



Biostatistik Modern

Rini Ernawati, S.Pd., M.Kes., Joko Susanto, S.Kep., Ns., M.Kes., apt. Retno Fitriati, S.Si., M.Kes., Ns. Rahayu Tri Nuritasari, M.Kep., Ns. Septi Viantri Kurdaningsih, M.Kep., Bangkit Ary Pratama, SKM., M.Kes., Onetusfisi Putra, M.K.M., Wahyu Dwi Lesmono, S.Si., M.Si., Rinco Siregar, S.Kep., Ns., M.N.S., Ph.D., dan Nurnianingsih A. Yasin, S.Kep., Ns., M.Kep.

Biostatistik Modern

Rini Ernawati, S.Pd., M.Kes.

Joko Susanto, S.Kep., Ns., M.Kes.

apt. Retno Fitriati, S.Si., M.Kes.

Ns. Rahayu Tri Nuritasari, M.Kep.

Ns. Septi Viantri Kurdaningsih, M.Kep.

Bangkit Ary Pratama, SKM., M.Kes.

Onetusfifi Putra, M.K.M.

Wahyu Dwi Lesmono, S.Si., M.Si.

Rinco Siregar, S.Kep., Ns., M.N.S., Ph.D.

Nurnianingsih A. Yasin, S.Kep., Ns., M.Kep.

PT BUKULOKA LITERASI BANGSA

Anggota IKAPI: No. 645/DKI/2024



Biostatistik Modern

Penulis : Rini Ernawati, S.Pd., M.Kes., Joko Susanto, S.Kep.,
Ns., M.Kes., apt. Retno Fitriati, S.Si., M.Kes., Ns.
Rahayu Tri Nuritasari, M.Kep., Ns. Septi Viantri
Kurdaningsih, M.Kep., Bangkit Ary Pratama, SKM.,
M.Kes., Onetusfisi Putra, M.K.M., Wahyu Dwi
Lesmono, S.Si., M.Si., Rinco Siregar, S.Kep., Ns.,
M.N.S., Ph.D., dan Nurnianingsih A. Yasin, S.Kep.,
Ns., M.Kep.

ISBN : 978-634-250-941-8 (PDF)

Penyunting Naskah : Rikhanatus Saliha, S.Sos

Tata Letak : Rikhanatus Saliha, S.Sos

Desain Sampul : Novikean Keysah Sanisri

Penerbit

Penerbit PT Bukuloka Literasi Bangsa

Distributor: PT Yapindo

Kompleks Business Park Kebon Jeruk Blok I No. 21, Jl. Meruya Ilir Raya
No. 88, Kelurahan Meruya Utara, Kecamatan Kembangan, Kota Adm.
Jakarta Barat, Provinsi DKI Jakarta, Kode Pos: 11620

Email : penerbit.blb@gmail.com

Whatsapp : 0878-3483-2315

Website : bukuloka.com

© Hak cipta dilindungi oleh undang-undang

Berlaku selama 50 (lima puluh) tahun sejak ciptaan tersebut pertama kali dilakukan pengumuman.

Dilarang mengutip atau memperbanyak sebagian atau seluruh isi buku ini tanpa izin tertulis dari penerbit. Ketentuan Pidana Sanksi Pelanggaran Pasal 2 UU Nomor 19 Tahun 2002 Tentang Hak Cipta.

Barang siapa dengan sengaja dan tanpa hak melakukan perbuatan sebagaimana dimaksud dalam Pasal 2 ayat (1) atau Pasal 49 ayat (1) dan ayat (2) dipidana dengan pidana penjara masing-masing paling singkat 1 (satu) bulan dan/atau denda paling sedikit Rp1.000.000,00 (satu juta rupiah), atau pidana penjara paling lama 7 (Tujuh) tahun dan/atau denda paling banyak Rp5.000.000.000,00 (lima miliar rupiah).

Barang siapa dengan sengaja menyerahkan, menyiarkan, memamerkan, mengedarkan atau menjual kepada umum suatu ciptaan atau barang hasil pelanggaran Hak Cipta atau Hak Terkait sebagaimana dimaksud pada ayat (1) dipidana dengan pidana penjara paling lama 5 (lima) tahun dan/atau denda paling banyak Rp500.000.000,00 (lima ratus juta rupiah).

KATA PENGANTAR

Puji syukur ke hadirat Tuhan Yang Maha Esa atas tersusunnya buku ajar yang berjudul *Biostatistik Modern*. Buku ini hadir untuk memberikan pemahaman sederhana mengenai konsep dan penerapan biostatistik dalam menganalisis data kesehatan, penelitian, maupun informasi ilmiah lainnya, sehingga pembaca dapat memahami pentingnya statistik dalam pengambilan keputusan yang tepat.

Buku ini ditujukan untuk masyarakat umum agar dapat menjadi bacaan yang mudah dipahami bagi siapa saja yang ingin mengenal dasar-dasar biostatistik dan penggunaannya dalam kehidupan sehari-hari maupun profesional. Dengan bahasa yang ringan dan penjelasan yang jelas, pembaca diajak memahami cara membaca data, menginterpretasikan hasil, dan menggunakan informasi statistik untuk membuat keputusan yang lebih akurat dan efektif.

.

Jakarta, Desember 2025

Tim Penyusun

DAFTAR ISI

KATA PENGANTAR.....	iii
DAFTAR ISI.....	iv
Bab 1: Pengenalan Biostatistik dalam Dunia Kesehatan	1
1.1 Pengertian dan Cakupan Biostatistik.....	1
1.2 Peran Biostatistik dalam Penelitian Kesehatan.....	3
1.3 Jenis Data dan Skala Pengukuran.....	6
1.4 Penerapan Biostatistik dalam Pelayanan dan Kebijakan Kesehatan	10
1.5 Latihan Soal.....	14
Bab 2: Konsep Dasar Data dan Statistik Deskriptif	16
2.1 Pengertian Data dan Statistik.....	16
2.2 Jenis dan Skala Data.....	20
2.3 Statistik Deskriptif.....	25
2.4 Penyajian Data.....	28
2.5 Latihan Soal.....	32
Bab 3: Distribusi Probabilitas dan Penerapan Peluang dalam Kesehatan.....	33
3.1 Konsep Dasar Probabilitas.....	33
3.2 Distribusi Probabilitas Diskret.....	36
3.3 Distribusi Probabilitas Kontinu	39
3.4 Penerapan Probabilitas dalam Kesehatan	41
3.5 Latihan Soal.....	44
Bab 4: Populasi, Sampel, dan Teknik Sampling	45
4.1 Pengertian Populasi dan Sampel.....	45
4.2 Jenis-jenis Populasi	50
4.3 Konsep Sampling dan Pentingnya dalam Penelitian	53
4.4 Teknik Sampling	57

4.5 Latihan Soal.....	60
Bab 5: Estimasi Parameter dan Uji Hipotesis	62
5.1 Pengertian Estimasi Parameter	62
5.2 Jenis Estimasi dan Tingkat Kepercayaan	66
5.3 Konsep Uji Hipotesis.....	70
5.4 Jenis Uji Statistik dalam Hipotesis	74
5.5 Latihan Soal.....	77
Bab 6: Perbandingan Data Kesehatan Berdasarkan Nilai dan Proporsi.....	78
6.1 Mengenal Perbandingan Data Kesehatan Berdasarkan Nilai dan Proporsi	78
6.2 Perbandingan Berdasarkan Nilai (Data Kontinu)	81
6.3 Perbandingan Berdasarkan Proporsi (Data Kategorikal).....	84
6.4 Contoh Praktik: Perbandingan Data Kesehatan.....	87
6.5 Pentingnya Uji Statistik dalam Data Kesehatan	90
6.6 Kesimpulan.....	93
Bab 7: Hubungan antar Variabel dalam Konteks Kesehatan .	96
7.1 Konsep Dasar Hubungan antar Variabel	96
7.2 Jenis Hubungan antar Variabel.....	99
7.3 Analisis Korelasi	102
7.4 Analisis Regresi.....	105
7.5 Aplikasi Hubungan antar Variabel dalam Penelitian Kesehatan .	109
7.6 Latihan Soal.....	112
Bab 8: Gizi, Lingkungan, dan Ketahanan Pangan: Tantangan dan Tren Pola Makan Modern	114
8.1 Gizi dan Perubahan Lingkungan Global	114
8.2 Pola Makan Berkelanjutan dan Planetary Health Diet	118
8.3 Ketahanan Pangan dan Perubahan Sosial Ekonomi	121
8.4 Tren Pola Makan Masyarakat Modern	124
8.5 Latihan Soal.....	129

Bab 9: Pengolahan Data Kesehatan Menggunakan R dan SPSS
..... **130**

 9.1 Menenal Analisis Data Kesehatan 130

 9.2 Langkah Dasar Pengolahan Data Menggunakan SPSS 133

 9.3 Langkah Dasar Pengolahan Data Menggunakan R 136

 9.4 Interpretasi dan Visualisasi Hasil Analisis 141

 9.5 Latihan Soal..... 144

Profil Penulis **145**

Daftar Pustaka..... **156**

Bab 1: Pengenalan

Biostatistik dalam Dunia Kesehatan

1.1 Pengertian dan Cakupan Biostatistik

Biostatistik merupakan cabang ilmu statistik yang menerapkan prinsip-prinsip analisis data dalam bidang biologi, kedokteran, dan kesehatan masyarakat. Ilmu ini berperan penting dalam mendeskripsikan fenomena kesehatan, mengevaluasi efektivitas intervensi medis, serta mendukung pengambilan keputusan berbasis bukti (*evidence-based decision making*). Menurut Sullivan (2023), biostatistik membantu mengubah data biologis dan medis yang kompleks menjadi informasi yang dapat digunakan untuk memahami determinan kesehatan, menilai risiko penyakit, dan memandu kebijakan publik. Dengan demikian, biostatistik menjadi landasan ilmiah bagi penelitian klinis, epidemiologi, dan kebijakan kesehatan masyarakat.

Ruang lingkup biostatistik meliputi berbagai kegiatan ilmiah, mulai dari perancangan penelitian, pengumpulan data, analisis inferensial, hingga interpretasi hasil untuk mendukung pengambilan keputusan medis. Rossi (2022) menjelaskan bahwa biostatistik mencakup penerapan metode deskriptif dan inferensial dalam

berbagai konteks, termasuk uji klinis, survei populasi, studi kohort, dan penelitian eksperimental. Dalam penelitian medis, biostatistik digunakan untuk mengidentifikasi hubungan antara faktor risiko dan penyakit, mengevaluasi efektivitas pengobatan, serta memprediksi tren kesehatan populasi. Peran ini semakin penting di era digital, ketika teknologi *big data* dan *machine learning* memungkinkan analisis data kesehatan secara lebih cepat dan akurat.

Selain itu, biostatistik berfungsi sebagai alat dalam pengembangan ilmu kedokteran berbasis bukti (*evidence-based medicine*). Melalui analisis statistik yang valid dan reliabel, biostatistik membantu menguji hipotesis ilmiah dan menilai signifikansi hasil penelitian medis. Dalam konteks kesehatan masyarakat, penerapan biostatistik meliputi analisis tren epidemiologi, pengukuran efektivitas program kesehatan, serta penilaian beban penyakit (*burden of disease*). Dengan kemajuan teknologi komputasi, ruang lingkup biostatistik kini juga mencakup bioinformatika dan *computational biology*, yang mengintegrasikan data genetik, molekuler, dan klinis untuk memperkuat pemahaman terhadap mekanisme biologis penyakit.

Dalam praktik profesional, biostatistik tidak hanya berfungsi sebagai alat analisis, tetapi juga sebagai pendekatan ilmiah untuk memastikan validitas dan reliabilitas hasil penelitian. Peran biostatistik semakin penting dalam memastikan bahwa kebijakan kesehatan, terapi medis, dan program intervensi berbasis pada bukti empiris yang dapat dipertanggungjawabkan secara ilmiah. Oleh karena itu, pemahaman yang kuat terhadap biostatistik menjadi

kebutuhan mendasar bagi tenaga kesehatan, peneliti, dan pengambil kebijakan untuk menciptakan sistem kesehatan yang efektif, efisien, dan berorientasi pada bukti.

1.2 Peran Biostatistik dalam Penelitian Kesehatan

Biostatistik merupakan cabang ilmu statistik yang diterapkan dalam bidang biologi dan kesehatan untuk mengumpulkan, mengolah, menganalisis, dan menafsirkan data yang berkaitan dengan fenomena kesehatan. Dalam penelitian kesehatan, biostatistik berperan penting dalam memastikan validitas, reliabilitas, serta generalisasi hasil penelitian. Melalui penerapan metode statistik yang tepat, peneliti dapat menarik kesimpulan ilmiah secara objektif dan menghindari kesalahan interpretasi data. Oleh karena itu, pemahaman biostatistik tidak hanya menjadi kebutuhan teknis, tetapi juga merupakan fondasi dalam pengambilan keputusan berbasis bukti (*evidence-based practice*).

1.2.1 Fungsi Biostatistik dalam Desain Penelitian

Tahapan awal dalam penelitian kesehatan adalah perancangan desain penelitian yang sistematis. Biostatistik berperan penting dalam menentukan jenis desain yang sesuai, seperti *cross-sectional*, *case-control*, *cohort*, atau *randomized controlled trial (RCT)*. Pemilihan desain ini didasarkan pada tujuan penelitian, jenis data, dan hipotesis yang ingin diuji.

Selain itu, biostatistik juga digunakan untuk menghitung ukuran sampel yang memadai agar hasil penelitian memiliki kekuatan statistik (*statistical power*) yang cukup. Kesalahan dalam menentukan ukuran sampel dapat mengakibatkan hasil penelitian menjadi tidak signifikan secara statistik atau tidak dapat digeneralisasi. Dengan demikian, biostatistik berfungsi memastikan bahwa rancangan penelitian efisien, valid, dan mampu menjawab pertanyaan ilmiah secara tepat (Rosner, 2016).

1.2.2 Analisis Data dan Interpretasi Hasil

Tahap analisis data merupakan inti dari penerapan biostatistik. Teknik analisis yang digunakan bervariasi tergantung pada jenis data (numerik atau kategorik) serta tujuan penelitian (deskriptif, komparatif, atau inferensial). Analisis deskriptif digunakan untuk menggambarkan karakteristik populasi penelitian melalui ukuran pemusatan dan penyebaran, seperti rata-rata, median, standar deviasi, dan rentang.

Sementara itu, analisis inferensial digunakan untuk menguji hipotesis dan menarik kesimpulan dari data sampel terhadap populasi. Uji statistik seperti *t-test*, *chi-square*, *ANOVA*, *regresi linear*, dan *regresi logistik* sering digunakan untuk menganalisis hubungan antarvariabel. Dalam penelitian epidemiologi, biostatistik juga digunakan untuk menghitung ukuran asosiasi seperti *relative risk (RR)*, *odds ratio (OR)*, dan *attributable risk (AR)*. Melalui analisis ini, peneliti dapat menentukan faktor risiko penyakit serta menilai efektivitas suatu intervensi medis (Sullivan, 2018).

1.2.3 Biostatistik dalam Penelitian Klinis dan Kesehatan Masyarakat

Dalam penelitian klinis, biostatistik menjadi alat utama untuk mengevaluasi keamanan dan efikasi intervensi medis, seperti obat, vaksin, atau prosedur terapi. Desain penelitian yang paling kuat untuk tujuan ini adalah *randomized controlled trial (RCT)*, di mana biostatistik digunakan untuk memastikan randomisasi berjalan baik, mengontrol bias, serta menilai signifikansi hasil.

Sementara itu, dalam kesehatan masyarakat, biostatistik berperan dalam surveilans penyakit, evaluasi program kesehatan, dan analisis tren epidemiologis. Data yang diperoleh dari sistem surveilans dianalisis untuk memantau pola penyakit dan menentukan intervensi yang tepat sasaran. Misalnya, analisis deret waktu (*time-series analysis*) digunakan untuk memprediksi lonjakan kasus penyakit menular, sedangkan analisis spasial membantu memetakan distribusi penyakit di wilayah tertentu.

1.2.4 Tantangan dan Relevansi Biostatistik di Era Digital

Perkembangan teknologi digital dan *big data* menghadirkan tantangan sekaligus peluang baru bagi biostatistik. Volume data kesehatan yang besar dari rekam medis elektronik, aplikasi kesehatan, dan data genomik menuntut metode analisis statistik yang lebih canggih. Integrasi antara biostatistik, *data science*, dan *machine learning* kini menjadi tren dalam analisis data kesehatan modern.

Meskipun demikian, penerapan biostatistik tetap harus memperhatikan prinsip-prinsip etika penelitian, terutama terkait

privasi data pasien dan interpretasi hasil. Penggunaan model statistik yang kompleks tanpa pemahaman yang memadai dapat menyebabkan kesalahan analisis dan keputusan yang menyesatkan. Oleh karena itu, kolaborasi antara peneliti, ahli statistik, dan tenaga kesehatan menjadi sangat penting dalam menjaga validitas ilmiah penelitian.

Dengan demikian, biostatistik bukan hanya sekadar alat analisis, tetapi juga merupakan kerangka berpikir ilmiah yang membantu peneliti memahami dinamika data kesehatan secara komprehensif. Penerapan biostatistik yang tepat akan menghasilkan penelitian yang lebih bermakna, beretika, dan bermanfaat dalam pengembangan kebijakan kesehatan berbasis bukti.

1.3 Jenis Data dan Skala Pengukuran

Dalam kajian *biostatistik*, pemahaman terhadap jenis data dan skala pengukuran merupakan langkah awal yang sangat penting sebelum melakukan analisis statistik. Setiap jenis data memiliki karakteristik dan perlakuan analitis yang berbeda, sehingga kesalahan dalam mengidentifikasi tipe data dapat menghasilkan interpretasi yang keliru. Oleh karena itu, klasifikasi data menjadi landasan metodologis dalam penelitian kesehatan, biologi, maupun ilmu sosial terapan (Rosner, 2016).

1.3.1 Jenis Data dalam Biostatistik

Secara umum, data dalam biostatistik dibedakan menjadi dua kelompok besar, yaitu data kualitatif (*qualitative data*) dan data kuantitatif (*quantitative data*).

1. Data Kualitatif (Kategorik)

Data kualitatif menggambarkan sifat, atribut, atau kategori yang tidak dapat dinyatakan dalam bentuk angka secara langsung. Jenis data ini biasanya digunakan untuk mengelompokkan individu atau objek ke dalam kategori tertentu. Contoh data kualitatif antara lain jenis kelamin (laki-laki atau perempuan), golongan darah (A, B, AB, O), dan status merokok (ya atau tidak).

Berdasarkan jumlah kategori dan hubungannya, data kualitatif dapat dibedakan menjadi:

- **Dikotomik (biner):** hanya memiliki dua kategori, misalnya “positif/negatif” atau “sehat/sakit”.
- **Politomik:** memiliki lebih dari dua kategori, misalnya “rendah”, “sedang”, dan “tinggi”.

Dalam analisis statistik, data kualitatif biasanya disajikan dalam bentuk distribusi frekuensi dan dianalisis menggunakan uji non-parametrik seperti *Chi-Square Test* atau *Fisher's Exact Test*.

2. Data Kuantitatif (Numerik)

Data kuantitatif menunjukkan besaran atau ukuran yang dapat dinyatakan dengan angka dan dapat dilakukan operasi matematika terhadapnya. Data ini sering digunakan untuk

menggambarkan variabel fisiologis seperti berat badan, tekanan darah, kadar glukosa, atau suhu tubuh.

Berdasarkan sifat nilainya, data kuantitatif dibagi menjadi dua jenis:

- **Data Diskrit (*Discrete Data*)**: diperoleh dari hasil perhitungan, hanya dapat mengambil nilai tertentu, misalnya jumlah anak atau jumlah sel darah putih.
- **Data Kontinu (*Continuous Data*)**: dapat mengambil nilai apa pun dalam rentang tertentu dan diperoleh dari hasil pengukuran, seperti tinggi badan atau kadar kolesterol.

Data kuantitatif biasanya dianalisis dengan uji parametrik seperti *t-test* atau *ANOVA* jika memenuhi asumsi distribusi normal.

1.3.2 Skala Pengukuran

Selain jenis data, skala pengukuran juga berperan penting dalam menentukan jenis analisis statistik yang sesuai. Menurut Stevens (1946), terdapat empat skala pengukuran utama yang digunakan dalam biostatistik, yaitu nominal, ordinal, interval, dan rasio.

1. Skala Nominal

Skala ini merupakan tingkat pengukuran paling sederhana, di mana data dikategorikan tanpa urutan atau peringkat tertentu. Contohnya adalah jenis kelamin, ras, dan golongan darah. Analisis statistik yang digunakan umumnya berupa penghitungan frekuensi atau persentase.

2. Skala Ordinal

Skala ini mengandung unsur urutan atau peringkat, namun jarak antar kategori tidak harus sama. Contohnya tingkat keparahan penyakit (ringan, sedang, berat) atau tingkat kepuasan pasien (puas, netral, tidak puas). Uji statistik yang sesuai untuk data ordinal antara lain *Mann–Whitney U Test* dan *Kruskal–Wallis Test*.

3. Skala Interval

Skala ini memiliki jarak antar nilai yang sama dan terukur, tetapi tidak memiliki titik nol absolut. Contohnya adalah suhu dalam derajat Celsius atau Fahrenheit. Operasi aritmetika seperti penjumlahan dan pengurangan dapat dilakukan, tetapi perbandingan rasio tidak bermakna.

4. Skala Rasio

Skala rasio memiliki semua karakteristik skala interval ditambah keberadaan titik nol absolut, yang memungkinkan perbandingan proporsional antar nilai. Contohnya berat badan, tinggi badan, atau tekanan darah. Skala ini mendukung hampir semua bentuk analisis statistik karena sifat datanya yang lengkap.

1.3.3 Implikasi Jenis dan Skala Data terhadap Analisis

Pemilihan metode statistik yang tepat bergantung pada jenis dan skala data yang digunakan. Misalnya, data berskala nominal tidak dapat dianalisis dengan uji parametrik karena tidak memenuhi asumsi distribusi normal, sedangkan data berskala rasio memungkinkan analisis inferensial yang lebih kompleks. Kesalahan

dalam mengklasifikasikan data dapat mengakibatkan kesimpulan yang tidak valid.

Selain itu, pemahaman terhadap skala pengukuran membantu peneliti dalam menentukan cara penyajian data yang sesuai—baik dalam bentuk tabel, diagram batang, maupun grafik histogram. Dengan demikian, identifikasi jenis dan skala data merupakan fondasi utama dalam proses penelitian biostatistik yang valid dan reliabel (Daniel & Cross, 2018).

1.4 Penerapan Biostatistik dalam Pelayanan dan Kebijakan Kesehatan

Biostatistik merupakan cabang ilmu yang mengombinasikan prinsip statistik dengan ilmu biomedis untuk menganalisis fenomena kesehatan secara kuantitatif. Dalam konteks pelayanan dan kebijakan kesehatan, biostatistik berfungsi sebagai instrumen ilmiah yang memastikan bahwa setiap keputusan didasarkan pada bukti empiris, bukan asumsi semata. Penggunaan analisis biostatistik yang tepat dapat meningkatkan efisiensi pelayanan kesehatan, memperkuat kebijakan publik, serta mendorong sistem kesehatan yang berorientasi pada hasil (*evidence-based health policy*) (Kurniawan & Yuliani, 2021).

1.4.1 Biostatistik dalam Analisis Epidemiologi dan Kebijakan Kesehatan

Salah satu penerapan paling mendasar dari biostatistik adalah dalam **analisis epidemiologi**, yang mencakup pengukuran insidensi,

prevalensi, mortalitas, dan morbiditas penyakit. Melalui analisis ini, pembuat kebijakan dapat menentukan prioritas intervensi dan alokasi sumber daya kesehatan.

Sebagai contoh, insidensi suatu penyakit digunakan untuk mengukur jumlah kasus baru dalam periode tertentu, yang menjadi dasar penentuan urgensi pencegahan. Sedangkan prevalensi memberikan gambaran beban penyakit dalam populasi, membantu perencanaan kebutuhan layanan kesehatan jangka panjang.

Biostatistik juga berperan dalam evaluasi efektivitas kebijakan kesehatan, seperti uji efektivitas vaksin, intervensi gizi, atau program skrining penyakit tidak menular. Analisis seperti *risk ratio*, *odds ratio*, dan *confidence interval* digunakan untuk menilai hubungan sebab-akibat antara faktor risiko dan outcome kesehatan.

Di tingkat nasional, hasil analisis statistik ini menjadi landasan penyusunan Rencana Pembangunan Kesehatan Nasional (RPKN) dan strategi prioritas Kementerian Kesehatan dalam menghadapi penyakit menular maupun kronik. Dengan demikian, biostatistik berperan sebagai jembatan antara data lapangan dan kebijakan publik berbasis bukti.

1.4.2 Biostatistik dalam Manajemen dan Evaluasi Pelayanan Kesehatan

Dalam konteks pelayanan kesehatan, biostatistik berfungsi untuk **meningkatkan mutu dan efisiensi pelayanan rumah sakit**. Setiap rumah sakit diwajibkan mengumpulkan data operasional seperti angka kunjungan pasien, lama rawat inap (*length of stay*),

tingkat keterisian tempat tidur (*bed occupancy rate*), angka infeksi nosokomial, serta mortalitas dan morbiditas pasien.

Analisis biostatistik terhadap indikator tersebut memungkinkan manajemen rumah sakit melakukan:

1. **Evaluasi kinerja unit pelayanan**, dengan membandingkan tren data antarperiode.
2. **Analisis mutu klinis**, melalui pengukuran *clinical outcomes* seperti tingkat keberhasilan tindakan operasi atau efektivitas terapi.
3. **Pengambilan keputusan berbasis data (data-driven decision making)**, misalnya dalam menentukan kebutuhan tenaga medis atau efisiensi penggunaan obat dan alat kesehatan.

Selain itu, survei biostatistik seperti **tingkat kepuasan pasien** membantu institusi kesehatan memahami persepsi masyarakat terhadap pelayanan yang diberikan. Hasil survei tersebut digunakan untuk memperbaiki standar pelayanan minimal (SPM) serta memastikan rumah sakit memenuhi akreditasi mutu pelayanan sesuai dengan ketentuan **Komisi Akreditasi Rumah Sakit (KARS)**.

1.4.3 Integrasi Biostatistik dalam Sistem Informasi dan Transformasi Digital Kesehatan

Transformasi digital di sektor kesehatan memperluas peran biostatistik melalui integrasi data dalam **Sistem Informasi Kesehatan (SIK)**. Sistem ini memungkinkan pengumpulan, analisis, dan pelaporan data kesehatan secara real-time, baik di fasilitas kesehatan primer maupun rujukan.

Penerapan **big data analytics** dan **machine learning** berbasis biostatistik kini digunakan dalam pemodelan penyakit, prediksi wabah, serta perencanaan sumber daya manusia kesehatan. Misalnya, model prediksi berbasis *regresi logistik* atau *time series analysis* membantu memprediksi pola penyebaran penyakit dan kebutuhan logistik medis di masa depan (Pratama & Lestari, 2020).

Selain itu, pemerintah melalui **Satu Data Kesehatan** berupaya menyatukan seluruh data layanan kesehatan dari berbagai instansi agar dapat digunakan untuk analisis lintas sektor. Dengan biostatistik sebagai fondasi analitik, data tersebut dapat diolah menjadi informasi strategis bagi pengambilan kebijakan nasional, baik di bidang pembiayaan kesehatan, distribusi fasilitas, maupun pemerataan pelayanan.

1.4.4 Tantangan dan Prospek Biostatistik dalam Kebijakan Publik

Meskipun potensinya besar, penerapan biostatistik dalam kebijakan kesehatan masih menghadapi beberapa tantangan, seperti kualitas data yang tidak seragam, kurangnya tenaga ahli statistik kesehatan, serta keterbatasan infrastruktur teknologi. Oleh karena itu, perlu dilakukan:

- **Peningkatan kapasitas tenaga kesehatan dalam analisis statistik**, melalui pelatihan atau kolaborasi dengan institusi akademik.
- **Penerapan standar nasional pengumpulan data kesehatan**, agar hasil analisis dapat dibandingkan antarwilayah.

- **Pemanfaatan teknologi informasi yang adaptif dan aman**, untuk mendukung integrasi data antar lembaga.

Dengan mengatasi tantangan tersebut, biostatistik dapat berfungsi optimal sebagai alat ilmiah dalam mendukung pembangunan kesehatan nasional yang berbasis bukti, efisien, dan berorientasi pada hasil jangka panjang.

1.4.5 Kesimpulan

Biostatistik memegang peranan sentral dalam sistem pelayanan dan kebijakan kesehatan modern. Melalui analisis epidemiologi, evaluasi mutu pelayanan, dan integrasi data digital, biostatistik tidak hanya membantu pengambilan keputusan klinis, tetapi juga memastikan kebijakan publik didasarkan pada bukti yang valid dan terukur. Dengan penguatan sumber daya manusia dan transformasi digital, biostatistik akan semakin menjadi fondasi utama dalam pembangunan sistem kesehatan nasional yang efektif, akuntabel, dan berkeadilan.

1.5 Latihan Soal

1. Jelaskan pengertian biostatistik dan perannya dalam penelitian kesehatan!
2. Sebutkan dan jelaskan empat skala pengukuran dalam biostatistik beserta contohnya!
3. Mengapa biostatistik penting dalam pembuatan kebijakan kesehatan masyarakat?

4. Berikan contoh penerapan biostatistik dalam bidang pelayanan rumah sakit!
5. Bedakan antara data kualitatif dan data kuantitatif dalam konteks penelitian kesehatan!

Bab 2: Konsep Dasar Data dan Statistik Deskriptif

2.1 Pengertian Data dan Statistik

Data dan statistik merupakan dua konsep yang saling berkaitan dan menjadi fondasi utama dalam proses ilmiah, khususnya dalam bidang biostatistik dan penelitian kesehatan. Data dapat didefinisikan sebagai sekumpulan informasi mentah yang diperoleh melalui hasil observasi, pengukuran, eksperimen, atau survei, yang kemudian diolah untuk memperoleh makna tertentu. Dalam konteks biostatistik, data memiliki peran strategis sebagai dasar bagi pengambilan keputusan berbasis bukti (*evidence-based decision making*), baik dalam penelitian ilmiah, kebijakan kesehatan, maupun praktik klinis. Statistik, di sisi lain, adalah ilmu yang mempelajari cara mengumpulkan, menyajikan, menganalisis, dan menafsirkan data sehingga dapat digunakan untuk memahami dan menjelaskan fenomena secara objektif. Dengan kata lain, statistik berfungsi sebagai alat untuk mengubah data mentah menjadi informasi yang bermakna dan dapat digunakan sebagai dasar pengambilan keputusan yang rasional dan terukur.

Secara epistemologis, data dalam biostatistik tidak hanya dipandang sebagai angka, tetapi juga sebagai representasi dari fenomena biologis, sosial, dan medis yang kompleks. Setiap data

memiliki konteks yang berkaitan dengan sumber, metode pengumpulan, dan tujuan analisisnya. Misalnya, data dalam penelitian epidemiologi dapat berupa jumlah kasus penyakit, tingkat mortalitas, atau hasil laboratorium, sedangkan dalam penelitian fisioterapi dapat mencakup pengukuran kekuatan otot, rentang gerak sendi, dan tingkat nyeri pasien. Kualitas data menjadi faktor penentu dalam validitas hasil penelitian; data yang akurat, reliabel, dan representatif akan menghasilkan kesimpulan yang dapat dipercaya. Oleh karena itu, proses pengumpulan data harus dilakukan dengan metode ilmiah yang sistematis, termasuk penentuan sampel yang tepat, penggunaan instrumen yang terstandar, serta pengendalian terhadap bias dan kesalahan pengukuran (Sullivan, 2018).

Statistik sebagai ilmu memiliki dua cabang utama, yaitu statistik deskriptif dan statistik inferensial. Statistik deskriptif berfungsi untuk menyajikan dan meringkas data dalam bentuk yang mudah dipahami, seperti tabel, grafik, rata-rata, median, modus, serta ukuran dispersi seperti varians dan simpangan baku. Tujuannya adalah memberikan gambaran umum tentang karakteristik data tanpa melakukan generalisasi terhadap populasi yang lebih luas. Sebagai contoh, dalam penelitian tentang efektivitas intervensi fisioterapi terhadap pasien pascaoperasi lutut, statistik deskriptif dapat digunakan untuk menggambarkan usia rata-rata pasien, tingkat pemulihan, serta distribusi skor mobilitas. Sementara itu, statistik inferensial digunakan untuk menarik kesimpulan atau membuat generalisasi tentang populasi berdasarkan data sampel. Melalui teknik seperti *t-test*, *ANOVA*, regresi linier, atau uji korelasi, peneliti

dapat menilai hubungan antarvariabel dan menguji hipotesis secara ilmiah. Dengan demikian, statistik inferensial berperan penting dalam menjawab pertanyaan penelitian dan mendukung pengambilan keputusan klinis yang berbasis data (Peat & Barton, 2020).

Dalam bidang biostatistik, data dapat diklasifikasikan menjadi beberapa jenis berdasarkan sifat dan skala pengukurannya. Berdasarkan sifatnya, data dibagi menjadi data kualitatif (non-numerik) dan data kuantitatif (numerik). Data kualitatif menggambarkan kategori atau karakteristik, seperti jenis kelamin, kelompok usia, atau status kesehatan, sedangkan data kuantitatif menunjukkan nilai numerik yang dapat diukur, seperti berat badan, tekanan darah, atau kadar glukosa darah. Berdasarkan skala pengukuran, data diklasifikasikan ke dalam empat tingkat: nominal, ordinal, interval, dan rasio. Skala nominal menunjukkan kategori tanpa urutan, seperti golongan darah, sedangkan skala ordinal menunjukkan urutan tetapi tidak memiliki jarak yang tetap, misalnya tingkat nyeri (ringan, sedang, berat). Skala interval memiliki jarak yang sama antar nilai tetapi tidak memiliki titik nol absolut, seperti suhu dalam derajat Celsius, sedangkan skala rasio memiliki semua karakteristik tersebut dengan titik nol absolut, misalnya berat badan atau tinggi badan. Pemahaman terhadap jenis dan skala data sangat penting karena menentukan jenis analisis statistik yang dapat digunakan.

Keterampilan dalam menggunakan data dan statistik menjadi aspek penting bagi tenaga kesehatan, termasuk fisioterapis, karena

pengambilan keputusan klinis yang efektif harus didasarkan pada bukti ilmiah yang dapat diukur dan diverifikasi. Misalnya, dalam menentukan efektivitas suatu modalitas terapi, fisioterapis harus mampu menganalisis hasil pengukuran pasien secara statistik untuk memastikan bahwa perubahan yang terjadi bukan semata akibat kebetulan, tetapi merupakan hasil nyata dari intervensi yang diberikan. Penggunaan statistik yang tepat memungkinkan interpretasi data secara objektif, sehingga mengurangi bias subjektif dalam penilaian klinis. Selain itu, statistik juga membantu dalam menilai validitas dan reliabilitas instrumen pengukuran, serta memastikan bahwa hasil penelitian dapat diterapkan secara luas dalam praktik profesional.

Di era digital dan *big data*, peran statistik dalam bidang kesehatan menjadi semakin penting. Teknologi informasi memungkinkan pengumpulan data dalam jumlah besar melalui sistem elektronik seperti *electronic health records* (EHR) dan aplikasi kesehatan berbasis sensor. Tantangannya terletak pada bagaimana mengelola, menganalisis, dan menginterpretasikan data tersebut agar dapat memberikan wawasan yang bermanfaat. Pendekatan modern seperti *data mining*, *machine learning*, dan *predictive analytics* kini digunakan untuk mengidentifikasi pola, tren, dan prediksi risiko kesehatan secara lebih akurat. Dalam konteks ini, statistik menjadi jembatan antara data mentah dan pengambilan keputusan berbasis bukti, yang dapat meningkatkan efisiensi diagnosis, efektivitas terapi, dan kualitas pelayanan kesehatan secara keseluruhan.

Dengan demikian, data dan statistik memiliki peran esensial dalam membangun dasar ilmiah bagi penelitian dan praktik di bidang kesehatan. Data memberikan gambaran empiris tentang fenomena yang diamati, sedangkan statistik menyediakan alat untuk memahami dan menafsirkan makna di balik data tersebut. Keduanya saling melengkapi dalam menghasilkan pengetahuan yang valid dan dapat diandalkan. Pemahaman yang baik tentang konsep, jenis, serta teknik analisis data menjadi syarat mutlak bagi setiap peneliti dan praktisi kesehatan yang ingin berkontribusi dalam pengembangan ilmu dan peningkatan kualitas pelayanan. Oleh karena itu, penguasaan biostatistik bukan hanya keterampilan teknis, tetapi juga bagian dari tanggung jawab profesional dalam memastikan bahwa setiap keputusan medis didasarkan pada bukti ilmiah yang kuat dan objektif.

2.2 Jenis dan Skala Data

Dalam biostatistik, pemahaman mengenai jenis dan skala data merupakan langkah awal yang fundamental sebelum melakukan analisis statistik. Klasifikasi data menentukan teknik analisis yang tepat serta interpretasi hasil yang valid. Data yang digunakan dalam penelitian kesehatan atau fisioterapi sering kali berasal dari pengukuran, observasi, maupun wawancara, sehingga memiliki karakteristik yang beragam baik secara kualitatif maupun kuantitatif (Pagano & Gauvreau, 2018).

Secara umum, data dalam biostatistik dibagi menjadi dua kategori besar — **data kualitatif** dan **data kuantitatif** — dengan empat skala pengukuran utama yang menentukan metode analisis yang dapat digunakan, yaitu **nominal**, **ordinal**, **interval**, dan **rasio**.

2.2.1 Jenis Data dalam Biostatistik

1. Data Kualitatif (*Qualitative Data*)

Data kualitatif menggambarkan karakteristik atau kategori yang tidak dinyatakan dalam bentuk angka. Data ini digunakan untuk mengelompokkan individu atau objek berdasarkan atribut tertentu, seperti jenis kelamin, golongan darah, atau tingkat kepuasan pasien.

Data kualitatif dibagi menjadi dua bentuk utama:

- **Data Nominal:** Data bersifat kategorikal tanpa urutan logis. Contoh: jenis kelamin (laki-laki/perempuan), status merokok (ya/tidak), atau golongan darah (A, B, AB, O).
- **Data Ordinal:** Data kategorikal yang memiliki urutan atau tingkatan, tetapi jarak antar kategori tidak dapat diukur secara kuantitatif. Contoh: tingkat nyeri (ringan, sedang, berat) atau status gizi (kurang, normal, lebih).

Data kualitatif banyak digunakan dalam penelitian kesehatan masyarakat dan fisioterapi untuk menggambarkan profil populasi atau persepsi pasien terhadap intervensi tertentu. Analisis yang digunakan umumnya bersifat deskriptif, seperti distribusi frekuensi dan persentase.

2. Data Kuantitatif (*Quantitative Data*)

Data kuantitatif menggambarkan hasil pengukuran yang dinyatakan dalam angka dan memungkinkan dilakukan perhitungan matematis. Jenis data ini sering digunakan untuk menganalisis hubungan atau perbedaan antar variabel dalam penelitian kesehatan.

Data kuantitatif dibagi menjadi dua bentuk utama:

- **Data Diskret (*Discrete Data*):** Data berupa bilangan bulat hasil penghitungan. Contoh: jumlah pasien, jumlah denyut jantung per menit, atau jumlah anak dalam keluarga.
- **Data Kontinu (*Continuous Data*):** Data hasil pengukuran yang dapat memiliki nilai dalam rentang tertentu. Contoh: berat badan, tinggi badan, tekanan darah, dan kadar hemoglobin.

Data kuantitatif memungkinkan penerapan analisis statistik inferensial seperti uji *t*, ANOVA, atau regresi linier.

2.2.2 Skala Pengukuran Data

Skala pengukuran adalah sistem klasifikasi yang menunjukkan hubungan matematis antara nilai data. Pemilihan skala yang tepat penting untuk menentukan metode analisis statistik yang sesuai (Dawson & Trapp, 2020).

1. Skala Nominal

Skala nominal digunakan untuk mengelompokkan data ke dalam kategori tanpa urutan atau nilai numerik. Data pada skala ini

hanya dapat dianalisis melalui penghitungan frekuensi atau proporsi.

- **Contoh:** jenis kelamin (1 = laki-laki, 2 = perempuan), status vaksinasi (ya/tidak).
- **Analisis yang sesuai:** uji *chi-square* atau tabel kontingensi.

2. Skala Ordinal

Skala ordinal menunjukkan urutan atau peringkat, tetapi jarak antar kategori tidak harus sama. Misalnya, selisih antara “ringan” dan “sedang” tidak dapat diasumsikan sama dengan antara “sedang” dan “berat.”

- **Contoh:** tingkat kepuasan pasien (puas, cukup puas, tidak puas).
- **Analisis yang sesuai:** uji nonparametrik seperti *Mann-Whitney U* atau *Kruskal-Wallis*.

3. Skala Interval

Skala interval memiliki urutan dan jarak yang sama antara nilai, tetapi tidak memiliki nol absolut. Nol pada skala ini tidak berarti “tidak ada.”

- **Contoh:** suhu tubuh dalam °C atau °F. Nol derajat bukan berarti tidak ada suhu.
- **Analisis yang sesuai:** uji *t*, ANOVA, dan analisis korelasi Pearson.

4. Skala Rasio

Skala rasio mencakup semua karakteristik skala interval, tetapi memiliki nol absolut yang menunjukkan ketiadaan nilai. Oleh karena itu, perbandingan antar nilai dapat dilakukan.

- **Contoh:** berat badan, tinggi badan, tekanan darah, kadar gula darah.
- **Analisis yang sesuai:** semua metode statistik parametrik dapat digunakan.

2.2.3 Hubungan antara Jenis dan Skala Data

Hubungan antara jenis dan skala data bersifat hierarkis. Data kualitatif umumnya menggunakan skala nominal dan ordinal, sedangkan data kuantitatif menggunakan skala interval dan rasio. Dalam penelitian fisioterapi dan kesehatan, pemahaman terhadap hubungan ini membantu dalam:

- Menentukan alat ukur yang tepat;
- Memilih metode analisis statistik yang sesuai;
- Menghindari kesalahan interpretasi hasil penelitian.

Sebagai contoh, dalam studi rehabilitasi pasien stroke, data ordinal dapat digunakan untuk menilai tingkat kemampuan fungsional dengan skala *Functional Independence Measure (FIM)*, sedangkan data rasio dapat digunakan untuk mengukur kekuatan otot menggunakan dinamometer.

2.2.4 Implikasi Praktis dalam Penelitian Kesehatan

Pemilihan jenis dan skala data yang tepat sangat memengaruhi validitas dan reliabilitas hasil penelitian. Kesalahan dalam pengklasifikasian data dapat menyebabkan pemilihan uji

statistik yang tidak sesuai, sehingga hasil penelitian menjadi bias atau tidak dapat digeneralisasikan.

Oleh karena itu, peneliti harus memahami bahwa setiap variabel memiliki karakteristik unik yang menentukan metode pengukuran dan analisisnya. Dalam praktik fisioterapi berbasis bukti, keakuratan data menjadi dasar pengambilan keputusan klinis yang rasional dan ilmiah.

2.3 Statistik Deskriptif

Statistik deskriptif merupakan cabang dari ilmu statistik yang berfungsi untuk menggambarkan, meringkas, dan menyajikan data agar mudah dipahami. Tujuan utama dari statistik deskriptif adalah memberikan gambaran umum mengenai karakteristik data yang dikumpulkan, baik melalui tabel, grafik, maupun ukuran numerik. Dengan demikian, statistik deskriptif berperan sebagai langkah awal yang penting sebelum peneliti melakukan analisis lanjutan menggunakan metode inferensial.

Melalui penerapan statistik deskriptif, peneliti dapat memperoleh informasi dasar seperti ukuran pemusatan data (rata-rata, median, dan modus), ukuran penyebaran (simpangan baku dan varians), serta bentuk distribusi data. Analisis ini membantu dalam memahami kecenderungan data dan mendeteksi adanya keanehan atau *outlier* yang dapat memengaruhi hasil penelitian secara keseluruhan (Field, 2018).

2.3.1 Ukuran Pemusatan Data

Ukuran pemusatan data menggambarkan nilai yang mewakili keseluruhan kumpulan data. Tiga ukuran utama yang sering digunakan adalah rata-rata (*mean*), median, dan modus.

Rata-rata menunjukkan nilai tengah aritmetika dari seluruh data, yang diperoleh dengan menjumlahkan seluruh nilai dan membaginya dengan jumlah observasi. Median adalah nilai tengah yang memisahkan data menjadi dua bagian sama besar, sehingga tidak terpengaruh oleh nilai ekstrem. Sementara itu, modus adalah nilai yang paling sering muncul dalam suatu himpunan data dan sering digunakan dalam data kategorik.

Ketiga ukuran ini berfungsi untuk memahami kecenderungan umum dari data, dan pemilihannya bergantung pada jenis data serta distribusinya. Sebagai contoh, pada data yang terdistribusi normal, *mean* merupakan ukuran pemusatan terbaik, sedangkan pada data yang memiliki sebaran tidak simetris, median lebih tepat digunakan untuk menggambarkan nilai tengah populasi.

2.3.2 Ukuran Penyebaran dan Variabilitas

Selain pemusatan, statistik deskriptif juga mencakup ukuran penyebaran yang menggambarkan seberapa jauh data menyebar dari nilai pusatnya. Ukuran penyebaran yang umum digunakan meliputi rentang (*range*), varians, dan simpangan baku (*standard deviation*).

Rentang menunjukkan perbedaan antara nilai tertinggi dan terendah dalam data, memberikan gambaran kasar tentang lebar distribusi. Varians mengukur rata-rata kuadrat dari deviasi setiap data terhadap *mean*, sedangkan simpangan baku merupakan akar

kuadrat dari varians yang menyatakan penyebaran data dalam satuan yang sama dengan data aslinya (Gravetter & Wallnau, 2020).

Nilai simpangan baku yang kecil menunjukkan bahwa data cenderung homogen dan berdekatan dengan nilai rata-rata, sedangkan simpangan baku yang besar mengindikasikan bahwa data lebih bervariasi. Analisis ini penting dalam menentukan tingkat keandalan data dan relevansi generalisasi hasil penelitian.

2.3.3 Penyajian Data dalam Bentuk Visual

Selain ukuran numerik, statistik deskriptif juga menekankan pentingnya penyajian data dalam bentuk visual, seperti tabel, diagram batang, diagram lingkaran, atau histogram. Penyajian ini membantu peneliti mengidentifikasi pola, tren, dan distribusi data secara cepat dan intuitif.

Histogram, misalnya, sering digunakan untuk menampilkan distribusi frekuensi data kuantitatif dan memudahkan pengamatan bentuk distribusi, apakah normal, miring ke kiri, atau miring ke kanan. Sementara itu, diagram batang dan diagram lingkaran efektif digunakan untuk data kategorik dalam menunjukkan perbandingan antar kelompok.

Dengan memanfaatkan visualisasi data, peneliti tidak hanya dapat memperjelas interpretasi hasil, tetapi juga meningkatkan keterbacaan dan validitas penyampaian informasi kepada audiens penelitian.

2.3.4 Peran Statistik Deskriptif dalam Penelitian

Statistik deskriptif memiliki peran penting dalam setiap tahap penelitian kuantitatif. Pada tahap awal, ia berfungsi untuk

memeriksa kualitas data, seperti mendeteksi kesalahan entri atau data yang hilang. Selanjutnya, hasil deskriptif membantu peneliti memahami struktur dan pola umum data yang menjadi dasar untuk memilih metode analisis inferensial yang tepat.

Dalam laporan penelitian, penyajian statistik deskriptif menjadi bagian penting untuk memberikan konteks terhadap hasil analisis lanjutan. Dengan demikian, penggunaan statistik deskriptif bukan sekadar tahap awal analisis, tetapi juga merupakan fondasi utama dalam proses interpretasi ilmiah terhadap fenomena yang diteliti

2.4 Penyajian Data

Penyajian data merupakan tahap penting dalam proses analisis statistik karena berfungsi untuk menyederhanakan, mengilustrasikan, dan menginterpretasikan hasil penelitian secara sistematis dan mudah dipahami. Data yang telah dikumpulkan dari hasil observasi, eksperimen, atau survei perlu disusun dalam bentuk yang terstruktur sehingga dapat menunjukkan pola, hubungan, dan kecenderungan yang signifikan. Dengan penyajian yang baik, informasi kompleks dapat diubah menjadi representasi visual yang informatif dan komunikatif, memudahkan pembaca atau peneliti lain dalam memahami hasil penelitian secara cepat dan akurat.

Pemilihan bentuk penyajian data harus disesuaikan dengan jenis data (kualitatif atau kuantitatif) dan skala pengukurannya (nominal, ordinal, interval, atau rasio). Kesalahan dalam

menentukan metode penyajian dapat mengaburkan makna data dan menurunkan validitas interpretasi statistik.

2.4.1 Tabel Distribusi Frekuensi

Tabel distribusi frekuensi merupakan bentuk penyajian data paling dasar dan banyak digunakan dalam statistik deskriptif. Tabel ini menampilkan jumlah kejadian (frekuensi) dari setiap kategori atau interval data, sehingga pola distribusi dapat diamati dengan mudah. Untuk data kuantitatif, kelas-kelas interval digunakan agar distribusi lebih teratur dan mudah dibandingkan.

Sebagai contoh, dalam penelitian mengenai indeks massa tubuh (IMT) mahasiswa, tabel distribusi frekuensi dapat digunakan untuk menampilkan jumlah responden berdasarkan kategori berat badan — kurus, normal, gemuk, dan obesitas. Penyajian semacam ini memungkinkan peneliti untuk melihat kecenderungan umum tanpa harus menelaah seluruh data mentah (Gravetter & Wallnau, 2017).

Selain itu, tabel distribusi dapat dilengkapi dengan kolom frekuensi relatif dan persentase untuk menunjukkan proporsi setiap kategori terhadap keseluruhan sampel. Penggunaan tabel ini membantu peneliti dalam membandingkan kelompok data dan mengidentifikasi pola penyebaran nilai.

2.4.2 Diagram Batang dan Diagram Lingkaran

Diagram batang (*bar chart*) digunakan untuk menyajikan data kategorik atau ordinal dengan menampilkan panjang batang yang proporsional terhadap frekuensi masing-masing kategori.

Diagram ini efektif untuk membandingkan antarvariabel, seperti jumlah peserta berdasarkan jenis kelamin atau tingkat pendidikan.

Sementara itu, diagram lingkaran (*pie chart*) cocok untuk menampilkan data proporsi dalam satu kesatuan total 100%. Setiap sektor pada diagram lingkaran mewakili persentase dari kategori tertentu. Meskipun populer, diagram lingkaran sebaiknya digunakan hanya untuk data dengan jumlah kategori yang terbatas agar tidak membingungkan pembaca (Weiss, 2018).

Kedua bentuk diagram ini memberikan visualisasi yang intuitif dan sering digunakan dalam laporan penelitian, publikasi ilmiah, maupun penyajian data populasi.

2.4.3 Histogram dan Pola Distribusi

Histogram digunakan untuk menyajikan data kuantitatif berskala interval atau rasio. Berbeda dengan diagram batang, histogram menggambarkan data dalam bentuk batang yang bersebelahan tanpa jarak, yang menandakan kesinambungan antarinterval. Histogram sangat berguna untuk menunjukkan bentuk distribusi data — apakah normal, miring ke kiri, atau miring ke kanan.

Sebagai contoh, histogram nilai ujian dapat memperlihatkan apakah distribusi data mengikuti pola normal atau terdapat outlier yang signifikan. Bentuk histogram juga membantu dalam menentukan langkah analisis lanjutan, seperti uji normalitas atau transformasi data. Dengan demikian, histogram tidak hanya berfungsi sebagai alat visualisasi, tetapi juga sebagai alat diagnostik dalam analisis statistik.

2.4.4 Boxplot dan Analisis Sebaran Data

Boxplot atau diagram kotak-garis merupakan salah satu bentuk visualisasi yang efektif untuk menggambarkan penyebaran dan sebaran data kuantitatif. Diagram ini menampilkan nilai minimum, kuartil pertama (Q1), median, kuartil ketiga (Q3), dan nilai maksimum dalam satu tampilan. Dengan demikian, *boxplot* dapat digunakan untuk mendeteksi keberadaan *outlier* atau anomali dalam data.

Keunggulan *boxplot* adalah kemampuannya menampilkan perbandingan antar kelompok secara kompak. Misalnya, dalam penelitian fisioterapi, *boxplot* dapat digunakan untuk membandingkan hasil kekuatan otot antara dua kelompok perlakuan. Melalui perbedaan panjang kotak dan posisi median, peneliti dapat melihat variasi data antar kelompok secara visual.

2.4.5 Prinsip Pemilihan Bentuk Penyajian Data

Pemilihan bentuk penyajian data harus memperhatikan kesesuaian antara tujuan analisis dan jenis data yang digunakan. Untuk data kategorik, diagram batang dan lingkaran lebih efektif; sedangkan untuk data kontinu, histogram dan *boxplot* memberikan gambaran yang lebih informatif.

Selain itu, aspek estetika dan keterbacaan juga perlu dipertimbangkan. Penggunaan warna, skala, dan label yang konsisten dapat meningkatkan kejelasan informasi. Data yang disajikan harus tetap netral dan tidak dimanipulasi secara visual untuk mendukung kesimpulan tertentu. Dengan demikian, integritas

penyajian data menjadi bagian dari etika ilmiah yang harus dijunjung tinggi dalam setiap penelitian.

2.4.6 Kesimpulan

Penyajian data merupakan langkah krusial dalam mengomunikasikan hasil penelitian secara efektif. Melalui tabel, diagram, histogram, atau *boxplot*, peneliti dapat menggambarkan distribusi, pola, dan hubungan antarvariabel secara jelas. Pemilihan bentuk penyajian harus disesuaikan dengan karakteristik data dan tujuan analisis agar hasil interpretasi tetap akurat dan objektif. Dengan penerapan prinsip penyajian yang rasional dan etis, data dapat menjadi dasar yang kuat dalam mendukung kesimpulan ilmiah dan pengambilan keputusan berbasis bukti.

2.5 Latihan Soal

1. Jelaskan perbedaan antara data kualitatif dan kuantitatif!
2. Sebutkan dan jelaskan empat skala pengukuran data
3. Apa fungsi utama statistik deskriptif dalam penelitian kesehatan?
4. Berikan contoh bentuk penyajian data untuk variabel kategorik!
5. Mengapa pemilihan bentuk grafik penting dalam analisis data?

Bab 3: Distribusi

Probabilitas dan Penerapan

Peluang dalam Kesehatan

3.1 Konsep Dasar Probabilitas

Probabilitas merupakan konsep fundamental dalam statistika yang digunakan untuk mengukur tingkat kemungkinan terjadinya suatu peristiwa acak. Dalam konteks biostatistik, probabilitas memiliki peran yang sangat penting karena banyak fenomena dalam bidang kesehatan yang bersifat tidak pasti atau acak, seperti timbulnya suatu penyakit, hasil dari pemeriksaan laboratorium, atau keberhasilan suatu terapi. Probabilitas dinyatakan dalam rentang antara 0 hingga 1, di mana nilai 0 menunjukkan peristiwa yang tidak mungkin terjadi, sedangkan nilai 1 menunjukkan peristiwa yang pasti terjadi. Konsep ini memungkinkan pengambilan keputusan yang lebih objektif dan terukur dalam menghadapi ketidakpastian di dunia medis maupun penelitian ilmiah. Melalui pendekatan probabilistik, profesional kesehatan dapat mengembangkan pemahaman yang lebih akurat terhadap risiko dan ketidakpastian yang melekat pada setiap tindakan medis.

Dalam praktiknya, probabilitas digunakan sebagai dasar berbagai metode analisis statistik yang umum diterapkan dalam

biostatistik, seperti pengujian hipotesis, estimasi parameter populasi, serta model prediktif. Misalnya, ketika menilai efektivitas suatu obat, peneliti tidak hanya mengamati berapa banyak pasien yang sembuh, tetapi juga menghitung kemungkinan keberhasilan terapi dibandingkan dengan kelompok kontrol. Konsep probabilitas membantu menjawab pertanyaan apakah perbedaan yang diamati bersifat nyata atau hanya disebabkan oleh fluktuasi acak. Pengujian hipotesis, seperti uji t atau uji chi-square, didasarkan pada probabilitas terjadinya hasil pengamatan jika asumsi nol (*null hypothesis*) benar. Dengan demikian, probabilitas menjadi dasar logika inferensial yang memungkinkan generalisasi dari sampel ke populasi secara ilmiah (Pagano & Gauvreau, 2018).

Lebih jauh lagi, pemahaman tentang probabilitas membantu tenaga kesehatan dalam membuat keputusan klinis berbasis bukti. Dalam dunia kedokteran dan kesehatan masyarakat, informasi mengenai risiko relatif, angka kejadian, dan prediksi hasil terapi sangat bergantung pada analisis probabilistik. Misalnya, ketika seorang dokter mempertimbangkan apakah akan memberikan terapi tertentu kepada pasien, ia harus menimbang probabilitas manfaat dibandingkan dengan kemungkinan efek samping. Dengan alat statistik yang didasarkan pada teori probabilitas, seperti *likelihood ratio*, *predictive value*, dan *Bayesian analysis*, pengambilan keputusan dapat dilakukan secara lebih rasional dan sistematis. Oleh karena itu, keterampilan dalam memahami dan mengaplikasikan konsep probabilitas menjadi kompetensi penting bagi para profesional kesehatan di era data saat ini (Pepe, 2003).

Pentingnya konsep probabilitas juga tercermin dalam pengembangan model epidemiologi dan prediksi penyebaran penyakit. Dengan menggunakan data historis dan algoritma statistik, para ahli kesehatan masyarakat dapat memperkirakan kemungkinan wabah di masa mendatang dan mengembangkan strategi pencegahan yang lebih efektif. Sebagai contoh, dalam pemodelan penyebaran penyakit menular seperti influenza atau COVID-19, digunakan probabilitas transmisi antarindividu untuk memproyeksikan jumlah kasus yang mungkin terjadi dalam periode tertentu. Probabilitas juga digunakan untuk menilai akurasi tes diagnostik, seperti sensitivitas dan spesifisitas, yang keduanya bergantung pada perhitungan probabilistik. Hal ini menunjukkan bahwa pemahaman terhadap probabilitas tidak hanya relevan dalam konteks akademik, tetapi juga dalam praktik lapangan yang berdampak langsung terhadap kesehatan masyarakat.

Secara keseluruhan, probabilitas bukan sekadar angka yang menggambarkan kemungkinan, melainkan kerangka berpikir yang memungkinkan para profesional kesehatan mengelola ketidakpastian secara sistematis. Dengan memahami konsep ini, tenaga medis dan peneliti dapat menilai risiko secara lebih akurat, menyusun strategi intervensi yang efektif, serta meningkatkan kualitas pengambilan keputusan klinis dan kebijakan kesehatan. Dalam era di mana data menjadi sumber utama informasi, kemampuan untuk menginterpretasikan probabilitas menjadi sangat krusial untuk mendukung pendekatan berbasis bukti dan praktik kesehatan yang lebih aman, efisien, dan tepat sasaran.

3.2 Distribusi Probabilitas Diskret

Distribusi probabilitas diskret merupakan fondasi penting dalam statistika kesehatan karena mampu menggambarkan peluang dari suatu kejadian yang dapat dihitung secara eksak. Dalam konteks biostatistik, distribusi ini digunakan untuk mewakili fenomena-fenomena seperti jumlah pasien yang menunjukkan gejala klinis tertentu, jumlah kematian akibat penyakit infeksi dalam satu minggu, atau jumlah kelahiran bayi dengan berat badan rendah. Ciri utama dari distribusi diskret adalah bahwa variabel acaknya hanya dapat mengambil nilai-nilai bulat yang terpisah, dan masing-masing nilai memiliki probabilitas yang dapat dihitung.

Tiga bentuk distribusi probabilitas diskret yang paling banyak digunakan dalam biostatistik adalah distribusi binomial, distribusi Poisson, dan distribusi hipergeometrik. Masing-masing distribusi memiliki asumsi dan kondisi penggunaan yang berbeda, tergantung pada struktur dan konteks data yang dianalisis.

3.2.1 Distribusi Binomial

Distribusi binomial digunakan ketika suatu eksperimen memiliki dua kemungkinan hasil yang saling eksklusif, yaitu berhasil atau gagal. Dalam konteks kesehatan, distribusi ini sering digunakan untuk menganalisis jumlah keberhasilan dalam sejumlah percobaan tetap, seperti jumlah pasien yang sembuh setelah diberi obat tertentu dari total pasien yang dirawat.

Distribusi ini bergantung pada dua parameter utama: jumlah percobaan (n) dan probabilitas keberhasilan dalam satu percobaan

(p). Nilai probabilitas dari suatu jumlah keberhasilan tertentu dalam n percobaan dihitung menggunakan rumus binomial yang melibatkan kombinasi dan pangkat dari p dan $(1-p)$. Menurut Pagano dan Gauvreau (2018), distribusi binomial sangat berguna untuk menguji hipotesis proporsi populasi dalam uji statistik, terutama jika ukuran sampel tidak terlalu besar dan asumsi independensi antar percobaan terpenuhi.

Contoh penggunaannya adalah untuk menghitung kemungkinan terdapat lima pasien positif dari sepuluh pasien yang diperiksa, dengan asumsi probabilitas infeksi adalah 0,3 per pasien. Interpretasi hasil distribusi ini sangat penting dalam pengambilan keputusan klinis dan evaluasi efektivitas intervensi medis.

3.2.2 Distribusi Poisson

Distribusi Poisson digunakan untuk memodelkan jumlah kejadian yang jarang terjadi dalam suatu interval waktu atau ruang tertentu. Dalam biostatistik, distribusi ini sering diterapkan untuk menghitung jumlah infeksi, komplikasi pascaoperasi, atau kematian mendadak yang terjadi dalam periode waktu terbatas.

Distribusi ini hanya memiliki satu parameter utama, yaitu λ (lambda), yang menyatakan rata-rata jumlah kejadian dalam satu interval. Karakteristik unik dari distribusi ini adalah bahwa variansnya sama dengan nilai rata-ratanya, sehingga distribusi ini cocok digunakan ketika data menunjukkan tingkat penyebaran yang serupa dengan nilai tengahnya.

Contoh penerapannya adalah menghitung jumlah kasus demam berdarah di satu wilayah dalam waktu satu bulan. Jika rata-

rata kasus per bulan adalah 4, maka distribusi Poisson dapat digunakan untuk menghitung peluang terjadinya tepat lima kasus dalam bulan tertentu. Menurut Rosner (2020), distribusi Poisson sangat efektif untuk pengamatan epidemiologi yang bersifat sporadis dan independen dari satu kejadian ke kejadian lain.

3.2.3 Distribusi Hipergeometrik

Distribusi hipergeometrik digunakan dalam situasi di mana pengambilan sampel dilakukan tanpa pengembalian dari populasi terbatas, dan tujuan analisis adalah menghitung probabilitas dari jumlah item dengan karakteristik tertentu dalam sampel. Dalam praktik biostatistik, distribusi ini dapat digunakan untuk menghitung jumlah pasien dengan penyakit tertentu dalam sampel yang diambil dari populasi rumah sakit, tanpa mengulang pemilihan individu.

Berbeda dengan distribusi binomial yang mengasumsikan percobaan independen dan probabilitas tetap, distribusi hipergeometrik memperhitungkan perubahan probabilitas akibat pengambilan tanpa pengembalian. Oleh karena itu, distribusi ini lebih realistis dalam konteks survei populasi kecil, audit klinis, atau pengujian batch vaksin yang jumlahnya terbatas.

Aplikasi distribusi ini penting dalam evaluasi mutu, misalnya untuk mengetahui kemungkinan jumlah cacat dari sejumlah vial obat dalam satu paket yang dipilih secara acak. Karena distribusi ini memperhitungkan seluruh ukuran populasi dan sampel, interpretasinya harus disesuaikan dengan konteks pengambilan sampel.

3.3 Distribusi Probabilitas Kontinu

3.3.1 Konsep Dasar dan Karakteristik Distribusi Kontinu

Distribusi probabilitas kontinu adalah distribusi yang menggambarkan peluang terjadinya suatu nilai dari variabel acak yang bersifat kontinu, yakni variabel yang dapat mengambil nilai dari suatu rentang tak terhingga, termasuk bilangan pecahan. Berbeda dengan distribusi diskrit yang menghitung probabilitas pada nilai-nilai tertentu, distribusi kontinu mengukur probabilitas dalam suatu interval. Oleh karena itu, probabilitas tepat pada satu titik dalam distribusi kontinu bernilai nol, sementara peluang dihitung melalui area di bawah kurva distribusi.

Contoh variabel yang mengikuti distribusi kontinu antara lain tekanan darah, kadar kolesterol, tinggi badan, dan suhu tubuh. Dalam konteks biostatistik, distribusi ini sangat penting untuk memodelkan fenomena fisiologis yang bervariasi secara alami dalam populasi. Distribusi kontinu ditentukan oleh fungsi kerapatan probabilitas (*probability density function*) yang menggambarkan bentuk dan karakteristik sebaran data (Pagano & Gauvreau, 2018).

3.3.2 Distribusi Normal dan Sifat-Sifatnya

Distribusi normal merupakan bentuk distribusi kontinu yang paling fundamental dan sering dijumpai dalam ilmu kesehatan. Ciri khas distribusi ini adalah bentuknya yang simetris menyerupai lonceng, dengan puncak di tengah dan meruncing di kedua ujung. Rata-rata (mean), median, dan modus pada distribusi normal berada

pada titik yang sama, menunjukkan keseimbangan nilai di kedua sisi kurva.

Distribusi normal ditentukan oleh dua parameter utama, yaitu rata-rata (μ) dan simpangan baku (σ). Parameter ini menentukan lokasi pusat dan lebar sebaran data. Dalam distribusi normal standar, nilai $\mu = 0$ dan $\sigma = 1$, yang kemudian digunakan sebagai acuan dalam transformasi skor z . Skor z memungkinkan kita membandingkan nilai dari distribusi yang berbeda atau menilai posisi relatif suatu nilai terhadap populasi.

Banyak variabel biologis dan perilaku manusia memiliki distribusi yang mendekati normal karena pengaruh berbagai faktor kecil yang bersifat independen dan akumulatif. Oleh karena itu, pemahaman terhadap distribusi normal menjadi sangat penting dalam biostatistik dan epidemiologi, terutama dalam perumusan hipotesis dan penarikan kesimpulan berdasarkan sampel (Rosner, 2015).

3.3.3 Aplikasi Distribusi Normal dalam Analisis Inferensial

Distribusi normal menjadi dasar utama dalam berbagai teknik analisis statistik inferensial. Salah satu contohnya adalah uji t , yang digunakan untuk membandingkan rata-rata dua kelompok populasi. Uji ini mengasumsikan bahwa data mengikuti distribusi normal, terutama ketika ukuran sampel relatif kecil. Jika asumsi ini tidak terpenuhi, maka hasil uji dapat menjadi bias dan tidak valid.

Selain itu, distribusi normal juga digunakan dalam analisis regresi linier, di mana kesalahan (*residuals*) dari model regresi diharapkan berdistribusi normal agar estimasi parameter tetap valid.

Pemodelan data biologis dengan distribusi normal juga memudahkan dalam menghitung interval kepercayaan dan mengestimasi probabilitas kejadian yang ekstrem.

Namun demikian, tidak semua data biologis secara alami berdistribusi normal. Oleh sebab itu, penting untuk melakukan uji normalitas sebelum menerapkan analisis berbasis asumsi distribusi normal. Jika data tidak berdistribusi normal, peneliti dapat menggunakan transformasi data atau memilih uji non-parametrik sebagai alternatif.

Distribusi normal tidak hanya memiliki nilai praktis dalam analisis statistik, tetapi juga menjadi dasar dalam teori *sampling distribution* dan *central limit theorem*—prinsip yang menjelaskan bahwa distribusi rata-rata sampel dari populasi manapun akan mendekati normal jika ukuran sampel cukup besar.

3.4 Penerapan Probabilitas dalam Kesehatan

3.4.1 Peran Probabilitas dalam Diagnosis dan Evaluasi Klinis

Probabilitas memiliki peranan penting dalam dunia medis, terutama dalam proses diagnosis dan penilaian risiko pasien. Seorang dokter tidak hanya mengandalkan gejala yang tampak secara kasat mata, tetapi juga mempertimbangkan kemungkinan terjadinya suatu penyakit berdasarkan data epidemiologi, riwayat medis, dan hasil pemeriksaan penunjang. Konsep probabilitas memungkinkan tenaga medis untuk menghitung *pre-test probability*

dan *post-test probability*, yaitu kemungkinan penyakit sebelum dan sesudah dilakukan tes diagnostik.

Misalnya, jika seorang pasien mengalami nyeri dada, dokter dapat memperkirakan probabilitas pasien mengalami infark miokard akut berdasarkan faktor risiko seperti usia, tekanan darah, kolesterol, dan hasil elektrokardiogram. Dengan pendekatan ini, keputusan klinis menjadi lebih berbasis data dan bukan semata intuisi (Steyerberg, 2019).

3.4.2 Probabilitas dalam Perencanaan dan Analisis Studi Kesehatan

Dalam ranah penelitian kesehatan, probabilitas digunakan untuk merancang uji klinis dan menentukan ukuran sampel yang representatif. Perhitungan probabilitas menentukan besarnya peluang suatu hasil terjadi secara kebetulan, sehingga digunakan untuk menguji hipotesis dan menentukan signifikansi statistik (*p-value*). Selain itu, konsep probabilitas juga mendasari analisis risiko dalam studi kohort, studi kasus-kontrol, dan uji klinis acak terkendali.

Sebagai contoh, probabilitas digunakan untuk mengukur kemungkinan terjadinya efek samping dari suatu intervensi medis, atau untuk membandingkan tingkat kejadian suatu penyakit antara kelompok intervensi dan kontrol. Pemahaman yang baik mengenai probabilitas membantu peneliti menafsirkan data dengan lebih tepat dan menghindari kesalahan pengambilan kesimpulan dari hasil yang bersifat kebetulan atau bias (Greenland et al., 2016).

3.4.3 Model Prediksi Klinis dan Pengambilan Keputusan

Model prediksi klinis merupakan bentuk lanjutan dari penerapan probabilitas dalam dunia kesehatan. Model ini, seperti *risk score* dan *prognostic model*, dirancang untuk memperkirakan kemungkinan terjadinya suatu hasil klinis berdasarkan kombinasi variabel-variabel yang relevan. Contohnya adalah *Framingham Risk Score* yang memperkirakan risiko penyakit jantung koroner dalam 10 tahun ke depan, berdasarkan faktor seperti usia, tekanan darah, merokok, dan kadar kolesterol.

Model-model prediktif ini dibangun menggunakan teknik statistik seperti regresi logistik atau regresi Cox, dan disempurnakan melalui validasi eksternal. Nilai probabilitas yang dihasilkan model dapat membantu dokter dalam membuat keputusan yang lebih objektif, misalnya apakah perlu dilakukan intervensi intensif atau cukup pemantauan rutin. Penggunaan probabilitas juga membantu menyampaikan informasi risiko kepada pasien secara kuantitatif, sehingga pasien lebih memahami kondisi dan pilihan terapinya.

3.4.4 Probabilitas dalam Analisis Laboratorium dan Epidemiologi

Dalam laboratorium diagnostik, probabilitas digunakan untuk mengevaluasi kinerja alat tes melalui parameter seperti sensitivitas, spesifisitas, nilai prediktif positif (PPV), dan nilai prediktif negatif (NPV). Semua parameter tersebut bergantung pada prevalensi penyakit dalam populasi dan mempengaruhi interpretasi hasil tes terhadap individu. Dengan bantuan probabilitas, tenaga laboratorium dapat menilai seberapa besar kemungkinan hasil positif

palsu atau negatif palsu terjadi, dan bagaimana hal tersebut memengaruhi diagnosis klinis.

Di bidang epidemiologi, probabilitas digunakan untuk memperkirakan laju penularan penyakit, menghitung angka kejadian (*incidence rate*), dan menyusun model penyebaran penyakit menular. Selama pandemi COVID-19, misalnya, model prediksi penyebaran virus berbasis probabilistik digunakan untuk membantu merancang intervensi kesehatan masyarakat dan memperkirakan dampak intervensi tersebut terhadap penurunan jumlah kasus.

3.5 Latihan Soal

1. Jelaskan perbedaan antara probabilitas diskret dan kontinu beserta contohnya dalam bidang kesehatan.
2. Sebutkan dan jelaskan dua jenis distribusi probabilitas diskret yang sering digunakan dalam biostatistik.
3. Mengapa distribusi normal sangat penting dalam analisis statistik kesehatan?
4. Bagaimana konsep probabilitas dapat digunakan untuk menilai risiko penyakit menular di masyarakat?
5. Berikan contoh penerapan probabilitas dalam menentukan efektivitas suatu intervensi medis.

Bab 4: Populasi, Sampel, dan Teknik Sampling

4.1 Pengertian Populasi dan Sampel

Dalam penelitian biostatistik, konsep populasi dan sampel merupakan dasar utama yang menentukan validitas dan generalisasi hasil penelitian. Populasi didefinisikan sebagai keseluruhan individu, objek, atau unit analisis yang memiliki karakteristik tertentu dan menjadi fokus perhatian peneliti. Populasi dapat berupa manusia, hewan, tumbuhan, peristiwa, atau bahkan fenomena yang ingin dikaji secara ilmiah. Dalam konteks penelitian kesehatan, misalnya, populasi dapat merujuk pada seluruh ibu hamil di suatu wilayah, pasien dengan penyakit tertentu, atau kelompok masyarakat dengan karakteristik demografis tertentu. Populasi dapat dibedakan menjadi dua jenis, yaitu *populasi target* (*target population*), yaitu keseluruhan kelompok yang menjadi tujuan akhir generalisasi penelitian, dan *populasi terjangkau* (*accessible population*), yaitu bagian dari populasi target yang secara realistis dapat dijangkau oleh peneliti untuk dilakukan pengumpulan data. Pembagian ini penting karena tidak semua populasi target dapat diobservasi secara langsung akibat keterbatasan waktu, biaya, atau sumber daya manusia (Polit & Beck, 2017).

Sampel, di sisi lain, adalah bagian dari populasi yang diambil untuk dijadikan objek penelitian dengan tujuan menggambarkan karakteristik populasi secara keseluruhan. Penggunaan sampel dalam penelitian biostatistik menjadi penting karena penelitian terhadap seluruh populasi sering kali tidak mungkin dilakukan, terutama jika jumlah populasi sangat besar atau tersebar luas secara geografis. Melalui teknik pengambilan sampel yang tepat, peneliti dapat memperoleh hasil yang akurat dan representatif tanpa harus mengumpulkan data dari seluruh anggota populasi. Sampel yang baik harus memenuhi dua kriteria utama, yaitu *representatif* dan *random*. Representatif berarti bahwa karakteristik sampel mencerminkan karakteristik populasi secara keseluruhan, sedangkan *random* menunjukkan bahwa setiap anggota populasi memiliki kesempatan yang sama untuk terpilih menjadi sampel. Keduanya berfungsi untuk meminimalkan *bias* dan meningkatkan reliabilitas hasil penelitian.

Pemilihan sampel dalam penelitian biostatistik dilakukan melalui berbagai metode, yang secara umum terbagi menjadi dua kategori besar, yaitu *probability sampling* dan *non-probability sampling*. Dalam *probability sampling*, setiap elemen populasi memiliki peluang yang diketahui dan sama besar untuk dipilih menjadi sampel. Metode ini mencakup *simple random sampling*, *systematic sampling*, *stratified random sampling*, dan *cluster sampling*. Misalnya, dalam *simple random sampling*, peneliti dapat menggunakan tabel angka acak atau perangkat komputer untuk memilih anggota sampel tanpa adanya campur tangan subjektif.

Metode *stratified random sampling* digunakan ketika populasi terdiri dari subkelompok yang berbeda (misalnya berdasarkan usia, jenis kelamin, atau status sosial ekonomi), sehingga setiap strata terwakili secara proporsional. Sementara itu, *cluster sampling* cocok digunakan dalam populasi yang luas dan tersebar, seperti memilih beberapa rumah sakit atau desa secara acak untuk mewakili keseluruhan wilayah penelitian (Creswell & Creswell, 2018).

Berbeda dengan *probability sampling*, metode *non-probability sampling* tidak memberikan kesempatan yang sama bagi setiap anggota populasi untuk terpilih menjadi sampel. Pemilihan dilakukan berdasarkan pertimbangan tertentu, kemudahan akses, atau kesesuaian dengan tujuan penelitian. Metode ini meliputi *convenience sampling*, *purposive sampling*, *quota sampling*, dan *snowball sampling*. Meskipun metode ini lebih mudah dan cepat dilakukan, kelemahannya terletak pada rendahnya kemampuan generalisasi hasil penelitian karena adanya potensi bias seleksi. Dalam penelitian biostatistik, metode non-probabilitas sering digunakan pada studi eksploratif, penelitian kualitatif, atau situasi di mana peneliti menghadapi keterbatasan sumber daya dan waktu. Namun, untuk penelitian kuantitatif dengan tujuan inferensial, metode *probability sampling* tetap menjadi pilihan utama karena memberikan hasil yang lebih dapat digeneralisasi terhadap populasi yang lebih luas.

Ukuran sampel (*sample size*) juga merupakan aspek penting dalam penelitian biostatistik, karena berhubungan langsung dengan kekuatan statistik (*statistical power*) dan keakuratan estimasi hasil

penelitian. Sampel yang terlalu kecil dapat menyebabkan hasil penelitian tidak signifikan secara statistik, sedangkan sampel yang terlalu besar dapat menyebabkan pemborosan sumber daya tanpa peningkatan substansial dalam validitas hasil. Oleh karena itu, penentuan ukuran sampel harus dilakukan berdasarkan pertimbangan ilmiah yang mencakup tingkat kesalahan yang dapat diterima (*margin of error*), tingkat kepercayaan (*confidence level*), dan variasi data dalam populasi. Dalam penelitian kesehatan, misalnya, perhitungan ukuran sampel sering menggunakan rumus *Cochran's formula* atau perangkat lunak statistik seperti *GPower** untuk memastikan bahwa jumlah responden yang digunakan cukup untuk menghasilkan kesimpulan yang valid.

Selain aspek teknis, pengambilan sampel juga harus memperhatikan pertimbangan etika penelitian, terutama ketika melibatkan manusia sebagai subjek. Peneliti harus menjamin bahwa proses pemilihan sampel dilakukan secara adil, menghormati hak-hak peserta, serta memastikan kerahasiaan dan persetujuan yang diinformasikan (*informed consent*). Dalam penelitian biostatistik yang melibatkan data kesehatan, prinsip kerahasiaan menjadi sangat penting untuk melindungi privasi individu dan mencegah penyalahgunaan data. Etika ini tidak hanya mencerminkan integritas ilmiah peneliti, tetapi juga membangun kepercayaan publik terhadap hasil penelitian yang dihasilkan.

Hubungan antara populasi dan sampel bersifat fundamental dalam analisis biostatistik karena hasil penelitian yang diperoleh dari sampel akan digunakan untuk melakukan *inferensi statistik* terhadap

populasi. Artinya, temuan dari sampel digunakan untuk memperkirakan parameter populasi melalui teknik estimasi dan uji hipotesis. Oleh karena itu, kualitas data yang diperoleh dari sampel sangat menentukan keandalan hasil penelitian secara keseluruhan. Dalam konteks kesehatan masyarakat, misalnya, hasil penelitian mengenai prevalensi anemia pada sampel ibu hamil di suatu wilayah dapat digunakan untuk memperkirakan prevalensi anemia di seluruh populasi ibu hamil di daerah tersebut. Proses ini menunjukkan bagaimana pengambilan sampel yang benar memungkinkan penelitian berkontribusi terhadap perumusan kebijakan kesehatan yang berbasis bukti (*evidence-based policy*).

Dengan demikian, pemahaman tentang konsep populasi dan sampel merupakan fondasi utama dalam penelitian biostatistik. Pemilihan sampel yang representatif dan pengumpulan data yang valid menjadi prasyarat untuk menghasilkan kesimpulan yang dapat diandalkan dan diterapkan secara luas. Penggunaan metode sampling yang tepat tidak hanya meningkatkan kualitas ilmiah penelitian, tetapi juga memperkuat relevansi temuan bagi pengambilan keputusan di bidang kesehatan. Oleh karena itu, setiap peneliti harus memiliki pengetahuan mendalam tentang prinsip-prinsip sampling dan keterampilan dalam menerapkannya secara etis dan ilmiah agar hasil penelitian dapat memberikan kontribusi nyata terhadap pengembangan ilmu biostatistik dan praktik kesehatan masyarakat.

4.2 Jenis-jenis Populasi

Dalam penelitian ilmiah, populasi merupakan keseluruhan subjek atau objek yang memiliki karakteristik tertentu dan menjadi perhatian peneliti. Pemahaman terhadap jenis-jenis populasi sangat penting karena memengaruhi teknik pengambilan sampel (*sampling method*), strategi analisis, serta interpretasi hasil penelitian. Menurut Sugiyono (2019), populasi dapat dibedakan berdasarkan jumlah dan batasannya, yakni menjadi **populasi terbatas** dan **populasi tak terbatas**.

Jenis populasi ini juga membantu peneliti menentukan pendekatan yang tepat dalam pengumpulan data, terutama dalam penelitian pendidikan, sosial, dan ilmu terapan. Dengan mengidentifikasi jenis populasi sejak awal, peneliti dapat memastikan bahwa sampel yang diambil benar-benar representatif terhadap keseluruhan populasi yang diteliti.

4.2.1 Populasi Terbatas (*Finite Population*)

Populasi terbatas adalah populasi yang jumlah elemennya dapat diketahui secara pasti. Artinya, semua anggota populasi dapat dihitung, diidentifikasi, dan dijangkau dalam konteks penelitian. Jenis populasi ini umumnya digunakan dalam penelitian yang ruang lingkupnya jelas dan terdefinisi dengan baik, seperti lembaga pendidikan, perusahaan, atau wilayah administratif tertentu.

Contohnya, populasi mahasiswa di Fakultas Ekonomi sebuah universitas, jumlah guru di sekolah menengah kota tertentu, atau pegawai negeri sipil dalam satu instansi. Dalam kasus ini, peneliti

dapat mengetahui total populasi dengan pasti, misalnya 200 mahasiswa atau 50 guru.

Penelitian dengan populasi terbatas memungkinkan peneliti menggunakan **metode sensus**, yaitu mengumpulkan data dari seluruh anggota populasi. Namun, bila jumlahnya terlalu besar, peneliti dapat menggunakan **sampling** untuk mengambil sebagian anggota populasi sebagai wakil. Karena ukuran populasi diketahui, maka proses pengambilan sampel dapat dilakukan dengan lebih terukur dan probabilitas setiap elemen untuk terpilih dapat dihitung dengan akurat (Creswell & Creswell, 2018).

Dalam bidang pendidikan, penelitian dengan populasi terbatas sering digunakan untuk menilai persepsi siswa terhadap kurikulum, efektivitas strategi pembelajaran, atau tingkat kepuasan mahasiswa terhadap layanan akademik. Populasi terbatas mempermudah peneliti dalam menentukan ukuran sampel dan memastikan hasil penelitian dapat digeneralisasikan ke seluruh kelompok sasaran.

4.2.2 Populasi Tak Terbatas (*Infinite Population*)

Populasi tak terbatas adalah populasi yang jumlah elemennya tidak dapat dihitung secara pasti karena terlalu besar atau tidak teridentifikasi dengan jelas. Populasi ini bersifat abstrak dan umumnya digunakan dalam penelitian yang melibatkan fenomena alam, proses biologis, atau situasi sosial yang terus berubah.

Sebagai contoh, populasi bakteri dalam air, jumlah pengguna media sosial di seluruh dunia, atau jumlah partikel udara di suatu wilayah. Dalam konteks pendidikan, populasi tak terbatas bisa

merujuk pada “seluruh siswa sekolah dasar di Indonesia” atau “seluruh mahasiswa di perguruan tinggi negeri,” di mana jumlah pastinya sulit diketahui karena terus berubah setiap waktu.

Dalam penelitian yang melibatkan populasi tak terbatas, peneliti tidak dapat menggunakan metode sensus. Oleh karena itu, digunakan teknik pengambilan sampel berdasarkan prinsip probabilitas atau nonprobabilitas. Menurut Sekaran dan Bougie (2016), pendekatan statistik inferensial menjadi penting untuk mengestimasi karakteristik populasi berdasarkan data dari sampel yang diambil.

Populasi tak terbatas sering ditemukan pada penelitian sosial berskala besar, survei nasional, atau studi longitudinal di mana peneliti hanya dapat mengambil sebagian kecil dari populasi untuk dianalisis. Tantangan utama dalam populasi ini adalah memastikan bahwa sampel benar-benar mewakili keragaman populasi secara keseluruhan. Oleh karena itu, proses perencanaan sampling harus mempertimbangkan ukuran, variabilitas, serta keterjangkauan populasi.

4.2.3 Pentingnya Identifikasi Jenis Populasi dalam Penelitian

Menentukan jenis populasi sejak awal menjadi langkah penting dalam perancangan penelitian. Populasi terbatas memungkinkan pengendalian yang lebih ketat terhadap data dan hasil penelitian yang lebih presisi, sementara populasi tak terbatas memberikan cakupan yang lebih luas dan generalisasi yang lebih tinggi.

Selain itu, jenis populasi berpengaruh langsung terhadap pemilihan metode pengambilan sampel. Pada populasi terbatas, teknik seperti *simple random sampling*, *stratified sampling*, atau *systematic sampling* dapat diterapkan dengan lebih akurat. Sebaliknya, untuk populasi tak terbatas, peneliti lebih sering menggunakan *cluster sampling*, *quota sampling*, atau *purposive sampling* untuk memperoleh representasi yang layak dari fenomena yang diamati.

Dalam penelitian pendidikan, misalnya, peneliti yang meneliti efektivitas pembelajaran daring perlu menentukan apakah populasi yang dituju adalah seluruh mahasiswa di universitas tertentu (populasi terbatas) atau seluruh mahasiswa di Indonesia (populasi tak terbatas). Penentuan ini akan memengaruhi desain penelitian, teknik sampling, serta kemampuan generalisasi hasil penelitian (Hair et al., 2020).

Dengan demikian, klasifikasi jenis populasi bukan hanya aspek teknis, tetapi juga strategis dalam perencanaan penelitian. Identifikasi yang tepat akan membantu peneliti memperoleh data yang valid, hasil yang akurat, serta kesimpulan yang dapat dipertanggungjawabkan secara ilmiah.

4.3 Konsep Sampling dan Pentingnya dalam Penelitian

Dalam penelitian ilmiah, *sampling* merupakan langkah metodologis yang sangat penting dalam menentukan

representativitas hasil penelitian. Istilah *sampling* mengacu pada proses pemilihan sebagian unit atau elemen dari populasi yang dijadikan sumber data penelitian. Melalui proses ini, peneliti dapat memperoleh informasi yang relevan tanpa harus meneliti seluruh populasi, yang sering kali tidak memungkinkan karena keterbatasan sumber daya seperti waktu, biaya, dan tenaga.

Tujuan utama *sampling* adalah untuk memperoleh data yang dapat menggambarkan kondisi populasi secara akurat. Dengan demikian, hasil penelitian dapat digeneralisasi dan digunakan untuk menarik kesimpulan ilmiah yang sah. Selain itu, teknik *sampling* yang tepat juga berperan penting dalam meminimalkan bias, meningkatkan validitas eksternal, serta memperkuat keandalan hasil penelitian. Menurut Creswell dan Creswell (2018), pemilihan sampel yang tepat merupakan dasar dari rancangan penelitian yang efektif, karena kualitas sampel menentukan sejauh mana temuan dapat dipercaya dan diterapkan secara luas.

4.3.1 Pengertian dan Tujuan *Sampling*

Sampling adalah proses sistematis dalam menentukan sebagian unit dari populasi yang akan dijadikan sumber data penelitian. Proses ini tidak hanya bersifat teknis, tetapi juga strategis, karena berkaitan langsung dengan validitas data yang dihasilkan. Peneliti harus memahami struktur dan karakteristik populasi sebelum menentukan cara pengambilan sampel yang sesuai.

Tujuan utama dari *sampling* adalah memperoleh informasi yang dapat digeneralisasi terhadap populasi dengan tingkat kesalahan minimal. Selain itu, *sampling* membantu peneliti dalam

mengefisienkan sumber daya tanpa mengurangi keakuratan hasil penelitian. Etikan dan Bala (2017) menyebutkan bahwa rancangan *sampling* yang baik mampu menyeimbangkan antara keandalan data dan keterbatasan operasional penelitian. Dengan demikian, perencanaan *sampling* bukan hanya aspek teknis, melainkan juga bagian dari strategi ilmiah untuk memastikan kualitas temuan.

4.3.2 Jenis-Jenis *Sampling*

Secara garis besar, *sampling* terbagi menjadi dua jenis utama, yaitu *probability sampling* dan *non-probability sampling*.

Dalam *probability sampling*, setiap anggota populasi memiliki peluang yang sama untuk terpilih menjadi sampel. Teknik ini umum digunakan dalam penelitian kuantitatif karena memungkinkan generalisasi hasil secara objektif. Contohnya meliputi *simple random sampling*, *stratified sampling*, dan *cluster sampling*. Misalnya, *stratified sampling* diterapkan jika populasi memiliki subkelompok berbeda, sehingga setiap kelompok memperoleh proporsi yang adil dalam sampel. Pendekatan ini memastikan bahwa karakteristik populasi tercermin secara proporsional dalam hasil penelitian (Taherdoost, 2016).

Sebaliknya, *non-probability sampling* tidak memberikan peluang yang sama bagi setiap elemen populasi. Teknik ini banyak digunakan dalam penelitian kualitatif, di mana tujuan utama bukan generalisasi, melainkan pemahaman mendalam terhadap fenomena tertentu. Beberapa jenisnya meliputi *purposive sampling*, *convenience sampling*, dan *snowball sampling*. Misalnya, *purposive sampling* memungkinkan peneliti memilih responden berdasarkan

kriteria relevan yang sesuai dengan fokus penelitian. Meskipun hasilnya tidak dapat digeneralisasi secara luas, teknik ini efektif untuk menggali konteks dan makna sosial yang kompleks (Palinkas et al., 2015).

4.3.3 Pentingnya *Sampling* dalam Penelitian

Pemilihan teknik *sampling* yang tepat memiliki peran krusial dalam menjaga kualitas penelitian. Kesalahan dalam menentukan metode atau ukuran sampel dapat menimbulkan *sampling bias*, yakni kondisi di mana sampel tidak mewakili populasi dengan benar. *Bias* semacam ini dapat menurunkan validitas eksternal penelitian, sehingga hasilnya tidak dapat digeneralisasi. Oleh karena itu, dalam tahap perancangan metodologi, peneliti perlu mempertimbangkan dengan matang karakteristik populasi, variabel penelitian, serta tujuan analisis (Creswell & Creswell, 2018).

Selain itu, *sampling* juga berperan dalam meningkatkan efisiensi pelaksanaan penelitian. Dengan ukuran sampel yang optimal, peneliti dapat memperoleh hasil yang valid tanpa mengorbankan sumber daya yang berlebihan. Menurut Etikan dan Bala (2017), keseimbangan antara ketepatan analisis dan efisiensi sumber daya merupakan prinsip utama dalam perancangan *sampling* yang efektif.

4.3.4 Implikasi terhadap Validitas Penelitian

Validitas hasil penelitian sangat dipengaruhi oleh representativitas sampel yang digunakan. Dalam penelitian kuantitatif, *probability sampling* umumnya dianggap lebih unggul dalam menjaga validitas eksternal karena karakter acakannya. Namun,

dalam penelitian kualitatif, validitas lebih menitikberatkan pada kedalaman dan relevansi data yang dihasilkan melalui *non-probability sampling*.

Pemilihan teknik *sampling* harus disesuaikan dengan desain penelitian dan tujuan analisis. Integrasi yang tepat antara pendekatan metodologis, karakteristik populasi, dan teknik pengambilan sampel akan menghasilkan penelitian yang tidak hanya efisien, tetapi juga memiliki kekuatan ilmiah yang tinggi. Oleh sebab itu, pemahaman mendalam terhadap konsep dan praktik *sampling* menjadi kunci keberhasilan dalam setiap penelitian ilmiah.

4.4 Teknik Sampling

Dalam penelitian ilmiah, teknik *sampling* merupakan langkah metodologis yang menentukan bagaimana peneliti memilih sebagian anggota dari suatu populasi untuk dijadikan responden atau unit analisis. Tujuan utama *sampling* adalah memperoleh data yang representatif dari populasi tanpa harus meneliti seluruh anggotanya. Pemilihan teknik *sampling* yang tepat berpengaruh langsung terhadap validitas dan reliabilitas hasil penelitian. Oleh karena itu, pemahaman mendalam terhadap prinsip dan jenis teknik *sampling* menjadi aspek penting dalam perancangan penelitian yang sistematis.

Menurut Creswell dan Creswell (2018), *sampling* berfungsi sebagai jembatan antara populasi dan hasil penelitian dengan cara memastikan bahwa sampel yang dipilih benar-benar mencerminkan

karakteristik populasi yang ingin dikaji. Kesalahan dalam menentukan teknik *sampling* dapat menimbulkan bias, menurunkan akurasi data, dan menghambat generalisasi temuan penelitian. Oleh sebab itu, pemilihan teknik *sampling* harus dilakukan secara rasional berdasarkan tujuan penelitian, sifat populasi, serta ketersediaan sumber daya penelitian.

4.4.1 Pengantar Teknik Sampling

Teknik *sampling* berperan penting dalam menghasilkan data yang valid dan dapat dipertanggungjawabkan secara ilmiah. Dalam praktiknya, peneliti tidak selalu mampu menjangkau seluruh populasi karena keterbatasan waktu, biaya, dan tenaga. Oleh karena itu, pengambilan sampel menjadi solusi efektif untuk memperoleh gambaran yang cukup akurat terhadap keseluruhan populasi. Sugiyono (2019) menjelaskan bahwa *sampling* merupakan teknik penentuan sampel yang dilakukan apabila peneliti tidak mungkin meneliti seluruh elemen populasi. Dengan demikian, *sampling* tidak hanya berfungsi sebagai strategi efisiensi, tetapi juga sebagai metode ilmiah untuk menjaga keandalan dan objektivitas data penelitian.

4.4.2 Jenis-Jenis Teknik Sampling

Secara garis besar, teknik *sampling* dibedakan menjadi dua kelompok besar, yaitu *sampling probabilitas (probability sampling)* dan *sampling non-probabilitas (non-probability sampling)*.

Pada *probability sampling*, setiap anggota populasi memiliki peluang yang sama untuk terpilih menjadi sampel. Teknik ini digunakan terutama dalam penelitian kuantitatif yang menuntut generalisasi hasil terhadap populasi. Jenis-jenis *probability sampling*

meliputi *simple random sampling*, *stratified sampling*, *cluster sampling*, dan *systematic sampling*. Misalnya, *stratified sampling* digunakan ketika populasi terdiri dari beberapa lapisan (strata) dengan karakteristik berbeda yang harus terwakili secara proporsional.

Sementara itu, *non-probability sampling* tidak memberikan kesempatan yang sama kepada setiap anggota populasi untuk terpilih. Pemilihan sampel dilakukan berdasarkan pertimbangan subjektif atau kemudahan akses terhadap responden. Jenis ini meliputi *purposive sampling*, *quota sampling*, dan *convenience sampling*. Teknik ini lebih sering digunakan dalam penelitian kualitatif yang menekankan kedalaman pemahaman terhadap fenomena tertentu, bukan pada representasi statistik (Etikan & Bala, 2017).

4.4.3 Pertimbangan dalam Pemilihan Teknik Sampling

Pemilihan teknik *sampling* harus disesuaikan dengan tujuan penelitian, karakteristik populasi, dan sumber daya yang tersedia. Dalam penelitian kuantitatif, *probability sampling* lebih disarankan karena mampu menghasilkan sampel yang representatif dan mengurangi potensi bias. Sebaliknya, dalam penelitian kualitatif yang berfokus pada eksplorasi fenomena dan makna sosial, *non-probability sampling* seperti *purposive sampling* lebih sesuai karena memungkinkan peneliti memilih responden dengan karakteristik tertentu yang relevan dengan topik penelitian.

Faktor praktis seperti waktu, biaya, dan ketersediaan data juga harus menjadi pertimbangan utama. Menurut Creswell dan

Creswell (2018), keputusan pemilihan metode *sampling* yang tepat harus mempertimbangkan keseimbangan antara validitas metodologis dan efisiensi pelaksanaan penelitian.

4.4.4 Implikasi Teknik Sampling terhadap Validitas Penelitian

Teknik *sampling* memiliki pengaruh langsung terhadap validitas hasil penelitian. Validitas eksternal, yakni sejauh mana hasil penelitian dapat digeneralisasi ke populasi yang lebih luas, sangat bergantung pada representativitas sampel yang digunakan. *Probability sampling* mendukung validitas eksternal yang tinggi karena memberikan peluang yang sama bagi seluruh anggota populasi untuk terpilih.

Sebaliknya, *non-probability sampling* lebih menitikberatkan pada kedalaman data daripada perluasan hasil. Dalam penelitian sosial dan pendidikan, pendekatan ini sering dipilih untuk memahami konteks dan fenomena secara komprehensif. Dengan demikian, peneliti harus mampu menyesuaikan antara kebutuhan akan generalisasi dan kedalaman analisis agar hasil penelitian tetap valid dan relevan (Etikan & Bala, 2017; Sugiyono, 2019).

4.5 Latihan Soal

1. Jelaskan pengertian populasi dan sampel dalam konteks biostatistik!
2. Apa perbedaan antara populasi terbatas dan tak terbatas?
3. Mengapa sampling diperlukan dalam penelitian?
4. Sebutkan dan jelaskan dua kategori utama teknik sampling!

5. Berikan contoh penerapan simple random sampling dalam penelitian kesehatan!

Bab 5: Estimasi Parameter dan Uji Hipotesis

5.1 Pengertian Estimasi Parameter

Estimasi parameter merupakan salah satu konsep sentral dalam analisis inferensial biostatistik yang bertujuan untuk menarik kesimpulan tentang populasi berdasarkan data yang diperoleh dari sampel. Dalam konteks ini, parameter populasi adalah nilai yang menggambarkan karakteristik tertentu dari populasi, seperti rata-rata (μ), proporsi (p), atau varians (σ^2), yang umumnya tidak diketahui karena keterbatasan dalam mengobservasi seluruh anggota populasi. Oleh karena itu, peneliti menggunakan teknik estimasi untuk memperoleh nilai perkiraan parameter tersebut dari data sampel. Proses ini sangat penting dalam penelitian kesehatan dan biostatistik, karena hampir semua keputusan dan kebijakan berbasis data diambil berdasarkan hasil estimasi, bukan nilai sebenarnya dari populasi. Estimasi parameter memungkinkan peneliti untuk menyimpulkan karakteristik populasi dengan tingkat keyakinan tertentu, sehingga hasil penelitian dapat digunakan secara valid dan dapat digeneralisasi.

Secara umum, terdapat dua jenis utama estimasi parameter, yaitu **estimasi titik** (*point estimation*) dan **estimasi interval** (*interval estimation*). Estimasi titik adalah metode yang

menghasilkan satu nilai tunggal sebagai perkiraan parameter populasi. Misalnya, rata-rata berat badan dari 100 responden dapat digunakan sebagai estimasi titik dari rata-rata berat badan seluruh populasi. Nilai ini dihitung menggunakan statistik sampel, seperti *sample mean* ($\bar{x}$), *sample proportion* ($\hat{p}$), atau *sample variance* (s^2), yang masing-masing digunakan untuk mengestimasi parameter populasi μ , p , dan σ^2 . Estimasi titik mudah dihitung dan sering digunakan dalam laporan statistik deskriptif, namun memiliki kelemahan karena tidak memberikan informasi tentang tingkat ketidakpastian dari hasil estimasi tersebut. Dalam kenyataannya, setiap estimasi berdasarkan sampel mengandung kesalahan (*sampling error*), sehingga nilai estimasi tidak pernah sama persis dengan parameter populasi (Wackerly, Mendenhall, & Scheaffer, 2014).

Untuk mengatasi keterbatasan estimasi titik, digunakan **estimasi interval**, yang memberikan rentang nilai di mana parameter populasi diperkirakan berada dengan tingkat kepercayaan tertentu. Rentang ini dikenal sebagai **interval kepercayaan** (*confidence interval*), dan dinyatakan dalam bentuk: $\text{Estimasi titik} \pm \text{Margin of Error}$. $\pm$ menunjukkan sejauh mana nilai estimasi mungkin berbeda dari parameter populasi sebenarnya. Misalnya, jika hasil penelitian menunjukkan bahwa rata-rata tekanan darah sistolik adalah 120 mmHg dengan interval kepercayaan 95% antara 118–122 mmHg, maka dapat diartikan bahwa peneliti yakin

95% bahwa nilai rata-rata tekanan darah populasi berada dalam rentang tersebut. Interval kepercayaan 95% merupakan konvensi yang umum digunakan dalam penelitian kesehatan, karena memberikan keseimbangan antara ketepatan (*precision*) dan keyakinan (*confidence*). Semakin besar ukuran sampel, semakin sempit interval kepercayaan yang dihasilkan, menunjukkan estimasi yang lebih presisi.

Estimasi parameter dilakukan dengan menggunakan berbagai rumus tergantung pada jenis data dan ukuran sampel yang digunakan. Untuk mengestimasi rata-rata populasi, digunakan distribusi *t-Student* jika ukuran sampel kecil ($n < 30$) dan varians populasi tidak diketahui, atau distribusi *z-normal* jika ukuran sampel besar. Sementara itu, untuk estimasi proporsi digunakan rumus berbasis distribusi normal dengan penyesuaian terhadap proporsi sukses ($\hat{p}$) dan jumlah sampel (n). Dalam biostatistik, penting juga untuk memperhatikan *standard error* (SE), yaitu ukuran yang menggambarkan seberapa besar variasi estimasi dari satu sampel ke sampel lain. *Standard error* berfungsi sebagai dasar dalam menghitung margin kesalahan dan menentukan lebar interval kepercayaan. Nilai SE yang kecil menunjukkan bahwa estimasi lebih stabil dan dapat dipercaya.

Konsep estimasi parameter tidak hanya bersifat matematis, tetapi juga memiliki implikasi praktis yang signifikan dalam penelitian kesehatan. Misalnya, dalam studi epidemiologi, estimasi proporsi prevalensi suatu penyakit digunakan untuk menggambarkan besarnya masalah kesehatan di masyarakat.

Estimasi interval pada penelitian ini membantu menentukan tingkat keyakinan terhadap hasil survei, sehingga dapat digunakan sebagai dasar dalam perencanaan intervensi kesehatan masyarakat. Demikian pula, dalam uji klinis, estimasi rata-rata efektivitas obat atau intervensi medis memberikan dasar bagi penilaian efikasi dan keamanan terapi sebelum diimplementasikan secara luas. Oleh karena itu, akurasi dan presisi estimasi parameter menjadi kunci dalam memastikan validitas ilmiah hasil penelitian biostatistik.

Selain itu, dalam penerapannya, peneliti harus memahami perbedaan antara **estimasi parameter** dan **pengujian hipotesis**, meskipun keduanya saling berkaitan dalam inferensi statistik. Estimasi parameter bertujuan untuk memperkirakan nilai dari suatu parameter, sedangkan pengujian hipotesis bertujuan untuk menilai apakah parameter populasi memiliki nilai tertentu berdasarkan bukti sampel. Keduanya menggunakan distribusi probabilitas sebagai dasar analisis, tetapi estimasi lebih menekankan pada pengukuran ketidakpastian dari nilai parameter itu sendiri. Dengan demikian, estimasi parameter berfungsi sebagai dasar untuk memahami struktur data, sedangkan pengujian hipotesis digunakan untuk membuat kesimpulan keputusan.

Dalam pelaksanaannya, estimasi parameter juga harus memenuhi beberapa asumsi penting, seperti independensi antarobservasi, distribusi data yang mendekati normal (khususnya untuk estimasi rata-rata), dan ukuran sampel yang memadai. Pelanggaran terhadap asumsi-asumsi ini dapat menyebabkan hasil estimasi menjadi bias atau tidak valid. Oleh karena itu, peneliti perlu

melakukan analisis pendahuluan seperti uji normalitas, deteksi pencilan (*outlier detection*), dan pemeriksaan homogenitas varians sebelum melakukan estimasi. Di era modern, perangkat lunak statistik seperti *SPSS*, *R*, dan *Python (SciPy)* telah mempermudah proses estimasi parameter dengan hasil yang cepat dan akurat, asalkan pengguna memahami prinsip-prinsip dasar di balik perhitungannya.

Dengan demikian, estimasi parameter merupakan alat penting dalam biostatistik yang berfungsi menjembatani antara data sampel dan kesimpulan tentang populasi. Melalui penggunaan estimasi titik dan estimasi interval, peneliti dapat menggambarkan nilai parameter populasi secara lebih informatif dengan mempertimbangkan unsur ketidakpastian. Estimasi yang baik harus didasarkan pada metode yang valid, ukuran sampel yang memadai, dan pemenuhan asumsi statistik agar hasil penelitian dapat diinterpretasikan secara tepat. Oleh karena itu, penguasaan konsep estimasi parameter menjadi keterampilan esensial bagi peneliti di bidang biostatistik dan kesehatan masyarakat untuk menghasilkan bukti ilmiah yang akurat, terpercaya, dan bermanfaat dalam pengambilan keputusan berbasis data.

5.2 Jenis Estimasi dan Tingkat Kepercayaan

Estimasi merupakan proses dalam analisis statistik yang digunakan untuk memperkirakan parameter populasi berdasarkan data yang diperoleh dari sampel. Dalam penelitian, terutama yang

bersifat kuantitatif, estimasi menjadi dasar dalam pengambilan keputusan dan generalisasi hasil penelitian. Menurut Walpole et al. (2017), terdapat dua jenis estimasi yang umum digunakan, yaitu **estimasi titik** (*point estimation*) dan **estimasi interval** (*interval estimation*).

Kedua jenis estimasi tersebut memiliki tujuan yang sama, yakni memperkirakan nilai parameter populasi, tetapi berbeda dalam cara dan tingkat keyakinan hasilnya. Estimasi titik menghasilkan satu nilai tunggal yang dianggap paling mewakili parameter populasi, sedangkan estimasi interval menghasilkan rentang nilai yang mungkin mencakup nilai parameter sebenarnya dengan tingkat kepercayaan tertentu.

5.2.1 Estimasi Titik (*Point Estimation*)

Estimasi titik adalah pendekatan yang menggunakan satu nilai hasil dari sampel sebagai perkiraan terhadap parameter populasi. Misalnya, jika peneliti ingin mengetahui rata-rata nilai ujian mahasiswa, maka rata-rata hasil dari sampel mahasiswa yang diambil akan menjadi estimasi titik bagi rata-rata populasi.

Sebagai contoh, jika rata-rata nilai sampel mahasiswa adalah 78, maka nilai 78 tersebut dianggap sebagai estimasi terbaik dari rata-rata populasi sebenarnya.

Keunggulan metode ini terletak pada kesederhanaannya dan kemudahan dalam perhitungan. Namun, karena estimasi titik hanya memberikan satu nilai tanpa memperhitungkan ketidakpastian, peneliti tidak dapat mengetahui sejauh mana nilai tersebut mungkin menyimpang dari parameter populasi. Oleh sebab itu, estimasi titik

sering dilengkapi dengan estimasi interval untuk memberikan ukuran ketelitian yang lebih baik (Triola, 2018).

5.2.2 Estimasi Interval (*Interval Estimation*)

Estimasi interval digunakan untuk memberikan rentang nilai yang kemungkinan besar mencakup parameter populasi. Rentang ini disebut **interval kepercayaan** (*confidence interval*), yang diikuti oleh tingkat kepercayaan tertentu, seperti 90%, 95%, atau 99%. Interval kepercayaan menyatakan bahwa peneliti yakin dengan persentase tertentu bahwa nilai sebenarnya dari parameter populasi berada dalam rentang yang dihitung.

Sebagai contoh, jika penelitian menunjukkan bahwa rata-rata berat badan sampel adalah 60 kilogram dengan simpangan baku tertentu, maka interval kepercayaan 95% dapat dinyatakan sebagai “antara 59,2 kilogram dan 60,8 kilogram.” Artinya, peneliti yakin sebesar 95% bahwa rata-rata berat badan populasi sesungguhnya berada di antara kedua batas tersebut.

Semakin tinggi tingkat kepercayaan yang digunakan, semakin besar pula kemungkinan bahwa interval mencakup nilai parameter sebenarnya. Namun, peningkatan tingkat kepercayaan juga menyebabkan interval menjadi lebih lebar karena memperhitungkan ketidakpastian yang lebih besar (Bluman, 2018).

Estimasi interval lebih informatif daripada estimasi titik karena memberikan gambaran mengenai derajat ketelitian hasil estimasi. Hal ini menjadikannya lebih disukai dalam penelitian yang menekankan pada akurasi dan reliabilitas hasil analisis.

5.2.3 Tingkat Kepercayaan (*Confidence Level*)

Tingkat kepercayaan menggambarkan sejauh mana peneliti yakin bahwa interval yang dihitung mencakup parameter populasi yang sebenarnya. Umumnya, tingkat kepercayaan yang sering digunakan dalam penelitian adalah 90%, 95%, dan 99%.

Sebagai ilustrasi, tingkat kepercayaan 95% menunjukkan bahwa apabila penelitian dilakukan berulang kali dengan prosedur yang sama, sekitar 95 dari 100 interval yang dihasilkan akan mencakup nilai parameter populasi sesungguhnya.

Namun, perlu dipahami bahwa semakin tinggi tingkat kepercayaan yang digunakan, interval yang dihasilkan akan semakin lebar. Hal ini disebabkan oleh peningkatan toleransi terhadap ketidakpastian. Sebaliknya, jika tingkat kepercayaan diturunkan (misalnya menjadi 90%), interval menjadi lebih sempit tetapi risiko tidak mencakup nilai parameter sebenarnya menjadi lebih besar (Walpole et al., 2017).

Dalam konteks penelitian pendidikan dan sosial, tingkat kepercayaan 95% dianggap ideal karena memberikan keseimbangan antara ketelitian dan efisiensi. Misalnya, jika peneliti menyatakan bahwa rata-rata nilai siswa berada dalam interval 78 hingga 82 dengan tingkat kepercayaan 95%, maka peneliti yakin bahwa nilai rata-rata populasi sesungguhnya berada di antara angka tersebut dengan kemungkinan kesalahan sebesar 5%.

5.2.4 Pentingnya Estimasi dalam Penelitian

Estimasi memiliki peranan penting dalam penelitian karena membantu peneliti menarik kesimpulan tentang populasi

berdasarkan hasil sampel. Estimasi titik memberikan nilai terbaik yang mewakili parameter populasi, sedangkan estimasi interval dan tingkat kepercayaan memberikan ukuran keyakinan terhadap hasil tersebut.

Dalam penelitian pendidikan, estimasi digunakan, misalnya, untuk memperkirakan rata-rata kemampuan siswa, tingkat kepuasan terhadap pembelajaran daring, atau efektivitas program pelatihan guru. Dengan menggunakan tingkat kepercayaan yang sesuai, peneliti dapat mengkomunikasikan hasilnya secara lebih hati-hati dan dapat dipertanggungjawabkan.

Selain itu, pemahaman terhadap jenis estimasi membantu peneliti dalam menyusun laporan ilmiah yang akurat, transparan, dan kredibel. Penggunaan estimasi interval dengan tingkat kepercayaan yang jelas juga memudahkan pembaca memahami batas ketepatan hasil penelitian dan memperkuat dasar pengambilan keputusan berbasis data (Triola, 2018).

5.3 Konsep Uji Hipotesis

Dalam penelitian ilmiah, *uji hipotesis* merupakan prosedur statistik yang digunakan untuk menilai apakah data sampel memberikan cukup bukti guna menerima atau menolak suatu pernyataan mengenai populasi. Proses ini menjadi inti dari analisis inferensial karena memungkinkan peneliti membuat kesimpulan yang didasarkan pada data empiris, bukan sekadar asumsi teoritis. *Uji hipotesis* membantu memastikan bahwa hasil penelitian

memiliki dasar statistik yang kuat dan dapat dipertanggungjawabkan secara ilmiah.

Menurut Gravetter dan Wallnau (2017), *uji hipotesis* berfungsi sebagai mekanisme untuk menguji klaim ilmiah dengan membandingkan hasil observasi dari sampel terhadap harapan teoretis dalam populasi. Dengan demikian, proses ini memberikan jembatan antara data empiris dan teori, memungkinkan peneliti menentukan apakah perbedaan atau hubungan yang ditemukan bersifat signifikan secara statistik atau hanya muncul karena kebetulan pengambilan sampel.

5.3.1 Pengertian dan Jenis Hipotesis

Secara umum, hipotesis dapat didefinisikan sebagai dugaan sementara yang dirumuskan berdasarkan teori dan digunakan sebagai dasar pengujian empiris. Dalam konteks *uji hipotesis*, terdapat dua jenis hipotesis utama, yaitu hipotesis nol (*null hypothesis*, H_0) dan hipotesis alternatif (*alternative hypothesis*, H_1).

Hipotesis nol (H_0) menyatakan bahwa tidak terdapat perbedaan atau pengaruh antara variabel yang diteliti. Sebaliknya, hipotesis alternatif (H_1) menyatakan adanya perbedaan atau pengaruh tersebut. Misalnya, dalam penelitian tentang efektivitas metode pembelajaran, H_0 menyatakan bahwa metode baru tidak lebih efektif dibanding metode lama, sedangkan H_1 menyatakan bahwa metode baru memberikan peningkatan signifikan dalam hasil belajar.

Hipotesis berfungsi sebagai kerangka pengujian dalam penelitian kuantitatif. Dengan menetapkan H_0 dan H_1 , peneliti

memiliki dasar untuk menentukan arah analisis dan interpretasi hasil. Selain itu, pengujian hipotesis membantu memastikan bahwa keputusan ilmiah dibuat secara objektif berdasarkan data, bukan preferensi subjektif peneliti (Field, 2018).

5.3.2 Proses Uji Hipotesis

Proses *uji hipotesis* melibatkan beberapa langkah sistematis yang bertujuan untuk mengevaluasi bukti empiris terhadap hipotesis yang diajukan.

Langkah pertama adalah merumuskan hipotesis nol (H_0) dan hipotesis alternatif (H_1) dengan jelas. Selanjutnya, peneliti menetapkan tingkat signifikansi (*alpha*, α), yaitu probabilitas maksimal untuk melakukan kesalahan tipe I, yakni menolak H_0 padahal H_0 benar. Nilai α yang umum digunakan adalah 0,05 atau 0,01, tergantung pada tingkat ketelitian yang diinginkan.

Langkah berikutnya adalah menentukan statistik uji yang sesuai, seperti *t-test*, *chi-square*, atau *ANOVA*, tergantung pada jenis data dan tujuan analisis. Statistik uji ini digunakan untuk menghitung nilai p (*p-value*), yang menunjukkan probabilitas memperoleh hasil seperti data sampel apabila H_0 benar. Jika nilai p lebih kecil dari α , maka H_0 ditolak dan H_1 diterima. Sebaliknya, jika p lebih besar dari α , maka tidak terdapat cukup bukti untuk menolak H_0 (Gravetter & Wallnau, 2017).

5.3.3 Pentingnya Tingkat Signifikansi dan Nilai p

Tingkat signifikansi (α) merupakan ambang batas yang digunakan peneliti untuk menentukan apakah hasil pengujian bersifat signifikan secara statistik. Pemilihan nilai α harus

mempertimbangkan konteks penelitian, karena semakin kecil α , semakin ketat kriteria penolakan terhadap H_0 . Nilai α yang terlalu besar dapat meningkatkan risiko kesimpulan keliru, sementara nilai yang terlalu kecil dapat menyebabkan hasil signifikan diabaikan.

Nilai p (p -value) berperan sebagai ukuran probabilistik dalam pengambilan keputusan. Nilai ini menunjukkan peluang bahwa hasil penelitian yang diperoleh terjadi secara kebetulan. Ketika $p \leq \alpha$, maka hasil penelitian dianggap signifikan, dan H_0 ditolak. Namun, penting diingat bahwa signifikansi statistik tidak selalu berarti signifikansi praktis. Oleh karena itu, peneliti juga perlu mempertimbangkan ukuran efek (*effect size*) dan konteks substantif dari hasil penelitian agar interpretasi lebih komprehensif (Field, 2018).

5.3.4 Implikasi Uji Hipotesis dalam Penelitian

Uji hipotesis memiliki peran sentral dalam memastikan keandalan dan objektivitas penelitian ilmiah. Melalui proses ini, peneliti dapat menilai apakah temuan empiris mendukung teori yang ada atau memerlukan revisi. Selain itu, *uji hipotesis* juga membantu dalam pengambilan keputusan berbasis bukti (*evidence-based decision making*), baik dalam bidang sosial, ekonomi, maupun kesehatan.

Pemahaman yang tepat terhadap konsep dan prosedur *uji hipotesis* memungkinkan peneliti menghindari kesalahan interpretasi data. Kesalahan seperti *Type I error* (menolak H_0 yang benar) dan *Type II error* (gagal menolak H_0 yang salah) dapat diminimalkan melalui perencanaan penelitian yang matang,

pemilihan ukuran sampel yang sesuai, serta penerapan metode analisis statistik yang relevan. Dengan demikian, *uji hipotesis* tidak hanya menjadi alat analisis, tetapi juga fondasi bagi keabsahan ilmiah dalam penelitian kuantitatif modern.

5.4 Jenis Uji Statistik dalam Hipotesis

Dalam penelitian kuantitatif, uji statistik merupakan prosedur penting untuk menentukan apakah hipotesis yang diajukan dapat diterima atau ditolak berdasarkan data yang diperoleh. Uji statistik berfungsi sebagai alat analisis yang membantu peneliti dalam menilai hubungan, perbedaan, atau pengaruh antarvariabel secara objektif. Pemilihan jenis uji statistik harus disesuaikan dengan karakteristik data, desain penelitian, serta asumsi yang mendasarinya. Dengan demikian, kesesuaian antara teknik analisis dan sifat data menjadi faktor krusial dalam memastikan validitas hasil penelitian.

Menurut Field (2018), uji statistik tidak hanya bertujuan untuk menguji hipotesis, tetapi juga untuk mengontrol kesalahan inferensial, seperti *Type I error* dan *Type II error*. Oleh karena itu, pemilihan jenis uji yang tepat memungkinkan peneliti menarik kesimpulan ilmiah yang sahih dan reliabel.

5.4.1 Pengantar Uji Statistik dalam Hipotesis

Uji statistik merupakan tahap analisis inferensial yang digunakan untuk menguji dugaan (hipotesis) terhadap populasi berdasarkan data sampel. Hipotesis nol (*null hypothesis*) umumnya

menyatakan tidak adanya perbedaan atau hubungan antarvariabel, sedangkan hipotesis alternatif (*alternative hypothesis*) menyatakan sebaliknya. Proses pengujian dilakukan dengan menghitung nilai statistik tertentu dan membandingkannya dengan nilai kritis pada tingkat signifikansi yang telah ditentukan (biasanya 0,05).

Dalam konteks ini, penting bagi peneliti untuk memahami distribusi data sebelum menentukan jenis uji yang akan digunakan. Data yang berdistribusi normal umumnya dianalisis menggunakan uji parametrik, sedangkan data yang tidak berdistribusi normal lebih sesuai dianalisis dengan uji nonparametrik (Gravetter & Wallnau, 2020).

5.4.2 Uji Parametrik

Uji parametrik digunakan untuk menganalisis data yang memenuhi asumsi tertentu, terutama distribusi normal, homogenitas varians, dan skala pengukuran interval atau rasio. Uji ini memiliki kekuatan statistik yang tinggi karena mempertimbangkan parameter populasi, seperti rata-rata dan simpangan baku.

Beberapa contoh uji parametrik yang umum digunakan antara lain:

1. **Uji t (*t-test*)**, digunakan untuk menguji perbedaan rata-rata antara dua kelompok, baik yang saling bebas (*independent sample t-test*) maupun berpasangan (*paired sample t-test*).
2. **Analisis varians (*ANOVA*)**, digunakan untuk menguji perbedaan rata-rata lebih dari dua kelompok. Jika ditemukan perbedaan yang signifikan, dapat dilanjutkan dengan uji lanjutan (*post hoc test*).

3. **Uji korelasi Pearson (*Pearson correlation test*)**, digunakan untuk mengukur kekuatan dan arah hubungan linear antara dua variabel interval atau rasio.

Uji parametrik memberikan hasil yang lebih akurat ketika asumsi dasar terpenuhi. Namun, apabila data tidak memenuhi asumsi normalitas atau homogenitas, penggunaan uji ini dapat menghasilkan kesimpulan yang menyesatkan (Field, 2018).

5.4.3 Uji Nonparametrik

Berbeda dengan uji parametrik, uji nonparametrik digunakan untuk menganalisis data yang tidak memenuhi asumsi distribusi normal atau memiliki skala pengukuran ordinal dan nominal. Uji ini lebih fleksibel dan dapat diterapkan pada data dengan sebaran yang tidak homogen.

Jenis-jenis uji nonparametrik antara lain:

1. **Uji Mann-Whitney**, sebagai alternatif dari *independent sample t-test*, digunakan untuk membandingkan median dua kelompok independen.
2. **Uji Kruskal-Wallis**, digunakan untuk membandingkan median lebih dari dua kelompok independen dan merupakan alternatif dari *ANOVA*.
3. **Uji Chi-square**, digunakan untuk menguji hubungan antarvariabel kategorik dalam bentuk frekuensi.

Meskipun uji nonparametrik memiliki kekuatan statistik yang lebih rendah dibandingkan uji parametrik, keunggulannya terletak pada fleksibilitas dan ketahanannya terhadap pelanggaran asumsi statistik (Gravetter & Wallnau, 2020).

5.4.4 Pertimbangan Pemilihan Jenis Uji Statistik

Pemilihan jenis uji statistik harus disesuaikan dengan karakteristik data dan tujuan penelitian. Jika data memenuhi asumsi normalitas dan homogenitas, maka uji parametrik merupakan pilihan terbaik. Namun, apabila data tidak berdistribusi normal atau berjenis ordinal/kategorik, maka uji nonparametrik lebih tepat digunakan. Selain itu, ukuran sampel juga berpengaruh; uji parametrik cenderung lebih stabil pada sampel besar, sedangkan uji nonparametrik lebih sesuai untuk sampel kecil.

Menurut Field (2018), peneliti juga perlu memperhatikan desain penelitian (komparatif, korelasional, atau eksperimental) karena setiap rancangan membutuhkan jenis uji statistik yang berbeda. Dengan mempertimbangkan semua aspek tersebut, peneliti dapat memilih uji yang paling relevan dan menjaga integritas hasil penelitian secara metodologis.

5.5 Latihan Soal

1. Jelaskan perbedaan antara estimasi titik dan estimasi interval!
2. Apa arti tingkat kepercayaan 95% dalam estimasi interval?
3. Sebutkan langkah-langkah dalam melakukan uji hipotesis!
4. Apa yang dimaksud dengan hipotesis nol (H_0) dan hipotesis alternatif (H_1)?
5. Berikan contoh situasi di mana digunakan uji nonparametrik!

Bab 6: Perbandingan Data Kesehatan Berdasarkan Nilai dan Proporsi

6.1 Mengenal Perbandingan Data Kesehatan Berdasarkan Nilai dan Proporsi

6.1.1 Latar Belakang

Kemajuan ilmu kesehatan modern sangat ditentukan oleh kemampuan melakukan analisis data yang sistematis dan tepat. Informasi kesehatan yang bersumber dari survei populasi, rekam medis elektronik, maupun laporan laboratorium berfungsi sebagai dasar penting dalam pengambilan keputusan berbasis bukti. Aktivitas membandingkan data kesehatan menjadi langkah kunci untuk menilai perbedaan antar kelompok, mengidentifikasi pola perubahan dari waktu ke waktu, serta menilai efektivitas intervensi.

Perbandingan data kesehatan terbagi menjadi dua bentuk utama, yaitu data kontinu dan data proporsi. Data kontinu berupa angka dengan rentang nilai tertentu, misalnya tekanan darah, kadar glukosa, atau indeks massa tubuh (BMI). Data proporsi menggambarkan perbandingan jumlah kejadian terhadap total populasi, contohnya proporsi kasus penyakit tertentu dalam

masyarakat. Pemahaman yang baik terhadap kedua jenis data ini menjadi syarat penting agar hasil analisis dapat ditafsirkan secara akurat (Field, 2018).

6.1.2 Tujuan Analisis Perbandingan Data

Analisis perbandingan bertujuan mengidentifikasi perbedaan signifikan yang mendukung pengambilan keputusan dalam kebijakan maupun praktik klinis. Pada penelitian kesehatan masyarakat, teknik ini digunakan untuk menilai dampak intervensi, mengukur kesenjangan antarwilayah, serta memantau perkembangan penyakit dari waktu ke waktu.

Contoh penerapannya adalah analisis rata-rata tekanan darah berdasarkan kelompok usia atau jenis kelamin untuk menilai risiko hipertensi. Perbandingan proporsi penderita diabetes antara wilayah perkotaan dan pedesaan juga dapat menjadi dasar strategi intervensi yang lebih tepat sasaran. Pemilihan metode analisis yang sesuai menentukan keabsahan hasil penelitian (Kirkwood & Sterne, 2017).

6.1.3 Jenis dan Karakteristik Data

Jenis data menentukan metode statistik yang digunakan. Data kontinu, seperti kadar kolesterol atau berat badan, biasanya dianalisis dengan uji parametrik seperti *t-test* atau ANOVA. Data proporsi, misalnya jumlah pasien sembuh dibandingkan total pasien, lebih tepat dianalisis menggunakan uji chi-square atau uji z.

Setiap metode memiliki asumsi yang harus dipenuhi. *T-test*, misalnya, mengharuskan data berdistribusi normal dan varians antar kelompok homogen. Pelanggaran terhadap asumsi dapat menghasilkan kesimpulan yang keliru. Oleh karena itu, peneliti perlu

melakukan uji normalitas, uji homogenitas, serta memastikan ukuran sampel memadai.

6.1.4 Tantangan dalam Analisis Data Kesehatan

Analisis data kesehatan sering menghadapi kendala metodologis. Data yang hilang akibat kesalahan pencatatan atau ketidakhadiran responden menjadi salah satu masalah umum. Bias seleksi juga dapat muncul ketika sampel tidak mewakili populasi. Perbedaan standar pengukuran antar lembaga kesehatan menambah variasi hasil yang sulit dikendalikan.

Pendekatan modern seperti multiple imputation dan sensitivity analysis dapat membantu mengurangi dampak data hilang. Prinsip *data cleaning* serta pemanfaatan perangkat lunak statistik seperti R, SPSS, atau Python memungkinkan analisis lebih komprehensif dan efisien.

6.1.5 Kesimpulan Sementara

Pemahaman terhadap teknik perbandingan data kesehatan merupakan fondasi penting dalam epidemiologi dan kebijakan publik. Baik data kontinu maupun proporsi memiliki metode analisis tersendiri yang harus dipahami agar hasil penelitian valid. Kemampuan memilih teknik analisis yang tepat mencerminkan kompetensi metodologis sekaligus mendukung kebijakan kesehatan berbasis bukti.

6.2 Perbandingan Berdasarkan Nilai (Data Kontinu)

6.2.1 Definisi dan Karakteristik Data Kontinu

Data kontinu adalah data kuantitatif yang dapat diukur secara numerik dan memiliki rentang nilai berkesinambungan. Contoh umum meliputi berat badan, tinggi badan, tekanan darah, kadar glukosa, dan kolesterol. Berbeda dengan data diskrit yang hanya memiliki nilai tertentu, data kontinu memungkinkan nilai di antara dua angka. Penggunaan data kontinu dalam penelitian memungkinkan perhitungan rerata (*mean*), simpangan baku (*standard deviation*), dan interval kepercayaan (*confidence interval*). Informasi ini sangat penting dalam pengambilan keputusan berbasis bukti (*evidence-based decision making*) (Field, 2018).

6.2.2 Uji Normalitas Data

Distribusi normal menjadi syarat utama bagi banyak uji parametrik, seperti pada analisis perbandingan, peneliti perlu memastikan apakah data kontinu yang dikumpulkan mengikuti distribusi normal. Uji normalitas bertujuan untuk menentukan apakah asumsi kenormalan terpenuhi, karena banyak uji statistik parametrik seperti *t-test* dan *ANOVA* mensyaratkan distribusi normal. Dua uji yang paling umum digunakan adalah Uji Shapiro–Wilk dan Uji Kolmogorov–Smirnov.

Uji Shapiro–Wilk lebih disarankan untuk ukuran sampel kecil hingga menengah ($n < 50$), sedangkan Uji Kolmogorov–Smirnov lebih sesuai untuk ukuran sampel besar. Hasil uji

dinyatakan dalam nilai p , di mana $p > 0,05$ menandakan bahwa data berdistribusi normal, sementara $p < 0,05$ menunjukkan sebaliknya. Jika data tidak berdistribusi normal, maka diperlukan transformasi data atau penggunaan uji nonparametrik sebagai alternatif (Ghasemi & Zahediasl, 2012).

6.2.3 Uji Perbandingan Dua Kelompok

Penelitian yang melibatkan dua kelompok independen dengan data kontinu biasanya dianalisis menggunakan *Independent Sample t-test*. Uji ini mengasumsikan bahwa data berdistribusi normal dan memiliki varians yang homogen. Contoh penerapannya adalah membandingkan rata-rata kadar kolesterol antara kelompok perokok dan bukan perokok.

Jika asumsi normalitas tidak terpenuhi, alternatif yang lebih tepat adalah Uji Mann–Whitney U. Uji ini tidak bergantung pada distribusi data dan membandingkan median kedua kelompok, bukan rerata. Keunggulan uji Mann–Whitney U terletak pada ketahanannya terhadap data yang memiliki pencilan (*outlier*) serta ketidaknormalan distribusi, sehingga lebih fleksibel untuk data biologis atau klinis yang cenderung bervariasi.

6.2.4 Uji Perbandingan Lebih dari Dua Kelompok

Penelitian yang melibatkan lebih dari dua kelompok dengan data kontinu biasanya dianalisis menggunakan ANOVA (*Analysis of Variance*) untuk menilai apakah terdapat perbedaan signifikan antar rerata kelompok. ANOVA bekerja dengan membandingkan variasi antar kelompok dan dalam kelompok untuk menentukan adanya pengaruh yang bermakna secara statistik. Misalnya, ANOVA dapat

diterapkan untuk membandingkan kadar tekanan darah pada tiga kelompok usia berbeda.

Jika data tidak memenuhi asumsi kenormalan atau homogenitas varians, maka uji nonparametrik seperti Uji Kruskal–Wallis digunakan sebagai alternatif. Uji ini menganalisis perbedaan median di antara beberapa kelompok tanpa memerlukan distribusi normal. Jika hasil uji menunjukkan perbedaan yang signifikan, maka diperlukan uji lanjut (*post-hoc test*) seperti Uji Dunn atau Uji Bonferroni untuk menentukan pasangan kelompok mana yang berbeda secara signifikan.

6.2.5 Implikasi dalam Penelitian

Pemilihan jenis uji statistik yang tepat memiliki implikasi penting terhadap validitas hasil penelitian. Kesalahan dalam menentukan uji dapat menyebabkan interpretasi yang bias dan kesimpulan yang tidak akurat. Oleh karena itu, peneliti perlu memahami karakteristik data, asumsi statistik, serta konteks penelitian sebelum menentukan metode analisis yang akan digunakan.

Hasil analisis juga perlu dilaporkan secara transparan termasuk nilai p , rerata, simpangan baku, serta ukuran efek (*effect size*), agar penelitian dapat direplikasi dan dievaluasi secara kritis. Uji perbandingan berbasis data kontinu memiliki peran lebih dari sekadar langkah teknis, karena penerapannya merupakan wujud komitmen terhadap integritas ilmiah dalam penelitian.

6.3 Perbandingan Berdasarkan Proporsi (Data Kategorikal)

6.3.1 Definisi Data Kategorikal

Data kategorikal merupakan jenis data yang disusun dalam kategori atau kelas yang tidak memiliki urutan tertentu. Jenis data ini berfungsi untuk menggambarkan karakteristik atau atribut yang bersifat deskriptif dan tidak dapat diukur secara numerik. Contoh umum data kategorikal mencakup jenis kelamin (laki-laki/perempuan), status perokok (ya/tidak), dan status imunisasi (lengkap/tidak lengkap). Pada analisis statistik, data kategorikal umumnya terbagi menjadi dua jenis utama, yaitu *nominal* dan *ordinal*. Data *nominal* tidak memiliki urutan logis seperti warna mata atau golongan darah, sedangkan data *ordinal* menunjukkan tingkatan tertentu, seperti tingkat pendidikan atau tingkat kepuasan (*low*, *medium*, *high*) (Agresti, 2019). Analisis terhadap data kategorikal biasanya bertujuan untuk memahami hubungan, perbandingan proporsi, atau distribusi antar-kelompok, yang sering digunakan dalam penelitian sosial, ekonomi, maupun kesehatan masyarakat.

6.3.2 Uji Perbandingan Dua Proporsi

Uji perbandingan dua proporsi bertujuan untuk menilai apakah terdapat perbedaan yang signifikan antara dua kelompok dalam suatu variabel kategorikal. Metode yang paling umum digunakan untuk tujuan ini adalah uji *Chi-Square* (*Chi-Square Test of Independence*). Uji ini digunakan untuk mengukur sejauh mana

frekuensi observasi berbeda dari frekuensi harapan berdasarkan asumsi bahwa tidak ada hubungan atau perbedaan antara dua variabel yang dibandingkan. Hipotesis nol (H_0) dalam uji ini menyatakan bahwa tidak terdapat perbedaan proporsi antara dua kelompok, sedangkan hipotesis alternatif (H_1) menyatakan adanya perbedaan proporsi yang signifikan. Nilai statistik *Chi-Square* dihitung dengan rumus $\chi^2 = \sum \frac{(O_i - E_i)^2}{E_i}$, di mana O_i merupakan frekuensi observasi dan E_i merupakan frekuensi harapan. Hasil uji kemudian dibandingkan dengan distribusi *Chi-Square* pada derajat kebebasan tertentu untuk menentukan tingkat signifikansi.

Hasil analisis dapat diinterpretasikan dengan benar jika memenuhi beberapa asumsi, antara lain: data harus disajikan dalam bentuk frekuensi, observasi antar-sampel bersifat independen, serta setiap sel dalam tabel kontingensi memiliki frekuensi harapan minimal lima (McHugh, 2013). Contoh implementasinya, jika peneliti ingin mengetahui apakah terdapat perbedaan proporsi perokok antara laki-laki dan perempuan, maka data dikumpulkan dalam tabel dua arah, kemudian diuji menggunakan *Chi-Square*. Jika hasil menunjukkan nilai $p < 0,05$, dapat disimpulkan bahwa terdapat perbedaan proporsi yang signifikan antara kedua kelompok. Penerapan uji *Chi-Square* memungkinkan peneliti mengidentifikasi perbedaan proporsi dalam data kategorikal dua kelompok secara lebih meyakinkan.

6.3.3 Uji Perbandingan Lebih dari Dua Proporsi

Pada situasi di mana jumlah kelompok yang dibandingkan lebih dari dua, uji *Chi-Square* tetap dapat digunakan untuk menilai apakah terdapat perbedaan proporsi yang signifikan antar-kelompok. Prinsip dasarnya sama, yaitu membandingkan frekuensi observasi dengan frekuensi harapan untuk menilai apakah distribusi kategori bersifat seragam di seluruh kelompok. Sebagai contoh, penelitian yang ingin menilai perbedaan tingkat kepatuhan imunisasi anak di tiga wilayah berbeda dapat menggunakan uji *Chi-Square* untuk mengetahui apakah proporsi anak dengan imunisasi lengkap berbeda secara signifikan di antara wilayah tersebut. Perhitungan derajat kebebasan menggunakan rumus $(r-1)(c-1)$, dengan r menunjukkan jumlah baris sebagai kategori variabel pertama dan c menunjukkan jumlah kolom sebagai kategori variabel kedua. Nilai *Chi-Square* yang besar menandakan bahwa terdapat deviasi besar antara nilai observasi dan nilai harapan, sehingga mengindikasikan adanya perbedaan proporsi yang signifikan antar-kelompok.

Namun, dalam praktiknya, jika terdapat sel dengan frekuensi harapan kurang dari lima, penggunaan uji *Chi-Square* menjadi kurang tepat karena asumsi dasar tidak terpenuhi. Dalam kondisi tersebut, peneliti disarankan menggunakan alternatif seperti *Fisher's Exact Test* untuk tabel 2×2 atau *Likelihood Ratio Test* untuk ukuran sampel kecil. Kedua alternatif ini dapat memberikan estimasi signifikansi yang lebih akurat dalam situasi dengan distribusi data yang tidak seimbang. Kesimpulan yang valid dan reliabel hanya dapat dicapai apabila pemilihan metode uji dilakukan dengan

memperhatikan karakteristik data, ukuran sampel, dan struktur tabel kontingensi.

6.3.4 Interpretasi dan Implikasi

Hasil uji perbandingan proporsi harus ditafsirkan secara kontekstual. Nilai p yang signifikan menunjukkan adanya perbedaan proporsi antar-kelompok, namun tidak menunjukkan hubungan sebab-akibat. Oleh karena itu, analisis tambahan seperti ukuran efek (*effect size*)—misalnya *Cramer's V* atau *Phi coefficient*—diperlukan untuk memahami kekuatan hubungan antar-variabel. Interpretasi yang tepat akan membantu peneliti memahami relevansi temuan dalam konteks empiris, sehingga hasil analisis proporsi tidak hanya menunjukkan perbedaan statistik, tetapi juga memiliki makna praktis bagi pengambilan keputusan di bidang sosial, ekonomi, maupun kesehatan masyarakat.

6.4 Contoh Praktik: Perbandingan Data Kesehatan

Analisis data kesehatan sering kali memerlukan penerapan uji statistik yang berbeda, tergantung pada jenis data dan tujuan penelitian. Metode statistik dapat dipahami lebih jelas melalui tiga contoh kasus yang menunjukkan penerapannya dalam membandingkan dan menelaah fenomena kesehatan masyarakat. Setiap contoh memberikan gambaran mengenai pendekatan analitis yang sesuai dengan karakteristik data dan asumsi statistik yang mendasarinya.

Contoh pertama berkaitan dengan perbandingan tekanan darah antara pria dan wanita menggunakan *t-test* untuk dua sampel independen. Analisis ini bertujuan untuk menentukan apakah terdapat perbedaan signifikan secara statistik antara rata-rata tekanan darah kedua kelompok tersebut. Pemenuhan asumsi normalitas dan homogenitas varians harus dipastikan peneliti agar uji dapat dilakukan dengan tepat. Jika kedua asumsi ini terpenuhi, maka hasil *t-test* akan menunjukkan perbedaan rata-rata tekanan darah yang tidak semata-mata disebabkan oleh kebetulan. Hasil penelitian oleh Omelchenko *et al.* (2020) menunjukkan bahwa pria cenderung memiliki tekanan darah lebih tinggi dibandingkan wanita, terutama pada usia produktif, yang disebabkan oleh faktor hormonal serta kebiasaan gaya hidup. Apabila nilai $p < 0,05$ diperoleh, maka dapat disimpulkan bahwa terdapat perbedaan signifikan antara kedua kelompok, yang penting untuk perencanaan program pencegahan hipertensi berbasis gender.

Contoh kedua adalah analisis prevalensi penyakit diabetes berdasarkan usia menggunakan uji *Chi-Square*. Uji ini digunakan untuk menentukan apakah terdapat hubungan yang signifikan antara dua variabel kategori, yaitu kelompok usia dan status diabetes (positif atau negatif). Variabel usia biasanya dikelompokkan dalam beberapa rentang, seperti di bawah 30 tahun, 30–50 tahun, dan di atas 50 tahun. Hasil perhitungan *Chi-Square* memberikan gambaran tentang kekuatan hubungan antara usia dan risiko diabetes. Pada penelitian epidemiologis, semakin bertambah usia seseorang, semakin tinggi pula risiko mengalami gangguan metabolik.

Penelitian oleh Huang *et al.* (2020) mengonfirmasi bahwa prevalensi diabetes meningkat secara signifikan pada individu berusia di atas 45 tahun, yang menunjukkan adanya korelasi kuat antara faktor usia dan peningkatan risiko penyakit tersebut. Jika hasil uji menunjukkan nilai $p < 0,05$, maka hal ini menandakan hubungan yang bermakna antara variabel usia dan status diabetes. Temuan seperti ini menjadi dasar bagi pemerintah dalam merancang kebijakan skrining dan edukasi kesehatan masyarakat untuk pencegahan dini.

Contoh ketiga adalah pengujian pengaruh vaksinasi pada anak-anak menggunakan uji *Fisher's Exact*. Uji ini digunakan ketika ukuran sampel kecil dan data tidak memenuhi asumsi uji *Chi-Square*. Relevansi metode ini tampak jelas pada penelitian kesehatan anak, khususnya ketika menganalisis hubungan antara status vaksinasi (ya/tidak) dan kejadian penyakit menular (terinfeksi/tidak). Uji *Fisher's Exact* menghitung probabilitas eksak dari distribusi data yang diamati, sehingga hasilnya lebih akurat dibandingkan pendekatan berbasis distribusi asimtotik pada ukuran sampel kecil. Apabila hasil analisis menghasilkan nilai signifikansi $< 0,05$, maka dapat disimpulkan bahwa vaksinasi memiliki pengaruh signifikan terhadap penurunan kejadian penyakit menular. Hasil ini memperkuat urgensi program imunisasi sebagai strategi pencegahan penyakit menular di kalangan anak-anak, terutama di wilayah dengan cakupan vaksinasi yang masih rendah. Ketiga contoh tersebut menegaskan betapa pentingnya pemilihan metode statistik yang tepat dalam penelitian kesehatan guna menghasilkan

kesimpulan yang sah serta dapat dipertanggungjawabkan secara ilmiah.

6.5 Pentingnya Uji Statistik dalam Data Kesehatan

6.5.1 Konsep Dasar Uji Statistik

Uji statistik merupakan prosedur ilmiah yang digunakan untuk menentukan ada atau tidaknya perbedaan, hubungan, atau pengaruh antarvariabel berdasarkan data yang dikumpulkan secara sistematis. Uji statistik berperan penting pada penelitian kesehatan karena menjamin kesimpulan tidak semata-mata didasarkan pada pengamatan subjektif, melainkan diperkuat oleh bukti empiris yang terukur secara matematis. Uji ini membantu peneliti menilai sejauh mana hasil penelitian dapat digeneralisasi terhadap populasi yang lebih luas.

Penggunaan uji statistik memungkinkan peneliti untuk mengendalikan kemungkinan terjadinya kesalahan tipe I (menolak hipotesis nol yang benar) dan tipe II (menerima hipotesis nol yang salah). Analisis statistik yang akurat menghasilkan keputusan dengan tingkat kepercayaan yang jelas terukur. Relevansi hal ini dalam bidang kesehatan sangat tinggi, sebab kesalahan interpretasi data dapat berdampak pada kebijakan publik, praktik klinis, hingga evaluasi terapi.

6.5.2 Jenis dan Pemilihan Uji Statistik

Pemilihan jenis uji statistik harus mempertimbangkan karakteristik data dan tujuan analisis. Uji statistik dibagi menjadi dua kategori besar, yaitu uji parametrik dan nonparametrik. Uji parametrik, seperti *t-test* dan *ANOVA*, digunakan apabila data memenuhi asumsi distribusi normal dan homogenitas varians. Ketika data tidak sesuai dengan asumsi parametrik atau berbentuk ordinal, analisis dapat dilakukan menggunakan uji nonparametrik seperti *Mann-Whitney U test* dan *Kruskal-Wallis test*.

Pada penelitian kesehatan, pemilihan uji statistik yang tepat juga bergantung pada desain penelitian. Misalnya, untuk menguji efektivitas suatu intervensi medis, uji *paired t-test* dapat digunakan jika data diambil dari sampel yang sama sebelum dan sesudah perlakuan. Sebaliknya, jika peneliti ingin mengetahui hubungan antara dua variabel kategorikal, uji *chi-square* merupakan pilihan yang lebih sesuai. Ketepatan dalam memilih uji statistik ini menentukan validitas hasil penelitian serta keandalan rekomendasi yang dihasilkan (Kim, 2017).

6.5.3 Peran Uji Statistik dalam Validitas dan Reliabilitas Penelitian

Validitas dan reliabilitas merupakan dua komponen penting dalam penilaian kualitas penelitian kesehatan. Uji statistik membantu memastikan bahwa data yang dianalisis benar-benar mengukur apa yang seharusnya diukur (validitas), serta memberikan hasil yang konsisten jika penelitian diulang (reliabilitas). Penggunaan metode statistik yang tepat memungkinkan peneliti

mengidentifikasi sumber variabilitas data, baik yang berasal dari perbedaan antarindividu, instrumen pengukuran, maupun kondisi lingkungan penelitian.

Uji statistik juga berperan dalam mengendalikan *confounding variables* yang dapat memengaruhi hasil penelitian. Misalnya, dalam studi tentang pengaruh pola makan terhadap tekanan darah, variabel seperti usia, jenis kelamin, dan tingkat aktivitas fisik perlu dikontrol secara statistik agar hasil yang diperoleh benar-benar merefleksikan hubungan kausal antara pola makan dan tekanan darah. Interpretasi hasil menjadi lebih akurat sehingga dapat dimanfaatkan sebagai dasar pengambilan keputusan klinis maupun kebijakan kesehatan masyarakat (Peat, 2021).

6.5.4 Tantangan dan Etika dalam Penggunaan Uji Statistik

Uji statistik memberikan kekuatan analitis yang signifikan, tetapi penggunaannya juga menuntut tanggung jawab etis dan kompetensi metodologis yang tinggi. Kesalahan umum yang sering terjadi adalah pemilihan uji statistik tanpa mempertimbangkan asumsi dasar atau tujuan penelitian yang sebenarnya. Hal ini dapat mengarah pada kesimpulan yang keliru dan menyesatkan, terutama jika hasil tersebut dijadikan dasar untuk kebijakan kesehatan publik.

Kecenderungan melakukan data dredging atau *p-hacking* masih ditemukan pada sebagian peneliti, di mana hasil signifikan dicari tanpa adanya hipotesis yang terdefinisi dengan baik. Praktik ini tidak hanya melanggar etika penelitian, tetapi juga merusak integritas ilmiah karena menghasilkan temuan yang tidak dapat direplikasi. Pemahaman mendalam mengenai prinsip-prinsip

statistik serta penerapan uji yang tepat merupakan keharusan bagi setiap peneliti di bidang kesehatan.

6.6 Kesimpulan

6.6.1 Pentingnya Perbandingan Data Kesehatan

Perbandingan data kesehatan berdasarkan nilai dan proporsi merupakan salah satu pendekatan yang krusial dalam analisis epidemiologis dan penelitian kesehatan masyarakat. Teknik ini memungkinkan peneliti untuk memahami sejauh mana perbedaan antara kelompok baik berdasarkan usia, jenis kelamin, wilayah, maupun status sosial ekonomi yang mempengaruhi indikator kesehatan tertentu. Perbandingan nilai absolut maupun proporsi relatif memungkinkan peneliti menilai keberagaman distribusi penyakit, pola perilaku kesehatan, serta efektivitas intervensi pada populasi yang berbeda (Friis & Sellers, 2021).

Pada praktiknya, perbandingan tersebut sering digunakan untuk mengidentifikasi kesenjangan kesehatan (*health disparities*). Misalnya, analisis proporsi penderita hipertensi antara kelompok penduduk perkotaan dan pedesaan dapat menggambarkan ketimpangan akses terhadap layanan kesehatan. Pendekatan ini tidak hanya bersifat deskriptif, tetapi juga memiliki nilai strategis dalam perumusan kebijakan kesehatan berbasis bukti. Keandalan data dan ketepatan metode pengukuran menjadi faktor penentu validitas hasil perbandingan yang dilakukan.

6.6.2 Metode Analisis Nilai dan Proporsi

Pada konteks metodologis, terdapat dua pendekatan utama dalam perbandingan data kesehatan, yakni perbandingan berdasarkan nilai rata-rata (*mean comparison*) dan perbandingan berdasarkan proporsi (*proportion comparison*). Pendekatan pertama digunakan ketika variabel yang dianalisis bersifat kuantitatif, seperti tekanan darah, kadar gula darah, atau indeks massa tubuh. Sementara itu, perbandingan proporsi digunakan ketika variabel yang diteliti bersifat kategorikal, seperti status merokok, kepatuhan terhadap pengobatan, atau kejadian penyakit tertentu (Kirkwood & Sterne, 2017).

Penggunaan uji statistik seperti *t-test* atau *ANOVA* untuk perbandingan nilai rata-rata, serta *chi-square test* untuk perbandingan proporsi, menjadi praktik umum dalam penelitian kesehatan. Uji-uji ini membantu menentukan apakah perbedaan yang ditemukan secara statistik signifikan atau sekadar terjadi secara kebetulan. Selain itu, pemilihan metode analisis harus mempertimbangkan ukuran sampel, distribusi data, serta asumsi yang melekat pada setiap teknik statistik. Pemahaman metodologis yang tepat memungkinkan hasil perbandingan diinterpretasikan secara akurat sehingga mendukung kesimpulan ilmiah yang valid.

6.6.3 Implikasi terhadap Kebijakan dan Praktik Kesehatan

Hasil perbandingan data kesehatan memiliki implikasi luas terhadap kebijakan publik dan praktik medis. Ketika perbedaan signifikan ditemukan antara kelompok tertentu, hal tersebut dapat menjadi dasar bagi pemerintah atau lembaga kesehatan untuk

menetapkan prioritas intervensi. Sebagai contoh, bila proporsi kejadian diabetes lebih tinggi pada kelompok dengan pendapatan rendah, maka intervensi yang berfokus pada edukasi gizi dan akses terhadap makanan sehat menjadi lebih relevan.

Hasil dari perbandingan nilai dan proporsi dapat digunakan sebagai alat evaluasi terhadap efektivitas program kesehatan. Program imunisasi, misalnya, dapat dinilai keberhasilannya dengan membandingkan proporsi anak yang telah mendapatkan vaksinasi sebelum dan sesudah intervensi dilakukan. Dengan demikian, analisis ini berperan penting dalam siklus kebijakan kesehatan yang berorientasi pada hasil (*evidence-based policy*). Lebih jauh, pemanfaatan data komparatif juga mendorong terwujudnya sistem kesehatan yang lebih adil dan efisien.

6.6.4 Kesimpulan Umum

Secara keseluruhan, perbandingan data kesehatan berdasarkan nilai dan proporsi merupakan fondasi analisis yang memungkinkan peneliti dan pembuat kebijakan memahami dinamika kesehatan populasi secara lebih komprehensif. Berbagai determinan sosial dan biologis dapat diidentifikasi serta dikaitkan dengan outcome kesehatan tertentu dengan menggunakan teknik ini. Akurasi dan reliabilitas data menjadi prasyarat utama agar hasil perbandingan dapat diandalkan dan digunakan sebagai dasar pengambilan keputusan. Oleh karena itu, peningkatan kapasitas analitik dan integritas data perlu menjadi perhatian dalam setiap kegiatan penelitian maupun perencanaan kebijakan di bidang kesehatan.

Bab 7: Hubungan antar Variabel dalam Konteks Kesehatan

7.1 Konsep Dasar Hubungan antar Variabel

Dalam penelitian kesehatan, pemahaman terhadap hubungan antar variabel merupakan aspek fundamental yang menentukan validitas dan kekuatan kesimpulan ilmiah yang dihasilkan. Analisis hubungan antar variabel tidak hanya membantu menjelaskan fenomena kesehatan yang diamati, tetapi juga menjadi dasar dalam merancang intervensi dan kebijakan kesehatan yang berbasis bukti (*evidence-based*). Melalui analisis ini, peneliti dapat mengidentifikasi faktor-faktor risiko, penyebab penyakit, serta menilai efektivitas suatu tindakan atau program kesehatan. Oleh karena itu, pemahaman konsep dasar mengenai jenis-jenis variabel, bentuk hubungan yang mungkin terjadi di antara variabel tersebut, serta metode analisis yang sesuai, merupakan fondasi penting dalam metodologi penelitian epidemiologi dan biostatistik.

Secara umum, variabel dalam penelitian kesehatan dapat diklasifikasikan menjadi tiga kategori utama: variabel bebas (*independent variable*), variabel terikat (*dependent variable*), dan variabel pengganggu (*confounding variable*). Variabel bebas

merupakan faktor yang diduga memengaruhi atau menyebabkan perubahan pada variabel lain. Contohnya, kebiasaan merokok dapat dianggap sebagai variabel bebas dalam penelitian yang meneliti hubungan antara perilaku merokok dan penyakit jantung koroner. Sementara itu, variabel terikat adalah hasil atau efek yang diukur untuk melihat pengaruh variabel bebas, misalnya tingkat kejadian penyakit jantung. Variabel pengganggu atau *confounder* adalah faktor lain yang dapat memengaruhi baik variabel bebas maupun variabel terikat, sehingga berpotensi menimbulkan bias dalam interpretasi hubungan antar variabel. Sebagai contoh, dalam studi mengenai hubungan antara aktivitas fisik dan obesitas, pola makan dapat berperan sebagai variabel pengganggu yang perlu dikendalikan agar hubungan yang diamati benar-benar mencerminkan pengaruh aktivitas fisik terhadap obesitas.

Hubungan antar variabel dalam penelitian kesehatan dapat dibedakan menjadi dua bentuk utama, yaitu hubungan kausal (*causal relationship*) dan hubungan asosiasi (*associative relationship*). Hubungan kausal menunjukkan adanya hubungan sebab-akibat yang nyata, di mana perubahan pada variabel bebas secara langsung menyebabkan perubahan pada variabel terikat. Untuk membuktikan hubungan kausal, peneliti harus memenuhi beberapa kriteria, seperti temporalitas (sebab terjadi sebelum akibat), kekuatan asosiasi, konsistensi hasil penelitian, dan adanya mekanisme biologis yang mendukung. Di sisi lain, hubungan asosiasi hanya menunjukkan adanya keterkaitan antara dua variabel tanpa menunjukkan sebab-akibat langsung. Misalnya, ditemukan hubungan antara konsumsi

makanan cepat saji dengan obesitas, tetapi hubungan tersebut belum tentu kausal tanpa memperhitungkan faktor lain seperti aktivitas fisik atau status sosial ekonomi.

Analisis statistik menjadi alat utama dalam memahami dan menguji hubungan antar variabel. Dalam konteks penelitian kesehatan, analisis biostatistik digunakan untuk menentukan kekuatan, arah, dan signifikansi hubungan yang ditemukan. Teknik analisis yang digunakan bergantung pada jenis data dan desain penelitian. Untuk data numerik, dapat digunakan uji korelasi Pearson atau Spearman guna mengukur kekuatan hubungan antara dua variabel. Sementara itu, untuk menilai perbedaan rata-rata antar kelompok, digunakan uji *t* atau ANOVA. Dalam penelitian yang melibatkan lebih dari satu variabel bebas, analisis regresi, baik linier maupun logistik, digunakan untuk memprediksi hubungan dan mengontrol pengaruh variabel pengganggu. Dalam studi epidemiologi yang lebih kompleks, analisis *multivariate modeling* menjadi penting untuk memahami interaksi antar faktor risiko dan outcome kesehatan secara simultan (Kirkwood & Sterne, 2003).

Penting untuk dicatat bahwa interpretasi hasil analisis hubungan antar variabel harus dilakukan secara hati-hati. Hubungan yang signifikan secara statistik belum tentu menunjukkan hubungan kausal secara biologis. Peneliti perlu mempertimbangkan konteks teoritis, desain penelitian, dan kemungkinan adanya bias atau kesalahan pengukuran. Selain itu, konsep *confounding*, *effect modification*, dan *interaction* harus diperhatikan agar interpretasi hasil tidak menyesatkan. Dalam penelitian kesehatan masyarakat,

pemahaman terhadap hubungan antar variabel juga memiliki implikasi praktis, misalnya dalam menentukan prioritas intervensi dan merancang kebijakan pencegahan yang efektif.

Dengan demikian, konsep dasar hubungan antar variabel merupakan fondasi penting dalam metodologi penelitian kesehatan. Pemahaman yang tepat mengenai jenis variabel, bentuk hubungan, serta penerapan analisis statistik yang sesuai akan membantu menghasilkan temuan yang valid, reliabel, dan bermakna. Melalui pendekatan ilmiah yang sistematis, hubungan antar variabel dapat diinterpretasikan secara komprehensif, sehingga hasil penelitian dapat digunakan untuk memperbaiki kesehatan masyarakat, meningkatkan efektivitas program intervensi, dan memperkuat kebijakan kesehatan berbasis bukti ilmiah (Szklo & Nieto, 2019).

7.2 Jenis Hubungan antar Variabel

Dalam penelitian epidemiologi maupun ilmu kesehatan masyarakat, pemahaman mengenai hubungan antar variabel menjadi dasar penting untuk menafsirkan fenomena yang terjadi dalam populasi. Variabel didefinisikan sebagai karakteristik yang dapat berubah-ubah nilainya, baik secara kuantitatif maupun kualitatif. Hubungan antar variabel mencerminkan pola keterkaitan antara satu faktor dengan faktor lainnya, yang dapat bersifat sebab-akibat atau sekadar korelasional. Analisis hubungan ini membantu peneliti mengidentifikasi determinan penyakit, menilai kekuatan asosiasi,

serta merancang intervensi kesehatan berbasis bukti (Szklo & Nieto, 2019).

7.2.1 Hubungan Linier dan Non-Linier

Hubungan antar variabel dapat digolongkan menjadi dua bentuk utama, yaitu **hubungan linier** dan **hubungan non-linier**.

Hubungan linier menggambarkan kondisi di mana perubahan pada satu variabel sebanding dengan perubahan variabel lainnya. Dalam hubungan ini, jika satu variabel meningkat secara konstan, maka variabel lainnya juga akan meningkat atau menurun dalam proporsi yang tetap. Secara matematis, hubungan linier dapat dinyatakan dalam bentuk persamaan garis lurus $Y = a + bX$, di mana b menunjukkan arah dan kekuatan hubungan. Contohnya, hubungan antara indeks massa tubuh (IMT) dan tekanan darah sering menunjukkan pola linier positif; semakin tinggi IMT seseorang, semakin besar pula kemungkinan peningkatan tekanan darah.

Sebaliknya, hubungan non-linier menggambarkan keterkaitan yang tidak proporsional, di mana perubahan pada satu variabel tidak selalu diikuti perubahan konstan pada variabel lainnya. Hubungan ini sering kali berbentuk kurva atau fungsi eksponensial. Dalam konteks kesehatan, hubungan non-linier dapat ditemukan pada hubungan antara dosis obat dan respons biologis, di mana efek terapi meningkat hanya hingga titik tertentu sebelum akhirnya menurun karena efek toksik. Pemahaman terhadap pola non-linier penting agar peneliti tidak salah menafsirkan kekuatan hubungan dan implikasi praktisnya (Kirkwood & Sterne, 2010).

7.2.2 Hubungan Positif dan Negatif

Selain dilihat dari bentuknya, hubungan antar variabel juga dapat dikategorikan berdasarkan arah korelasi, yaitu **hubungan positif** dan **hubungan negatif**.

Hubungan positif terjadi ketika peningkatan satu variabel diikuti oleh peningkatan variabel lainnya. Misalnya, semakin tinggi tingkat aktivitas fisik, maka semakin besar pula pengeluaran energi tubuh. Hubungan ini menunjukkan arah yang sejalan. Sementara itu, hubungan negatif menggambarkan kondisi sebaliknya, di mana peningkatan satu variabel menyebabkan penurunan variabel lain. Contohnya, semakin tinggi tingkat vaksinasi dalam suatu populasi, semakin rendah tingkat kejadian penyakit menular.

Arah hubungan ini dapat diukur melalui koefisien korelasi (r), yang nilainya berkisar antara -1 hingga +1. Nilai positif menunjukkan korelasi positif, nilai negatif menunjukkan korelasi negatif, sedangkan nilai mendekati nol menunjukkan tidak adanya hubungan linear yang kuat antara kedua variabel. Pemahaman arah dan kekuatan korelasi menjadi dasar dalam menginterpretasikan hasil penelitian, terutama dalam menentukan apakah hubungan yang ditemukan bersifat kausal atau hanya kebetulan statistik.

7.2.3 Implikasi Analisis Hubungan antar Variabel

Analisis hubungan antar variabel tidak hanya bertujuan untuk mendeskripsikan keterkaitan, tetapi juga untuk mengidentifikasi faktor risiko dan penyebab potensial dari suatu kondisi kesehatan. Dalam epidemiologi, hubungan yang kuat antara variabel paparan (seperti merokok) dan variabel akibat (seperti

kanker paru) dapat menjadi dasar bagi hipotesis kausal. Namun, interpretasi hubungan harus dilakukan dengan hati-hati karena korelasi tidak selalu berarti sebab-akibat. Faktor perancu (*confounding factors*) sering kali memengaruhi kekuatan hubungan yang tampak.

Oleh karena itu, pemahaman tentang jenis hubungan antar variabel membantu peneliti dalam memilih metode analisis yang tepat, seperti regresi linier, korelasi Pearson, atau regresi logistik, sesuai dengan sifat data dan tujuan penelitian. Pendekatan analisis yang akurat memastikan hasil penelitian epidemiologi dapat digunakan secara valid dalam pengambilan keputusan kebijakan kesehatan masyarakat.

7.3 Analisis Korelasi

7.3.1 Konsep dan Prinsip Dasar Analisis Korelasi

Analisis korelasi merupakan salah satu metode statistik yang digunakan untuk mengukur kekuatan dan arah hubungan antara dua variabel kuantitatif. Dalam konteks penelitian ilmiah, analisis ini membantu peneliti memahami sejauh mana perubahan pada satu variabel berkaitan dengan perubahan pada variabel lain. Korelasi tidak menunjukkan hubungan sebab-akibat secara langsung, melainkan hanya menggambarkan keterkaitan yang terjadi di antara variabel yang diamati.

Secara matematis, korelasi diukur menggunakan **koefisien korelasi**, yang nilainya berkisar antara -1 hingga +1. Nilai positif

menunjukkan hubungan searah, di mana peningkatan satu variabel diikuti peningkatan variabel lainnya. Sebaliknya, nilai negatif menunjukkan hubungan berlawanan arah, sedangkan nilai mendekati nol menunjukkan bahwa kedua variabel tidak memiliki hubungan yang signifikan. Analisis korelasi penting untuk memahami pola hubungan yang mendasari fenomena, misalnya dalam bidang pendidikan, kesehatan, atau ekonomi.

Dalam praktiknya, pemilihan metode korelasi sangat bergantung pada jenis dan distribusi data. Untuk data yang berdistribusi normal, digunakan **koefisien korelasi Pearson** (*Pearson's correlation coefficient*), sedangkan untuk data yang tidak berdistribusi normal atau berskala ordinal digunakan **koefisien korelasi Spearman** (*Spearman's rank correlation coefficient*). Pemahaman terhadap asumsi-asumsi ini penting agar hasil analisis valid dan dapat diinterpretasikan secara tepat (Field, 2018).

7.3.2 Jenis dan Teknik Penghitungan Korelasi

Koefisien korelasi Pearson dihitung berdasarkan hubungan linear antara dua variabel kuantitatif. Rumusnya mempertimbangkan kovarians kedua variabel yang dinormalisasi terhadap produk simpangan baku masing-masing. Nilai +1 menandakan hubungan linear positif sempurna, -1 menunjukkan hubungan linear negatif sempurna, dan 0 berarti tidak ada hubungan linear yang signifikan. Misalnya, terdapat korelasi positif yang kuat antara tinggi badan dan berat badan pada populasi tertentu, sedangkan antara lama belajar dan tingkat stres mungkin menunjukkan korelasi negatif.

Sementara itu, koefisien korelasi Spearman digunakan untuk data ordinal atau data yang tidak memenuhi asumsi normalitas. Korelasi ini didasarkan pada peringkat (ranking) nilai variabel, bukan nilai mentahnya. Pendekatan ini berguna dalam situasi di mana hubungan antarvariabel bersifat monoton, yakni meningkat atau menurun secara konsisten tetapi tidak harus linear. Misalnya, dalam penelitian pendidikan, hubungan antara motivasi belajar dan prestasi akademik sering kali lebih cocok dianalisis menggunakan korelasi Spearman karena data cenderung tidak berdistribusi normal.

Selain Pearson dan Spearman, terdapat juga metode korelasi lainnya seperti **Kendall's tau**, yang sering digunakan untuk ukuran sampel kecil atau data dengan banyak nilai yang sama (*ties*). Pemilihan jenis korelasi yang tepat sangat krusial karena kesalahan pemilihan metode dapat menyebabkan interpretasi yang keliru terhadap hubungan antarvariabel. Oleh karena itu, peneliti perlu melakukan uji asumsi terlebih dahulu sebelum menentukan teknik korelasi yang akan digunakan (Dancey & Reidy, 2017).

7.3.3 Interpretasi dan Keterbatasan Analisis Korelasi

Interpretasi hasil analisis korelasi harus dilakukan dengan hati-hati karena korelasi tidak selalu menunjukkan adanya **kausalitas**. Dua variabel dapat memiliki hubungan yang kuat tanpa saling memengaruhi secara langsung. Fenomena ini dikenal sebagai *spurious correlation* atau korelasi semu. Sebagai contoh, peningkatan konsumsi es krim mungkin berkorelasi dengan peningkatan kasus tenggelam, tetapi keduanya disebabkan oleh

variabel ketiga, yaitu suhu udara yang lebih tinggi pada musim panas.

Selain itu, kekuatan hubungan yang ditunjukkan oleh nilai korelasi perlu dipahami dalam konteks bidang kajian. Dalam ilmu sosial, nilai korelasi sebesar 0,5 mungkin dianggap cukup kuat, sedangkan dalam ilmu eksakta nilai tersebut dianggap sedang. Peneliti juga harus memperhatikan ukuran sampel karena sampel kecil dapat menghasilkan nilai korelasi yang tidak stabil.

Untuk meningkatkan keandalan interpretasi, analisis korelasi sering dikombinasikan dengan metode lain seperti regresi linier, yang dapat menjelaskan arah hubungan dan besarnya pengaruh variabel bebas terhadap variabel terikat. Dengan demikian, analisis korelasi berfungsi sebagai langkah awal dalam eksplorasi hubungan antarvariabel sebelum dilakukan analisis yang lebih mendalam.

Secara keseluruhan, analisis korelasi merupakan alat penting dalam penelitian kuantitatif untuk memahami hubungan antarfenomena. Namun, penggunaannya harus disertai pemahaman terhadap asumsi, keterbatasan, dan konteks penelitian agar hasilnya dapat diinterpretasikan secara valid dan ilmiah.

7.4 Analisis Regresi

7.4.1 Konsep Dasar Analisis Regresi

Analisis regresi merupakan salah satu metode statistik yang paling penting dalam penelitian kuantitatif, terutama dalam bidang kesehatan masyarakat, biostatistika, dan ilmu sosial. Tujuan utama

analisis regresi adalah untuk memodelkan dan memprediksi hubungan antara satu atau lebih variabel bebas (*independent variables*) dengan variabel terikat (*dependent variable*). Secara umum, regresi digunakan untuk memperkirakan pengaruh faktor-faktor tertentu terhadap suatu hasil yang diukur.

Model paling sederhana dari analisis regresi adalah regresi linier sederhana (*simple linear regression*), yang hanya melibatkan satu variabel bebas. Persamaan matematisnya ditulis sebagai:

$$Y = a + bX + e$$

di mana Y adalah variabel terikat, X adalah variabel bebas, a adalah konstanta (intersep), b adalah koefisien regresi yang menunjukkan arah dan besarnya pengaruh X terhadap Y , serta e merupakan komponen kesalahan atau residu. Dalam praktiknya, nilai b memberikan informasi mengenai seberapa besar perubahan pada Y yang diharapkan terjadi jika X meningkat satu satuan.

Apabila terdapat lebih dari satu variabel bebas, digunakan model regresi berganda (*multiple linear regression*), yang ditulis dalam bentuk:

$$Y = a + b_1X_1 + b_2X_2 + \cdots + b_nX_n + e$$

Model ini memungkinkan peneliti untuk menilai pengaruh simultan dari beberapa faktor terhadap variabel terikat, sekaligus

mengendalikan pengaruh variabel lain yang mungkin berperan sebagai pengganggu (*confounding variables*).

7.4.2 Penerapan Analisis Regresi dalam Penelitian Kesehatan

Dalam penelitian kesehatan, analisis regresi memiliki peran penting dalam mengidentifikasi dan mengukur hubungan antara faktor risiko dan kejadian penyakit. Misalnya, regresi linier sederhana dapat digunakan untuk meneliti hubungan antara indeks massa tubuh (*body mass index/BMI*) dan tekanan darah sistolik. Hasil analisis dapat menunjukkan apakah peningkatan BMI berhubungan dengan peningkatan tekanan darah, serta seberapa besar pengaruhnya secara kuantitatif.

Sementara itu, regresi berganda digunakan ketika peneliti ingin mengontrol berbagai faktor sekaligus. Misalnya, dalam penelitian mengenai kadar kolesterol, model regresi berganda dapat melibatkan variabel bebas seperti usia, jenis kelamin, aktivitas fisik, dan pola makan. Dengan demikian, peneliti dapat mengisolasi pengaruh masing-masing faktor terhadap kadar kolesterol, sekaligus memperkirakan risiko relatif terhadap penyakit kardiovaskular.

Selain regresi linier, terdapat pula model regresi lain yang digunakan untuk variabel dependen non-linier. Misalnya, regresi logistik (*logistic regression*) digunakan untuk memprediksi kejadian biner seperti “sakit” atau “tidak sakit”, sedangkan regresi Poisson atau regresi negatif binomial digunakan untuk memodelkan data dengan hasil berbentuk hitungan (*count data*). Model-model ini sangat berguna dalam analisis epidemiologi dan studi klinis (Kleinbaum et al., 2014).

7.4.3 Interpretasi dan Validasi Model Regresi

Keberhasilan analisis regresi tidak hanya bergantung pada hasil perhitungan koefisien, tetapi juga pada kemampuan peneliti dalam menginterpretasikan makna hasil tersebut secara tepat. Koefisien regresi yang signifikan secara statistik menunjukkan adanya hubungan yang dapat dipercaya antara variabel bebas dan variabel terikat, namun hal ini tidak selalu berarti adanya hubungan sebab-akibat. Oleh karena itu, interpretasi hasil regresi harus mempertimbangkan desain penelitian, kualitas data, serta kemungkinan adanya variabel pengganggu.

Selain interpretasi, langkah penting lainnya adalah validasi model. Validasi dilakukan untuk memastikan bahwa model yang dibangun memiliki kemampuan prediktif yang baik dan tidak hanya sesuai untuk data sampel tertentu. Teknik validasi dapat meliputi analisis *residual*, pengujian asumsi klasik (seperti normalitas, homoskedastisitas, dan independensi), serta metode *cross-validation* untuk menilai generalisasi model pada data baru.

Dalam konteks kesehatan, model regresi yang tervalidasi dapat digunakan untuk membangun sistem prediksi risiko penyakit. Sebagai contoh, model Framingham Risk Score yang terkenal dibangun menggunakan regresi berganda untuk memprediksi risiko penyakit jantung koroner berdasarkan variabel usia, tekanan darah, kadar kolesterol, dan kebiasaan merokok (D'Agostino et al., 2008). Model semacam ini menjadi dasar penting bagi pengambilan keputusan klinis dan kebijakan kesehatan preventif.

7.4.4 Tantangan dan Pengembangan Model Regresi Modern

Meskipun analisis regresi merupakan alat statistik yang kuat, penerapannya memiliki keterbatasan, terutama dalam menghadapi data kompleks dengan hubungan non-linier atau interaksi tinggi antarvariabel. Untuk mengatasi hal ini, pendekatan regresi modern seperti *ridge regression*, *lasso regression*, dan *random forest regression* mulai banyak digunakan. Model-model ini dikembangkan untuk meningkatkan akurasi prediksi dan mengatasi masalah *multicollinearity* serta *overfitting*.

Selain itu, integrasi antara analisis regresi dan kecerdasan buatan (*artificial intelligence*) juga membuka peluang baru dalam epidemiologi dan kedokteran prediktif. Dengan bantuan algoritma *machine learning*, regresi kini dapat diaplikasikan pada *big data* kesehatan untuk mendeteksi pola kompleks yang sebelumnya sulit diidentifikasi dengan pendekatan konvensional.

Dengan demikian, analisis regresi bukan hanya alat analitik tradisional, tetapi juga fondasi penting bagi pengembangan model prediktif modern yang mendukung pengambilan keputusan berbasis data di bidang kesehatan.

7.5 Aplikasi Hubungan antar Variabel dalam Penelitian Kesehatan

7.5.1 Pentingnya Analisis Hubungan antar Variabel

Analisis hubungan antar variabel merupakan komponen fundamental dalam penelitian kesehatan, khususnya dalam bidang

epidemiologi. Analisis ini berfungsi untuk memahami keterkaitan antara faktor-faktor risiko dan kejadian penyakit, sehingga memungkinkan peneliti mengidentifikasi pola yang relevan untuk pengendalian dan pencegahan penyakit. Hubungan antar variabel tidak hanya menunjukkan adanya korelasi, tetapi juga dapat digunakan untuk menilai potensi kausalitas antara paparan (*exposure*) dan hasil kesehatan (*health outcome*) tertentu. Dalam konteks ini, pemilihan desain penelitian yang tepat menjadi kunci agar interpretasi hubungan antar variabel bersifat valid dan dapat digeneralisasi (Setia, 2016).

7.5.2 Jenis Desain Penelitian Epidemiologi

7.5.2.1 Studi Kohort

Studi kohort merupakan salah satu desain yang paling efektif dalam menganalisis hubungan sebab-akibat antara paparan dan kejadian penyakit. Dalam studi ini, sekelompok individu diikuti secara prospektif dari waktu ke waktu untuk mengamati munculnya penyakit atau kondisi kesehatan tertentu. Keunggulan studi kohort terletak pada kemampuannya untuk menghitung risiko relatif dan insidensi penyakit secara langsung. Namun, desain ini juga memiliki kelemahan, antara lain biaya tinggi dan durasi penelitian yang panjang. Meskipun demikian, dalam konteks kesehatan masyarakat, studi kohort tetap menjadi pilihan utama untuk memahami determinan penyakit kronis seperti diabetes melitus, penyakit jantung koroner, dan kanker (Wang & Attia, 2017).

7.5.2.2 Studi Kasus-Kontrol

Berbeda dengan studi kohort, studi kasus-kontrol bersifat retrospektif dan digunakan untuk menilai hubungan antara paparan dan kejadian penyakit dengan lebih efisien, terutama ketika penyakit yang diteliti jarang terjadi. Studi ini dimulai dengan mengidentifikasi individu yang telah mengalami penyakit (kasus) dan membandingkannya dengan individu yang tidak mengalami penyakit (kontrol), kemudian menelusuri riwayat paparannya. Studi kasus-kontrol dapat memberikan *odds ratio* sebagai ukuran asosiasi, yang meskipun tidak sekuat *relative risk*, tetap memberikan informasi berharga untuk memperkirakan hubungan kausal dalam populasi. Fleksibilitas dan efisiensi waktu menjadi keunggulan desain ini, meskipun risiko *recall bias* dan *selection bias* perlu diwaspadai dalam interpretasinya (Setia, 2016).

7.5.2.3 Studi Intervensi atau Eksperimental

Studi intervensi, termasuk *randomized controlled trial* (RCT), merupakan desain penelitian yang digunakan untuk mengevaluasi efektivitas suatu terapi atau program kesehatan. Dalam studi ini, peserta secara acak dibagi menjadi kelompok intervensi dan kelompok kontrol, sehingga variabel perancu dapat diminimalkan. Keunggulan utama studi intervensi adalah kemampuannya menghasilkan bukti yang paling kuat mengenai hubungan kausal antara variabel bebas dan variabel terikat. Namun, desain ini sering kali menghadapi kendala etika dan logistik, terutama ketika intervensi yang diteliti dapat menimbulkan risiko

bagi peserta. Oleh karena itu, penerapannya harus mengikuti prinsip-prinsip etika penelitian kesehatan yang ketat (Wang & Attia, 2017).

7.5.3 Implikasi Analisis Hubungan dalam Kebijakan Kesehatan

Analisis hubungan antar variabel dalam penelitian kesehatan memiliki dampak strategis terhadap penyusunan kebijakan berbasis bukti (*evidence-based policy*). Data yang diperoleh dari studi epidemiologi menjadi dasar bagi perancang kebijakan dalam menentukan prioritas intervensi kesehatan masyarakat. Misalnya, ketika studi menunjukkan hubungan yang signifikan antara konsumsi tembakau dan kejadian penyakit paru kronis, kebijakan pengendalian tembakau dapat diperkuat melalui peraturan dan kampanye publik. Selain itu, pemahaman terhadap hubungan antar variabel juga mendukung evaluasi program kesehatan, seperti menilai keberhasilan vaksinasi atau intervensi gizi di komunitas tertentu. Dengan demikian, hasil analisis hubungan variabel tidak hanya memperluas pengetahuan ilmiah, tetapi juga memperkuat kapasitas pengambilan keputusan yang berbasis data dan ilmiah (Setia, 2016).

7.6 Latihan Soal

1. Jelaskan perbedaan antara variabel bebas, terikat, dan pengganggu dalam penelitian kesehatan.
2. Apa arti nilai korelasi $r = -0,85$ pada hubungan antara aktivitas fisik dan kadar kolesterol?

3. Sebutkan perbedaan antara korelasi Pearson dan Spearman.
4. Bagaimana interpretasi hasil analisis regresi dengan koefisien $b = 0,75$ dalam penelitian tekanan darah?
5. Berikan contoh penerapan hubungan antar variabel dalam studi epidemiologi kohort.

Bab 8: Gizi, Lingkungan, dan Ketahanan Pangan: Tantangan dan Tren Pola Makan Modern

8.1 Gizi dan Perubahan Lingkungan Global

Perubahan lingkungan global, terutama akibat perubahan iklim, memiliki dampak yang signifikan terhadap ketersediaan, kualitas, dan keamanan pangan di seluruh dunia. Fenomena seperti peningkatan suhu bumi, ketidakstabilan cuaca, serta bencana alam berdampak langsung pada sistem produksi pangan dan pola konsumsi masyarakat. Dalam konteks ini, isu gizi tidak hanya berkaitan dengan pemenuhan kebutuhan nutrisi individu, tetapi juga dengan keberlanjutan ekosistem yang mendukung kehidupan manusia.

Perubahan iklim global menyebabkan gangguan pada ketahanan pangan melalui penurunan produktivitas pertanian, degradasi lahan, serta berkurangnya sumber air bersih. Akibatnya, banyak populasi, terutama di negara berkembang, menghadapi risiko kekurangan gizi (*malnutrition*) baik dalam bentuk kekurangan

maupun kelebihan gizi. Dalam situasi ini, penting untuk mengembangkan sistem pangan berkelanjutan yang mampu menjaga keseimbangan antara kebutuhan nutrisi manusia dan kelestarian lingkungan (FAO, 2019).

8.1.1 Dampak Perubahan Iklim terhadap Ketahanan dan Kualitas Pangan

Ketahanan pangan global ditentukan oleh empat pilar utama, yaitu ketersediaan, akses, pemanfaatan, dan stabilitas pangan. Keempat aspek ini kini menghadapi tekanan besar akibat perubahan iklim. Peningkatan suhu rata-rata global mengakibatkan berkurangnya hasil pertanian, terutama tanaman pokok seperti padi, gandum, dan jagung. Selain itu, pola hujan yang tidak menentu dan kekeringan berkepanjangan menurunkan kesuburan tanah dan mempercepat proses desertifikasi di berbagai wilayah tropis.

Kualitas pangan juga menurun karena peningkatan kadar karbon dioksida di atmosfer mempengaruhi kandungan nutrisi tanaman. Penelitian menunjukkan bahwa konsentrasi protein, zat besi, dan seng pada tanaman sereal dan legum cenderung menurun seiring meningkatnya kadar CO_2 di udara (Myers et al., 2017). Hal ini menimbulkan ancaman serius terhadap pemenuhan gizi mikro masyarakat yang bergantung pada sumber pangan nabati.

Selain aspek produksi, perubahan iklim juga meningkatkan risiko kontaminasi pangan akibat kenaikan suhu dan kelembaban yang mempercepat pertumbuhan mikroba patogen. Kondisi tersebut memperburuk masalah *food insecurity* dan meningkatkan beban kesehatan masyarakat, terutama di negara berpendapatan rendah.

8.1.2 Degradasi Lingkungan dan Pola Konsumsi Global

Degradasi lingkungan, seperti deforestasi, pencemaran air, dan penggunaan pestisida berlebihan, turut memperburuk ketidakseimbangan ekosistem pangan. Sektor pertanian modern yang bergantung pada input kimia dan energi fosil menyumbang emisi gas rumah kaca dalam jumlah besar. Ironisnya, aktivitas yang dimaksudkan untuk memenuhi kebutuhan pangan justru mempercepat perubahan iklim global.

Pola konsumsi masyarakat dunia juga berkontribusi terhadap krisis gizi dan lingkungan. Konsumsi berlebihan terhadap daging merah, makanan ultra-proses, serta produk dengan jejak karbon tinggi memperparah tekanan terhadap sistem pangan global. Di sisi lain, masyarakat di negara berkembang masih menghadapi tantangan kekurangan kalori dan protein. Ketimpangan ini menunjukkan bahwa permasalahan gizi tidak hanya bersifat biologis, tetapi juga sosial-ekologis.

Untuk mengatasi persoalan tersebut, diperlukan perubahan menuju pola makan rendah karbon (*low-carbon diet*), yaitu konsumsi makanan yang ramah lingkungan seperti sayuran, buah, biji-bijian, serta sumber protein nabati. Pola makan ini tidak hanya mendukung kesehatan individu, tetapi juga berkontribusi pada pengurangan emisi gas rumah kaca dan pelestarian sumber daya alam.

8.1.3 Strategi Pembangunan Gizi Berkelanjutan di Era Perubahan Global

Upaya meningkatkan ketahanan gizi dan pangan global harus mengedepankan pendekatan keberlanjutan (*sustainable nutrition*). Pendekatan ini mencakup perbaikan sistem produksi, distribusi, dan konsumsi pangan dengan mempertimbangkan efisiensi sumber daya serta dampak ekologisnya. Prinsip utama pembangunan gizi berkelanjutan meliputi efisiensi energi, pengelolaan limbah pangan (*food waste management*), serta diversifikasi sumber pangan lokal yang tahan terhadap perubahan iklim.

Selain itu, kebijakan publik perlu diarahkan untuk memperkuat ketahanan komunitas lokal melalui pendidikan gizi, teknologi pertanian ramah lingkungan, serta dukungan terhadap petani kecil. Integrasi ilmu gizi dan lingkungan menjadi kunci dalam merancang kebijakan yang menyeluruh, karena keduanya saling terkait dalam menciptakan masyarakat yang sehat sekaligus menjaga kelestarian planet.

Dalam jangka panjang, keberhasilan sistem gizi global tidak hanya diukur dari tingkat kecukupan kalori, tetapi juga dari keseimbangan antara kesehatan manusia dan bumi. Perubahan paradigma menuju konsumsi dan produksi pangan berkelanjutan merupakan langkah penting untuk menjawab tantangan gizi di tengah krisis iklim yang semakin nyata.

8.2 Pola Makan Berkelanjutan dan Planetary Health Diet

8.2.1 Konsep Pola Makan Berkelanjutan

Pola makan berkelanjutan adalah sistem konsumsi yang bertujuan menjaga keseimbangan antara kesehatan manusia dan keberlanjutan lingkungan. Pendekatan ini berupaya mengoptimalkan asupan gizi tanpa membahayakan sumber daya alam yang menopang kehidupan. Dalam praktiknya, pola makan berkelanjutan mendorong konsumsi pangan lokal, berbasis nabati, dan minim pengolahan untuk mengurangi jejak karbon dari produksi serta distribusi makanan.

Menurut penelitian terbaru oleh Willett et al. (2019), sistem pangan global saat ini menyumbang hampir 30% dari total emisi gas rumah kaca dunia. Jika tidak ada perubahan pola konsumsi, risiko degradasi lahan dan penurunan keanekaragaman hayati akan meningkat secara signifikan. Oleh karena itu, pola makan berkelanjutan menjadi salah satu solusi strategis untuk mencapai keseimbangan antara pemenuhan kebutuhan gizi dan konservasi ekosistem. Pendekatan ini menuntut perubahan paradigma konsumsi masyarakat dari orientasi kuantitas menuju kualitas dan keberlanjutan.

8.2.2 Prinsip-Prinsip *Planetary Health Diet*

Konsep *Planetary Health Diet* (PHD) diperkenalkan oleh *EAT-Lancet Commission* pada tahun 2019 sebagai model diet global yang mengintegrasikan kesehatan manusia dan kelestarian bumi.

Pola ini menekankan konsumsi pangan nabati dalam proporsi besar, sementara asupan daging merah, gula tambahan, dan makanan olahan dikurangi seminimal mungkin. Secara umum, PHD merekomendasikan sekitar 50% dari total asupan kalori harian berasal dari buah, sayuran, biji-bijian utuh, dan kacang-kacangan, sedangkan sisanya berasal dari sumber protein sehat seperti ikan, unggas, dan produk susu dalam jumlah moderat (Willett et al., 2019).

Model ini juga menyesuaikan pola konsumsi dengan daya dukung lingkungan dan kondisi sosial-ekonomi lokal. Misalnya, di wilayah dengan ketergantungan tinggi pada sumber protein hewani, adaptasi dilakukan melalui substitusi bertahap dengan sumber protein nabati seperti kedelai, lentil, dan kacang-kacangan. Hal ini bertujuan untuk mengurangi tekanan terhadap lahan dan air, sekaligus menjaga keberlanjutan rantai pangan global.

8.2.3 Manfaat Kesehatan dan Lingkungan

Penerapan pola makan berkelanjutan memberikan manfaat ganda. Dari segi kesehatan, peningkatan konsumsi buah, sayuran, dan biji-bijian utuh terbukti menurunkan risiko penyakit tidak menular seperti diabetes tipe 2, hipertensi, obesitas, dan penyakit jantung koroner (Springmann et al., 2020). Kandungan serat, vitamin, dan antioksidan dalam bahan nabati membantu menjaga sistem imun, menurunkan kadar kolesterol, dan memperlambat proses penuaan sel.

Sementara itu, dari perspektif lingkungan, pergeseran konsumsi dari daging merah ke pangan nabati mampu mengurangi emisi gas rumah kaca hingga 49% dan menurunkan penggunaan air

bersih secara global (Clark et al., 2022). Peternakan intensif diketahui menghasilkan metana dan nitrous oxide dalam jumlah besar, dua jenis gas yang berkontribusi signifikan terhadap perubahan iklim. Dengan demikian, penerapan *Planetary Health Diet* dapat menjadi salah satu strategi mitigasi iklim yang efektif melalui perubahan perilaku konsumsi individu dan kolektif.

8.2.4 Tantangan Implementasi di Berbagai Wilayah

Meskipun manfaatnya luas, implementasi pola makan berkelanjutan menghadapi sejumlah tantangan. Salah satunya adalah perbedaan budaya makan di berbagai negara, terutama di masyarakat yang menganggap konsumsi daging sebagai simbol status sosial atau kemakmuran. Di sisi lain, keterbatasan akses terhadap bahan pangan segar dan bergizi di beberapa wilayah membuat penerapan *Planetary Health Diet* tidak mudah dilakukan secara merata.

Penelitian oleh Fanzo et al. (2021) menekankan bahwa keberhasilan penerapan pola makan berkelanjutan bergantung pada sinergi antara kebijakan publik, edukasi masyarakat, dan sistem produksi pangan lokal. Pemerintah perlu menyediakan insentif bagi petani yang menerapkan praktik ramah lingkungan, memperluas akses pangan sehat melalui pasar lokal, serta mengintegrasikan pendidikan gizi berkelanjutan dalam kurikulum sekolah.

8.2.5 Peran Edukasi dan Kebijakan Pangan

Untuk memperkuat implementasi pola makan berkelanjutan, diperlukan pendekatan interdisipliner yang melibatkan sektor pendidikan, kesehatan, dan pertanian. Edukasi masyarakat tentang pentingnya memilih makanan yang sehat dan ramah lingkungan

harus diperluas melalui kampanye publik dan pelatihan gizi berbasis komunitas. Selain itu, kebijakan pemerintah yang mendorong reformasi sistem pangan—seperti pembatasan iklan makanan olahan dan pengenaan pajak karbon pada industri peternakan intensif—dapat mempercepat transisi menuju sistem pangan yang lebih hijau dan adil.

Dalam jangka panjang, perubahan pola makan menuju *Planetary Health Diet* dapat membantu menciptakan generasi yang lebih sehat, lingkungan yang lebih lestari, dan sistem pangan global yang tangguh menghadapi perubahan iklim. Dengan kesadaran kolektif dan dukungan kebijakan yang memadai, pola makan berkelanjutan bukan hanya pilihan gaya hidup, melainkan tanggung jawab moral terhadap kelestarian bumi dan kehidupan masa depan.

8.3 Ketahanan Pangan dan Perubahan Sosial Ekonomi

Ketahanan pangan tidak hanya bergantung pada ketersediaan komoditas, tetapi juga pada dimensi akses ekonomi, distribusi yang efisien, serta kondisi sosial yang membentuk preferensi konsumsi dan kemampuan rumah tangga memenuhi kebutuhan gizi. Dalam konteks urbanisasi dan perubahan struktur kerja, sistem pangan menghadapi tekanan baru: volatilitas harga, kerentanan mata pencaharian, dan ketimpangan akses yang memperlebar *food insecurity*. Di sisi lain, inovasi teknologi pangan dan praktik *urban farming* membuka peluang kemandirian dan diversifikasi sumber

pangan, khususnya di wilayah perkotaan. Bagian ini menguraikan keterkaitan faktor ekonomi-sosial dengan distribusi, lalu memaparkan peran teknologi dan *urban farming* sebagai penguat ketahanan pangan kontemporer (FAO, 2023).

8.3.1 Akses Ekonomi, Harga, dan Modal Sosial

Akses ekonomi menentukan kemampuan rumah tangga membeli pangan bergizi. Pendapatan yang fluktuatif—akibat *informal employment*, guncangan makroekonomi, atau bencana iklim, mengurangi daya beli dan mendorong konsumsi pangan murah namun rendah gizi. Instrumen perlindungan sosial seperti bantuan tunai, *e-voucher*, dan subsidi terarah membantu menstabilkan akses saat harga bergejolak, tetapi efektivitasnya bergantung pada akurasi penargetan dan interoperabilitas data kependudukan. Modal sosial juga berperan: jejaring komunitas, koperasi, dan *community pantry* memperkecil biaya transaksi dan menjadi bantalan saat pendapatan menurun.

Di tingkat mikro, literasi finansial dan akses *fintech* pertanian, misalnya *pay-as-you-harvest* atau asuransi indeks iklim meningkatkan daya tahan produsen kecil. Di hilir, perubahan preferensi ke pangan siap saji mendorong pergeseran pola pengeluaran; intervensi gizi perlu menyasar edukasi konsumsi sehat agar peningkatan daya beli benar-benar diikuti peningkatan kualitas gizi (FAO, 2023).

8.3.2 Distribusi, Infrastruktur, dan Dinamika Sosial

Distribusi yang adil menuntut *value chain* yang tangguh. Kualitas jalan, *first-mile logistics*, dan kapasitas *cold chain*

menentukan kehilangan pascapanen, harga di konsumen, serta stabilitas pasokan. Investasi pada gudang pendingin komunal, *milk run* logistik, dan digitalisasi *supply chain* (pelacakan asal, *demand forecasting*) menekan *food loss* dan memperbaiki margin produsen.

Dimensi sosial seperti kesenjangan gender dalam akses lahan dan input mempengaruhi produktivitas dan keputusan distribusi di tingkat rumah tangga. Kebijakan yang mengakui peran perempuan petani, memperkuat hak atas lahan, serta menyediakan layanan penyuluhan inklusif mempercepat perbaikan produktivitas dan ketahanan pangan. Di perkotaan, *last-mile delivery* melalui *platform e-commerce* pangan memperluas akses, tetapi berisiko meminggirkan pedagang kecil jika tidak ada regulasi persaingan dan skema kemitraan yang adil.

8.3.3 Teknologi Pangan, *Urban Farming*, dan Kemandirian

Kemajuan teknologi mempercepat transformasi sistem pangan dari hulu ke hilir. Di produksi, *precision agriculture* memanfaatkan sensor, *Internet of Things (IoT)*, dan citra satelit untuk optimasi input, penghematan air, serta pengelolaan hama berbasis data. Pendekatan *climate-smart agriculture* mengombinasikan varietas toleran cekaman, praktik konservasi tanah, dan manajemen risiko iklim. Pada pengolahan, inovasi *minimal processing*, pengemasan cerdas, dan fortifikasi mikro-nutrien memperpanjang umur simpan sekaligus menjaga kualitas gizi.

Di wilayah padat penduduk, *urban farming*—meliputi hidroponik, akuaponik, dan vertikultur mengubah ruang sempit

menjadi sumber sayuran segar sepanjang tahun. Praktik ini mempersingkat rantai pasok, menekan jejak karbon transportasi, dan memperkuat *food literacy* warga. Bukti kajian mutakhir menunjukkan *urban farming* berkontribusi pada ketahanan pangan rumah tangga, penciptaan kerja hijau, dan kohesi komunitas, terutama bila didukung kebijakan tata ruang, insentif air-listrik, serta skema pasar lokal yang pasti (Kaur, Dhir, Talwar, & Ghuman, 2022).

Transformasi digital melengkapi upaya tersebut melalui *marketplace* hasil tani, *crowdfunding* pertanian, dan dasbor analitik untuk koperasi. Namun, adopsi teknologi harus memperhatikan literasi digital, keamanan data, dan kelayakan finansial agar tidak memperlebar kesenjangan. Sinergi kebijakan—inklusif, berbasis bukti, dan lintas sektor—menjamin teknologi dan *urban farming* bermuara pada kemandirian, gizi yang lebih baik, serta ketahanan pangan yang berkelanjutan.

8.4 Tren Pola Makan Masyarakat Modern

8.4.1 Perubahan Pola Konsumsi di Era Modern

Perubahan pola konsumsi masyarakat modern dalam satu dekade terakhir menunjukkan peningkatan signifikan pada konsumsi makanan cepat saji (*fast food*) dan makanan olahan tinggi (*ultra-processed foods* atau UPF). Perkembangan teknologi, urbanisasi, serta meningkatnya aktivitas kerja membuat masyarakat lebih memilih makanan yang praktis dan cepat saji. Fenomena ini juga

didukung oleh kemudahan layanan pesan antar makanan serta pemasaran agresif dari industri makanan global. Makanan cepat saji kini bukan sekadar kebutuhan, tetapi telah menjadi bagian dari gaya hidup masyarakat urban yang menekankan kecepatan dan efisiensi.

Tren tersebut membawa dampak pada pergeseran budaya makan, masyarakat beralih dari makanan tradisional yang berbasis bahan segar dan alami ke arah makanan kemasan, beku, dan siap konsumsi. Pergeseran ini berdampak pada menurunnya kualitas gizi karena UPF umumnya tinggi gula, lemak jenuh, dan garam, namun rendah serat serta mikronutrien penting. Kondisi ini meningkatkan risiko obesitas, penyakit jantung, diabetes tipe 2, serta berbagai gangguan metabolik lainnya (González & Martínez, 2023).

8.4.2 Faktor Sosial dan Ekonomi yang Mempengaruhi

Tingginya konsumsi makanan cepat saji dan UPF tidak terlepas dari pengaruh sosial dan ekonomi. Faktor ekonomi menjadi salah satu pendorong utama, di mana harga makanan cepat saji relatif terjangkau dibandingkan dengan makanan sehat berbahan segar. Selain itu, gaya hidup urban menyebabkan berkurangnya waktu memasak di rumah, khususnya bagi masyarakat pekerja yang padat aktivitas. Di sisi lain, perubahan struktur keluarga, seperti meningkatnya partisipasi perempuan dalam dunia kerja, turut memperkuat ketergantungan pada makanan instan atau siap saji.

Konsumsi makanan cepat saji telah melonjak secara global, menimbulkan kekhawatiran kritis bagi kesehatan masyarakat. Tinjauan sistematis ini menganalisis hubungan antara asupan makanan cepat saji dan dampak metabolik serta kardiovaskularnya,

terutama obesitas (Lestari, 2023)
https://journal.sinergi.or.id/index.php/jhlqr/article/view/542?utm_source=chatgpt.com

Selain itu, promosi masif di media digital dan media sosial membentuk persepsi bahwa makanan cepat saji identik dengan modernitas dan kepraktisan. Generasi muda yang tumbuh di era digital menjadi sasaran utama industri ini. Promosi yang menekankan kecepatan layanan dan rasa yang menarik memperkuat kebiasaan konsumsi tanpa mempertimbangkan dampak jangka panjang terhadap kesehatan. Akibatnya, makanan cepat saji menjadi bagian dari rutinitas, bukan sekadar pilihan sesekali. Paparan pemasaran makanan dan minuman tidak sehat merupakan faktor risiko yang diakui secara luas untuk perkembangan obesitas dan penyakit tidak menular pada generasi muda. Pemasaran makanan melibatkan penggunaan berbagai teknik persuasif untuk memengaruhi sikap, preferensi, dan konsumsi makanan generasi muda (Smith, et al, 2019) https://www.mdpi.com/2072-6643/11/4/875?utm_source=chatgpt.com

8.4.3 Dampak Kesehatan dari Konsumsi *Ultra-Processed Foods*

Konsumsi berlebihan terhadap UPF memiliki implikasi kesehatan yang serius. Produk-produk ini mengandung bahan tambahan seperti pengawet, pewarna, pemanis buatan, dan penstabil yang dalam jangka panjang dapat memengaruhi fungsi metabolisme tubuh. Sejumlah penelitian mutakhir menunjukkan bahwa peningkatan konsumsi UPF hingga 10% dari total asupan energi harian dapat meningkatkan risiko kematian akibat semua penyebab

sebesar 10–15%. Dampak jangka panjangnya juga mencakup gangguan kesehatan mental, seperti kecemasan dan depresi, akibat ketidakseimbangan gizi dan ketergantungan terhadap makanan tinggi gula (Frontini et al., 2024).

Kandungan kalori tinggi dalam UPF seringkali tidak diimbangi dengan kandungan nutrisi esensial. Konsumsi berlebihan dapat menyebabkan kelebihan energi, obesitas, dan resistensi insulin. Sementara itu, kandungan aditif tertentu, seperti natrium tinggi, turut meningkatkan risiko hipertensi dan gangguan kardiovaskular. Dengan demikian, pergeseran pola makan masyarakat modern menuntut perhatian serius karena berpotensi menurunkan kualitas kesehatan populasi secara umum. Peningkatan konsumsi makanan cepat saji dikaitkan dengan asupan makanan yang buruk dan beberapa faktor risiko PJK pada orang dewasa Iran (Wang, 2020) https://www.mdpi.com/2072-6643/12/10/2934?utm_source=chatgpt.com

8.4.4 Strategi Edukasi dan Promosi Pola Makan Sehat

Untuk menghadapi meningkatnya konsumsi makanan cepat saji, strategi edukasi dan promosi pola makan sehat menjadi kebutuhan mendesak. Edukasi gizi perlu ditanamkan sejak dini melalui sistem pendidikan, dengan menekankan pentingnya konsumsi makanan alami seperti buah, sayuran, biji-bijian utuh, serta protein rendah lemak. Pemerintah juga perlu memperkuat kebijakan pelabelan gizi yang mudah dipahami agar masyarakat dapat membuat keputusan konsumsi yang lebih sadar dan sehat.

Edukasi Pola Makan Sehat Bagi generasi muda dengan “Mengurangi Konsumsi Makanan Cepat Saji Untuk Hidup Lebih Sehat” bertujuan untuk meningkatkan kesadaran generasi muda akan pentingnya pola makan sehat. Inisiatif ini berfokus pada pengurangan konsumsi makanan tinggi natrium, pedas, dan manis, seperti ayam geprek dan es teh, yang umum dikonsumsi karena praktis dan terjangkau. (Efendy dkk, 2024)
https://pakisjournal.com/index.php/mestaka/article/view/516?utm_source=chatgpt.com

Kampanye publik melalui media sosial dapat menjadi sarana efektif untuk menjangkau generasi muda, menggunakan pendekatan yang menarik seperti *challenge* memasak sehat atau kompetisi menu tradisional bergizi. Selain itu, kolaborasi lintas sektor antara pemerintah, industri pangan, dan masyarakat sipil dapat memperkuat upaya menciptakan lingkungan yang mendukung gaya hidup sehat. Penyediaan pilihan makanan sehat di sekolah, kantor, dan ruang publik merupakan langkah konkret yang dapat mengubah kebiasaan konsumsi secara bertahap.

8.4.5 Arah Kebijakan dan Inovasi Masa Depan

Kebijakan masa depan perlu berfokus pada pengurangan konsumsi UPF melalui pendekatan berbasis bukti ilmiah dan teknologi pangan berkelanjutan. Inovasi dalam pengembangan produk pangan sehat, praktis, dan terjangkau dapat menjadi solusi untuk menggantikan ketergantungan masyarakat terhadap produk olahan tinggi. Pemerintah perlu memperkuat pengawasan terhadap iklan dan distribusi makanan cepat saji, serta mendorong riset dan

pengembangan produk pangan lokal yang bergizi. Instruksi Presiden Nomor 1 Tahun 2017 tentang Gerakan Masyarakat Hidup Sehat dan Peraturan Menteri Kesehatan Nomor 39 Tahun 2016 tentang Pedoman Penyelenggaraan Program Indonesia Sehat dengan Pendekatan Keluarga , sekaligus menjadi dasar operasional terkait aspek-pelaksanaan hidup sehat (https://extranet.who.int/ncdccs/Data/IDN_B27_s21_Presidential%20Instruction%20No.%201%20Year%202017%20on%20GERMA S.pdf?utm_source=chatgpt.com)

Di sisi lain, perubahan perilaku konsumsi juga harus didukung oleh peningkatan kesadaran individu terhadap pentingnya *mindful eating*, yaitu kebiasaan makan dengan memperhatikan kebutuhan tubuh dan kualitas makanan yang dikonsumsi. Dengan demikian, strategi edukasi, kebijakan publik, dan inovasi pangan yang sinergis dapat mengarahkan masyarakat menuju pola makan yang lebih sehat, berkelanjutan, dan seimbang antara aspek praktis dan gizi (González & Martínez, 2023; Frontini et al., 2024).

8.5 Latihan Soal

1. Jelaskan pengaruh perubahan iklim terhadap ketahanan pangan!
2. Apa itu pola makan Planetary Health Diet?
3. Sebutkan tantangan pola makan masyarakat modern!
4. Bagaimana urban farming membantu ketahanan pangan?
5. Mengapa penting mengurangi konsumsi ultra-processed food?

Bab 9: Pengolahan Data Kesehatan Menggunakan R dan SPSS

9.1 Menenal Analisis Data Kesehatan

Analisis data kesehatan merupakan fondasi penting dalam pengambilan keputusan berbasis bukti (*evidence-based practice*) di bidang keperawatan, kedokteran, dan kesehatan masyarakat. Melalui analisis data yang sistematis dan akurat, tenaga kesehatan dapat menilai efektivitas intervensi klinis, mengidentifikasi faktor risiko penyakit, serta mengembangkan kebijakan pelayanan yang lebih efisien. Di era digital, meningkatnya volume data kesehatan dari rekam medis elektronik, surveilans penyakit, dan riset klinis menuntut kemampuan analisis data yang kuat agar informasi yang terkumpul dapat diubah menjadi pengetahuan yang bermakna.

9.1.1 Peran Analisis Data dalam Praktik Kesehatan

Analisis data berfungsi sebagai jembatan antara observasi klinis dan kebijakan kesehatan. Dengan melakukan pengolahan data yang benar, tenaga kesehatan dapat memahami pola kejadian penyakit, mengevaluasi hasil terapi, serta mengidentifikasi kelompok populasi berisiko tinggi. Sebagai contoh, analisis data epidemiologis membantu dalam menentukan prevalensi penyakit

kronis, sementara analisis statistik inferensial digunakan untuk menguji efektivitas intervensi kesehatan tertentu.

Dalam keperawatan, analisis data digunakan untuk memantau mutu pelayanan, mengukur kepuasan pasien, serta menilai efektivitas program asuhan keperawatan. Menurut Jensen et al. (2020), penerapan analisis data dalam praktik klinis meningkatkan akurasi diagnosis keperawatan dan membantu perawat merancang intervensi yang lebih terarah. Dengan demikian, kemampuan menganalisis data bukan sekadar keterampilan teknis, melainkan bagian integral dari proses ilmiah dalam peningkatan mutu layanan kesehatan.

9.1.2 Perangkat Lunak dalam Analisis Data Kesehatan

Dalam praktik analisis data kesehatan, dua perangkat lunak yang paling umum digunakan adalah **SPSS (Statistical Package for the Social Sciences)** dan **R**. SPSS dikenal karena antarmukanya yang intuitif dan kemampuannya untuk melakukan analisis statistik deskriptif dan inferensial dengan mudah. Perangkat ini banyak digunakan di institusi pendidikan dan rumah sakit karena cocok untuk pengguna dengan latar belakang non-statistik. SPSS mendukung berbagai uji statistik seperti *t-test*, *ANOVA*, *regresi linier*, serta analisis korelasi yang membantu tenaga kesehatan memahami hubungan antarvariabel.

Sementara itu, perangkat lunak **R** menawarkan fleksibilitas lebih tinggi karena berbasis *open-source* dan mendukung analisis statistik kompleks serta visualisasi data yang interaktif. R juga banyak digunakan dalam riset biomedis dan epidemiologi karena

kemampuannya mengolah data besar (*big data*) serta mendukung *machine learning* untuk prediksi pola penyakit. Menurut Wang et al. (2021), penggunaan R dalam penelitian kesehatan mempercepat proses analisis data longitudinal dan meningkatkan akurasi pemodelan risiko klinis.

Perawat dan tenaga kesehatan dianjurkan untuk memahami dasar penggunaan kedua perangkat ini agar dapat memilih alat yang sesuai dengan kebutuhan dan sumber daya yang tersedia. Selain itu, kolaborasi dengan analis data atau ahli biostatistik sering kali diperlukan untuk memastikan validitas hasil dan interpretasi yang tepat.

9.1.3 Tantangan dan Arah Masa Depan Analisis Data Kesehatan

Meskipun teknologi analisis data semakin berkembang, tantangan utama tetap terletak pada kualitas data, keamanan informasi, dan keterampilan pengguna. Data yang tidak lengkap, bias pencatatan, atau tidak terstandar dapat mengganggu validitas hasil analisis. Oleh karena itu, integrasi sistem rekam medis elektronik yang akurat dan pelatihan tenaga kesehatan dalam pengelolaan data menjadi hal krusial.

Selain itu, aspek etika juga harus diperhatikan dalam setiap tahap analisis. Privasi pasien harus dijaga sesuai dengan regulasi seperti *Health Insurance Portability and Accountability Act* (HIPAA) dan standar etika profesi. Penerapan prinsip transparansi dan akuntabilitas dalam analisis data menjadi keharusan agar hasilnya dapat dipercaya dan dimanfaatkan untuk kepentingan publik.

Ke depan, integrasi *artificial intelligence (AI)* dan *machine learning* diharapkan semakin memperkuat kemampuan analisis data kesehatan, memungkinkan deteksi dini penyakit, prediksi tren epidemi, dan personalisasi terapi secara lebih efisien. Oleh karena itu, literasi data dan kemampuan berpikir analitis menjadi kompetensi wajib bagi tenaga kesehatan di era transformasi digital ini.

9.2 Langkah Dasar Pengolahan Data Menggunakan SPSS

9.2.1 Pengenalan dan Fungsi SPSS

Statistical Package for the Social Sciences (SPSS) merupakan perangkat lunak statistik yang banyak digunakan dalam bidang kesehatan, psikologi, pendidikan, dan ilmu sosial lainnya untuk menganalisis data secara efisien. Program ini dirancang untuk mempermudah pengguna yang tidak memiliki latar belakang pemrograman statistik karena antarmuka (*interface*)-nya bersifat intuitif dan berbasis menu.

Dalam konteks penelitian kesehatan, SPSS sering digunakan untuk mengelola data hasil survei, catatan medis, uji klinis, serta penelitian epidemiologi. Kelebihan utamanya adalah kemampuannya menangani data dalam jumlah besar, menjalankan berbagai uji statistik dengan cepat, dan menghasilkan output yang mudah diinterpretasikan. Selain itu, SPSS menyediakan fasilitas visualisasi data berupa grafik, tabel, dan diagram yang membantu

peneliti memahami pola dan hubungan antarvariabel sebelum melakukan analisis lanjutan (Pallant, 2020).

9.2.2 Langkah-Langkah Dasar Pengolahan Data

Proses analisis data di SPSS umumnya diawali dengan **pemasukan data ke dalam lembar kerja (data view)**. Setiap baris mewakili satu kasus atau responden, sedangkan setiap kolom merepresentasikan variabel. Sebelum memasukkan data, pengguna perlu mendefinisikan variabel terlebih dahulu melalui **variable view**, dengan menentukan nama variabel, tipe data (numerik atau string), label, nilai kategori (*value labels*), serta skala pengukuran (nominal, ordinal, interval, atau rasio).

Langkah berikutnya adalah melakukan **analisis deskriptif** untuk memperoleh gambaran umum data, seperti nilai rata-rata, median, simpangan baku, serta distribusi frekuensi. Analisis ini membantu memastikan data tidak memiliki kesalahan input dan telah memenuhi asumsi dasar sebelum dilakukan uji statistik inferensial.

Setelah itu, dilakukan **uji hipotesis** yang bertujuan menguji hubungan, perbedaan, atau pengaruh antarvariabel. Uji yang sering digunakan meliputi *t-test*, *chi-square test*, *ANOVA*, dan *regresi linear*. Pemilihan jenis uji tergantung pada bentuk data dan tujuan analisis. Misalnya, *t-test* digunakan untuk membandingkan dua kelompok, sementara *regresi* digunakan untuk menilai pengaruh variabel bebas terhadap variabel terikat.

SPSS juga menyediakan fitur **visualisasi data**, seperti histogram, diagram batang, *pie chart*, dan *scatter plot*. Visualisasi

ini berfungsi untuk memperjelas hasil analisis serta membantu peneliti mengomunikasikan temuan secara efektif. Hasil uji statistik dapat langsung ditampilkan dalam bentuk tabel output dengan nilai signifikansi (*p-value*), koefisien korelasi, atau persamaan regresi yang siap digunakan dalam laporan penelitian.

9.2.3 Interpretasi dan Validasi Hasil Analisis

Tahapan penting setelah pengolahan data adalah **interpretasi hasil analisis**. Peneliti harus mampu membaca output SPSS dengan cermat, memahami arti nilai signifikansi, serta menafsirkan implikasi hasil terhadap hipotesis yang diajukan. Misalnya, jika nilai $p < 0,05$, maka hasil dianggap signifikan secara statistik, artinya ada hubungan atau perbedaan yang tidak terjadi secara kebetulan.

Selain itu, validitas hasil analisis perlu diperiksa melalui uji asumsi statistik seperti normalitas, linearitas, dan homoskedastisitas. Uji-uji tersebut memastikan bahwa model analisis memenuhi syarat untuk interpretasi yang sahih. SPSS menyediakan menu khusus untuk menguji asumsi ini secara otomatis, sehingga pengguna dapat mengidentifikasi data yang menyimpang dan melakukan transformasi bila diperlukan.

Keunggulan lain SPSS adalah kemampuan ekspor hasil ke format lain, seperti Excel, Word, atau PDF, yang memudahkan integrasi dengan laporan akhir. Namun, meskipun SPSS sangat membantu dalam proses analisis, pemahaman konseptual terhadap teori statistik tetap diperlukan agar hasil tidak disalahartikan. Oleh karena itu, penggunaan SPSS sebaiknya disertai dengan landasan

metodologis yang kuat agar hasil penelitian memiliki keabsahan ilmiah dan dapat dipertanggungjawabkan (Field, 2018).

9.3 Langkah Dasar Pengolahan Data Menggunakan R

R merupakan bahasa pemrograman dan lingkungan komputasi statistik yang dikembangkan untuk keperluan analisis data, eksplorasi statistik, dan visualisasi. Sebagai perangkat lunak *open-source*, *R* banyak digunakan oleh akademisi, analis data, dan profesional di berbagai bidang karena kemampuannya yang fleksibel serta dukungan komunitas yang luas. Dalam konteks analisis data kesehatan, sosial, dan ekonomi, *R* memungkinkan pengguna untuk mengolah data dalam jumlah besar dengan efisien, mengembangkan model statistik, dan menghasilkan grafik visual yang informatif.

9.3.1 Impor dan Eksplorasi Data

Langkah pertama dalam pengolahan data menggunakan *R* adalah mengimpor data dari berbagai sumber, seperti file *CSV*, *Excel*, *SQL*, atau sumber daring. Fungsi dasar yang sering digunakan adalah `read.csv()` untuk membaca data dalam format *comma-separated values*. Setelah data berhasil diimpor, pengguna dapat melakukan eksplorasi awal untuk memahami struktur dan karakteristik dataset.

Fungsi `summary()` digunakan untuk menampilkan statistik deskriptif seperti nilai minimum, maksimum, rata-rata, dan median dari setiap variabel, sedangkan `str()` memberikan informasi tentang

struktur data, termasuk tipe variabel dan jumlah observasi. Langkah ini penting untuk mendeteksi kesalahan input, data kosong (*missing values*), atau ketidaksesuaian tipe data sebelum analisis lebih lanjut dilakukan (Kabacoff, 2022).

Selain itu, fungsi seperti `head()` dan `tail()` sering digunakan untuk meninjau beberapa baris pertama atau terakhir dari dataset. Tahapan eksplorasi ini membantu peneliti memastikan bahwa data yang diolah sudah lengkap dan siap untuk tahap berikutnya, yaitu pembersihan dan transformasi data.

9.3.2 Pembersihan dan Transformasi Data

Pembersihan data (*data cleaning*) merupakan tahapan krusial dalam analisis, karena kualitas hasil sangat bergantung pada keakuratan data yang digunakan. Dalam R, paket populer seperti *dplyr* dan *tidyr* digunakan untuk mempermudah manipulasi data secara efisien dan intuitif.

Fungsi `filter()`, `mutate()`, `select()`, dan `arrange()` dari paket *dplyr* membantu pengguna menyaring, menambah variabel baru, memilih kolom tertentu, dan mengurutkan data dengan cepat. Sementara itu, *tidyr* digunakan untuk merapikan data agar sesuai dengan prinsip *tidy data*, yaitu setiap kolom mewakili variabel dan setiap baris mewakili observasi.

Misalnya, fungsi `pivot_longer()` dan `pivot_wider()` memungkinkan pengguna mengubah format tabel agar lebih mudah dianalisis. Proses pembersihan juga mencakup pengisian nilai yang hilang, penghapusan duplikasi, serta standarisasi nama variabel. Dengan sintaks yang sederhana dan berbasis *pipe operator* (`%>%`),

R memungkinkan transformasi data dilakukan secara berurutan dan mudah dibaca (Wickham et al., 2019).

Selain pembersihan, *R* juga mendukung integrasi dengan basis data besar melalui paket seperti *DBI* dan *RSQLite*, memungkinkan pengguna melakukan analisis langsung pada data dalam jumlah besar tanpa harus memuat seluruh dataset ke dalam memori.

9.3.3 Analisis dan Visualisasi Data

Tahap analisis dan visualisasi data merupakan inti dari proses pengolahan data menggunakan *R*. Setelah proses impor, eksplorasi, dan pembersihan dilakukan, langkah berikutnya adalah menafsirkan makna dari data melalui analisis statistik serta menyajikannya dalam bentuk visual yang mudah dipahami. Pada tahap ini, *R* menunjukkan keunggulannya sebagai lingkungan analitik yang komprehensif karena mampu memadukan kekuatan statistik dan visualisasi dalam satu platform terpadu.

Analisis data dalam *R* dapat dilakukan mulai dari tingkat dasar hingga kompleks, bergantung pada tujuan penelitian dan jenis data yang digunakan. Analisis deskriptif merupakan langkah awal yang umum dilakukan untuk memahami karakteristik data. Fungsi-fungsi seperti `summary()`, `table()`, dan `aggregate()` digunakan untuk menghitung ukuran pemusatan dan penyebaran data, seperti rata-rata, median, modus, dan simpangan baku. Pada tingkat lanjutan, *R* menyediakan berbagai paket seperti *MASS*, *car*, dan *caret* untuk melakukan uji hipotesis, analisis regresi, serta pemodelan prediktif secara efisien (Kabacoff, 2022).

Keunggulan lain dari *R* adalah fleksibilitasnya dalam berbagai konteks analisis. Dalam bidang kesehatan masyarakat, misalnya, *R* digunakan untuk menganalisis hubungan antara faktor risiko dan kejadian penyakit dengan model regresi logistik. Di bidang ekonomi, *R* mampu mengolah data *time series* untuk memprediksi tren jangka panjang seperti inflasi atau pertumbuhan produk domestik bruto. Setiap analisis yang dilakukan melalui *R* dapat direplikasi dengan mudah karena seluruh langkah tercatat dalam bentuk skrip, menjadikannya transparan dan dapat diaudit secara ilmiah.

Hasil analisis kemudian disajikan melalui visualisasi data yang berfungsi untuk memperjelas hubungan antarvariabel serta memudahkan pemahaman terhadap pola yang muncul. Paket *ggplot2* merupakan salah satu alat visualisasi paling populer di *R*. Prinsip *Grammar of Graphics* yang digunakan dalam *ggplot2* memungkinkan pengguna membangun grafik berdasarkan komponen visual yang terstruktur, seperti sumbu, warna, bentuk titik, serta label. Dengan pendekatan ini, visualisasi tidak hanya menjadi alat bantu komunikasi, tetapi juga sarana eksplorasi yang membantu menemukan pola tersembunyi di dalam data (Wickham et al., 2019).

Visualisasi yang dihasilkan *ggplot2* sangat fleksibel. Grafik batang dapat digunakan untuk membandingkan kelompok data, *scatter plot* untuk melihat korelasi antarvariabel, dan *histogram* untuk menggambarkan distribusi data. Pengguna juga dapat menyesuaikan elemen estetika seperti warna, skala, dan legenda

untuk meningkatkan kejelasan interpretasi. Dengan demikian, *R* tidak hanya mampu menghasilkan grafik informatif, tetapi juga dapat memenuhi standar estetika visual dalam publikasi ilmiah.

Lebih lanjut, *R* juga mendukung pembuatan grafik interaktif melalui paket seperti *plotly*, *shiny*, dan *highcharter*. Grafik interaktif ini memungkinkan pengguna berinteraksi langsung dengan data, misalnya dengan menyorot nilai tertentu, memperbesar area grafik, atau menampilkan informasi tambahan saat kursor diarahkan ke titik data. Fitur semacam ini sangat berguna dalam penyusunan laporan dinamis dan *dashboard* analitik yang digunakan dalam presentasi atau pengambilan keputusan cepat.

Visualisasi data yang baik bukan hanya memperindah tampilan hasil analisis, melainkan juga memperkuat kredibilitas temuan ilmiah. Karena seluruh grafik dalam *R* dapat direproduksi melalui kode yang terdokumentasi, setiap visualisasi yang dihasilkan memiliki transparansi penuh dan bebas dari manipulasi subjektif. Hal ini sejalan dengan semangat *open science*, yaitu keterbukaan dan replikabilitas dalam penelitian ilmiah.

Secara keseluruhan, analisis dan visualisasi data menggunakan *R* merupakan perpaduan antara kekuatan statistik dan kejelasan komunikasi visual. Dengan penguasaan keterampilan ini, pengguna dapat menghasilkan interpretasi data yang akurat, komunikatif, dan berdaya guna dalam mendukung pengambilan keputusan berbasis bukti di berbagai bidang, mulai dari kesehatan, pendidikan, ekonomi, hingga kebijakan publik.

9.4 Interpretasi dan Visualisasi Hasil Analisis

Interpretasi dan visualisasi hasil analisis merupakan tahap penting dalam proses analisis data kesehatan. Kedua tahap ini berperan dalam menjembatani antara temuan statistik dan pemahaman klinis yang aplikatif. Hasil analisis yang tidak diinterpretasikan dengan tepat berisiko menimbulkan kesimpulan keliru, sementara penyajian data yang tidak komunikatif dapat menghambat pemanfaatannya dalam pengambilan keputusan kesehatan. Oleh karena itu, kemampuan untuk menginterpretasikan dan memvisualisasikan data secara efektif menjadi kompetensi utama bagi tenaga kesehatan dan analis data.

9.4.1 Interpretasi Hasil Analisis

Interpretasi data kesehatan harus berorientasi pada makna klinis, bukan sekadar angka atau hasil uji statistik. Nilai p , *confidence interval*, dan ukuran efek (*effect size*) perlu dikaji dalam konteks relevansinya terhadap kondisi pasien, kebijakan kesehatan, dan sistem pelayanan. Sebagai contoh, hasil uji *t-test* yang menunjukkan perbedaan signifikan antara dua kelompok harus disertai penjelasan mengenai signifikansi klinisnya—apakah perbedaan tersebut bermakna secara praktis atau hanya secara statistik.

Selain makna klinis, interpretasi juga harus memperhatikan keterbatasan data, seperti ukuran sampel kecil, bias seleksi, atau variabel pengganggu yang tidak terukur. Penggunaan analisis multivariat membantu mengontrol faktor-faktor tersebut, tetapi

interpretasi tetap harus disertai penilaian kritis terhadap validitas internal dan eksternal. Di bidang kesehatan masyarakat, interpretasi hasil analisis dapat menjadi dasar untuk menyusun strategi intervensi, menetapkan prioritas program, atau mengevaluasi efektivitas kebijakan yang telah diterapkan (Bland & Altman, 2015).

Dalam praktiknya, perangkat lunak seperti *R* dan *SPSS* menyediakan berbagai fitur yang memudahkan pengguna dalam melakukan analisis inferensial maupun deskriptif. Namun, hasil keluaran perangkat tersebut tidak serta-merta dapat disimpulkan tanpa pemahaman terhadap konteks data. Oleh karena itu, penting bagi pengguna untuk menggabungkan pemahaman statistik dengan wawasan epidemiologis agar kesimpulan yang diambil benar-benar mendukung pengambilan keputusan berbasis bukti (*evidence-based decision making*).

9.4.2 Visualisasi Data Kesehatan

Visualisasi data berfungsi untuk menyederhanakan informasi kompleks agar mudah dipahami oleh berbagai pihak, mulai dari tenaga medis hingga pembuat kebijakan. Dalam bidang kesehatan, visualisasi yang baik dapat mempercepat deteksi tren penyakit, mengidentifikasi kelompok berisiko, dan mengevaluasi dampak program intervensi. Bentuk visualisasi umum mencakup grafik batang, diagram garis, *heat map*, dan peta distribusi geografis (*choropleth map*).

Penyajian data harus mengikuti prinsip kejelasan dan keakuratan. Misalnya, grafik tren penyakit kronis harus menampilkan sumbu waktu yang konsisten dan skala numerik yang

proporsional agar tidak menimbulkan persepsi salah. Penggunaan warna juga berperan penting dalam memperjelas pesan visual; misalnya, warna merah untuk risiko tinggi dan hijau untuk kondisi aman. Namun, pemilihan warna harus memperhatikan aksesibilitas bagi individu dengan gangguan penglihatan warna.

Perangkat lunak *R* menawarkan pustaka seperti *ggplot2*, *plotly*, dan *shiny* yang memungkinkan pengguna membuat visualisasi interaktif dan dinamis. Sementara itu, *SPSS* menyediakan *Chart Builder* yang intuitif untuk pengguna nonprogrammer. Hasil visualisasi dari kedua perangkat dapat diekspor ke berbagai format seperti *Word*, *Excel*, atau *PDF* guna mempermudah pelaporan dan publikasi hasil analisis (Kassambara, 2018).

Visualisasi juga memainkan peran strategis dalam komunikasi publik kesehatan. Misalnya, grafik sebaran kasus COVID-19 berdasarkan wilayah memberikan gambaran cepat kepada masyarakat mengenai tingkat risiko daerahnya. Dengan demikian, visualisasi bukan hanya alat bantu teknis, tetapi juga media edukatif yang efektif untuk meningkatkan kesadaran dan partisipasi masyarakat dalam upaya pencegahan penyakit.

9.4.3 Integrasi Interpretasi dan Visualisasi

Kekuatan utama analisis data kesehatan terletak pada integrasi antara interpretasi dan visualisasi. Data yang diinterpretasikan tanpa visualisasi cenderung sulit dipahami, sedangkan visualisasi tanpa interpretasi dapat menyesatkan. Oleh karena itu, kedua elemen ini harus saling melengkapi untuk menghasilkan laporan yang informatif dan akurat.

Laporan hasil analisis idealnya memuat narasi interpretatif yang menjelaskan makna klinis setiap grafik atau tabel. Misalnya, tren peningkatan angka kejadian hipertensi perlu disertai analisis mengenai faktor gaya hidup, usia, dan kebiasaan diet yang mungkin berkontribusi. Dengan demikian, visualisasi berfungsi sebagai pendukung argumentasi ilmiah, bukan sekadar hiasan grafis.

Di era digital, integrasi ini semakin diperkuat oleh munculnya *dashboard analytics* berbasis *web* yang memungkinkan pemangku kepentingan mengakses data real-time. Melalui *dashboard*, pengguna dapat memantau indikator kesehatan secara langsung dan mengambil keputusan cepat berdasarkan informasi visual yang jelas.

9.5 Latihan Soal

1. Apa perbedaan utama antara R dan SPSS dalam pengolahan data kesehatan?
2. Sebutkan langkah-langkah dasar analisis data menggunakan SPSS!
3. Mengapa visualisasi data penting dalam laporan hasil penelitian kesehatan?
4. Jelaskan manfaat paket *ggplot2* dalam analisis data menggunakan R.
5. Bagaimana hasil analisis dari R dapat digunakan untuk mendukung keputusan klinis?

Profil Penulis



Rini Ernawati, S.Pd., M.Kes., lahir di Bukittinggi pada 2 September 1969 dan saat ini berdomisili di Samarinda, Kalimantan Timur. Sejak kecil, ia telah menumbuhkan minat yang besar terhadap dunia pendidikan dan kesehatan, yang membawanya menekuni karier di bidang tersebut. Selain profesinya, Rini memiliki hobi membaca, menulis, dan memasak. Kegemarannya membaca dan menulis turut membentuk cara pandangnya yang kritis dan reflektif, sementara kegiatan memasak menjadi sarana untuk menyalurkan kreativitas dan ketekunan dalam keseharian. Dalam pesannya kepada pembaca, Rini menekankan pentingnya membaca sebagai kunci untuk memperluas pengetahuan dan wawasan. Ia percaya bahwa melalui membaca, seseorang dapat terus belajar, mengembangkan diri, dan memahami dunia di sekitarnya dengan lebih baik.



Joko Susanto, S.Kep., Ns., M.Kes. lahir di Bojonegoro pada 26 Juni 1981 dan saat ini berdomisili di Lamongan, Jawa Timur. Ia memiliki minat dalam sepak bola. Dalam pesannya kepada pembaca, Joko menekankan pentingnya perpaduan antara seni dan sains dalam pendidikan: “*Dua*

pendidikan yang memengaruhi pendidikan manusia adalah seni dan sains, dan keduanya bertemu dalam buku.” Joko Susanto memulai pendidikan keperawatannya di SPK Pemkab Bojonegoro dan lulus pada tahun 2000. Ia melanjutkan studi D3 Keperawatan di Akademi Keperawatan Pemkab Lamongan (lulus 2003), memperoleh gelar Sarjana dan Profesi Ners di PSIK FK Universitas Gadjah Mada Yogyakarta (lulus 2011), serta menyelesaikan Magister Kedokteran Keluarga di Universitas Sebelas Maret Surakarta (lulus 2014). Saat ini, ia sedang menempuh studi doktoral keperawatan di Fakultas Keperawatan Universitas Airlangga Surabaya. Dalam kariernya, Joko Susanto pernah menjadi dosen tetap di Akademi Keperawatan Pemerintah Kabupaten Lamongan (2004–2017) dan sejak 2018 menjadi dosen tetap di Program D3 Keperawatan Fakultas Vokasi Universitas Airlangga Surabaya, mengampu mata kuliah Keperawatan Komunitas, Keluarga, dan Gerontik. Ia juga aktif sebagai Sekretaris DPD PPNI Kab. Lamongan, pengurus IPEGRI Provinsi Jawa Timur, dan pengurus serta survei akreditasi fasilitas kesehatan primer di LAFKESFRI. Beberapa karya tulisnya antara lain *Buku Ajar Ilmu Keperawatan Dasar (Buku 1 dan 2)*, *Standar*

Asuhan Keperawatan dan Prosedur Tetap dalam Praktik Keperawatan, serta Buku Saku Pelibatan Keluarga Dalam Diet Hipertensi: Kenali dan Obati Hipertensi.



apt. Retno Fitriati, S.Si., M.Kes. lahir di Semarang pada 4 Februari 1973 dan saat ini berdomisili di Jakarta Timur. Sejak muda, ia telah menumbuhkan ketertarikan yang kuat pada dunia ilmu pengetahuan dan kesehatan, yang membawanya menekuni profesi di bidang farmasi. Selain sebagai profesional, Retno juga memiliki minat yang luas dalam membaca dan kuliner. Kegiatan membaca membantunya terus memperluas wawasan dan mengikuti perkembangan ilmu terbaru, sementara hobi kuliner menjadi cara untuk menyalurkan kreativitas dan menikmati sisi lain kehidupan sehari-hari. Dalam pesannya kepada pembaca, Retno menekankan nilai penting dari pembelajaran sepanjang hayat. Dengan kata-kata sederhana, “Belajar sepanjang hayat,” ia mengajak setiap orang untuk terus menuntut ilmu, mengembangkan diri, dan memanfaatkan kesempatan belajar tanpa terbatas oleh usia maupun latar belakang.



Ns. Rahayu Tri Nuritasari, M.Kep. lahir di Palembang pada 9 Maret 1995 dan saat ini berdomisili di kota kelahirannya. Sejak muda, ia telah menumbuhkan minat yang kuat dalam bidang kesehatan, khususnya keperawatan, yang membawanya menekuni pendidikan dan profesi di bidang tersebut. Selain profesionalisme di bidang keperawatan, Rahayu juga memiliki hobi memasak dan travelling. Kegiatan ini membantunya menyeimbangkan kehidupan profesional dan personal, sambil terus menumbuhkan kreativitas serta rasa ingin tahu tentang dunia sekitar. Dalam pesannya kepada pembaca, Rahayu menekankan pentingnya menikmati proses belajar. Dengan kata-kata, “Setiap halaman yang kamu baca adalah langkah menuju pemahaman baru. Nikmati setiap perjalanannya,” ia mengajak pembaca untuk menghargai setiap kesempatan membaca sebagai sarana memperluas wawasan dan pengalaman hidup.



Ns. Septi Viantri Kurdaningsih, M.Kep. lahir di Palembang pada 23 September 1987 dan saat ini berdomisili di Palembang, Sumatera Selatan. Ia memiliki minat dalam *travelling* dan aktif di bidang keperawatan. Dalam pesannya kepada pembaca, Septi Viantri menyampaikan bahwa buku ini disusun sebagai panduan bagi mahasiswa dan praktisi keperawatan untuk memahami konsep, prinsip, dan praktik keperawatan yang berfokus pada pelayanan di lingkungan rumah. Ia berharap buku ini dapat memperluas wawasan serta meningkatkan keterampilan dalam memberikan asuhan yang humanis, berkelanjutan, dan berpusat pada pasien beserta keluarganya di komunitas.



Bangkit Ary Pratama, SKM., M.Kes., lahir di Sukoharjo pada 20 Mei 1989 dan saat ini berdomisili di Sukoharjo, Jawa Tengah. Sejak muda, ia telah menumbuhkan ketertarikan pada kegiatan travelling, yang membantunya memperluas wawasan dan memahami berbagai pengalaman kehidupan secara langsung. Selain gemar menjelajah, Bangkit juga aktif mengembangkan diri melalui ilmu pengetahuan dan pengalaman praktis yang dimilikinya. Minatnya pada pengembangan diri mendorongnya untuk terus belajar dan berbagi pengetahuan dengan lingkungan sekitarnya. Dalam pesannya kepada pembaca, Bangkit menekankan pentingnya ilmu sebagai kunci pengambilan keputusan dan kualitas hidup. Dengan kata-kata, “Pengetahuan yang tepat membuka jalan bagi kemajuan. Wawasan luas menuntun kita pada keputusan bijak dan kualitas hidup yang lebih baik,” ia mengajak pembaca untuk menghargai setiap kesempatan belajar sebagai sarana untuk membuat keputusan yang bijak dan meningkatkan kualitas hidup secara keseluruhan.



Onetusfisi Putra, M.K.M., lahir di Pelangai Kecil pada 2 November 1994 dan saat ini berdomisili di Aceh Barat, Aceh. Sejak muda, ia menumbuhkan ketertarikan yang kuat

terhadap olahraga, khususnya badminton, yang membantunya mengembangkan disiplin, fokus, dan semangat kompetitif. Selain aktif dalam olahraga, Onetusfisi juga menekuni bidang akademik dengan penuh dedikasi, membekali diri dengan pengetahuan dan kemampuan untuk berkontribusi melalui karya ilmiah. Minatnya pada penelitian dan pengembangan ilmu mendorongnya terus mencari peluang untuk memperluas wawasan dan berbagi pengetahuan. Dalam pesannya kepada pembaca, Onetusfisi menekankan pentingnya pengetahuan sebagai alat untuk perubahan. Dengan kata-kata, “Knowledge is a tool for change,” ia berharap tulisannya dapat menjadi referensi ilmiah yang bermanfaat, memperluas wawasan, serta mendorong munculnya diskusi dan penelitian lebih lanjut.



Wahyu Dwi Lesmono, S.Si., M.Si., lahir di Bogor pada 16 Oktober 1994 dan saat ini berdomisili di Jalan Raya Bogor, Kampung Sawah, RT 01/RW 11 No. 32, Kelurahan Cibinong, Kecamatan Cibinong, Kabupaten Bogor. Ia menempuh pendidikan S1 Matematika di Universitas Pakuan dan S2 Matematika di Universitas Indonesia, serta mengabdikan diri sebagai dosen. Dalam pesannya kepada pembaca, Wahyu menyampaikan refleksi tentang keterkaitan antara disiplin ilmu: *“Farmasi dan matematika mungkin terlihat berbeda, tetapi keduanya berbicara dalam bahasa yang sama yaitu tentang akurasi, presisi, dan solusi bagi kehidupan.”* Pesan ini mengajak pembaca untuk melihat kesamaan esensial dalam berbagai bidang ilmu dan pentingnya ketelitian serta pemahaman mendalam dalam praktik profesional.



Rinco Siregar, S.Kep., Ns., M.N.S., Ph.D. lahir di Napasingkam pada 9 September 1976 dan berdomisili di Jln. Nusa Indah 2, No. 119 LK IX Blok II, Kelurahan Helvetia Tengah, Kecamatan Helvetia, Medan, Sumatera Utara. Ia menempuh pendidikan Doctor of Philosophy in Nursing di Prince of Songkla University (2020–2024) dengan fokus pada Keperawatan Gerontik, Keluarga, dan Komunitas, Master of Nursing Science di Kasertsart University (2011–2013) dengan fokus pada Keperawatan Keluarga dan Komunitas, Pendidikan Profesi Ners di STIKes Mutiara Indonesia (2001–2004), serta pendidikan awal di Akademi Perawat Sari Mutiara (1996–1999). Rinco memiliki pengalaman sebagai perawat pelaksana di RSU Sari Mutiara Lubuk Pakam pada tahun 1999–2001, dan sejak 2004 mengabdikan sebagai dosen keperawatan di Universitas Sari Mutiara Indonesia. Ia mengajar berbagai mata kuliah, termasuk Keperawatan Keluarga, Keperawatan Komunitas, Keperawatan Gerontik, Psikososial Budaya dalam Keperawatan, Metodologi Penelitian, dan Bahasa Inggris Keperawatan. Selain itu, Rinco menjabat sebagai Preceptor Akademik dan Koordinator pada Stase Keperawatan Keluarga dan Komunitas serta Keperawatan Gerontik Pendidikan Profesi Ners sejak 2010. Ia pernah menjabat Ketua Program Studi Keperawatan dan Profesi Ners (2016–2020) dan saat ini menjabat Kepala Kantor Urusan Internasional Universitas Sari Mutiara Indonesia (2024–sekarang).



Nurnianingsih A. Yasin, S.Kep., Ns., M.Kep., lahir di Telaga pada 15 Januari 1996 dan saat ini berdomisili di Gorontalo. Sejak muda, ia telah menaruh minat yang besar pada penelitian keperawatan medikal bedah, serta aktif dalam kegiatan travelling dan aktivitas sosial di masyarakat, yang memperluas wawasan serta pengalaman praktisnya. Selain mengembangkan keahlian profesional, Nurnianingsih juga berkomitmen untuk berbagi ilmu dan pengalaman melalui tulisan. Ia percaya bahwa pengetahuan yang dibagikan akan memberikan manfaat lebih luas bagi pendidikan dan praktik keperawatan. Dalam pesannya kepada pembaca, Nurnianingsih berharap tulisannya dapat menjadi referensi yang bermanfaat dalam pembelajaran dan penelitian, sekaligus menjadi kontribusi kecil namun berarti dalam upaya meningkatkan kualitas pendidikan serta pelayanan kesehatan di Indonesia.

Daftar Pustaka

- Agresti, A. (2019). *Statistical Methods for the Social Sciences* (5th ed.). Pearson.
- Bland, J. M., & Altman, D. G. (2015). Statistics notes: Interpreting data. *BMJ*, 351, h4831.
- Bluman, A. G. (2018). *Elementary statistics: A step by step approach*. McGraw-Hill Education.
- Clark, M. A., Springmann, M., Hill, J., & Tilman, D. (2022). Multiple health and environmental impacts of foods. *Proceedings of the National Academy of Sciences*, 119(3), e2117728119.
- Creswell, J. W., & Creswell, J. D. (2018). *Research design: Qualitative, quantitative, and mixed methods approaches*. SAGE Publications.
- D'Agostino, R. B., Vasan, R. S., Pencina, M. J., Wolf, P. A., Cobain, M., Massaro, J. M., & Kannel, W. B. (2008). *General cardiovascular risk profile for use in primary care: The Framingham Heart Study*. *Circulation*, 117(6), 743–753.
- Dancey, C. P., & Reidy, J. (2017). *Statistics without maths for psychology* (7th ed.). Pearson Education.
- Daniel, W. W., & Cross, C. L. (2018). *Biostatistics: A foundation for analysis in the health sciences* (11th ed.). Hoboken, NJ: Wiley.

- Dawson, B., & Trapp, R. G. (2020). *Basic and clinical biostatistics* (5th ed.). McGraw-Hill Education.
- Etikan, I., & Bala, K. (2017). Sampling and sampling methods. *Biometrics & Biostatistics International Journal*, 5(6), 215–217.
- Fanzo, J., Haddad, L., Schneider, K. R., Béné, C., Covic, N., Guarin, A., & Herrero, M. (2021). *Viewpoint: The importance of food systems for food security and nutrition*. *Food Security*, 13(3), 627–648.
- FAO. (2023). *The state of food security and nutrition in the world 2023: Urbanization, agrifood systems transformation and healthy diets across the rural–urban continuum*. Rome: Food and Agriculture Organization.
- Field, A. (2018). *Discovering statistics using IBM SPSS Statistics* (5th ed.). SAGE Publications.
- Pallant, J. (2020). *SPSS survival manual: A step-by-step guide to data analysis using IBM SPSS* (7th ed.). Routledge.
- Frontini, R., Smith, L., & et al. (2024). Ultra-processed food consumption and health-related outcomes. *Frontiers in Nutrition*.
- 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.
- González, K., & Martínez, A. (2023). Ultra-processed foods: How bad are they for your health? *BMJ*, 383, 2023–075294.

- Gravetter, F. J., & Wallnau, L. B. (2017). *Statistics for the behavioral sciences* (10th ed.). Cengage Learning.
- Greenland, S., Pearl, J., & Robins, J. M. (2016). *Causal diagrams for epidemiologic research*. *Epidemiology*, 10(1), 37–48.
- Steyerberg, E. W. (2019). *Clinical prediction models: A practical approach to development, validation, and updating* (2nd ed.). New York: Springer.
- Hair, J. F., Hult, G. T. M., Ringle, C. M., & Sarstedt, M. (2020). *A primer on partial least squares structural equation modeling (PLS-SEM)*. Sage Publications.
- Huang, Y., Cai, X., Mai, W., Li, M., & Hu, Y. (2020). Association between prediabetes and risk of cardiovascular disease and all-cause mortality: systematic review and meta-analysis. *BMJ*, 370, m2297.
- Jensen, C. S., Vestergaard, M., & Møller, S. (2020). *Data-driven nursing practice: The role of analytics in clinical decision-making*. *Journal of Nursing Management*, 28(4), 823–832.
- Kabacoff, R. I. (2022). *R in action: Data analysis and graphics with R* (3rd ed.). Manning Publications.
- Kassambara, A. (2018). *Practical guide to principal component methods in R: PCA, M (CA), FAMD, MFA, HCPC, factoextra*. STHDA.
- Kaur, J., Dhir, A., Talwar, S., & Ghuman, K. (2022). Urban farming and sustainable food systems: A systematic review and future research agenda. *Sustainable Cities and Society*, 78, 103610.

- Kim, T. K. (2017). Understanding one-way *ANOVA* using conceptual figures. *Korean Journal of Anesthesiology*, 70(1), 22–26.
- Kirkwood, B. R., & Sterne, J. A. C. (2003). *Essential Medical Statistics* (2nd ed.). Blackwell Science.
- Szklo, M., & Nieto, F. J. (2019). *Epidemiology: Beyond the Basics* (4th ed.). Jones & Bartlett Learning.
- Kleinbaum, D. G., Kupper, L. L., Nizam, A., & Rosenberg, E. S. (2014). *Applied regression analysis and other multivariable methods* (5th ed.). Cengage Learning.
- Kurniawan, A., & Yuliani, N. (2021). Penerapan Analisis Biostatistik dalam Kebijakan dan Evaluasi Program Kesehatan. *Jurnal Kebijakan dan Administrasi Kesehatan Indonesia*, 9(2), 157–174.
- McHugh, M. L. (2013). The chi-square test of independence. *Biochemia Medica*, 23(2), 143–149.
- Omelchenko, D. M., Zhuravleva, E. N., & Karaseva, N. V. (2020). Gender differences in arterial blood pressure among adults: A population-based study. *Journal of Hypertension*, 38(5), 889–896.
- Pagano, M., & Gauvreau, K. (2018). *Principles of biostatistics* (2nd ed.). CRC Press.
- Rosner, B. (2015). *Fundamentals of biostatistics* (8th ed.). Cengage Learning.
- Palinkas, L. A., Horwitz, S. M., Green, C. A., Wisdom, J. P., Duan, N., & Hoagwood, K. (2015). Purposeful sampling for qualitative data collection and analysis in mixed method

- implementation research. *Administration and Policy in Mental Health and Mental Health Services Research*, 42(5), 533–544.
- Peat, J. (2021). *Health science research: A handbook of quantitative methods*. McGraw-Hill Education.
- Pepe, M. S. (2003). *The statistical evaluation of medical tests for classification and prediction*. Oxford University Press.
- Polit, D. F., & Beck, C. T. (2017). *Nursing research: Generating and assessing evidence for nursing practice* (10th ed.). Philadelphia: Wolters Kluwer.
- Pratama, R., & Lestari, D. (2020). Big Data dan Analisis Biostatistik dalam Transformasi Sistem Kesehatan Indonesia. *Indonesian Journal of Public Health Informatics*, 4(1), 67–83.
- Rosner, B. (2016). *Fundamentals of biostatistics* (8th ed.). Boston, MA: Cengage Learning.
- Rossi, R. J. (2022). *Applied Biostatistics for the Health Sciences*. Cambridge University Press.
- Sekaran, U., & Bougie, R. (2016). *Research methods for business: A skill-building approach*. Wiley.
- Setia, M. S. (2016). Methodology series module 3: Cross-sectional studies. *Indian Journal of Dermatology*, 61(3), 261–264.
- Springmann, M., Clark, M., Rayner, M., Scarborough, P., & Webb, P. (2020). The global and regional costs of healthy and sustainable dietary patterns: A modelling study. *The Lancet Planetary Health*, 4(1), e84–e91.

- Sugiyono. (2019). *Metode penelitian kuantitatif, kualitatif, dan R&D*. Alfabeta.
- Sullivan, L. M. (2023). *Essentials of Biostatistics in Public Health*. Jones & Bartlett Learning.
- Szklo, M., & Nieto, F. J. (2019). *Epidemiology: Beyond the basics* (4th ed.). Jones & Bartlett Learning.
- Taherdoost, H. (2016). Sampling methods in research methodology; how to choose a sampling technique for research. *International Journal of Academic Research in Management*, 5(2), 18–27.
- Triola, M. F. (2018). *Elementary statistics*. Pearson.
- Wackerly, D. D., Mendenhall, W., & Scheaffer, R. L. (2014). *Mathematical statistics with applications* (7th ed.). Boston: Cengage Learning.
- Walpole, R. E., Myers, R. H., Myers, S. L., & Ye, K. (2017). *Probability and statistics for engineers and scientists*. Pearson Education.
- Wang, L., Zhao, Y., & Chen, J. (2021). *Applications of R software in health data analytics and predictive modeling*. *BMC Medical Research Methodology*, 21(1), 135.
- Wickham, H., François, R., Henry, L., & Müller, K. (2019). *dplyr: A grammar of data manipulation*. R Foundation for Statistical Computing.

Willett, W., Rockström, J., Loken, B., Springmann, M., Lang, T., Vermeulen, S., & Murray, C. J. L. (2019). Food in the Anthropocene: The EAT–Lancet Commission on healthy diets from sustainable food systems. *The Lancet*, 393(10170), 447–492

Buku ajar yang berjudul **Biostatistik Modern** disusun untuk masyarakat umum agar lebih memahami bagaimana data kesehatan dan ilmiah dianalisis untuk mendukung pengambilan keputusan yang tepat. Dengan bahasa yang sederhana dan mudah dipahami, buku ini menjelaskan konsep dasar biostatistik, teknik analisis data, serta penerapannya dalam berbagai bidang, termasuk kesehatan, penelitian, dan kehidupan sehari-hari.

Buku ini ditujukan untuk disebarluaskan kepada masyarakat luas, agar masyarakat dapat memahami pentingnya data dan analisis statistik, serta mampu menggunakan informasi tersebut untuk keputusan yang lebih cerdas dan bermanfaat dalam kehidupan sehari-hari. Pembaca akan diajak mengenal cara membaca, menafsirkan, dan memanfaatkan data secara efektif untuk menilai kondisi, membuat keputusan, atau mengevaluasi hasil kegiatan secara lebih akurat.

