlu dilihat adalah uji asumsi terkait distribusi dan homogenitas data; data harus terdistribusi secara normal dan homogen. Pada contoh tabel hasil uji hipotesis yang diperlihatkan pada Tabel 18.1, diketahui bahwa *self-efficacy* pada kedua kelompok terdistribusi normal (kelompok laki-laki $p = .16$; perempuan $p = .07$; $p > .05$). Artinya, sebaran data *self-efficacy* partisipan pada kedua kelompok tidak berbeda secara signifikan dibandingkan sebaran data kurva normal.

Tabel 18.1: Contoh hasil uji normalitas (*Shapiro-Wilk*)

Variabel	Kelompok	W	p
<i>Self-efficacy</i>	Laki-laki	0.961	0.155
	Perempuan	0.961	0.066

Uji homogenitas berdasarkan skor *Levene's test* juga tidak signifikan ($F = 1.53$, $p = .22 > .05$), menunjukkan bahwa kedua kelompok tidak menunjukkan varians

yang berbeda, atau data bersifat homogen (Tabel 18.2). Dengan demikian, data penelitian memenuhi kedua uji asumsi sehingga dapat dilanjutkan untuk dianalisis dengan menggunakan *independent samples t-test*. Jika data tidak memenuhi uji asumsi tersebut, maka uji parametrik tidak dapat dilanjutkan, perlu dilakukan uji non-parametrik (*Mann-Withney test*).

Tabel 18.2: Contoh hasil uji homogenitas *Levene*

Variabel	F	df ₁	df ₂	p
<i>Self-efficacy</i>	1.527	1	98	.220

2. Uji Hipotesis

Selanjutnya, hal yang perlu dilihat adalah nilai p pada tabel *independent samples t-test*. Nilai *p* = .006 (*p* < .05) menunjukkan terdapat perbedaan signifikan tingkat *self-efficacy* yang dilaporkan oleh partisipan laki-laki dan perempuan. Pada tabel *output* uji *independent t-test* (lihat Tabel 18.3) juga dapat dilihat nilai *Cohen's d* yang menunjukkan seberapa besar efek dari perbedaan jenis kelamin terhadap perbedaan *self-efficacy* pada kedua kelompok (*effect size*). Nilai *t*, *df*, *p* dan *Cohen's d* diperlukan dalam penulisan interpretasi hasil uji beda antar kelompok, yang dapat dilihat pada contoh interpretasi hasil *independent samples t-test*.

Tabel 18.3: Contoh hasil analisis *independent samples t-test*

Variabel	<i>t</i>	<i>df</i>	<i>p</i>	<i>Cohen's D</i>	<i>SE Cohen's D</i>
<i>Self-efficacy</i>	2.280	98	.006	0.570	0211

3. Data Deskriptif Kelompok

Terakhir, informasi mengenai *mean* dan *SD* diperlukan untuk mengetahui kelompok mana yang menunjukkan rata-rata skor paling tinggi (lihat Tabel 18.4). Contoh tabel menunjukkan laki-laki memiliki rata-rata *self-efficacy* yang lebih tinggi dibandingkan perempuan.

Tabel 18.4: Contoh statistik deskriptif *independent samples t-test*

Variabel	Kelompok	N	Mean	SD	SE	<i>Coefficient of variation</i>
<i>Self-efficacy</i>	Laki-laki	43	8.209	1.922	0.293	0.234
	Perempuan	57	7.193	1.674	0.222	0.233

Pelaporan hasil uji beda *independent samples t-test* dengan mengikuti format penulisan APA, adalah: Hasil *independent samples t-test* menunjukkan terdapat perbedaan tingkat *self-efficacy* yang signifikan (*t* (98) = 2.82, *p* = .006, *d* = .57) antara kelompok partisipan laki-laki (*M* = 8.21, *SD* = 1.92) dan partisipan perempuan (*M* = 7.19, *SD* = 1.67). Atau,

partisipan laki-laki ($M = 8.21, SD = 1.92$) memiliki *self-efficacy* yang lebih tinggi ($t (98) = 2.82, p = .006, d = .57$) dibandingkan partisipan perempuan ($M = 7.19, SD = 1.67$).

Prinsip yang sama juga dapat diterapkan untuk interpretasi dan melaporkan hasil uji beda non-parametrik (lihat Tabel 18.5), dengan melihat hasil uji *Mann-Whitney test (U)*, dan *mean rank* sebagai pengganti *mean*. Berikut adalah contoh *output Mann-Whitney*:

Tabel 18.5: Contoh hasil analisis Mann-Whitney

Variabel	<i>U</i>	<i>df</i>	<i>p</i>	<i>Rank-Biserial Correlation</i>	<i>SE Rank-Biserial Correlation</i>
<i>Self-efficacy</i>	1576.000		0.013	0.286	0.117

Hasil *Mann-Whitney test* menunjukkan bahwa terdapat perbedaan tingkat *self-efficacy* yang signifikan ($U = 1576, p = .013$) antara kelompok partisipan laki-laki ($Mdn = 58.65$) dan partisipan perempuan ($Mdn = 44.35$). Atau, partisipan laki-laki ($Mdn = 58.65$) memiliki *self-efficacy* yang lebih tinggi ($U = 1576, p = .013$) dibandingkan partisipan perempuan ($Mdn = 44.35$) (lihat Tabel 18.6).

Tabel 18.6: Contoh statistik deskriptif *Mann-Whitney*

Variabel	Kelompok	<i>N</i>	<i>Mean</i>	<i>SD</i>	<i>SE</i>	<i>Coefficient of Variation</i>	<i>Mean Rank</i>	<i>Sum Rank</i>
<i>Self-efficacy</i>	Laki-laki	43	8.209	1.922	0.293	0.234	58.651	2522
	Perempuan	57	7.193	1.674	0.222	0.233	44.351	2528

Bab 19

Paired Samples t-test: Parametrik dan Non-Parametrik

Asumsi dasar yang perlu dipenuhi untuk melakukan uji beda menggunakan *paired samples t-test* adalah:

- *Dependent Variable* berupa data kontinum (data interval atau rasio).
- *Independent Variable* adalah dua kelompok partisipan yang sama dan memiliki data berpasangan (misalnya: *pre-test* dan *post-test*). Dengan kata lain, pengukuran dilakukan pada orang yang sama sebanyak dua kali.
- Data harus terdistribusi normal. Gunakan uji beda non parametrik jika data tidak terdistribusi dengan normal.
- Tidak ada *outliers*.

Uji *paired samples t-test* digunakan untuk menganalisis perbedaan dalam satu kelompok yang sama. Responden untuk kedua pengukuran dalam pasangan adalah orang yang sama, sehingga skor pada kedua pengukuran tersebut saling berkaitan. Bentuk yang umum ditemui adalah partisipan yang sama, memiliki dua skor atau hasil pengukuran yang dilakukan pada waktu yang berbeda. Misalnya, skor saat sebelum dan sesudah mendapatkan *treatment*, pengukuran pada pagi dan malam hari, dan sebagainya.

Interpretasi dari hasil uji beda menggunakan *paired samples t-test* juga perlu dilakukan secara bertahap. Misalnya, peneliti ingin mengetahui perbedaan waktu reaksi partisipan ketika melihat tulisan berwarna hitam dan kuning. Tahapan interpretasi hasil *paired samples t-test* dapat dipahami dengan mempelajari contoh berikut:

1. Uji asumsi

Hal pertama yang perlu dilihat adalah uji asumsi terkait distribusi data; data harus terdistribusi normal. Tabel 19.1 menunjukkan hasil uji normalitas data (*Shapiro-Wilk*). Diketahui bahwa skor waktu reaksi partisipan terdistribusi normal ($p = .34$; $p > .05$). Artinya, sebaran data waktu reaksi pada kedua kondisi (stimulus hitam dan kuning) tidak berbeda secara signifikan dibandingkan sebaran data kurva normal. Dengan demikian, uji beda dapat dilakukan dengan statistik parametrik, *paired samples t-test*. Jika data tidak memenuhi uji asumsi normalitas data, maka uji parametrik tidak dapat dilanjutkan, perlu dilakukan uji non parametrik (*Wilcoxon signed-rank*).

Tabel 19.1: Contoh hasil uji normalitas sampel berpasangan

	<i>W</i>	<i>p</i>
Hitam Kuning	0.974	0.344

2. Uji Hipotesis

Selanjutnya, hal yang perlu dilihat adalah nilai p pada tabel *paired samples t-test*. Nilai $p < .001$ ($p < .05$) menunjukkan terdapat perbedaan signifikan waktu reaksi antara stimulus berwarna hitam dan kuning (lihat Tabel 19.2). Pada tabel *output* uji *paired samples t-test* juga dapat dilihat nilai *Cohen's d* yang menunjukkan seberapa besar efek dari perbedaan warna stimulus terhadap perbedaan waktu reaksi pada kedua kelompok (*effect size*). Nilai t , df , p dan *Cohen's d* diperlukan dalam penulisan interpretasi hasil uji beda antar kelompok, yang dapat dilihat pada contoh interpretasi hasil uji *paired samples t-test*.

Tabel 19.2: Contoh hasil analisis *paired samples t-test*

Penguku- ran 1	Penguku- ran 2	<i>t</i>	<i>df</i>	<i>p</i>	<i>Cohen's d</i>	<i>SE Cohen's d</i>
Hitam	Kuning	5.546	49	< .001	0.784	0.224

3. Data deskriptif

Terakhir, informasi mengenai *Mean* dan *SD* diperlukan untuk mengetahui kelompok mana yang menunjukkan rata-rata skor paling tinggi. Contoh Tabel 19.3 menunjukkan stimulus hitam memiliki rata-rata waktu reaksi yang lebih lama dibandingkan stimulus kuning.

Tabel 19.3: Contoh statistik deskriptif *paired samples t-test*

	<i>N</i>	<i>Mean</i>	<i>SD</i>	<i>SE</i>	<i>Coefficient of Variation</i>
Hitam	50	67.840	9.303	1.316	0.137
Kuning	50	58.560	7.680	1.086	0.131

Pelaporan hasil uji beda *paired samples t-test* dengan mengikuti format penulisan APA, sama dengan pelaporan uji *independent samples t-test*. Hasil *paired samples t-test* menunjukkan terdapat perbedaan waktu reaksi yang signifikan (t (49) = 5.55, $p < .001$, d = .78) antara stimulus berwarna hitam (M = 67.84, SD = 9.30) dan stimulus kuning (M = 58.56, SD = 7.68). Atau, waktu reaksi partisipan (t (49) = 5.55, $p < .001$, d = .78) ketika melihat stimulus hitam (M = 67.84, SD = 9.30) lebih lama dibandingkan ketika melihat stimulus kuning (M = 58.56, SD = 7.68).

Prinsip yang sama juga dapat diterapkan untuk interpretasi dan melaporkan hasil uji beda non parametrik, dengan melihat hasil uji *Wilcoxon signed-rank*, seperti pada contoh di Tabel 19.4.

Hasil uji *Wilcoxon signed-rank* menunjukkan bahwa partisipan menunjukkan perbedaan waktu reaksi yang signifikan antara stimulus berwarna hitam dan kuning ($z= 4.45$, $p < .001$). Output JASP uji *Wilcoxon signed-rank* tidak menampilkan secara otomatis Median (Mdn) masing-masing kelompok, sehingga jika diperlukan Median masing-masing kelompok, dapat dihitung dengan menggunakan menu Descriptive pada *software* JASP.

Tabel 19.4: Contoh hasil analisis *Wilcoxon signed-rank*

Pen-gukuran 1	Pen-gukuran 2	<i>W</i>	<i>z</i>	<i>df</i>	<i>p</i>	Rank-Biserial <i>Correla-tion</i>	SE Rank-Biserial <i>Correla-tion</i>
Hitam	Kuning	1021.5	4.446		< .001	0.737	0.164

Soal Latihan

Penilaian mengenai hasil belajar merupakan bagian penting dalam sebuah proses pembelajaran. Ujilah pengetahuan dan pemahaman Anda mengenai materi Bagian 5 ini dengan menjawab sejumlah pertanyaan pada link di bawah ini:



*Scan QR untuk membuka versi interaktif: [https://www.yarsi.ac.id/fakultas-psikologi/
bukustatistik/latihan5/](https://www.yarsi.ac.id/fakultas-psikologi/bukustatistik/latihan5/)*

Bagian VI

MEMBANDINGKAN RATA-RATA TIGA ATAU LEBIH KELOMPOK

Pada bagian sebelumnya, kita telah mempelajari bagaimana membandingkan rata-rata dua kelompok menggunakan uji-t dan uji Mann-Whitney. Namun, dalam banyak situasi penelitian, kita dihadapkan pada lebih dari dua kelompok yang perlu dibandingkan sekaligus. Misalnya, seorang peneliti ingin mengetahui apakah tingkat stres berbeda antara mahasiswa tahun pertama, kedua, dan ketiga. Untuk keperluan ini, menggunakan serangkaian uji-t antar pasangan kelompok bukanlah pilihan yang tepat, karena meningkatkan risiko kesalahan Tipe I. Di sinilah Analisis Varians (*ANOVA*) menjadi penting. *ANOVA* memungkinkan kita membandingkan rata-rata tiga kelompok atau lebih secara serempak dalam satu kerangka uji statistik, sambil menjaga tingkat kesalahan tetap terkendali. Pendekatan ini jauh lebih efisien dan akurat dibandingkan mengandalkan uji-t berulang-ulang.

i Tujuan Pembelajaran

Bab 20. Tujuan Perbandingan Tiga Kelompok atau Lebih:

- Memahami alasan mengapa perbandingan rata-rata tiga kelompok atau lebih diperlukan dalam penelitian.

Bab 21. Prinsip-prinsip Dasar dan Asumsi-Asumsi ANOVA:

- Menjelaskan prinsip kerja *ANOVA* dalam membandingkan variasi antar dan dalam kelompok.
- Memahami dan mengenali asumsi-asumsi dasar *ANOVA* yang harus dipenuhi untuk analisis yang valid.
- Memahami perbedaan antara *planned contrasts* dan uji *post-hoc* dalam analisis data.

Bab 22. Jenis-Jenis ANOVA:

- Mengenali berbagai jenis *ANOVA* (*one-way*, *two-way*, *repeated measures*, dan *mixed-design*) serta aplikasinya.

Bab 23. Interpretasi Hasil Analisis:

- Membaca, menginterpretasikan, dan menuliskan hasil analisis *ANOVA* sesuai kaidah akademik.

Bab 24. Alternatif Non-Parametrik untuk ANOVA:

- Menginterpretasikan dan menuliskan hasil uji non-parametrik sebagai alternatif *ANOVA*.

Bab 20

Tujuan Perbandingan Tiga Kelompok atau Lebih

Dalam penelitian di bidang sosial dan perilaku, kita sering kali menghadapi kebutuhan untuk mengetahui apakah rata-rata suatu variabel berbeda di antara beberapa kelompok. Misalnya, seorang peneliti mungkin ingin mengetahui apakah tingkat stres mahasiswa berbeda antara tahun pertama, kedua, dan ketiga. Dalam situasi seperti ini, membandingkan rata-rata menggunakan uji-t berulang-ulang tidaklah efisien, karena dapat meningkatkan risiko kesalahan Tipe I. Oleh karena itu, **analisis varians (ANOVA)** digunakan sebagai pendekatan yang lebih tepat.

Tujuan utama dari membandingkan tiga kelompok atau lebih adalah untuk menguji apakah terdapat perbedaan rata-rata yang signifikan di antara kelompok-kelompok tersebut, tanpa meningkatkan tingkat kesalahan statistik. Dengan menggunakan **ANOVA**, peneliti dapat **menganalisis variasi yang terjadi antara kelompok serta variasi yang terjadi di dalam kelompok**, sehingga memungkinkan pengambilan kesimpulan yang lebih akurat. Teknik ini juga membantu **menghindari kesalahan dalam interpretasi data** yang dapat terjadi bila banyak uji-t dilakukan secara terpisah.

Sebagai contoh, seorang peneliti pendidikan dapat menggunakan **ANOVA** untuk membandingkan nilai rata-rata siswa yang mengikuti tiga metode pengajaran berbeda, yaitu metode konvensional, *blended learning*, dan pembelajaran daring penuh. Demikian pula, seorang psikolog klinis dapat menggunakan **ANOVA** untuk menguji perbedaan tingkat kecemasan antara pasien yang menjalani terapi kognitif, terapi perilaku, dan terapi kombinasi. Dalam kedua contoh tersebut, **ANOVA** memberikan cara yang terstruktur untuk menguji satu pertanyaan penting: apakah rata-rata dari setidaknya satu kelompok berbeda secara signifikan dibandingkan dengan kelompok lain.

Jika hasil **ANOVA** menunjukkan adanya perbedaan yang signifikan, langkah selanjutnya adalah melakukan analisis lanjutan, seperti uji *post-hoc* atau *planned contrasts* (Bab 22.5), untuk mengidentifikasi kelompok mana yang berbeda. Dengan demikian, memahami tujuan perbandingan tiga kelompok atau lebih merupakan landasan penting sebelum mempelajari prinsip-prinsip, jenis-jenis, serta cara penulisan laporan hasil analisis varians.

Bab 21

Prinsip-prinsip Dasar dan Asumsi-asumsi ANOVA

Analisis Varians, atau *ANOVA*, adalah metode statistik yang digunakan untuk menentukan apakah terdapat **perbedaan yang signifikan secara statistik antara rata-rata tiga atau lebih kelompok** penelitian. Teknik ini sangat berguna dalam penelitian eksperimental di mana beberapa kelompok dibandingkan dalam kondisi atau perlakuan yang berbeda. Dengan menganalisis varians dalam dan antar kelompok, *ANOVA* membantu peneliti memahami bagaimana faktor-faktor berbeda memengaruhi hasil tertentu, memberikan informasi penting tentang hubungan antar variabel.

ANOVA didasarkan pada prinsip **membandingkan variasi yang terjadi antar kelompok dengan variasi yang terjadi di dalam kelompok**. Ketika perbedaan antar kelompok lebih besar dibandingkan variasi di dalam kelompok, hal ini mengindikasikan bahwa ada kemungkinan rata-rata kelompok memang berbeda secara nyata. Dengan kata lain, *ANOVA* menilai apakah variabilitas yang diamati dalam data lebih banyak dijelaskan oleh perbedaan antar kelompok (*true variance*) daripada oleh variasi acak di dalam kelompok (*error variance*).

21.1 Statistik F: Rasio Variabilitas

Ide utama *ANOVA* adalah untuk membandingkan kedua jenis variabilitas ini melalui rasio yang disebut **Statistik F (F statistics)**, yang dapat digambarkan dalam persamaan berikut:

$$F = \frac{\text{varians antar kelompok}}{\text{varians dalam kelompok}}$$

Jika **nilai F besar**, ini berarti perbedaan antar kelompok lebih besar dibandingkan variasi dalam kelompok. Artinya, perbedaan rata-rata kelompok bukan kebetulan dan kemungkinan disebabkan oleh variabel independen. Dengan kata lain, efek perlakuan lebih kuat daripada variasi acak. Jika nilai F melebihi nilai kritis dari tabel distribusi F, hipotesis nol dapat ditolak.

Jika **statistik F kecil** (mendekati 1 atau kurang), ini berarti variabilitas antar kelompok serupa dengan variabilitas dalam kelompok. Ini menunjukkan bahwa perbedaan antara rata-rata kelompok mungkin hanya disebabkan oleh kebetulan atau variasi acak, dan variabel independen mungkin tidak memiliki efek yang signifikan.

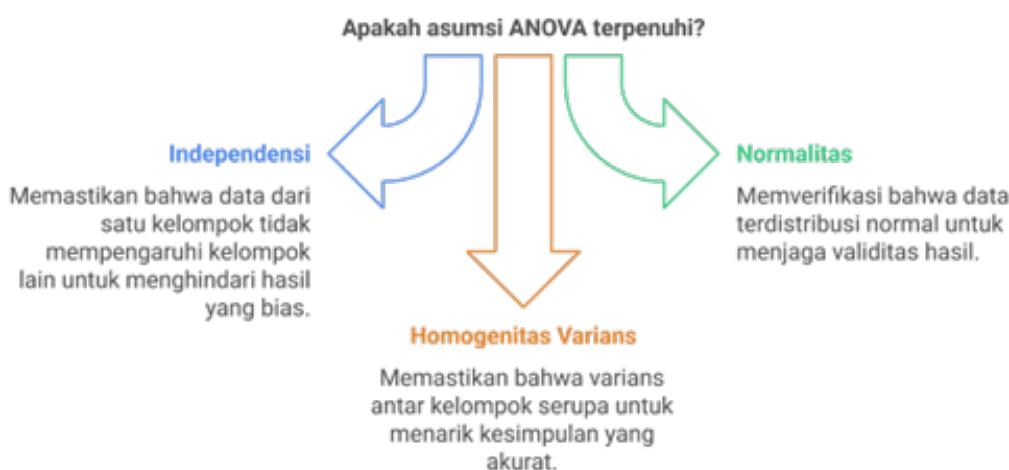
21.2 Asumsi-asumsi Dasar ANOVA

Untuk memastikan hasil *ANOVA* valid dan dapat diandalkan, terdapat beberapa asumsi dasar yang perlu dipenuhi (Gambar 21.1). Pertama, data dari setiap kelompok yang dibandingkan harus saling **independen**, artinya nilai-nilai dari satu kelompok tidak boleh memengaruhi nilai-nilai dari kelompok lain. Independensi ini biasanya dijaga melalui prosedur pengambilan sampel yang tepat atau penugasan acak dalam eksperimen.

Kedua, *ANOVA* mengasumsikan bahwa data dalam setiap kelompok mengikuti **distribusi normal**. Meskipun *ANOVA* cukup *robust* terhadap pelanggaran normalitas pada ukuran sampel besar, penyimpangan yang berat dari normalitas dapat mengurangi keakuratan hasil. Oleh karena itu, sebelum melakukan analisis, sebaiknya dilakukan pemeriksaan normalitas melalui visualisasi data, seperti histogram atau plot Q-Q, ataupun menggunakan uji statistik seperti *Shapiro-Wilk*.

Ketiga, variansi antar kelompok yang dibandingkan diharapkan **homogen** atau seragam. Homogenitas varians ini penting karena *ANOVA* membandingkan rata-rata dengan memperhitungkan variabilitas dalam kelompok. Apabila variansi antar kelompok sangat berbeda, maka kesimpulan yang dihasilkan dari *ANOVA* bisa menjadi bias. Untuk menguji asumsi ini, dapat digunakan uji *Levene* atau uji *Bartlett*. Uji *Levene* sering digunakan untuk menilai homogenitas ini.

Apabila asumsi-asumsi tersebut tidak sepenuhnya terpenuhi, ada beberapa strategi yang bisa ditempuh. Transformasi data, penggunaan *ANOVA* varian *robust* seperti *ANOVA Welch*, atau beralih ke pendekatan non-parametrik menjadi pilihan *untuk* menjaga validitas hasil analisis. Penting untuk diingat bahwa pengecekan asumsi bukan sekadar prosedur teknis, melainkan bagian integral dari proses analisis yang memastikan bahwa interpretasi hasil *ANOVA* merefleksikan kondisi data dengan akurat.



Gambar 21.1: Asumsi-asumsi *ANOVA*

Dengan memahami prinsip-prinsip dasar ini dan memastikan terpenuhinya asumsi-asumsi yang mendasarinya serta peran *plannedcontrasts* dan uji *post-hoc*, kita dapat menggunakan *ANOVA* secara efektif untuk menganalisis perbedaan antar kelompok dalam berbagai konteks penelitian. Prinsip-prinsip ini menjadi fondasi penting sebelum

melangkah lebih jauh ke dalam ragam jenis *ANOVA* yang lebih kompleks.

Bab 22

Jenis-jenis ANOVA dan Interpretasi Hasil Analisis

Jenis-jenis **ANOVA** mengacu pada model statistik berbeda yang digunakan untuk menganalisis varians di antara rata-rata kelompok dalam berbagai desain eksperimental. Jenis yang paling umum digunakan adalah **ANOVA Satu Arah** (*One-Way ANOVA*) dan **ANOVA Dua Arah** (*Two-Way ANOVA*). Setiap jenis memiliki tujuan berbeda dan sesuai untuk skenario penelitian tertentu. Memahami jenis-jenis ini sangat penting untuk memilih metode analisis yang tepat berdasarkan struktur data dan desain eksperimen Anda.

Pemilihan jenis **ANOVA** yang akan digunakan bergantung pada desain penelitian dan sifat data yang dianalisis. Pemahaman menyeluruh tentang jenis-jenis yang berbeda memungkinkan peneliti untuk membuat keputusan yang terinformasi, memastikan bahwa analisis statistik secara akurat mencerminkan hubungan dan perbedaan yang ada dalam data. Landasan ini penting tidak hanya untuk melakukan analisis statistik yang efektif tetapi juga untuk menafsirkan hasil dengan cara yang bermakna yang berkontribusi pada bidang studi yang lebih luas.

22.1 ANOVA Satu Arah

ANOVA Satu Arah digunakan ketika peneliti ingin membandingkan rata-rata tiga atau lebih kelompok independen berdasarkan satu variabel independen. Misalnya, jika seorang peneliti sedang memeriksa dampak metode pengajaran yang berbeda terhadap kinerja siswa, mereka dapat mengelompokkan siswa ke dalam tiga kelompok berdasarkan metode yang digunakan. **ANOVA** Satu Arah membantu menentukan apakah ada perbedaan yang signifikan secara statistik dalam skor kinerja rata-rata di antara kelompok-kelompok ini. Jenis **ANOVA** ini menjawab pertanyaan mendasar tentang apakah setidaknya satu rata-rata kelompok berbeda dari yang lain tanpa mendalamai interaksi antara beberapa variabel.

Menghitung hasil dalam ANOVA satu arah melibatkan beberapa langkah kunci yang memungkinkan peneliti untuk menarik kesimpulan bermakna dari data yang diperoleh. Tujuan utama **ANOVA** satu arah adalah untuk menentukan apakah ada perbedaan yang signifikan secara statistik di antara rata-rata tiga atau lebih kelompok independen. Proses ini dimulai dengan perumusan hipotesis nol, yang menyatakan bahwa semua rata-rata kelompok adalah sama, dan hipotesis alternatif, yang menunjukkan bahwa setidaknya satu rata-rata kelompok berbeda. Dengan menggunakan perangkat lunak statistik atau perhitungan manual, Anda dapat memperoleh statistik F, yang membandingkan varians antar kelompok dengan varians dalam kelompok. Tabel 22.1 memuat contoh hasil analisis **ANOVA** satu-arah dengan menggunakan perangkat lunak JASP.

Pada *ANOVA* satu arah, **interpretasi** dimulai dengan menilai nilai F dan *p-value*. Jika *p-value* lebih kecil dari batas signifikansi, misalnya 0,05, maka dapat disimpulkan bahwa ada perbedaan signifikan di antara kelompok yang dibandingkan. Interpretasi *ANOVA* satu arah hanya berfokus pada satu variabel independen sehingga lebih sederhana, namun penting untuk memperhatikan ukuran efek seperti *eta-squared* untuk memahami seberapa besar pengaruh faktor tersebut terhadap variabel dependen. Dalam pelaporan, hasil *ANOVA* satu arah disusun dengan mencantumkan nilai F, derajat kebebasan antar dan dalam kelompok, nilai *p*, dan ukuran efek. Misalnya: "Analisis varians satu arah menunjukkan perbedaan signifikan antara kelompok, $F(2, 87) = 4,76, p = 0,011, \eta^2 = 0,098$."

Tabel 22.1: Contoh tabel *output* hasil *ANOVA* satu-arah

	<i>Sum of squares</i>	<i>df1</i>	<i>Mean square</i>	F	<i>p</i>	η^2
Metode terapi	24.420	2	8.140	4.761	.011	.098
Residuals	613.139	87	2.813			

22.2 ANOVA Dua Arah

ANOVA Dua Arah memungkinkan pengujian dua variabel independen secara bersamaan dan efek interaksi kedua variabel tersebut terhadap variabel dependen. Misalnya, sebuah penelitian bertujuan mengeksplorasi bagaimana metode pengajaran dan jenis kelamin siswa memengaruhi kinerja. Dengan menggunakan *ANOVA* dua arah, peneliti dapat menilai tidak hanya efek utama dari setiap variabel independen tetapi juga apakah efek dari satu variabel berbeda tergantung pada tingkat variabel lainnya. Jenis ini sangat berguna dalam desain eksperimental yang lebih kompleks di mana interaksi antar faktor diharapkan memainkan peran penting.

Dalam *ANOVA* Dua Arah, **interpretasi** mencakup dua komponen utama: efek utama dari masing-masing variabel independen dan efek interaksi antara keduanya. Peneliti perlu menilai apakah setiap efek utama dan interaksi menunjukkan signifikansi. Memahami **efek interaksi** sangat penting untuk menginterpretasikan hasil *ANOVA* dua arah. Ketika peneliti menyelidiki pengaruh dua variabel independen terhadap variabel dependen, mereka tidak hanya tertarik pada efek utama dari masing-masing variabel independen tetapi juga bagaimana variabel-variabel ini dapat berinteraksi satu sama lain. Efek interaksi terjadi ketika pengaruh satu variabel independen terhadap variabel dependen berubah tergantung pada level variabel independen lainnya. Ini berarti bahwa pengaruh gabungan dari variabel-variabel tersebut tidak hanya bersifat aditif; sebaliknya, hubungannya lebih kompleks dan memerlukan pemeriksaan yang cermat.

! ANOVA desain faktorial

Penelitian yang melibatkan lebih dari satu variabel independen (faktor) untuk melihat pengaruhnya terhadap variabel dependen disebut juga sebagai **ANOVA desain faktorial**. Umumnya, desain ini dituliskan dalam bentuk "faktorial 2x2" (atau 3x2, 2x4, dan seterusnya). Notasi tersebut menunjukkan jumlah level dari masing-masing faktor; misalnya, desain 2x2 berarti ada dua faktor, dan masing-

masing faktor memiliki dua level.

Untuk mengilustrasikan efek interaksi, perhatikan sebuah penelitian yang menguji pengaruh metode pengajaran dan jenis kelamin siswa terhadap kinerja akademik. Jika kinerja siswa laki-laki dan perempuan berbeda secara signifikan berdasarkan metode pengajaran yang digunakan, ini menunjukkan efek interaksi (lihat Tabel 22.2). Misalkan, metode pengajaran tertentu mungkin lebih efektif untuk siswa perempuan se-mentara metode lain mungkin menguntungkan siswa laki-laki. Memahami interaksi ini memungkinkan peneliti untuk menyesuaikan strategi pendidikan untuk meningkatkan hasil bagi kelompok siswa yang berbeda, menekankan pentingnya mempertimbangkan bagaimana variabel-variabel bekerja bersama daripada secara terpisah.

Penulisan hasil mengikuti urutan pelaporan efek utama terlebih dahulu, baru kemudian efek interaksi. Sebagai contoh: "Terdapat efek utama yang signifikan dari metode pengajaran terhadap prestasi siswa, $F(2, 84) = 5,23, p = 0,007, \eta^2 = 0,111$, namun tidak ditemukan interaksi yang signifikan antara metode pengajaran dan jenis kelamin, $F(2, 84) = 1,04, p = 0,358$."

Tabel 22.2: Contoh tabel *output* hasil *ANOVA* dua-arah

	Sum of Squares	df	Mean Square	F	p	η^2
Metode	11.341	2	4.341	5.232	.007	.111
Jenis kelamin	2.234	1	0.653	2.331	.105	.012
Metode x Jenis kelamin	0.351	2	0.121	1.042	.358	.008
Residuals	6.748	84	2.548			

22.3 ANOVA Pengukuran Berulang

Selain Satu Arah dan Dua Arah, ada bentuk khusus *ANOVA* seperti ***ANOVA Pengukuran Berulang (Repeated Measures ANOVA)***, yang digunakan ketika partisipan yang sama diukur beberapa kali dalam kondisi berbeda. Metode ini dapat terjadi dalam studi longitudinal atau eksperimen di mana individu terpapar pada semua tingkat variabel independen.

Pada *ANOVA* Pengukuran Berulang, **interpretasi** memperhatikan signifikansi perubahan antar kondisi atau waktu. Karena pengukuran dilakukan berulang pada individu yang sama, penyesuaian untuk *sphericity* sering diperlukan, misalnya menggunakan koreksi *Greenhouse-Geisser*.

Sphericity mengacu pada kondisi di mana varians perbedaan antara setiap pasangan tingkat pengukuran berulang adalah sama. Dengan kata lain, kovarians antara setiap pasangan pengukuran berulang harus sama. Pelanggaran *sphericity* dapat menyebabkan peningkatan risiko kesalahan Tipe I (menolak H_0 padahal benar) dan Tipe II (gagal menolak H_0 padahal salah), sehingga hasil F-statistik menjadi tidak akurat.

Tabel 22.3 menunjukkan contoh hasil *ANOVA* pengukuran berulang mengenai efektivitas terapi meditasi dalam mengurangi kecemasan. Penulisan hasil mencantumkan statistik *F*,

derajat kebebasan yang telah disesuaikan (jika perlu), nilai p , dan ukuran efek, misalnya: "Analisis pengukuran berulang menunjukkan perubahan signifikan tingkat kecemasan antar sesi terapi, $F(2, 29) = 8,57, p = 0,006, \eta^2_{partial} = 0,228$."

Tabel 22.3: Contoh tabel *output* hasil ANOVA pengukuran berulang

	<i>Sum of squares</i>	<i>df1</i>	<i>Mean square</i>	F	<i>p</i>	η^2_p
Kece- masan	19.443	2	5.341	8.567	.006	.228
Residuals	8.458	29	2.547			

22.4 ANOVA Desain Campuran

Variasi lainnya adalah **ANOVA Desain Campuran** (*Mixed-Design ANOVA*), yang menggabungkan faktor independen dan pengukuran berulang, memungkinkan analisis data yang lebih terperinci. Memahami variasi-variasi ini memungkinkan peneliti untuk memilih metode statistik yang paling efektif untuk pertanyaan penelitian mereka yang spesifik.

Interpretasi hasil analisis ANOVA Desain Campuran lebih kompleks karena mencakup efek utama untuk faktor *between-subjects*, faktor *within-subjects*, serta interaksi keduanya. Peneliti perlu mengevaluasi apakah ada perbedaan antar kelompok, perubahan dalam kelompok dari waktu ke waktu, dan apakah pola perubahan berbeda antar kelompok. Contoh hasil analisisnya dapat dilihat di Tabel 22.4. Penulisan hasil memuat masing-masing efek secara sistematis. Misalnya: "Terdapat efek utama yang signifikan dari kelompok perlakuan terhadap tingkat stres, $F(1, 58) = 10,12, p = 0,002$, serta efek waktu yang signifikan, $F(2, 116) = 6,87, p = 0,002$. Interaksi antara kelompok dan waktu juga signifikan, $F(2, 116) = 4,21, p = 0,017$, menunjukkan bahwa perubahan stres antar waktu berbeda berdasarkan kelompok perlakuan."

Tabel 22.4: Contoh tabel output hasil ANOVA campuran

<i>Cases</i>	<i>Sum of squares</i>	<i>df</i>	<i>Mean square</i>	F	<i>p</i>	η^2_p
Waktu	28.212	1	7.311	6.871	.002	.311
Waktu x Treatment	23.553	2	6.431	4.212	.017	.124
Residuals	29.221	116	7.531			

Between-subjects effects

	<i>Sum of squares</i>	<i>df</i>	<i>Mean square</i>	F	<i>p</i>	η^2
Treatment	45.112	1	12.477	10.124	.002	.398
Residuals	60.211	58	13.415			

22.5 Uji Post-hoc vs. Planned Contrasts

Setelah uji *ANOVA* menunjukkan adanya perbedaan rata-rata yang signifikan, peneliti perlu mengidentifikasi lebih jauh kelompok mana yang berbeda. Di sinilah *planned contrasts* dan uji *post-hoc* berperan. *Planned contrasts* digunakan ketika peneliti sejak awal telah menetapkan hipotesis spesifik mengenai kelompok mana yang diharapkan berbeda. Misalnya, dalam studi tentang efektivitas tiga metode pengajaran, seorang peneliti mungkin sejak awal ingin menguji apakah metode A lebih efektif dibandingkan metode B dan C secara gabungan.

Sementara itu, uji *post-hoc* digunakan ketika peneliti tidak memiliki prediksi spesifik tentang kelompok mana yang berbeda sebelum analisis dilakukan. Setelah menemukan hasil *ANOVA* yang signifikan, uji *post-hoc* dijalankan untuk mengeksplorasi semua kemungkinan perbedaan antar pasangan kelompok.

Misalnya, kita melakukan uji *post hoc* dari data Tabel 22.1. Hasil *ANOVA* menunjukkan adanya perbedaan yang signifikan antar kelompok metode terapi, $F(2, 87) = 4,76, p = 0,011$. Namun, dari hasil tersebut kita belum bisa mengidentifikasi antar kelompok mana yang berbeda. Uji *post hoc* bisa membantu kita menyimpulkan perbedaan dari masing-masing kelompok yang diteliti. Dari hasil yang dipaparkan pada Tabel 22.6, data menunjukkan bahwa terdapat perbedaan yang signifikan antara kelompok CBT dan Journaling ($t(87) = 4.360, P_{tukey} = .002$). Begitu juga dengan perbedaan antara Journaling dan Placebo ($t(87) = 5.876, P_{tukey} < .001$). Namun, tidak ditemukan adanya perbedaan antara kelompok CBT dan Placebo ($t(87) = 1.516, P_{tukey} = .012$). Dari hasil ini kita dapat menyimpulkan bahwa metode Journaling adalah metode yang paling efektif dalam mereduksi stres pada partisipan, sedangkan metode CBT dan Placebo menunjukkan efek yang tidak berbeda.

Post hoc comparisons: Metode terapi

Tabel 22.6: Contoh hasil uji *post hoc*

		Mean difference	SE	df	t	P _{tukey}
CBT	Journaling	-0.766	0.176	87	-4.360	.002
	Placebo	0.267	0.176	87	1.516	.012
	Journaling	1.033	0.176	87	5.876	<.001

Beberapa uji *post-hoc* yang umum meliputi:

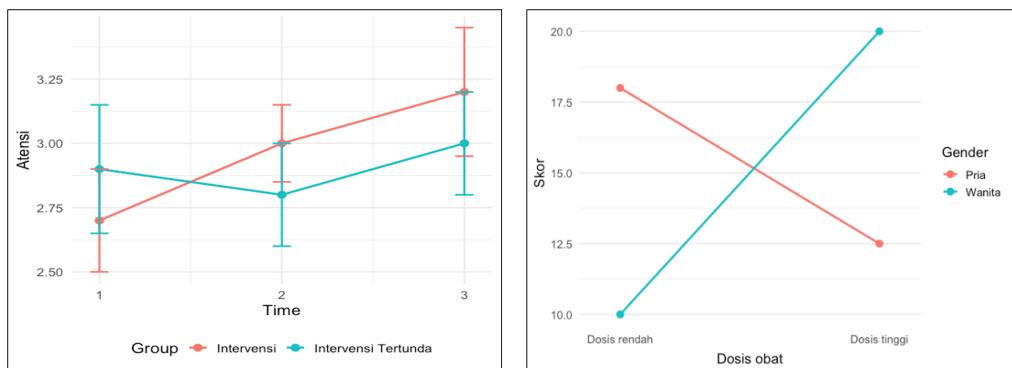
- **Uji Tukey HSD:** Uji ini digunakan ketika varians kelompok sama dan ukuran sampel sama.
- **Uji Bonferroni:** Uji ini digunakan ketika varians kelompok tidak sama atau ukuran sampel tidak sama.
- **Uji Scheffe:** Uji ini adalah uji yang paling konservatif dan digunakan ketika ada banyak perbandingan yang dibuat.

Memahami perbedaan antara *planned contrasts* dan uji *post-hoc* penting dalam desain analisis, karena keduanya menjadi metode yang menjawab kebutuhan penelitian yang berbeda. *Planned contrasts* lebih terfokus dan lebih kuat secara statistik untuk hipotesis

yang sudah direncanakan, sedangkan uji *post-hoc* memberikan fleksibilitas eksploratif dalam menganalisis data setelah hasil utama ditemukan.

22.6 Effect Size dan Visualisasi Hasil Analisis

Dalam semua jenis *ANOVA*, ukuran efek seperti *eta-squared* atau *partial eta-squared* sangat dianjurkan untuk dilaporkan karena memberikan informasi tambahan tentang besarnya pengaruh yang ditemukan. Selain itu, visualisasi hasil menggunakan diagram batang atau plot garis dapat memperjelas interpretasi perbedaan antar kelompok atau perubahan dalam waktu, seperti yang ditampilkan dalam Gambar 22.1.



Gambar 22.1: Contoh penggunaan visualisasi hasil analisis

Bab 23

Alternatif Non-Parametrik untuk ANOVA

Ketika data tidak memenuhi asumsi-asumsi parametrik dalam pengujian *ANOVA*, atau memiliki skala pengukuran yang berbeda, pendekatan non-parametrik menjadi alternatif yang tepat. Uji non-parametrik tidak memerlukan asumsi distribusi data yang ketat dan cocok untuk data ordinal atau nominal. Oleh karena itu, teknik ini memungkinkan analisis yang lebih fleksibel tanpa mengorbankan validitas hasil.

23.1 Uji Kruskal-Wallis: Alternatif Non-Parametrik untuk ANOVA Satu Arah

Salah satu uji non-parametrik yang paling umum digunakan untuk menggantikan *ANOVA* Satu Arah adalah **Uji Kruskal-Wallis**. Uji ini digunakan untuk menguji apakah ada perbedaan yang signifikan dalam median antar tiga kelompok atau lebih. Berbeda dengan *ANOVA* yang membandingkan rata-rata, uji *Kruskal-Wallis* membandingkan peringkat data dalam setiap kelompok.

Kapan menggunakan *Kruskal-Wallis*?

- Data variabel dependen tidak terdistribusi normal di setiap kelompok.
- Ukuran sampel kecil dan data menunjukkan penyimpangan yang jelas dari normalitas.
- Variabel dependen adalah ordinal (misalnya, peringkat, skala *Likert* yang dianggap ordinal).

Prosedur analisis:

1. Data diurutkan dan diberi peringkat (*ranking*).
2. Uji ini menguji apakah distribusi peringkat antar kelompok berbeda secara signifikan.
3. Jika *p-value* dari uji *Kruskal-Wallis* lebih kecil dari tingkat signifikansi (biasanya 0,05), kita menolak hipotesis nol, yang berarti ada perbedaan yang signifikan antar kelompok.
4. Jika uji *Kruskal-Wallis* signifikan, Anda perlu melakukan perbandingan berpasangan (*post-hoc*) untuk mengidentifikasi kelompok mana yang berbeda. Namun, tidak ada uji *post-hoc* standar tunggal seperti *Tukey* untuk *Kruskal-Wallis*. Umumnya, Anda dapat melakukan **uji Mann-Whitney U** secara berpasangan untuk setiap

kombinasi kelompok, dan menerapkan koreksi Bonferroni untuk mengontrol kesalahan Tipe I.

Dalam penulisan laporan, hasil uji non-parametrik harus dilaporkan dengan cara yang jelas dan sistematis. Contoh penulisan laporan untuk uji *Kruskal-Wallis*:

"Uji *Kruskal-Wallis* menunjukkan bahwa terdapat perbedaan yang signifikan dalam peringkat antara kelompok, $H(2) = 6,32, p = 0,042$. *Post-hoc* dengan uji *Mann-Whitney* menunjukkan bahwa kelompok A memiliki peringkat yang lebih tinggi dibandingkan kelompok B ($p = 0,012$)."

23.2 Uji *Friedman*: Alternatif Non-Parametrik untuk ANOVA Pengukuran Berulang

Untuk desain eksperimen dengan pengukuran berulang yang tidak memenuhi asumsi normalitas, **uji Friedman** dapat digunakan sebagai alternatif non-parametrik untuk *Repeated Measures ANOVA*. Uji *Friedman* membandingkan peringkat data pada lebih dari dua kondisi pengukuran yang dilakukan pada subjek yang sama.

Uji *Friedman* adalah perpanjangan non-parametrik dari uji *Wilcoxon Signed-Rank* untuk lebih dari dua pengukuran. Prosedur analisisnya meliputi:

- Memberikan peringkat pada observasi dalam setiap subjek di antara kondisi/titik waktu.
- Menjumlahkan peringkat untuk setiap kondisi/titik waktu.
- Menghitung statistik *Friedman* (χ^2), yang membandingkan jumlah peringkat yang diamati dengan jumlah peringkat yang diharapkan jika tidak ada perbedaan antar kondisi.

Jika uji *Friedman* signifikan, Anda perlu melakukan uji *post hoc* untuk mengidentifikasi pasangan kondisi/titik waktu mana yang berbeda secara signifikan. Anda dapat menggunakan **uji Wilcoxon Signed-Rank** berpasangan untuk setiap kombinasi dan menerapkan koreksi *Bonferroni* untuk mengontrol kesalahan Tipe I.

Contoh penulisan hasil (non-parametrik):

"Karena asumsi normalitas dilanggar, uji *Friedman* dilakukan untuk membandingkan peringkat kinerja tugas siswa pada tiga sesi pelatihan yang berbeda (Sesi 1, Sesi 2, Sesi 3). Ditemukan perbedaan yang signifikan secara statistik di antara sesi-sesi pelatihan, $\chi^2(2) = 9,87, p = .007$. Uji *post-hoc* *Wilcoxon Signed-Rank* dengan koreksi *Bonferroni* menunjukkan bahwa kinerja meningkat secara signifikan dari Sesi 1 ke Sesi 2 ($p = .008$) dan dari Sesi 2 ke Sesi 3 ($p = .015$)."

Soal Latihan

Penilaian mengenai hasil belajar merupakan bagian penting dalam sebuah proses pembelajaran. Ujilah pengetahuan dan pemahaman Anda mengenai materi Bagian 6 ini dengan menjawab sejumlah pertanyaan pada link di bawah ini:



Scan QR untuk membuka versi interaktif: [https://www.yarsi.ac.id/fakultas-psikologi/
bukustatistik/latihan6/](https://www.yarsi.ac.id/fakultas-psikologi/bukustatistik/latihan6/)

Bagian VII

ANALISIS KORELASI

Pada bagian ini akan dibahas mengenai hubungan atau korelasi antara dua variabel penelitian, yaitu melihat apakah dua variabel mengalami perubahan secara bersama-sama. Untuk melihat hubungan ini, setiap individu yang menjadi sampel atau yang diukur, harus memiliki skor kedua variabel (Variabel X dan Y). Uji korelasi umumnya digunakan untuk uji validitas alat ukur, reliabilitas aitem dalam sebuah alat ukur, dan untuk verifikasi atau menguji keterhubungan variabel-variabel dalam teori yang dikembangkan.

i Tujuan Pembelajaran

Bab 25. Karakteristik Hubungan Antara Dua Variabel:

- Memahami konsep-konsep dasar analisis korelasi.
- Memahami arah, bentuk, dan besaran hubungan antar variabel.

Bab 26. Asumsi-asumsi Korelasi, Interpretasi, dan Penulisan Hasil Analisis:

- Memahami asumsi yang perlu dipenuhi dalam hubungan antara dua variabel
- Mampu menginterpretasikan hasil analisis dan menulis laporan hasil korelasi sesuai standar akademik.

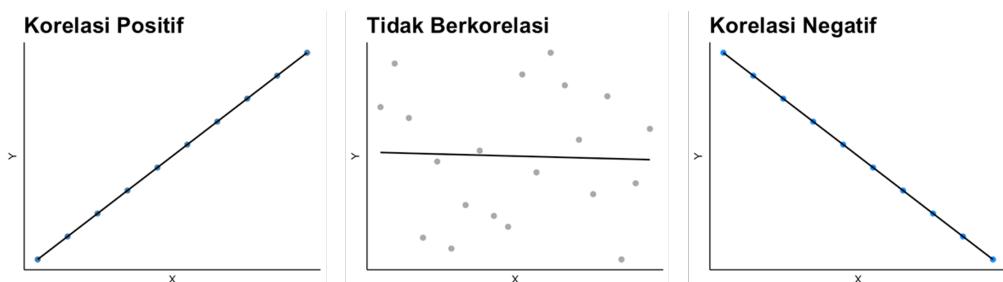
Bab 24

Karakteristik Hubungan Antara Dua Variabel

Hubungan antara dua variabel dalam statistik dapat dilihat dari uji korelasi, yaitu suatu teknik yang digunakan untuk mengukur dan menggambarkan sebuah keterkaitan antara dua variabel. Dalam uji korelasi, terdapat tiga karakteristik hubungan antara variabel X dan Y yaitu arah hubungan (positif atau negatif), bentuk hubungan (linear atau non linear) dan kekuatan hubungan.

24.1 Arah Hubungan Korelasi

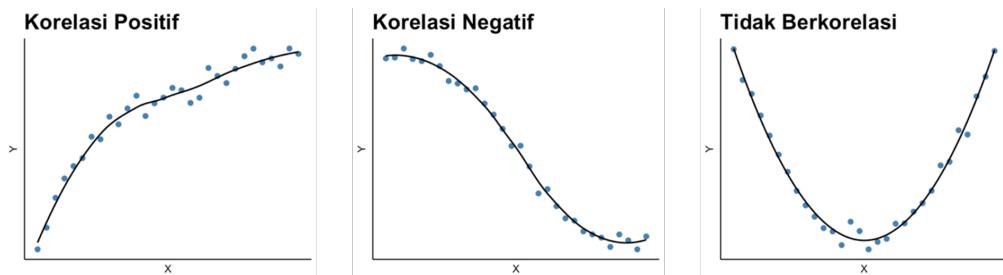
Berdasarkan arah hubungan, korelasi dapat dibedakan menjadi korelasi positif dan negatif. **Korelasi positif (+)** adalah ketika kedua variabel berubah dengan arah yang sama (X dan Y sama-sama naik/ meningkat atau X dan Y sama-sama turun). Dengan kata lain, jika variabel X mengalami peningkatan skor, maka variabel Y juga mengalami peningkatan skor. Atau, jika variabel X mengalami penurunan skor, maka variabel Y juga menurun. Misalnya, jika suhu di luar panas (terik), penjualan es teh manis meningkat; sedangkan jika cuaca mendung (suhu lebih dingin), penjualan es teh manis menurun. Hal ini menunjukkan bahwa suhu berkorelasi positif dengan penjualan es teh manis. Sementara itu, **korelasi negatif (-)** terjadi ketika kedua variabel berubah dengan arah yang berlawanan (X naik maka Y turun, atau X turun maka Y naik). Peningkatan skor pada variabel X diikuti dengan penurunan skor pada variabel Y. Misalnya, seorang dengan tingkat stres yang tinggi cenderung memiliki tingkat kebahagiaan yang rendah; stres berkorelasi negatif dengan kebahagiaan. Arah hubungan dapat divisualisasikan dalam *scatter plot* (lihat Gambar 24.1).



Gambar 24.1: Visualisasi arah korelasi

24.2 Bentuk Hubungan Korelasi

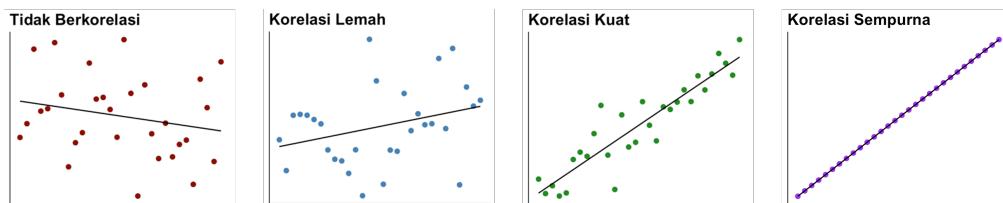
Bentuk hubungan antara dua variabel dibedakan menjadi hubungan linear dan non linear. Bentuk hubungan ini akan terlihat lebih jelas ketika data kedua variabel divisualisasikan dalam sebuah plot. **Hubungan linear** adalah ketika titik-titik dalam *scatter plot* cenderung membentuk garis lurus. Sedangkan, **hubungan non linear** terjadi jika titik-titik dalam *scatter plot* cenderung membentuk garis lengkung. Visualisasi hubungan linear dapat dilihat pada Gambar 7.1., sedangkan hubungan non linear dapat dilihat pada Gambar 24.2.



Gambar 24.2: Visualisasi korelasi non-linear

24.3 Besaran Hubungan Korelasi

Besarnya kekuatan korelasi ditentukan oleh angka hasil pengujian korelasi atau disebut dengan koefisien korelasi. Koefisien korelasi mencerminkan tingkat konsistensi hubungan antara dua variabel. Koefisien korelasi berada pada rentang 0.00 (tidak berkorelasi) dan 1.00 (korelasi sempurna), semakin mendekati 1.00 menunjukkan derajat hubungan yang semakin kuat. Koefisien korelasi .72 memiliki hubungan yang lebih kuat dibandingkan .55. Ketika divisualisasikan dalam *scatter plot*, kuat atau lemahnya korelasi dapat dilihat dari kepadatan sebaran data di sekitar garis korelasi. Visualisasi dapat dilihat pada Gambar 24.3, sedangkan kekuatan korelasi dapat diinterpretasikan dengan mengacu pada Tabel 24.1.



Gambar 24.3: Visualisasi kekuatan korelasi

Tabel 24.1: Kekuatan koefisien korelasi

Kekuatan korelasi	Koefisien korelasi Positif	Koefisien korelasi Negatif
Lemah	0 – .30	0 – -.30

Kekuatan korelasi	Koefisien korelasi Positif	Koefisien korelasi Negatif
Sedang	.31 – .50	-.31 – -.50
Kuat	.51 – .70	-.51 – -.70
Sangat kuat	.71 – 1.00	-.71 – -1.00

Bab 25

Asumsi Korelasi, Interpretasi, dan Penulisan Hasil Analisis

Tujuan analisis korelasi utamanya adalah untuk mengukur kekuatan dan arah hubungan antara dua variabel. Meskipun metode ini sering digunakan dalam berbagai disiplin ilmu, agar hasilnya valid, ada beberapa **asumsi dasar** yang harus dipenuhi, seperti normalitas distribusi data. Memahami asumsi-asumsi ini sangat penting agar peneliti dapat menarik kesimpulan yang akurat dan tidak bias. Selain itu, interpretasi hasil analisis korelasi perlu dilakukan dengan hati-hati, mengingat korelasi tidak membuktikan sebab-akibat, melainkan hanya mengukur derajat hubungan antar variabel. Sub-bab ini akan membahas **asumsi-asumsi yang perlu dipenuhi** dalam analisis korelasi, serta bagaimana **menginterpretasikan** hasil korelasi dan **menuliskan laporan** hasil analisis dengan cara yang jelas dan sistematis.

25.1 Asumsi dalam Hubungan Antara Dua Variabel

Uji asumsi hubungan antara dua variabel digunakan untuk menentukan teknik statistik yang digunakan dalam menghitung koefisien korelasi, apakah akan menggunakan uji korelasi parametrik (*Pearson correlation; r*) atau non-parametrik (*Spearman correlation; r_s*) (Azwar, 2019; Howell, 2013). Terdapat empat asumsi yang perlu dipenuhi jika ingin menggunakan uji korelasi parametrik (*Pearson correlation*), yaitu:

- a. Kedua variabel yang diteliti atau yang ingin dikorelasikan harus berupa skala kontinum.
- b. Kedua variabel memiliki hubungan yang linear yang dapat dipastikan dengan menggunakan *scatter plot*.
- c. Data tidak memiliki banyak *outliers*.
- d. Data terdistribusi dengan normal.

Jika terdapat asumsi yang tidak terpenuhi, maka dilakukan uji korelasi non-parametrik menggunakan analisis korelasi *Spearman*.

25.2 Interpretasi dan Penulisan Hasil Uji Korelasi

Dalam melakukan interpretasi, hal pertama yang perlu dilihat adalah nilai signifikansi (*p-value*). Uji hubungan antara dua variabel dikatakan signifikan jika nilai *p-value* lebih kecil dari .05 ($p < .05$), artinya terdapat hubungan yang signifikan antara variabel X dan variabel Y. Jika terdapat hubungan, maka interpretasi dilanjutkan dengan melihat arah korelasi (positif atau negatif) dan besaran koefisien korelasi. Jika tidak terdapat

hubungan yang signifikan ($p > .05$), maka interpretasi terkait arah dan derajat korelasi tidak perlu dilanjutkan. Interpretasi hasil uji korelasi dapat lebih dipahami melalui *output* dari analisis JASP (Gambar 25.1).

Spearman's Correlations				
Variable		Well-being	Learning engagement	Neuroticism
1. Well-being	n	—		
	Spearman's rho	—		
	p-value	—		
2. Learning engagement	n	267	—	
	Spearman's rho	0.669***	—	
	p-value	< .001	—	
3. Neuroticism	n	267	267	—
	Spearman's rho	-0.396***	-0.384***	—
	p-value	< .001	< .001	—
4. Opennes to experiences	n	267	267	267
	Spearman's rho	0.119	0.189**	-0.128*
	p-value	0.052	0.002	0.037

* p < .05, ** p < .01, *** p < .001

Gambar 25.1: Output hasil uji korelasi

Dari *output* yang ditampilkan pada Gambar 25.1 diketahui bahwa *well-being* berkorelasi secara signifikan dengan *learning engagement* dan aspek kepribadian *neuroticism* ($p < .001$), namun tidak berkorelasi dengan aspek kepribadian *opennes to experiences*, karena *p value* lebih besar dari .05 ($p = .052$). Interpretasi terkait arah dan besaran korelasi dapat dilanjutkan pada hasil yang signifikan, diketahui terdapat hubungan positif antara *well-being* dan *learning engagement* ($r_s = .669$), sementara itu, hubungan antara *well-being* dan *neuroticism* berkorelasi negatif ($r_s = -.396$). Derajat kekuatan korelasi antara antara *well-being* dan *learning engagement* tergolong kuat (rentang .51 hingga .70), dan korelasi antara *well-being* dan *neuroticism* tergolong sedang/cukup kuat (rentang .31 hingga .50).

Hasil uji korelasi umumnya dilaporkan dalam bentuk tabel, jika peneliti melakukan uji korelasi beberapa variabel sekaligus, atau dalam bentuk kalimat naratif. Pelaporan hasil uji korelasi dengan menggunakan tabel dan narasi merujuk kepada gaya penulisan *Publication Manual of the APA* edisi ke-7. Berikut adalah contoh pelaporan hasil uji korelasi menggunakan tabel (Tabel 25.1).

Tabel 25.1: Koefisien korelasi variabel penelitian ($N = 267$)

	1	2	3
1. Well-being	-		
2. Learning engagement	.669**	-	
3. Neuroticism	-.396**	-.384**	-
4. Opennes to experience	.119	1.89**	-.128

Catatan: * $p < .05$, ** $p < .01$, *** $p < .001$

Pelaporan hasil uji korelasi dalam bentuk naratif perlu mencantumkan beberapa unsur penting, yaitu: jenis uji korelasi yang digunakan (*Pearson correlation – r*, atau *Spearman correlation – r_s*), *degree of freedom* (n-2), nilai signifikansi, arah dan besaran korelasi. Format penulisannya adalah:

$r(df)$ = "koefisien korelasi", "p-value"

Contoh:

Terdapat hubungan positif yang sangat kuat antara *learning engagement* dengan *well-being* pada mahasiswa, $r_s(265) = .669, p <.001$.

Formulasi dan redaksi kalimat dalam menyampaikan hasil korelasi dapat disesuaikan dengan gaya bahasa peneliti, dengan memastikan bahwa seluruh unsur penting tetap disertakan dalam pelaporan. Hasil uji korelasi dapat dibuat dalam bentuk narasi seperti berikut:

"Hasil uji korelasi Pearson menunjukkan bahwa terdapat hubungan positif yang sangat kuat antara *learning engagement* dengan *well-being* pada mahasiswa ($r_s(265) = .669, p <.001$). Selain itu, semakin tinggi karakteristik kepribadian *neuroticism* yang dimiliki individu, semakin rendah *well-being* yang dimiliki. Korelasi antara kedua variabel tergolong sedang ($r_s(265) = -.396, p <.001$). Sementara itu, karakteristik kepribadian *openness to experiences* tidak berkorelasi dengan *well-being* individu ($r_s(265) = .119, p = .052$)."

Soal Latihan

Penilaian mengenai hasil belajar merupakan bagian penting dalam sebuah proses pembelajaran. Ujilah pengetahuan dan pemahaman Anda mengenai materi Bagian 7 ini dengan menjawab sejumlah pertanyaan pada link di bawah ini:



*Scan QR untuk membuka versi interaktif: [https://www.yarsi.ac.id/fakultas-psikologi/
bukustatistik/latihan7/](https://www.yarsi.ac.id/fakultas-psikologi/bukustatistik/latihan7/)*

Bagian VIII

MODEL REGRESI LINEAR

Pada bagian sebelumnya kita telah mempelajari analisis korelasi yang menunjukkan hubungan antara dua variabel. Di bagian ini kita akan membahas analisis regresi linear yang juga digunakan untuk memahami hubungan antar variabel. Namun, analisis regresi memungkinkan kita untuk tidak hanya mengukur hubungan, tetapi juga memodelkan hubungan tersebut dan memprediksi nilai variabel luaran (variabel dependen) berdasarkan variabel prediktor (variabel bebas).

Dalam regresi linear, kita akan fokus pada tiga jenis model regresi utama yang digunakan untuk menguji hubungan antara variabel-variabel ini:

1. Regresi linear sederhana (*simple linear regression*)
2. Regresi majemuk (*multiple linear regression*)
3. Regresi hierarkis (*hierarchical regression*)

Bab ini akan memperkenalkan Anda pada prinsip dasar regresi linear, menjelaskan asumsi yang perlu dipenuhi, serta bagaimana menginterpretasikan hasil analisis dan menuliskan laporan regresi. Dengan memahami teknik-teknik ini, Anda akan siap untuk mengaplikasikan regresi linear dalam berbagai konteks penelitian untuk menggali hubungan yang lebih kompleks antar variabel.

i Tujuan Pembelajaran

Bab 26. Prinsip Dasar Model Regresi Linear:

- Memahami prinsip dasar model regresi linear dan perbedaannya dengan analisis korelasi.

Bab 27. Regresi Linear Sederhana:

- Memahami konsep regresi sederhana.
- Mampu menginterpretasikan hasil analisis dan menulis laporan hasil regresi sederhana sesuai standar akademik.

Bab 28. Prinsip Dasar Model Regresi Linear:

- Memahami konsep regresi majemuk.
- Mampu menginterpretasikan hasil analisis dan menulis laporan hasil regresi majemuk sesuai standar akademik.

Bab 29. Regresi Linear Hierarkis:

- Memahami konsep regresi hierarkis.
- Mampu menginterpretasikan hasil analisis dan menulis laporan hasil regresi hierarkis sesuai standar akademik.

Bab 26

Prinsip Dasar Model Regresi Linear

26.1 Regresi vs. Korelasi

Dalam analisis statistik, kita sering kali tertarik untuk memahami hubungan antara dua variabel atau lebih. Korelasi adalah salah satu teknik yang digunakan untuk mengukur seberapa kuat dan sejauh mana dua variabel berhubungan. Namun, korelasi hanya mengukur sejauh mana dua variabel saling terkait tanpa membedakan mana yang berdampak terhadap mana. Misalnya, dalam studi pendidikan, kita bisa mengukur korelasi antara jumlah jam belajar dan skor ujian siswa untuk melihat apakah semakin banyak jam belajar, semakin tinggi skor ujian.

Dalam hal ini, korelasi tidak memberi kita gambaran tentang dampak atau arah hubungan. Di sinilah regresi linear menjadi lebih berguna. **Regresi** memungkinkan kita untuk **memodelkan hubungan** fungsional antara satu **variabel prediktor** (atau variabel independen) dengan satu **variabel luaran** (atau variabel dependen) (Aiken & West, 1991; Gravetter & Wallnau, 2017). Dengan kata lain, regresi memungkinkan kita untuk **memprediksi** nilai variabel luaran berdasarkan nilai variabel prediktor. Sebagai contoh, regresi dapat digunakan untuk memprediksi skor ujian berdasarkan jumlah jam belajar.

26.2 Persamaan dan Garis Regresi Linear

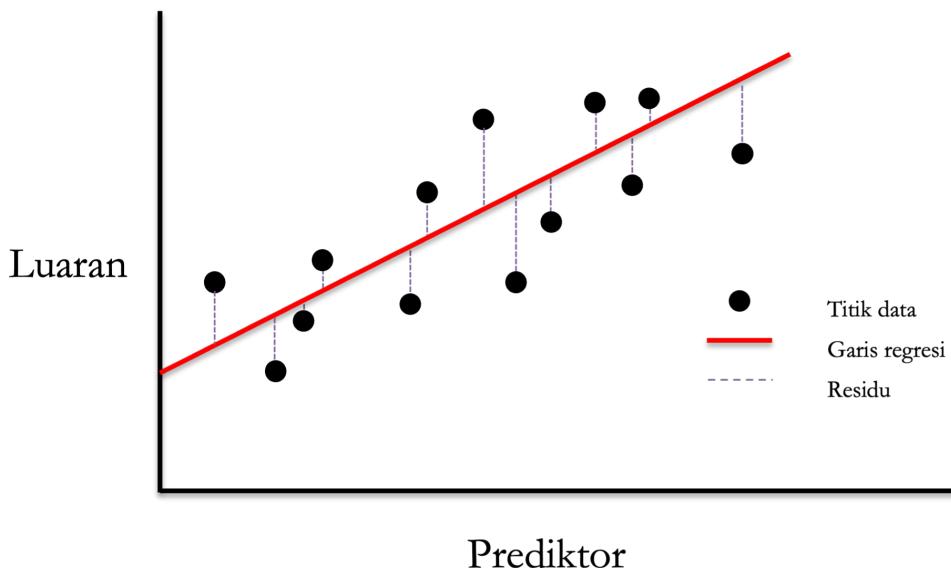
Dalam regresi linear sederhana, kita menganalisis hubungan antara satu variabel prediktor dan satu variabel *outcome*, menggunakan persamaan garis regresi untuk menggambarkan hubungan tersebut. Persamaan regresi biasanya berbentuk:

$$Y = \alpha + \beta X + \varepsilon$$

di mana:

- **Y** adalah variabel luaran (terikat),
- **X** adalah variabel prediktor (bebas),
- **α** adalah intercept atau konstanta (nilai Y saat $X = 0$),
- **β** adalah koefisien regresi yang menunjukkan perubahan rata-rata Y untuk setiap unit perubahan X ,
- **ε** adalah residu, yang menggambarkan perbedaan antara nilai yang diprediksi oleh model dan nilai observasi sebenarnya.

Garis persamaan regresi (Gambar 26.1) adalah representasi grafis dari persamaan regresi ini, di mana sumbu horizontal (X) mewakili variabel prediktor dan sumbu vertikal (Y) mewakili variabel luaran. Garis regresi menggambarkan hubungan linear antara X dan Y, dengan koefisien β_1 mengukur kemiringan garis tersebut. Semakin besar nilai β_1 , semakin curam garis regresi, menunjukkan hubungan yang lebih kuat antara X dan Y.



Gambar 26.1: Garis persamaan regresi

Sementara itu, **residu** adalah perbedaan antara nilai observasi dan nilai yang diprediksi oleh model regresi. Dalam kata lain, residu mengukur seberapa baik model regresi dapat memprediksi nilai variabel luaran untuk setiap pengamatan. Residu yang besar menunjukkan bahwa model kurang akurat dalam memprediksi nilai-nilai tersebut, sedangkan residu yang kecil menunjukkan model yang lebih akurat.

Regresi linear tidak hanya memberikan informasi tentang kekuatan hubungan, tetapi juga memberikan model prediktif yang dapat digunakan untuk memperkirakan nilai variabel *outcome* berdasarkan nilai variabel prediktor. Ini berbeda dengan korelasi yang hanya menunjukkan hubungan antara dua variabel tanpa memprediksi nilai-nilai spesifik.

Sebagai contoh lainnya, dalam dunia bisnis, kita dapat menggunakan regresi untuk memprediksi penjualan (variabel *outcome*) berdasarkan pengeluaran iklan (variabel prediktor). Regresi memberikan model prediktif yang dapat diterapkan dalam perencanaan dan pengambilan keputusan, sementara korelasi hanya menunjukkan apakah ada hubungan antara dua variabel tanpa memberikan model prediksi yang jelas.

Dengan regresi, kita juga bisa menganalisis hubungan lebih kompleks, seperti **menguji peran beberapa variabel prediktor** terhadap satu variabel luaran, yang dikenal dengan regresi majemuk. Hal ini membuka banyak peluang untuk menganalisis data yang lebih kompleks, baik dalam penelitian sosial, ekonomi, maupun ilmu kesehatan.

Bab 27

Regresi Linear Sederhana

Regresi sederhana adalah model analisis yang menguji hubungan antara satu variabel prediktor (X) dan satu variabel luaran (Y). Dalam regresi linear sederhana, kita berusaha memprediksi nilai Y berdasarkan nilai X, dengan anggapan bahwa hubungan di antara keduanya dapat digambarkan dengan garis lurus. Untuk dapat melakukan analisis regresi linear sederhana dengan tepat, penting untuk memahami asumsi dasar, bagaimana menginterpretasi hasilnya, dan cara melaporkan hasil analisisnya.

27.1 Asumsi-asumsi dalam Regresi Sederhana

Regresi sederhana didasarkan pada beberapa asumsi yang harus dipenuhi agar hasil analisis dapat diandalkan:

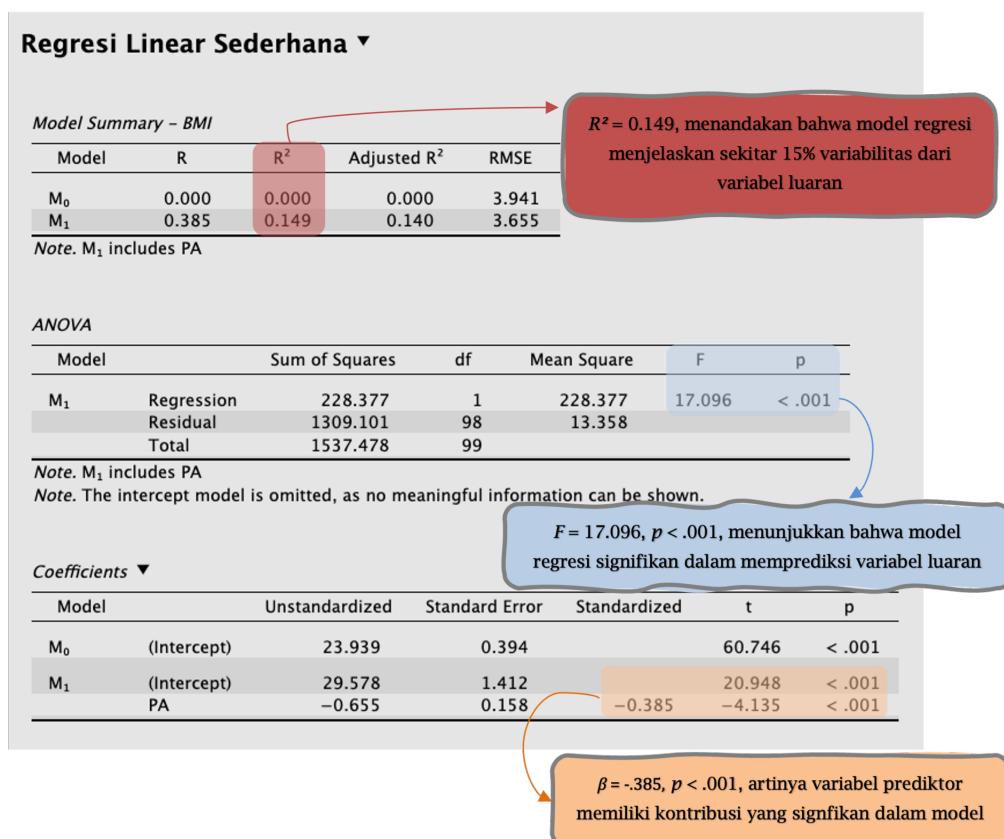
1. **Linieritas:** Asumsi pertama adalah hubungan antara variabel prediktor (X) dan variabel luaran (Y) harus linear, yang berarti hubungan dapat digambarkan dengan garis lurus. Untuk memeriksa linieritas, dapat digunakan *scatter plot* antara X dan Y, yang sebaiknya menunjukkan pola garis lurus.
2. **Independensi:** Data antar observasi harus independen, artinya nilai variabel pada satu pengamatan (misalnya, respons dari satu individu) tidak boleh bergantung atau dipengaruhi oleh pengamatan lainnya. Uji independensi bisa dilakukan dengan menggunakan uji *Durbin-Watson*.
3. **Normalitas residual:** Residual (selisih antara nilai yang diprediksi dan nilai observasi) harus terdistribusi normal. Pelanggaran normalitas dapat memengaruhi hasil uji statistik dan kesimpulan yang diambil. Normalitas dapat diuji dengan uji *Shapiro-Wilk* atau dengan melihat *Q-Q plot* residual.
4. **Homoskedastisitas:** Variabilitas residual harus seragam di sepanjang garis regresi. Jika residu lebih besar pada nilai-nilai prediktor tertentu, ini menunjukkan pelanggaran homoskedastisitas. Uji untuk homoskedastisitas dapat dilakukan dengan melihat plot residual vs. prediktor.

Jika asumsi-asumsi ini tidak terpenuhi, analisis regresi dapat menghasilkan kesimpulan yang bias. Dalam hal ini, peneliti bisa mempertimbangkan untuk menggunakan transformasi data atau beralih ke model regresi yang lebih *robust*.

27.2 Interpretasi Hasil Regresi Sederhana

Setelah menjalankan analisis regresi, hasil yang diperoleh harus diinterpretasikan untuk memahami hubungan antara X dan Y (lihat Gambar 27.1). Hal-hal yang perlu diperhatikan dalam interpretasi hasil adalah:

- **Koefisien Regresi (β):** Koefisien regresi menunjukkan seberapa besar perubahan **rata-rata Y** untuk setiap perubahan satu unit pada **X**. Interpretasi ini memberikan gambaran tentang peran variabel prediktor terhadap variabel luaran.
- **Intercept (a):** Intercept menunjukkan nilai Y ketika X = 0. Dalam beberapa kasus, intercept mungkin tidak memiliki makna praktis, terutama jika X = 0 tidak realistik dalam konteks penelitian.
- **Nilai p:** Nilai p digunakan untuk menguji apakah koefisien regresi signifikan secara statistik. Jika nilai p lebih kecil dari 0,05 (nilai signifikansi standar), maka kita dapat menyimpulkan bahwa peran X terhadap Y signifikan. Sebaliknya, jika nilai p lebih besar dari 0,05, maka kita tidak dapat menolak hipotesis nol (bahwa ada peran yang signifikan).
- **R-squared (R^2):** Nilai R^2 menunjukkan seberapa besar **proporsi variabilitas** dalam variabel *outcome* yang dapat dijelaskan oleh variabel prediktor. R^2 berkisar antara 0 dan 1, di mana semakin mendekati 1, semakin baik model regresi dalam memprediksi variabel *outcome*. R^2 yang rendah menunjukkan bahwa model tidak dapat menjelaskan variabilitas Y secara signifikan.



Gambar 27.1: Contoh output hasil analisis regresi sederhana

27.3 Penulisan Laporan Hasil Regresi Sederhana

Dalam laporan hasil regresi sederhana, peneliti harus menyajikan hasil regresi secara sistematis dan jelas. Sebagai contoh:

- **Laporan koefisien:** "Analisis regresi linear menunjukkan bahwa jumlah jam belajar berpengaruh signifikan terhadap skor ujian, $\beta = 2,45, p < 0,001$. Setiap tambahan satu jam belajar berhubungan dengan peningkatan skor ujian sebesar 2,45 poin."
- **Laporan R^2 dan nilai-p:** "Model regresi sederhana memiliki nilai R^2 sebesar 0,35, yang menunjukkan bahwa 35% variasi dalam skor ujian dapat dijelaskan oleh jumlah jam belajar."

Laporan seperti ini memberikan informasi yang jelas tentang hubungan antara variabel serta seberapa kuat model regresi dalam menjelaskan variabilitas data.

Bab 28

Regresi Linear Majemuk

Analisis regresi majemuk digunakan untuk memodelkan hubungan antara satu variabel *outcome* dan lebih dari satu variabel prediktor. Regresi majemuk memungkinkan kita untuk mengevaluasi bagaimana beberapa faktor berkontribusi terhadap variabel yang ingin diprediksi, dengan mengontrol pengaruh variabel lain dalam model.

28.1 Konsep Dasar

Persamaan dalam regresi majemuk adalah sebagai berikut:

$$Y = \alpha + \beta_1 X_1 + \beta_2 X_2 + \cdots + \beta_n X_n + \varepsilon$$

di mana:

- **Y** adalah variabel luaran,
- **X₁, X₂, dan X_n** adalah skor pada variabel-variabel prediktor,
- **α** adalah intersep atau konstanta (nilai Y saat semua variabel prediktor bernilai nol),
- **β₁, β₂, β_n** adalah koefisien regresi yang menunjukkan seberapa besar peran masing-masing variabel prediktor terhadap variabel luaran, dengan asumsi bahwa variabel lain tetap konstan
- **ε (error)** adalah residu, yang menggambarkan perbedaan antara nilai yang diprediksi oleh model dan nilai observasi sebenarnya.

Dalam regresi majemuk, yang paling penting adalah apakah nilai β masing-masing variabel prediktor signifikan atau tidak dalam model. Signifikansi ini diuji menggunakan uji *t* untuk menentukan apakah peran variabel prediktor tersebut cukup kuat untuk diterima sebagai kontribusi yang bermakna terhadap perubahan variabel luaran.

28.2 Asumsi-asumsi dalam Regresi Majemuk

Asumsi-asumsi dalam analisis regresi majemuk pada dasarnya sama dengan asumsi-asumsi dalam regresi sederhana, yaitu linearitas, independensi residual, normalitas residual, dan homoskedastisitas. Namun, dalam regresi majemuk, terdapat sebuah asumsi lainnya yang perlu diperhatikan, yaitu **multikolinearitas**. Multikolinearitas terjadi ketika dua atau lebih variabel prediktor sangat berkorelasi satu sama lain, yang dapat mengganggu interpretasi koefisien regresi dan memengaruhi keakuratan hasil analisis.

Kondisi ini dapat menyebabkan kesulitan dalam mengidentifikasi pengaruh individual masing-masing variabel prediktor terhadap variabel luaran, karena estimasi koefisien regresi menjadi tidak stabil dan sulit diinterpretasikan. Akibatnya, koefisien regresi dapat menjadi tidak signifikan meskipun secara teoritis seharusnya signifikan.

Cara untuk menguji multikolinearitas adalah dengan menggunakan **Variance Inflation Factor (VIF)**. VIF mengukur sejauh mana variabilitas koefisien regresi meningkat karena adanya multikolinearitas. Nilai VIF yang tinggi (umumnya lebih besar dari 5) menunjukkan adanya multikolinearitas yang signifikan antara variabel-variabel independen dalam model, yang perlu diperbaiki untuk memastikan kestabilan dan akurasi estimasi regresi.

Selain **VIF**, indikator lain yang dapat digunakan untuk mendeteksi multikolinearitas adalah **Tolerance**. **Tolerance** adalah kebalikan dari **VIF**, dihitung sebagai 1 dibagi dengan **VIF** ($Tolerance = 1/VIF$). **Tolerance** yang rendah (biasanya kurang dari 0,1) menunjukkan adanya multikolinearitas yang signifikan. Semakin rendah nilai **Tolerance**, semakin tinggi kemungkinan adanya multikolinearitas antara variabel prediktor.

28.3 Interpretasi Hasil Analisis Regresi Majemuk

Dalam regresi majemuk, evaluasi dimulai dengan uji F , yang menguji signifikansi model secara keseluruhan. Jika nilai $p < .05$, model regresi dianggap signifikan. Selanjutnya, koefisien regresi (b) masing-masing variabel prediktor diuji menggunakan uji t . Koefisien yang signifikan (nilai $p < .05$) menunjukkan bahwa variabel tersebut memiliki peran yang bermakna terhadap variabel luaran. Selain itu, nilai R^2 menggambarkan seberapa banyak variabilitas variabel luaran yang dapat dijelaskan oleh variabel prediktor dalam model. Semakin tinggi nilai R^2 , semakin baik model dalam menjelaskan variasi data. Interpretasi hasil uji ini membantu menentukan apakah model regresi memberikan prediksi yang valid dan apakah setiap variabel prediktor berkontribusi signifikan terhadap variabel luaran.

Meskipun model regresi secara keseluruhan dapat menunjukkan kemampuan prediksi yang signifikan berdasarkan hasil uji F dan nilai R^2 , tidak berarti bahwa semua variabel prediktor berkontribusi secara signifikan. Beberapa variabel mungkin tidak menunjukkan peran yang signifikan terhadap variabel luaran meskipun dimasukkan dalam model. Hal ini dapat dilihat dari hasil uji t untuk koefisien regresi, di mana variabel yang memiliki nilai $p > .05$ dianggap tidak berkontribusi secara signifikan. Oleh karena itu, meskipun model secara keseluruhan berhasil menjelaskan variasi dalam variabel luaran, tidak semua prediktor dalam model tersebut selalu memiliki peran yang berarti secara statistik.

28.4 Penulisan Laporan Hasil Regresi Majemuk

Meskipun tidak wajib dilaporkan dalam detail, penting untuk menyebutkan bahwa asumsi-asumsi regresi majemuk (seperti normalitas residual, homoskedastisitas, dan independensi residual) telah diperiksa dan tidak ada pelanggaran yang signifikan. Contoh: "Pemeriksaan residual menunjukkan bahwa asumsi normalitas dan homoskedastisitas terpenuhi, dan tidak ada bukti multikolinearitas yang signifikan antara variabel independen."

Dalam pelaporan hasil analisis regresi majemuk, penting untuk menyajikan informasi

secara jelas dan sistematis agar pembaca dapat memahami bagaimana analisis dilakukan dan hasil yang diperoleh (Gambar 28.1). Berikut adalah format umum untuk menulis laporan hasil analisis regresi majemuk:

"Analisis regresi majemuk menunjukkan bahwa model secara keseluruhan signifikan, $F(3, 220) = 9.237, p < 0.001$, dengan $R^2 = .205$, yang berarti 20% variabilitas tingkat stres dapat dijelaskan oleh jam kerja, dukungan sosial, dan kecemasan. Hasil regresi menunjukkan bahwa jam kerja ($b = .354, t(220) = 4.48, p < .001$) memiliki peran signifikan terhadap tingkat stres, sementara kecemasan ($b = .09, t(220) = 1.17, p = .25$) dan dukungan sosial ($b = .08, t(220) = .91, p = .362$) tidak signifikan."

Regresi Linear Majemuk

Model Summary - Stres					$R^2 = 0.205$, model regresi menjelaskan sekitar 20% variabilitas dari variabel luaran	
Model	R	R^2	Adjusted R ²	RMSE		
M ₀	0.000	0.000	0.000	0.779		
M ₁	0.452	0.205	0.194	0.700		

Note. M₁ includes Jam kerja, Kecemasan, Dukungan sosial

ANOVA						
Model		Sum of Squares	df	Mean Square	F	p
M ₁	Regression	27.712	3	9.237	18.861	< .001
	Residual	107.750	220	0.490		
	Total	135.463	223			

Note. M₁ includes Jam kerja, Kecemasan, Dukungan sosial

Note. The intercept model is omitted, as no meaningful information can be shown.

$F = 18.861, p < .001$, model regresi signifikan dalam memprediksi variabel luaran

Coefficients							Collinearity Statistics	
Model		Unstandardized	Standard Error	Standardized	t	p	Tolerance	VIF
M ₀	(Intercept)	2.635	0.052		50.604	< .001		
M ₁	(Intercept)	0.590	0.294		2.005	0.046		
	Jam kerja	0.169	0.035	0.354	4.749	< .001	0.649	1.540
	Kecemasan	0.045	0.039	0.087	1.166	0.245	0.645	1.550
	Dukungan sosial	0.034	0.038	0.075	0.914	0.362	0.539	1.855

Hanya variabel Jam kerja yang memiliki kontribusi signifikan dalam model ($\beta = .354, p < .001$), lainnya tidak ($p > .05$)

VIF < 5, Tolerance > .1, tidak terdapat gejala multikolinearitas antar variabel prediktor

Gambar 28.1: Contoh *output* hasil analisis regresi majemuk

Bab 29

Regresi Linear Hierarkis

Analisis regresi hierarkis digunakan untuk memodelkan hubungan antara satu variabel luaran dan lebih dari satu variabel prediktor, dengan memeriksa kontribusi tambahan dari variabel prediktor yang dimasukkan secara bertahap ke dalam model (Aiken & West, 1991; Hayes, 2022). Metode ini memungkinkan peneliti untuk mengevaluasi bagaimana faktor-faktor tertentu berkontribusi terhadap variabel luaran setelah mempertimbangkan variabel-variabel lain yang lebih dahulu dimasukkan.

29.1 Konsep Dasar

Dalam regresi hierarkis, analisis dilakukan dalam beberapa langkah. Pada setiap langkah, sejumlah variabel prediktor ditambahkan ke dalam model untuk menguji seberapa besar kontribusi tambahan yang diberikan oleh variabel tersebut terhadap variabilitas variabel luaran. Model pertama biasanya dimulai dengan satu atau lebih variabel kontrol (mis., data demografi, seperti jenis kelamin dan usia) yang dianggap penting, sementara variabel prediktor utama dimasukkan pada langkah berikutnya.

Contoh persamaan regresi hierarkis untuk dua langkah adalah:

$$Y = \alpha + \beta_1 X_1 + \varepsilon \quad (\text{Model 1: variabel kontrol } X_1)$$

$$Y = \alpha + \beta_1 X_1 + \beta_2 X_2 + \varepsilon \quad (\text{Model 2: variabel prediktor } X_2 \text{ ditambahkan})$$

Sama seperti pada regresi majemuk, pada analisis regresi hierarkis, yang penting adalah memeriksa apakah penambahan setiap variabel prediktor memberikan kontribusi yang signifikan terhadap model, yang diuji dengan uji F dan perubahan dalam nilai R^2 .

29.2 Asumsi-asumsi dalam Regresi Hierarkis

Asumsi-asumsi dalam analisis regresi hierarkis pada dasarnya sama dengan asumsi dalam regresi sederhana dan regresi majemuk, yaitu linearitas, independensi residual, normalitas residual, dan homoskedastisitas, serta multikolinearitas. Namun, regresi hierarkis juga memerlukan perhatian khusus terhadap urutan variabel yang dimasukkan ke dalam model, yang harus didasarkan pada teori atau logika yang kuat. Dengan memasukkan variabel secara bertahap, peneliti dapat menguji kontribusi tambahan dari setiap variabel terhadap model.

29.3 Interpretasi Hasil Analisis Regresi Hierarkis: Delta (Δ) R^2

Proses analisis dimulai dengan memasukkan variabel kontrol (jika ada) ke dalam model pada langkah pertama. Model kemudian diperluas dengan menambahkan variabel prediktor lainnya secara bertahap. Setelah setiap langkah, perubahan dalam nilai R^2 dihitung untuk menilai apakah variabel tambahan memberikan kontribusi signifikan terhadap model.

Kontribusi tambahan dari variabel-variabel yang dimasukkan ke dalam model secara bertahap ditunjukkan sebagai *delta R²* (ΔR^2). Setiap kali variabel baru ditambahkan, kita menghitung perubahan dalam R^2 untuk menilai seberapa banyak variabilitas variabel prediktor yang dapat dijelaskan oleh variabel tersebut setelah mempertimbangkan variabel lain yang sudah ada.

Delta R² yang signifikan menunjukkan bahwa variabel yang ditambahkan memberikan kontribusi yang bermakna dalam meningkatkan kemampuan model untuk menjelaskan variasi pada variabel luaran. Pengujian *delta R²* dilakukan dengan menggunakan uji *F* untuk perubahan R^2 , yang menguji apakah peningkatan dalam R^2 setelah penambahan variabel baru cukup besar dan signifikan secara statistik (Gambar 29.1).

Regresi Linear Hierarkis ▾									
Model Summary - Exam									
Model	R	R ²	Adjusted R ²	RMSE	R ² Change	F Change	df1	df2	p
M ₀	0.005	0.000	-0.010	26.068	0.000	0.002	1	101	0.963
M ₁	0.457	0.209	0.185	23.417	0.209	13.082	2	99	< .001

Note. M₀ includes Gender
Note. M₁ includes Gender, Revise, Anxiety

ANOVA ▾					
Model		Sum of Squares	df	Mean Square	F
M ₀	Regression	1.500	1	1.500	0.002
	Residual	68635.704	101	679.561	
	Total	68637.204	102		
M ₁	Regression	14348.948	3	4782.983	8.722
	Residual	54288.256	99	548.366	< .001
	Total	68637.204	102		

Note. M₀ includes Gender
Note. M₁ includes Gender, Revise, Anxiety

Coefficients					
Model		Unstandardized	Standard Error	Standardized	t
M ₀	(Intercept)	56.210	8.147		6.900
	Gender	0.241	5.137	0.005	0.047
M ₁	(Intercept)	85.907	19.169		4.482
	Gender	1.040	4.648	0.020	0.224
	Revise	0.246	0.182	0.172	1.349
	Anxiety	-0.481	0.192	-0.319	-2.506

Gambar 29.1: Contoh output hasil analisis regresi hierarkis

Sebagai contoh, jika perubahan dalam R^2 antara dua model adalah 0.12, ini berarti

bawa variabel yang ditambahkan dalam model kedua menjelaskan tambahan 12% variabilitas dalam variabel luaran. Jika perubahan tersebut diuji dengan uji F dan hasilnya signifikan ($p < .05$), maka kita dapat menyimpulkan bahwa penambahan variabel tersebut memperbaiki model secara signifikan.

29.4 Penulisan Laporan Hasil Regresi Hierarkis

Pelaporan hasil analisis regresi hierarkis mengikuti struktur yang jelas dan sistematis. Penulis harus menyajikan informasi tentang kontribusi setiap variabel prediktor, perubahan dalam R^2 , dan hasil uji F serta uji-t. Sebagai contoh, laporan hasil regresi hierarkis bisa disusun sebagai berikut:

"Analisis regresi hierarkis menunjukkan bahwa model pertama, dengan memasukkan pengalaman kerja sebagai variabel kontrol, tidak memberikan perubahan signifikan dalam $R^2 = .12$, $F(1, 150) = 3.56$, $p = .062$. Namun, setelah menambahkan stres kerja dan dukungan sosial sebagai variabel prediktor, ΔR^2 menjadi .45, $F(3, 148) = 15.89$, $p < .001$, yang menunjukkan bahwa kedua variabel ini memberikan kontribusi signifikan terhadap tingkat kepuasan kerja. Stres kerja ($\beta = .65$, $t(148) = 4.22$, $p < .001$) dan dukungan sosial ($\beta = -.35$, $t(148) = 2.94$, $p = .004$) berperan signifikan dalam meningkatkan kepuasan kerja, sementara pengalaman kerja ($\beta = .15$, $t(148) = 1.01$, $p = .31$) tidak menunjukkan kontribusi signifikan."

Soal Latihan

Penilaian mengenai hasil belajar merupakan bagian penting dalam sebuah proses pembelajaran. Ujilah pengetahuan dan pemahaman Anda mengenai materi Bagian 8 ini dengan menjawab sejumlah pertanyaan pada link di bawah ini:



*Scan QR untuk membuka versi interaktif: [https://www.yarsi.ac.id/fakultas-psikologi/
bukustatistik/latihan8/](https://www.yarsi.ac.id/fakultas-psikologi/bukustatistik/latihan8/)*

Glosarium

A

ANOVA (Analysis of Variance): Uji parametrik untuk membandingkan rata-rata tiga kelompok atau lebih.

ANOVA Dua Arah: ANOVA dengan dua faktor/variabel bebas.

ANOVA Pengukuran Berulang: ANOVA yang digunakan untuk data longitudinal atau *repeated measures*.

ANOVA Satu Arah: ANOVA dengan satu faktor/variabel bebas.

C

Central Limit Theorem: Teorema yang menyatakan bahwa distribusi rata-rata sampel akan mendekati normal meski populasi asal tidak normal.

Chi-square (χ^2): Uji non-parametrik untuk menganalisis hubungan antar variabel kategorikal.

Cohen's d: Ukuran effect size untuk perbedaan rata-rata antar kelompok dalam satuan SD.

Confidence Interval (CI): Rentang nilai yang mungkin mencakup parameter populasi dengan tingkat keyakinan tertentu.

D

Distribusi Normal: Distribusi data berbentuk lonceng simetris dengan mean = median = modus.

E

Effect Size: Ukuran kekuatan pengaruh atau perbedaan dalam analisis statistik.

Eta Squared (η^2): Ukuran effect size yang sering digunakan pada ANOVA.

F

Friedman Test: Uji non-parametrik untuk membandingkan tiga atau lebih kelompok berpasangan ordinal.

I

Interval: Skala pengukuran yang memiliki jarak sama antar satuan tetapi tanpa nol mutlak.

K

Kolmogorov-Smirnov Test: Uji statistik untuk menilai normalitas data, lebih cocok untuk sampel besar.

Korelasi ***Pearson:*** Ukuran hubungan linear antara dua variabel interval/rasio yang normal.

Korelasi ***Spearman:*** Ukuran hubungan monotonic antara dua variabel ordinal atau tidak normal.

Kruskal-Wallis Test: Uji non-parametrik untuk membandingkan tiga atau lebih kelompok independen ordinal.

Kuartil: Pembagian data menjadi empat bagian yang sama besar.

Kurtosis: Ukuran kelancipan distribusi data dibanding distribusi normal.

L

Levene Test: Salah satu uji homogenitas varians yang umum digunakan.

M

Mann-Whitney U Test: Uji non-parametrik untuk membandingkan dua kelompok independen ordinal.

Mean: Rata-rata aritmatika.

Median: Nilai tengah dari data yang sudah diurutkan.

Modus: Nilai yang paling sering muncul dalam distribusi data.

N

Nominal: Skala pengukuran yang hanya memberi label tanpa urutan atau jarak.

O

Ordinal: Skala pengukuran yang menunjukkan urutan, tetapi jarak antar kategori tidak sama.

Outlier: Nilai data yang jauh berbeda dari sebagian besar data lain.

P

Parameter: Ukuran yang dihitung dari populasi, misalnya mean populasi.

Partial Eta Squared ($\eta^2 p$): Ukuran effect size untuk desain ANOVA dengan beberapa faktor.

Persentil: Posisi data tertentu dalam distribusi, dinyatakan dalam persentase.

Populasi: Keseluruhan individu atau objek yang menjadi sasaran penelitian.

Post-hoc Test: Uji lanjutan setelah ANOVA signifikan, misalnya Tukey atau Bonferroni.

Probabilitas: Peluang terjadinya suatu peristiwa.

R

r^2 (Koefisien Determinasi): Proporsi varians pada variabel dependen yang dijelaskan oleh variabel independen.

Randomisasi: Proses pengacakan dalam pembagian kelompok penelitian.

Range: Selisih antara nilai maksimum dan minimum.

Rasio: Skala pengukuran dengan jarak sama dan nol mutlak, memungkinkan perhitungan rasio.

Regrisi Hierarkis: Regresi dengan memasukkan variabel secara bertahap (blok).

Regresi Linear Majemuk: Regresi dengan lebih dari satu variabel prediktor.

Regresi Linear Sederhana: Regresi dengan satu variabel prediktor.

Regresi Logistik: Regresi yang digunakan ketika variabel dependen bersifat kategorikal.

Reliabilitas: Konsistensi hasil pengukuran jika diulang pada kondisi yang sama.

S

Sampel: Bagian dari populasi yang digunakan dalam penelitian untuk mewakili populasi.

Sampling Error: Perbedaan antara parameter populasi dengan statistik sampel akibat pemilihan sampel.

Shapiro-Wilk Test: Uji statistik untuk menilai normalitas data, lebih sensitif pada sampel kecil.

Skewness: Ukuran kemencengan distribusi data; positif jika miring ke kanan, negatif jika miring ke kiri.

Standar Deviasi (SD): Akar kuadrat dari varians, menunjukkan seberapa jauh data menyebar dari mean.

Statistik Deskriptif: Teknik untuk meringkas dan menyajikan data dalam bentuk tabel, grafik, mean, median, modus, dsb.

Statistik Inferensial: Teknik untuk membuat kesimpulan atau generalisasi dari sampel ke populasi.

Statistik: Ukuran yang dihitung dari sampel, misalnya mean sampel.

T

t-test: Uji parametrik untuk membandingkan rata-rata dua kelompok atau satu kelompok dengan nilai tertentu.

U

Uji Homogenitas Varians: Uji untuk memeriksa kesamaan varians antar kelompok.

Uji Normalitas: Uji untuk memeriksa apakah data mengikuti distribusi normal.

V

Validitas Konstruk: Sejauh mana alat ukur benar-benar mengukur konstruk yang dimaksud.

Variabel: Karakteristik atau atribut yang dapat bervariasi, misalnya usia, jenis kelamin, skor tes.

variabel dependen: Variabel yang dipengaruhi atau dijelaskan dalam penelitian.

Variabel independen: Variabel atau kelompok yang tidak saling memengaruhi.

Varians: Ukuran penyebaran data, dihitung sebagai rata-rata kuadrat deviasi dari mean.

W

Wilcoxon Signed-Rank Test: Uji non-parametrik untuk dua kelompok berpasangan ordinal.

Daftar Pustaka

- Aiken, L. S., & West, S. G. (1991). *Multiple Regression: Testing and Interpreting Interactions*. Sage.
- Azwar, S. (2019). *Metode Penelitian Psikologi* (3rd ed.). Pustaka Pelajar.
- Buuren, S. van. (2018). *Flexible Imputation of Missing Data*. CRC Press.
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences*. Lawrence Erlbaum Associates.
- Conover, W. J. (1999). *Practical Nonparametric Statistics* (3rd ed.). Wiley.
- Coolican, H. (2014). *Research methods and statistics in psychology* (hlm. 773). Psychology Press, Taylor & Francis Group.
- Cowles, M. (2000). *Statistics in Psychology: An Historical Perspective*. Psychology Press.
- Cumming, G. (2011). *Understanding the New Statistics: Effect Sizes, Confidence Intervals, and Meta-Analysis*. Routledge.
- Dancey, C. P., & Reidy, J. (2017). *Statistics without maths for psychology* (7th ed.). Pearson Education Limited.
- Field, A. (2018). *Discovering Statistics Using IBM SPSS Statistics* (5th ed.). Sage.
- Gelfond, J. A. L., Klugman, C. M., Welty, L. J., Heitman, E., Louden, C., & Pollock, B. H. (2014). [How to Tell the Truth with Statistics: The Case for Accountable Data Analyses in Team-based Science](#). *Journal of translational medicine & epidemiology*, 2.
- Ghasemi, A., & Zahediasl, S. (2012). Normality tests for statistical analysis: A guide for non-statisticians. *International Journal of Endocrinology and Metabolism*, 10(2), 486–489. <https://doi.org/10.5812/ijem.3505>
- Gravetter, F. J., & Wallnau, L. B. (2017). *Statistics for the Behavioral Sciences*. Cengage Learning.
- Hayes, A. F. (2022). *Introduction to Mediation, Moderation, and Conditional Process Analysis: A Regression-Based Approach* (3rd ed.). Guilford Press.
- Howell, D. C. (2013). *Statistical Methods for Psychology* (8th ed.). Wadsworth.
- Lakens, D. (2013). Calculating and reporting effect sizes to facilitate cumulative science: A practical primer for t-tests and ANOVAs. *Frontiers in Psychology*, 4, 863. <https://doi.org/10.3389/fpsyg.2013.00863>
- Moore, D. S., Notz, W., & Fligner, A. (2018). *The basic practice of statistics* (7th ed.). W.H. Freeman.
- Oktaviani, M. A., & Notobroto, H. B. (2014). Perbandingan tingkat konsistensi normalitas distribusi metode kolmogorov-smirnov, lilliefors, shapiro-wilk, dan skewness-kurtosis. *Jurnal Biometrika dan kependudukan*, 3, 127–135.
- Razali, N., & Wah, Y. (2011). Power comparisons of Shapiro-Wilk , Kolmogorov-Smirnov , Lilliefors and Anderson-Darling test. *Journal of Statistical Modeling and Analytics*, 2, 21–33.
- Salsburg, D. (2002). *The lady tasting tea: how statistics revolutionized science in the twentieth century*. Henry Holt; Company.
- Stevens, S. S. (1946). On the theory of scales of measurement. *Science*, 103(2684), 677–680. <https://doi.org/10.1126/science.103.2684.677>
- Sugiyono. (2015). *Statistika untuk Penelitian*. Alfabeta.
- Tabachnick, B. G., & Fidell, L. S. (2019a). *Using Multivariate Statistics* (7th ed.). Pearson Education Limited.
- Tabachnick, B. G., & Fidell, L. S. (2019b). *Using Multivariate Statistics* (7th ed.). Pearson.
- Yap, B. W., & Sim, C. H. (2011). Comparisons of various types of normality tests. *Journal of Statistical Computation and Simulation*, 81, 2141–2155. <https://doi.org/10.1080/00949655.2010.520163>

Indeks

Data, [13](#)

Statistik, [7](#)

Statistik deskriptif, [9](#)

Statistik inferensial, [9](#)