

BIOSTATISTIK KESEHATAN



PENULIS :

**Arif Rachman
Iha Nursolihah
Nur Al-Faida
Wardiah Hamzah
Nasruddin Syam
Islamiyati
Gede Wirabuana Putra
M. Habibullah Arief**

BIOSTATISTIK KESEHATAN

**Arif Rachman
Iha Nursolihah
Nur Al-Faida
Wardiah Hamzah
Nasruddin Syam
Islamiyati
Gede Wirabuana Putra
M. Habibullah Arief**



GET PRESS INDONESIA

BIOSTATISTIK KESEHATAN

Penulis :

Arif Rachman
Iha Nursolihah
Nur Al-Faida
Wardiah Hamzah
Nasruddin Syam
Islamiyati
Gede Wirabuana Putra
M. Habibullah Arief

ISBN : 978-634-255-030-4

Editor : Dr. Oktavianis, M.Biomed.

Penyunting : Rantika Maida Sahara., S.Tr. Kes

Desain Sampul dan Tata Letak : Atyka Trianisa, S.Pd

Penerbit : GET PRESS INDONESIA

Anggota IKAPI No. 033/SBA/2022

Redaksi :

Jln. Palarik Air Pacah No 26 Kel. Air Pacah
Kec. Koto Tengah Kota Padang Sumatera Barat
Website : www.getpress.co.id
Email : adm.getpress@gmail.com

Cetakan pertama, September 2025

Hak cipta dilindungi undang-undang
Dilarang memperbanyak karya tulis ini dalam bentuk dan
dengan cara apapun tanpa izin tertulis dari penerbit.

KATA PENGANTAR

Segala Puji dan syukur atas kehadiran Allah SWT dalam segala kesempatan. Sholawat beriring salam dan doa kita sampaikan kepada Nabi Muhammad SAW. Alhamdulillah atas Rahmat dan Karunia-Nya penulis telah menyelesaikan Buku Biostatistik Kesehatan ini.

Buku Ini Membahas : Pengantar Biostatistik, Pengumpulan Data dalam Penelitian Kesehatan, Distribusi dan Deskripsi Data, Probabilitas dalam Biostatistik, Inferensi Statistik: Estimasi dan Pengujian Hipotesis, Analisis Regresi dalam Kesehatan, Metode Meta-analisis dalam Kesehatan, Analisis Data Spasial dalam Kesehatan Masyarakat.

Proses penulisan buku ini berhasil diselesaikan atas kerjasama tim penulis. Demi kualitas yang lebih baik dan kepuasan para pembaca, saran dan masukan yang membangun dari pembaca sangat kami harapkan.

Penulis ucapkan terima kasih kepada semua pihak yang telah mendukung dalam penyelesaian buku ini. Terutama pihak yang telah membantu terbitnya buku ini dan telah mempercayakan mendorong, dan menginisiasi terbitnya buku ini. Semoga buku ini dapat bermanfaat bagi masyarakat Indonesia.

Padang, Agustus 2025

Penulis

DAFTAR ISI

KATA PENGANTAR	i
DAFTAR ISI	ii
DAFTAR TABEL	v
DAFTAR GAMBAR	vi
BAB 1 PENGANTAR BIOSTATISTIK.....	1
1.1 Pendahuluan	1
1.2 Konsep Dasar Statistik	5
1.2.1 Variabel dan Data.....	6
1.2.2 Skala Pengukuran.....	7
1.2.3 Statistik deskriptif	9
1.3 Probabilitas dalam Biostatistik.....	10
1.4 Inferensi Statistik	11
1.5 Analisis Regresi dan Korelasi	13
1.6 Desain Penelitian dalam Biostatistik.....	14
1.7 Analisis Data dalam Biostatistik.....	15
DAFTAR PUSTAKA.....	18
BAB 2 PENGUMPULAN DATA DALAM	
PENELITIAN KESEHATAN	19
2.1 Pendahuluan	19
2.2 Angket/Kuesioner	20
2.2.1 Jenis-Jenis Kuesioner	21
2.2.2 Langkah Penyusunan Kuesioner	22
2.3 Wawancara	23
2.3.1 Jenis Wawancara.....	24
2.3.2 Panduan Wawancara	24
2.3.3 Teknik Wawancara.....	25
2.4 Observasi.....	26
2.4.1 Jenis Observasi	27
2.4.2 Instrumen Observasi	28
2.5 Pengukuran Klinis	29
2.6 Tes Laboratorium	29
DAFTAR PUSTAKA.....	31
BAB 3 DISTRIBUSI DAN DESKRIPSI DATA	33
3.1 Pendahuluan	33
3.2 Ditribusi Data.....	34
3.2.1 Jenis-jenis Distribusi	34
3.2.2 Karakteristik Distribusi Normal	36

3.3 Deskriptif Data	43
DAFTAR PUSTAKA	46
BAB 4 PROBABILITAS DALAM BIOSTATISTIK	47
4.1 Pendahuluan.....	47
4.2 Konsep Dasar Probabilitas dalam Biostatistik.....	48
4.2.1 Defenisi Probabilitas	49
4.2.2 Jenis-Jenis Probabilitas	49
4.2.3 Aturan Dasar Probabilitas.....	51
4.3 Distribusi Probabilitas dalam Biostatistik.....	52
4.3.1 Distribusi Probabilitas Diskrit	53
4.3.2 Distribusi Probabilitas Kontinu	54
4.4 Penerapan Probabilitas dalam Biostatistik.....	56
4.4.1 Analisis Risiko Penyakit	56
4.4.2 Uji Klinis dan Efektivitas Pengobatan	58
4.4.3 Probabilitas dalam Pengujian Diagnostik.....	59
4.4.3 Prediksi Kejadian Epidemiologi	60
4.5 Konsep Studi Kasus: Probabilitas Biostatistik.....	60
4.5.1 Prediksi Risiko Penyakit Jantung.....	60
4.5.2 Efektivitas Vaksin	61
4.5.3 Pengujian Diagnostik dan Nilai Prediktif	63
4.6 Kesimpulan dan Saran.....	64
4.6.1 Kesimpulan	64
4.6.2 Saran	64
DAFTAR PUSTAKA	66
BAB 5 INFERENSI STATISTIK ESTIMASI DAN	
PENGUJIAN HIPOTESIS.....	69
5.1 Pendahuluan.....	69
5.2 Konsep Dasar Inferensi Biostatistik Kesehatan	70
5.2.1 Definisi dan Ruang Lingkup.....	71
5.2.2 Jenis Data dalam Biostatistik.....	72
5.3 Estimasi dalam Inferensi Biostatistik.....	74
5.3.1 Pengertian dan Jenis-Jenis Estimasi.....	74
5.3.1 Estimasi Parameter dalam Biostatistik.....	76
5.4 Uji Hipotesis dalam Biostatistik.....	77
5.4.1 Konsep Dasar Uji Hipotesis	77
5.4.2 Jenis Uji Hipotesis dalam Biostatistik.....	79
5.4.3 Kesalahan Uji Hipotesis dan Implikasinya	80
5.5 Tantangan dan Keterbatasan Inferensi Biostatistik.....	81
5.5.1 Tantangan dalam Inferensi Biostatistik.....	82

5.5.2 Keterbatasan dalam Inferensi Biostatistik	83
5.6 Kesimpulan dan Rekomendasi	84
DAFTAR PUSTAKA.....	87
BAB 6 ANALISIS REGRESI DALAM KESEHATAN	89
6.1 Pendahuluan	89
6.2 Manfaat Analisis Regresi dalam Kesehatan	89
6.2.1 Mengidentifikasi Faktor Risiko	89
6.2.2 Mengevaluasi Intervensi	89
6.2.3 Memprediksi Hasil Kesehatan	90
6.2.4 Memahami Hubungan Antar Variabel	90
6.3 Jenis-jenis Analisis Regresi	90
6.3.1 Regresi Linear	90
6.3.2 Regresi Logistik.....	94
DAFTAR PUSTAKA.....	100
BAB 7 METODE META ANALISIS DALAM KESEHATAN.....	101
7.1 Konsep Dasar dan Prinsip Meta Analisis dalam Penelitian Kesehatan	101
7.2 Metodologi Pencarian dan Seleksi Studi untuk Meta Analisis.....	105
7.3 Analisis Statistik dan Interpretasi Hasil Meta Analisis.....	109
7.4 Aplikasi Meta Analisis dalam <i>Evidence-Based Medicine</i> dan Kebijakan Kesehatan	113
DAFTAR PUSTAKA.....	120
BAB 8 ANALISIS DATA SPASIAL DALAM KESEHATAN MASYARAKAT	123
8.1 Pendahuluan	123
8.2 Representasi Data Spasial.....	124
8.2.1 Representasi Data Diskrit.....	127
8.2.2 Representasi Data Kontinu	130
8.3 Implementasi Analisis Data Spasial dalam Bidang Kesehatan Masyarakat	132
DAFTAR PUSTAKA.....	137
BIODATA PENULIS	

DAFTAR TABEL

Tabel 2.1. Kuesioner Tertutup dalam Bentuk <i>Checklist</i>	21
Tabel 2.2. Menyusun Indikator Pengukuran Variabel	23
Tabel 7.1. Perbandingan Metode Sintesis Bukti Ilmiah .	102
Tabel 7.2. Karakteristik Basis Data Literatur Utama untuk Meta Analisis.....	106
Tabel 7.3. Instrumen Penilaian Kualitas Studi Berdasarkan Desain Penelitian	108
Tabel 7.4. Perbandingan Komprehensif Model Efek Tetap vs Efek Acak.....	110
Tabel 7.5. Aplikasi Meta Analisis dalam Berbagai Domain Kesehatan.....	114

DAFTAR GAMBAR

Gambar 3.1. Kurva Distribusi Normal	35
Gambar 3.2. Variabel Random Diskrit	38
Gambar 3.3. Distribusi Poisson.....	42
Gambar 6.1. Contoh Kurva estimasi Regresi Linier	91
Gambar 6.2. Kurva Persamaan Regresi Logistik.....	95
Gambar 7.1. Grafik Evolusi Jumlah Meta Analisis dalam Bidang Kesehatan (2000-2023)	104
Gambar 7.2. Grafik Distribusi Ukuran Efek dalam Meta Analisis Terapi Musik.....	112
Gambar 7.3. Grafik Proses Integrasi Meta Analisis dalam Kebijakan Kesehatan.....	117
Gambar 8.1. Beruang	129
Gambar 8.2. Danau di Helsinki, Finlandia.....	131
Gambar 8.3. Sebaran Kasus TBC di Kabupaten Jember 2020-2023	132
Gambar 8.4. Pemetaan Prediksi Sebaran Kasus TBC di Kabupaten Jember Tahun 2025	133
Gambar 8.5. Peta Keterjangkauan Puskesmas di Kab. Jember	134
Gambar 8.6. Peta Penyebaran Penyakit.....	134
Gambar 8.7. <i>Vaccine Landing Page of Jakarta COVID-19 official website</i>	136

BAB 1

PENGANTAR BIOSTATISTIK

Oleh Arif Rachman

1.1 Pendahuluan

Biostatistik adalah cabang ilmu statistik yang memfokuskan pada aplikasi metode statistik untuk menganalisis data yang berasal dari bidang-bidang ilmu kehidupan dan kesehatan, seperti kedokteran, biologi, epidemiologi, dan ilmu kesehatan masyarakat. Definisi ini menunjukkan betapa pentingnya penggunaan alat statistik untuk menginterpretasikan fenomena yang berkaitan dengan kehidupan dan kesehatan manusia dan makhluk hidup lainnya (Rosner, B., 2015).

Ruang Lingkup dan Tujuan Biostatistik adalah sebagai berikut (Daniel, W. W., & Cross, C. L., 2018) (Pagano, M., & Gauvreau, K., 2018) :

1. Ruang Lingkup:

- a. **Desain Penelitian:** Meliputi perencanaan dan pelaksanaan desain penelitian yang mencakup pemilihan sampel, pengumpulan data, dan pengaturan eksperimen.
- b. **Analisis Data:** Menggunakan berbagai teknik statistik untuk menganalisis data yang diperoleh dari penelitian, survei, atau percobaan.
- c. **Interpretasi Hasil:** Membantu dalam menginterpretasikan hasil analisis data untuk membuat kesimpulan yang informatif dan berdasarkan bukti.
- d. **Pengambilan Keputusan:** Memberikan dasar yang kuat bagi pengambilan keputusan dalam bidang ilmu kehidupan dan kesehatan berdasarkan bukti-bukti yang diperoleh dari analisis statistik.
- e. **Pengembangan Metode:** Mengembangkan dan memperbaiki metode statistik yang sesuai dengan

kebutuhan dalam bidang biologi, kedokteran, dan kesehatan.

2. Tujuan:

- a. **Meningkatkan Pengertian:** Memfasilitasi pemahaman yang lebih baik tentang fenomena biologis dan kesehatan melalui analisis data yang sistematis.
- b. **Mendukung Penemuan Ilmiah:** Memfasilitasi penemuan baru dan pemahaman yang lebih dalam tentang bagaimana variabel biologi dan variabel kesehatan berinteraksi satu sama lain.
- c. **Memastikan Validitas dan Keandalan:** Memastikan bahwa analisis statistik yang dilakukan mendukung validitas dan keandalan hasil penelitian.
- d. **Pengembangan Praktik Klinis dan Kebijakan Kesehatan:** Memberikan dukungan untuk pengembangan praktik klinis yang berbasis bukti dan kebijakan kesehatan yang efektif.

Peran Biostatistik dalam Ilmu Kesehatan sangat penting dalam berbagai aspek, termasuk penelitian, pengembangan kebijakan kesehatan, praktik klinis, dan pemantauan kesehatan masyarakat. Berikut adalah beberapa peran utama biostatistik dalam ilmu kesehatan (Kirkwood, B. R., & Sterne, J. A. C., 2019) (Altman, D. G. (Ed.), 2018):

1. **Desain Penelitian:** Biostatistik membantu dalam perencanaan dan pelaksanaan desain penelitian yang baik, termasuk pemilihan sampel yang representatif dan teknik pengumpulan data yang tepat.
2. **Analisis Data:** Biostatistik menganalisis data klinis, survei kesehatan masyarakat, dan penelitian kesehatan untuk menemukan pola, tren, dan korelasi. Ini dilakukan dengan menggunakan berbagai teknik statistik.
3. **Interpretasi Hasil:** Biostatistik membantu dalam menginterpretasikan hasil analisis data secara tepat dan membuat kesimpulan yang relevan, memberikan dasar untuk pengambilan keputusan yang berbasis bukti dalam bidang kesehatan.

4. **Pengembangan Model dan Prediksi:** Dalam epidemiologi dan penelitian kesehatan masyarakat, biostatistik digunakan untuk mengembangkan model matematika dan prediksi yang membantu memahami penyebaran penyakit, faktor risiko, dan dampak intervensi kesehatan.
5. **Evaluasi Efektivitas Intervensi:** Biostatistik digunakan untuk mengevaluasi efektivitas intervensi kesehatan, seperti vaksinasi, program pencegahan penyakit, atau terapi medis, dengan membandingkan kelompok intervensi dengan kelompok kontrol.
6. **Pengembangan Kebijakan Kesehatan:** Analisis statistik yang dilakukan oleh ahli biostatistik memberikan dukungan untuk pengembangan kebijakan kesehatan yang efektif, berdasarkan bukti-bukti yang diperoleh dari penelitian dan pemantauan kesehatan masyarakat.
7. **Pemantauan Kesehatan Masyarakat:** Biostatistik digunakan untuk mengumpulkan, menganalisis, dan memantau data kesehatan populasi, membantu dalam deteksi dini penyakit, peningkatan kualitas layanan kesehatan, dan pemantauan faktor-faktor risiko kesehatan.

Sejarah singkat Biostatistik meliputi perkembangan dan evolusi penggunaan metode statistik dalam ilmu kesehatan dan biologi. Berikut adalah beberapa titik penting dalam sejarah Biostatistik:

1. **Abad ke-17:** Awal Penggunaan Statistik dalam Ilmu Kesehatan
Penggunaan teknik statistik sederhana dalam survei epidemiologi awal, seperti penghitungan angka kematian dan prevalensi penyakit.
2. **Abad ke-19:** Perkembangan Statistik Medis
Perkembangan statistik medis oleh tokoh-tokoh seperti William Farr dan Florence Nightingale, yang menggunakan data untuk mengevaluasi kondisi kesehatan populasi dan efektivitas intervensi kesehatan.

3. **Abad ke-20:** Munculnya Biostatistik sebagai Disiplin Ilmiah
 - a. Pada awal abad ke-20, biostatistik mulai diakui sebagai disiplin ilmiah tersendiri dengan kontribusi para tokoh seperti Sir Ronald Fisher, yang mengembangkan banyak teknik statistik yang penting untuk penelitian biologi dan kesehatan.
 - b. Penggunaan metode statistik yang semakin canggih dalam penelitian kesehatan, seperti desain eksperimen klinis acak dan analisis regresi.
4. **Abad ke-21:** Kemajuan Teknologi dan Metode Statistik
Peningkatan penggunaan teknologi komputer dan perangkat lunak statistik telah memungkinkan analisis data yang lebih kompleks dan akurat dalam bidang biologi dan kesehatan.
5. Perkembangan metode statistik baru untuk menangani data besar (*big data*) dan analisis genomik dalam riset kesehatan.
6. **Masa Kini:** Biostatistik dalam Pengambilan Keputusan Kesehatan

Biostatistik memiliki peran yang semakin penting dalam pengambilan keputusan kesehatan, termasuk evaluasi efektivitas pengobatan, perencanaan kebijakan kesehatan, dan pemantauan kesehatan masyarakat.

Interdisiplineritas biostatistik dengan bidang-bidang lain seperti bioinformatika, epidemiologi, dan ilmu kesehatan masyarakat.

Sejarah Biostatistik mencerminkan evolusi penggunaan metode statistik dalam memahami dan mengatasi tantangan dalam bidang biologi, kedokteran, dan kesehatan masyarakat.

Peran Biostatistik dalam Ilmu Kesehatan merupakan bagian integral dalam mencapai tujuan utama ilmu kesehatan, yaitu meningkatkan kesehatan dan kesejahteraan masyarakat secara keseluruhan (Altman, D. G. (Ed.), 2018) (Pagano, M., & Gauvreau, K., 2018).

1.2 Konsep Dasar Statistik

Untuk memahami analisis data dan pengambilan keputusan berbasis data, konsep dasar statistik sangat penting. Berikut adalah beberapa konsep dasar statistik yang penting:

1. **Variabel:** Variabel adalah atribut atau karakteristik yang dapat diukur atau diamati. Variabel dapat bersifat kuantitatif (memiliki nilai numerik) atau kualitatif (berupa kategori atau jenis).
2. **Data:** Informasi yang diperoleh melalui pengukuran atau observasi dikenal sebagai data. Data dapat berupa angka (data numerik) atau kata-kata (data kualitatif). Data menjadi bahan utama dalam analisis statistik.
3. **Populasi dan Sampel:** Sampel adalah bagian dari populasi yang dipilih untuk diamati atau diuji. Populasi adalah himpunan lengkap dari semua orang atau objek yang menjadi subjek penelitian. Untuk menggeneralisasi hasil dari sampel ke populasi, sampel diambil.
4. **Skala Pengukuran:** Jenis informasi yang diberikan oleh variabel digambarkan oleh skala pengukuran, yang dapat berupa nominal, ordinal, interval, atau rasio, masing-masing dengan tingkat informasi yang berbeda.
5. **Statistik Deskriptif:** Karakteristik dasar set data, seperti median, mode, dan mean, dibahas dalam statistik deskriptif. Teknik ini mencakup ukuran penyebaran dan pemusatan data, seperti rentang, varians, dan simpangan baku.
6. **Probabilitas:** Probabilitas adalah ukuran dari tingkat keyakinan atau kepastian terhadap kejadian tertentu. Dalam statistik, probabilitas digunakan untuk mengukur peluang dari suatu kejadian atau hasil.
7. **Distribusi Probabilitas:** Distribusi probabilitas menggambarkan pola atau bentuk yang mungkin dari hasil suatu eksperimen acak. Contoh distribusi probabilitas termasuk distribusi normal, binomial, dan Poisson.
8. **Inferensi Statistik:** Inferensi statistik adalah proses menggunakan data sampel untuk membuat kesimpulan atau membuat perkiraan tentang populasi yang lebih luas. Ini

melibatkan estimasi parameter populasi dan pengujian hipotesis.

9. Analisis Regresi dan Korelasi: Analisis regresi dan korelasi digunakan untuk mempelajari bagaimana dua atau lebih variabel berinteraksi satu sama lain. Regresi menghitung nilai variabel dependen berdasarkan nilai variabel independen, sementara korelasi menghitung kekuatan dan arah hubungan antara variabel.

10. Penggunaan Software Statistik: Penggunaan perangkat lunak statistik, seperti R, Python, SPSS, atau SAS, mempermudah analisis data dengan menyediakan berbagai alat dan teknik statistik serta visualisasi data yang efisien.

Pemahaman konsep dasar statistik menjadi kunci untuk berhasil dalam menerapkan metode statistik dalam analisis data dan pengambilan keputusan yang berkualitas (Moore, D. S., & McCabe, G. P., 2018) (Daniel, W. W., & Cross, C. L., 2018).

1.2.1 Variabel dan Data

Variabel dan data adalah konsep dasar yang sangat penting dalam statistik. Mari kita bahas keduanya secara terpisah:

1. Variabel:

- a. Variabel adalah sifat atau fitur yang dapat diukur, diamati, atau diidentifikasi pada objek atau orang dalam suatu penelitian atau penyelidikan.
- b. Variabel dapat berupa kuantitatif (numerik) atau kualitatif (kategori atau jenis).
- c. Variabel kuantitatif dapat dikelompokkan menjadi dua jenis:
 - 1) Variabel diskrit: Variabel yang nilainya terbatas pada sejumlah nilai tertentu, biasanya berupa bilangan bulat. Contoh: jumlah anak dalam sebuah keluarga.
 - 2) Variabel kontinu: Variable yang memiliki kemampuan untuk mengambil nilai antara dua angka yang tersedia. Salah satu contohnya adalah tinggi badan seseorang.

- d. Variabel kualitatif juga dapat dikelompokkan menjadi dua jenis:
- 1) Variabel nominal: Variabel yang menunjukkan jenis atau kategori tanpa urutan tertentu. Salah satu contohnya adalah jenis kelamin.
 - 2) Variabel ordinal: Variabel yang memiliki urutan atau tingkatan, tetapi jarak antar tingkatan tidak konsisten. Contoh: tingkat akademik (SD, SMP, dan SMA).

2. Data:

- a. Data merupakan kumpulan fakta, angka, atau informasi yang diperoleh dari pengukuran atau observasi.
- b. Data dapat berasal dari banyak sumber, seperti survei, eksperimen, catatan medis, atau studi observasional.
- c. Data dapat dikelompokkan menjadi dua jenis utama:
 - 1) Data kualitatif: Data yang berupa deskripsi atau karakteristik non-angka. Contoh: jenis kelamin, warna mata.
 - 2) Data kuantitatif: Data yang berupa jumlah atau angka. Data kuantitatif dapat dikelompokkan menjadi dua jenis:
 - Data diskrit: Data yang nilainya terbatas pada sejumlah nilai tertentu, biasanya berupa bilangan bulat. Contoh: jumlah anak dalam sebuah keluarga.
 - Data kontinu: Data yang dapat mengambil nilai di antara dua angka yang diberikan. Contoh: tinggi badan, berat badan.

Variabel dan data merupakan komponen penting dalam analisis statistik, dan pemahaman yang baik tentang keduanya diperlukan untuk melakukan analisis yang akurat dan informatif (Triola, M. F., 2018) (Moore, D. S., & McCabe, G. P., 2018).

1.2.2 Skala Pengukuran

Skala pengukuran merujuk pada cara variabel diukur atau dikelompokkan dalam statistik. Terdapat empat jenis utama skala

pengukuran, yang masing-masing memiliki tingkat informasi yang berbeda. Berikut adalah penjelasan singkat tentang masing-masing jenis skala pengukuran:

1. Skala Nominal:

- a. Jenis skala yang paling sederhana adalah skala nominal.
- b. Kategori yang tidak memiliki urutan atau tingkat termasuk variabel yang diukur dengan skala nominal.
- c. Beberapa contoh variabel nominal adalah jenis kelamin (laki-laki, perempuan), warna (merah, biru, hijau), atau status perkawinan (belum menikah, menikah, cerai).

2. Skala Ordinal:

- a. Skala ordinal memiliki tingkat informasi yang lebih tinggi daripada skala nominal karena variabelnya memiliki urutan atau tingkat.
- b. Namun, jarak antara tingkatan tidak selalu konsisten.
- c. Tingkat kepuasan (sangat puas, puas, netral, tidak puas, sangat tidak puas) atau tingkat pendidikan (SD, SMP, SMA) adalah contoh dari variabel ordinal.

3. Skala Interval:

- a. Skala interval memiliki semua sifat skala ordinal, ditambah dengan jarak antar tingkatan yang konsisten dan definisi nol yang merupakan nilai arbitrari.
- b. Tidak ada titik nol mutlak dalam skala interval.
- c. Suhu Celsius atau Fahrenheit, waktu dalam jam atau menit, adalah contoh dari variabel interval..

4. Skala Rasio:

- a. Skala rasio memiliki semua karakteristik skala interval dan memiliki titik nol mutlak.
- b. Titik nol yang mutlak memungkinkan perbandingan proporsional yang lebih valid antar nilai.
- c. Berat badan, tinggi badan, atau penghasilan adalah contoh dari variabel rasio.

Sangat penting untuk memahami jenis skala pengukuran karena akan memengaruhi jenis analisis statistik yang tepat yang dapat dilakukan terhadap variabel tersebut. Misalnya, jenis skala pengukuran akan memengaruhi jenis statistik deskriptif dan

inferensial yang akan digunakan dalam analisis data. (Trochim, W. M., & Donnelly, J. P., 2008) (Gravetter, F. J., & Wallnau, L. B., 2014).

1.2.3 Statistik deskriptif

Cabang statistik yang disebut statistik deskriptif berfokus pada pengorganisasian, penggalan, dan penjelasan data secara ringkas dan informatif. Tujuan utama dari statistik deskriptif adalah untuk merangkum dan menggambarkan karakteristik dasar dari suatu set data. Berikut adalah beberapa konsep utama dalam statistik deskriptif:

1. **Ukuran Pemusatan Data:** Ukuran pemusatan data digunakan untuk menunjukkan nilai "tengah" atau "pusat" dari suatu distribusi data. Contoh ukuran pemusatan data meliputi:
 - a. Mean: Nilai rata-rata set data.
 - b. Median: Nilai tengah data adalah nilai yang ditentukan ketika data diurutkan dari yang terkecil hingga yang terbesar.
 - c. Mode: Nilai-nilai yang paling sering ditemukan dalam data.
2. **Ukuran Penyebaran Data:** Ukuran penyebaran data menggambarkan seberapa jauh data tersebar dari nilai tengah atau pusatnya. Contoh ukuran penyebaran data meliputi:
 - a. Rentang: Perbedaan antara nilai data maksimum dan minimum.
 - b. Varians: Ukuran yang menunjukkan seberapa jauh setiap titik data dari nilai rata-rata.
 - c. Simpangan Baku: Ukuran sebaran atau variasi dari nilai-nilai individu dalam data terhadap nilai rata-rata.
3. **Distribusi Data:** Distribusi data menggambarkan cara data terdistribusi di sekitar nilai-nilai pusatnya. Contoh distribusi data meliputi distribusi normal, distribusi simetris, atau distribusi bimodal.
4. **Grafik Statistik:** Grafik statistik adalah representasi visual dari data yang membantu dalam memahami pola dan karakteristik data. Contoh grafik statistik meliputi histogram,

diagram batang, diagram lingkaran, atau diagram kotak dan cambuk (*boxplot*).

Statistik deskriptif sangat penting karena memberikan gambaran umum tentang data yang diamati, sehingga memudahkan dalam pemahaman dan interpretasi. Ini juga merupakan langkah pertama yang sering dilakukan sebelum melakukan analisis statistik yang lebih kompleks (Triola, M. F., 2018) (Moore, D. S., & McCabe, G. P., 2018).

1.3 Probabilitas dalam Biostatistik

Probabilitas memainkan peran penting dalam biostatistik, terutama dalam interpretasi hasil penelitian, pengambilan keputusan kesehatan, dan pengembangan model matematika dalam epidemiologi dan kesehatan masyarakat. Berikut adalah beberapa konsep penting tentang probabilitas dalam biostatistik:

1. Definisi Probabilitas:

- a. Probabilitas adalah ukuran dari tingkat keyakinan atau kepastian terhadap terjadinya suatu kejadian.
- b. Dalam konteks biostatistik, probabilitas dapat digunakan untuk mengukur kemungkinan dari berbagai hasil, seperti keberhasilan suatu pengobatan, kejadian suatu penyakit, atau efek samping dari suatu tindakan medis.

2. Distribusi Probabilitas:

- a. Distribusi probabilitas adalah model matematika yang menggambarkan pola atau bentuk yang mungkin dari hasil suatu eksperimen acak.
- b. Beberapa distribusi probabilitas yang umum digunakan dalam biostatistik meliputi:
 - 1) Distribusi Binomial: Digunakan ketika variabel acak hanya memiliki dua hasil mungkin (berhasil atau gagal) dalam serangkaian uji coba independen.
 - 2) Distribusi Normal: Merupakan distribusi probabilitas yang paling umum digunakan, sering digunakan untuk memodelkan data kontinu dalam berbagai bidang ilmu kesehatan.

- 3) Distribusi Poisson: Digunakan untuk memodelkan jumlah kejadian jarang dalam interval waktu atau ruang tertentu, seperti jumlah penyakit dalam suatu populasi.

3. Aplikasi Probabilitas dalam Penelitian Kesehatan:

- a. Probabilitas digunakan dalam berbagai aspek penelitian kesehatan, termasuk dalam perencanaan dan analisis data.
- b. Penggunaan statistik inferensial, seperti uji hipotesis dan interval kepercayaan, bergantung pada gagasan probabilitas untuk mencapai kesimpulan tentang populasi berdasarkan data sampel.

4. Pengambilan Keputusan Kesehatan:

- a. Probabilitas membantu dalam pengambilan keputusan kesehatan dengan menyediakan dasar untuk memperkirakan risiko dan manfaat dari tindakan medis atau intervensi kesehatan.
- b. Dalam pengobatan, probabilitas dapat digunakan untuk memperkirakan efektivitas suatu pengobatan atau prosedur bedah, serta untuk memperkirakan risiko efek samping.

5. Pengembangan Model Epidemiologi: Probabilitas digunakan dalam pengembangan model matematika dalam epidemiologi untuk memahami penyebaran penyakit, mengukur dampak intervensi kesehatan, dan meramalkan tren epidemiologi di masa depan.

Probabilitas merupakan alat penting dalam analisis data dan pengambilan keputusan dalam konteks kesehatan dan ilmu biologi. Pemahaman yang baik tentang konsep probabilitas memungkinkan para peneliti dan praktisi kesehatan untuk membuat keputusan yang informasional dan berbasis bukti (Kleinbaum, D. G. et al.,2013) (Armitage, P. et al.,2001).

1.4 Inferensi Statistik

Inferensi statistik adalah proses menggunakan data sampel untuk membuat kesimpulan atau membuat perkiraan tentang

populasi yang lebih luas. Berikut adalah beberapa konsep penting dalam inferensi statistik, karena ini adalah bagian utama statistik yang memungkinkan kita untuk membuat generalisasi dari data yang diamati ke populasi yang lebih besar (Agresti, A., & Finlay, B., 2009) (Daniel, W. W., & Cross, C. L., 2018):

1. **Estimasi Parameter:** Estimasi parameter melibatkan penggunaan data sampel untuk membuat perkiraan tentang parameter yang tidak diketahui dalam populasi. Misalnya, kita dapat menggunakan data sampel untuk mengestimasi rata-rata populasi, proporsi, atau varians.
2. **Interval Kepercayaan:** Rentang nilai yang mungkin untuk parameter populasi disebut interval kepercayaan, yang didasarkan pada data sampel dan tingkat kepercayaan tertentu. Interval kepercayaan memberikan informasi tentang seberapa presisi estimasi kita dan seberapa yakin kita bahwa rentang tersebut berisi nilai parameter sebenarnya.
3. **Pengujian Hipotesis:** Pengujian hipotesis menggunakan data sampel untuk menguji hipotesis atau asumsi populasi. Hipotesis nol (H_0) menyatakan bahwa tidak ada efek atau perubahan yang signifikan, sedangkan hipotesis alternatif (H_1) menyatakan bahwa ada efek atau perubahan yang signifikan. Pengujian hipotesis melibatkan perbandingan antara data sampel dengan apa yang diharapkan jika hipotesis nol benar.
4. **Uji Parametrik dan Non-Parametrik:**
 - a. Uji parametrik menggunakan asumsi tentang distribusi data, seperti distribusi normal, untuk melakukan inferensi statistik.
 - b. Uji non-parametrik adalah metode alternatif yang tidak memerlukan asumsi tertentu tentang distribusi data. Mereka lebih fleksibel tetapi sering kali memiliki kekuatan statistik yang lebih rendah daripada uji parametrik dalam mengidentifikasi perbedaan yang signifikan.

Dalam berbagai bidang, seperti ilmu kesehatan, ilmu sosial, dan bisnis, inferensi statistik membantu kita membuat kesimpulan berdasarkan data sampel. Penting untuk memahami prinsip-prinsip dasar inferensi statistik agar dapat menginterpretasikan hasil analisis dengan benar dan membuat keputusan yang informasional.

1.5 Analisis Regresi dan Korelasi

Dua teknik statistik, analisis regresi dan korelasi, digunakan untuk memahami bagaimana variabel-variabel dalam data berinteraksi satu sama lain. Meskipun keduanya seringkali dikaitkan, korelasi dan regresi memiliki perbedaan dalam tujuan dan aplikasinya. Berikut adalah penjelasan singkat tentang masing-masing:

1. Korelasi:

- a. Metode untuk mengukur kekuatan dan arah hubungan antara dua variabel kuantitatif dikenal sebagai korelasi.
- b. Koefisien korelasi Pearson adalah ukuran yang biasa digunakan untuk mengukur bagaimana dua variabel berkorelasi satu sama lain. Ini juga dikenal sebagai koefisien korelasi.
- c. Koefisien korelasi memiliki nilai antara -1 dan 1; nilai positif menunjukkan hubungan positif, nilai negatif menunjukkan hubungan negatif, dan nilai yang hampir nol menunjukkan hubungan yang lemah atau tidak ada.

2. Regresi:

- a. Analisis regresi digunakan untuk mengetahui bagaimana satu atau lebih variabel independen, juga dikenal sebagai variabel penjelas, berhubungan dengan satu variabel dependen, juga dikenal sebagai variabel yang ingin diprediksi.
- b. Tujuan analisis regresi adalah untuk membuat model matematis yang dapat digunakan untuk memprediksi nilai variabel dependen menggunakan nilai variabel independen.
- c. Model regresi dalam regresi linier sederhana hanya memperhitungkan satu variabel independen; namun,

model regresi dalam regresi linier berganda memperhitungkan lebih dari satu variabel independen.

Berbeda dengan metode regresi, metode korelasi hanya melihat kekuatan dan arah hubungan antara variabel, sedangkan metode regresi berusaha untuk memprediksi nilai variabel dependen dengan mempertimbangkan variabel independen.

Kedua teknik ini sangat bermanfaat untuk analisis data dalam berbagai bidang, seperti ilmu kesehatan, ekonomi, psikologi, dan ilmu sosial. Untuk membuat perkiraan, mengidentifikasi faktor risiko, dan membuat keputusan yang lebih baik, kita perlu memahami bagaimana variabel-variabel tersebut berhubungan satu sama lain (Kutner, M, et al., 2004)(Field, A., 2013).

1.6 Desain Penelitian dalam Biostatistik

Dalam biostatistik, desain penelitian mengacu pada perencanaan dan struktur yang digunakan untuk mengumpulkan data yang diperlukan untuk menjawab pertanyaan penelitian. Desain penelitian yang baik sangat penting karena memengaruhi validitas, reliabilitas, dan generalisabilitas hasil penelitian. Berikut adalah beberapa jenis desain penelitian yang umum digunakan dalam biostatistik:

1. **Desain Penelitian Observasional:** Desain penelitian observasional melibatkan pengamatan terhadap individu atau kelompok tanpa intervensi yang ditetapkan. Studi potong lintang (*cross-sectional*), studi kohort, dan studi kasus-kontrol adalah contoh desain penelitian observasional. Design ini sering digunakan untuk memahami pola penyakit dalam populasi atau untuk mengidentifikasi hubungan antara faktor risiko dan hasil kesehatan.
2. **Desain Penelitian Eksperimental:** Uji klinis acak terkontrol, uji lapangan, dan percobaan laboratorium adalah beberapa contoh desain penelitian eksperimental. Mereka juga melibatkan memberikan intervensi atau perlakuan kepada kelompok eksperimen dan membandingkannya dengan kelompok kontrol yang tidak menerima intervensi.

Desain ini membantu dalam mengevaluasi efektivitas intervensi kesehatan atau pengobatan.

3. **Desain Penelitian Longitudinal:** Studi kohort dan studi intervensi dengan pengukuran sebelum dan sesudah intervensi adalah contoh desain penelitian longitudinal, yang melibatkan pengumpulan data dari subjek yang sama pada beberapa titik waktu yang berbeda. Desain seperti ini membantu memahami perubahan dalam variabel dari waktu ke waktu dan menemukan pola perubahan dalam jangka waktu yang lebih lama.
4. **Desain Penelitian Campuran:** Desain penelitian campuran menggabungkan elemen-elemen dari desain penelitian observasional dan eksperimental. Contoh desain penelitian campuran meliputi desain penelitian kohort prospektif dengan intervensi atau desain penelitian observasional dengan sub-grup intervensi. Hubungan sebab-akibat dan variabel yang memengaruhi hasil kesehatan dapat dipahami lebih baik dengan desain ini.

Pemilihan desain penelitian yang tepat sangat tergantung pada pertanyaan penelitian yang diajukan, sumber daya yang tersedia, dan etika penelitian yang relevan. Dengan merancang penelitian dengan cermat, peneliti dapat menghasilkan bukti yang kuat dan relevan untuk memajukan pemahaman dan praktik dalam bidang biostatistik dan ilmu kesehatan (Dawson, B., & Trapp, R. G., 2004)(Porta, M., 2014).

1.7 Analisis Data dalam Biostatistik

Analisis data dalam biostatistik merupakan proses penting dalam merangkum, menganalisis, dan menafsirkan data yang diperoleh dari penelitian kesehatan dan ilmu biologi. Tujuan utama dari analisis data adalah untuk memperoleh pemahaman yang lebih baik tentang pola, hubungan, dan perbedaan yang ditemukan dalam data yang diamati. Berikut adalah beberapa langkah umum dalam analisis data dalam biostatistik (Dawson, B., & Trapp, R. G., 2004) (Altman, D. G. (Ed.), 2018):

1. **Persiapan Data:** Langkah awal dalam analisis data adalah persiapan data, termasuk pembersihan data, pemrosesan, dan penyusunan data dalam format yang sesuai untuk analisis lebih lanjut. Identifikasi dan penanganan nilai yang hilang, data yang tidak valid, atau outlier juga termasuk dalam persiapan data.
2. **Deskripsi Data:** Karakteristik data dasar seperti pemusatan (seperti mode, mean, dan median) dan penyebaran (seperti rentang, varians, dan simpangan baku) digambarkan dengan statistik deskriptif. Grafik statistik, seperti histogram, diagram batang, atau diagram kotak dan cambuk, juga dapat digunakan untuk memberikan gambaran visual tentang distribusi data.
3. **Analisis Bivariat:** Analisis bivariat melihat bagaimana dua variabel dalam data berhubungan satu sama lain. Bergantung pada jenis data dan pertanyaan penelitian, metode analisis bivariat termasuk korelasi, uji t, dan analisis regresi.
4. **Analisis Multivariat:** Analisis multivariat melihat bagaimana lebih dari dua variabel dalam data berinteraksi satu sama lain. Metode yang digunakan dalam analisis multivariat termasuk analisis regresi berganda, analisis faktor, analisis komponen utama, dan analisis kluster.
5. **Uji Hipotesis dan Estimasi Parameter:** Berdasarkan data sampel, uji hipotesis digunakan untuk menguji klaim atau asumsi populasi. Sementara itu, estimasi parameter digunakan untuk membuat perkiraan parameter populasi berdasarkan data sampel.
6. **Interpretasi Hasil:** Interpretasi hasil adalah langkah akhir dalam analisis data, yang melibatkan penafsiran temuan dan pembuatan kesimpulan berdasarkan bukti yang diperoleh dari analisis statistik. Interpretasi hasil harus berhubungan kembali dengan pertanyaan penelitian yang diajukan dan relevan untuk tujuan penelitian.
7. **Pelaporan Hasil:** Hasil analisis data biasanya dilaporkan dalam bentuk laporan ilmiah atau artikel jurnal, yang mencakup deskripsi metodologi, temuan utama, analisis statistik, dan interpretasi hasil.

Analisis data biostatistik sangat penting untuk pemahaman yang lebih baik tentang fenomena biologi dan kesehatan serta untuk membuat keputusan yang berbasis bukti. Dengan menggunakan metode dan teknik statistik yang tepat, peneliti dapat menyajikan hasil yang akurat dan informatif untuk memajukan ilmu pengetahuan dan praktik kesehatan.

DAFTAR PUSTAKA

- Agresti, A., & Finlay, B. (2009). *Statistical Methods for the Social Sciences*. Pearson Education.
- Altman, D. G. (Ed.). (2018). *Practical Statistics for Medical Research*. CRC Press.
- Armitage, P., Berry, G., & Matthews, J. N. S. (2001). *Statistical Methods in Medical Research*. John Wiley & Sons.
- Daniel, W. W., & Cross, C. L. (2018). *Biostatistics: A Foundation for Analysis in the Health Sciences*. John Wiley & Sons.
- Dawson, B., & Trapp, R. G. (2004). *Basic & Clinical Biostatistics*. McGraw-Hill Medical.
- Field, A. (2013). *Discovering Statistics Using IBM SPSS Statistics*. Sage.
- Gravetter, F. J., & Wallnau, L. B. (2014). *Essentials of Statistics for the Behavioral Sciences*. Cengage Learning.
- Kirkwood, B. R., & Sterne, J. A. C. (2019). *Essential Medical Statistics*. John Wiley & Sons.
- Kleinbaum, D. G., Kupper, L. L., Nizam, A., & Muller, K. E. (2013). *Applied Regression Analysis and Other Multivariable Methods*. Cengage Learning.
- Kutner, M. H., Nachtsheim, C. J., Neter, J., & Li, W. (2004). *Applied Linear Statistical Models*. McGraw-Hill.
- Moore, D. S., & McCabe, G. P. (2018). *Introduction to the Practice of Statistics*. W. H. Freeman.
- Pagano, M., & Gauvreau, K. (2018). *Principles of Biostatistics*. Cengage Learning.
- Porta, M. (2014). *A Dictionary of Epidemiology*. Oxford University Press.
- Rosner, B. (2015). *Fundamentals of Biostatistics*. Cengage Learning.
- Triola, M. F. (2018). *Elementary Statistics*. Pearson.
- Trochim, W. M., & Donnelly, J. P. (2008). *The Research Methods Knowledge Base*. Atomic Dog Publishing.

BAB 2

PENGUMPULAN DATA DALAM PENELITIAN KESEHATAN

Oleh Iha Nursolihah

2.1 Pendahuluan

Penelitian merupakan proses sistematis untuk menjawab suatu permasalahan melalui pengumpulan, analisis, dan interpretasi data. Pengumpulan data dalam penelitian kesehatan adalah proses mengumpulkan informasi yang relevan, akurat, dan lengkap untuk menjawab pertanyaan penelitian atau menguji hipotesis terkait kesehatan individu maupun populasi. Tahap pengumpulan data dalam penelitian kesehatan bukan sekedar tahapan dalam memperoleh sebanyak mungkin data/informasi, tetapi peneliti perlu memastikan bahwa data tersebut valid, reliabel, etis, dan terstandarisasi.

Prinsip validitas perlu dipenuhi karena data yang dikumpulkan harus benar-benar mengukur variabel yang ingin diteliti. Misalnya, jika penelitian ingin mengukur berat badan bayi, instrumen yang digunakan harus relevan dan teruji untuk mengukur berat badan bayi, bukan instrumen untuk mengukur berat badan orang dewasa. Prinsip reliabilitas merupakan upaya dalam memastikan bahwa data yang diperoleh konsisten dari waktu ke waktu dan tidak berubah hanya karena faktor kebetulan. Artinya, jika pengukuran dilakukan ulang pada kondisi yang sama, hasilnya akan relatif sama. Hal ini penting agar temuan penelitian dapat dipercaya dan digunakan sebagai dasar pengambilan keputusan. Kemudian prinsip etis, etika dalam pengumpulan data tidak dapat diabaikan, terutama dalam konteks kesehatan. Peneliti wajib menghormati hak dan privasi responden salah satunya dengan meminta persetujuan (*informed consent*) sebelum pengumpulan data, serta menjaga kerahasiaan informasi yang diperoleh. Aspek ini tidak hanya melindungi peserta penelitian, tetapi juga menjaga reputasi dan

integritas peneliti. Terakhir prinsip standarisasi merupakan landasan untuk menjaga konsistensi. Semua prosedur pengumpulan data harus dilakukan dengan cara yang sama untuk setiap responden, mulai dari instruksi, waktu pengukuran, hingga alat yang digunakan. Dengan demikian, perbedaan hasil dapat dikaitkan dengan variabel penelitian, bukan karena perbedaan metode pengumpulan data.

Dengan menerapkan prinsip-prinsip dalam pengumpulan data secara konsisten, peneliti dapat memastikan bahwa data yang diperoleh memiliki kualitas tinggi, dapat dipercaya, dan relevan dengan tujuan penelitian. Prinsip-prinsip tersebut pada akhirnya akan memperkuat temuan dan kontribusi penelitian terhadap ilmu pengetahuan dan praktik kesehatan.

Pengumpulan data dalam penelitian kesehatan dapat dilakukan melalui dua jenis sumber, yaitu data primer dan data sekunder. Data primer merupakan data yang dikumpulkan langsung oleh peneliti dari responden atau subjek penelitian, sedangkan data sekunder adalah data yang diperoleh dari sumber yang sudah ada sebelumnya, seperti data dari rekam medis rumah sakit, laporan survei kesehatan dari instansi resmi, atau melalui basis data lainnya. Berbeda dengan data sekunder, data primer yaitu data yang dikumpul sendiri oleh peneliti dari subjek penelitian, memiliki berbagai metode seperti wawancara, kuesioner, observasi, pengukuran klinis dan tes laboratorium.

2.2 Angket/Kuesioner

Kuesioner adalah instrumen penelitian yang berisi serangkaian pertanyaan tertulis atau pernyataan yang disusun secara sistematis untuk memperoleh data dari responden. Responden memberikan jawaban secara langsung pada setiap pertanyaan, sekarang ini jawaban bukan hanya dalam bentuk tulisan langsung tetapi juga memungkinkan dalam bentuk elektronik. Dalam penelitian kesehatan, kuesioner digunakan untuk mengumpulkan informasi terkait kondisi kesehatan, perilaku, pengetahuan, sikap, atau persepsi responden terhadap suatu topik.

Tujuan penggunaan kuesioner dalam suatu penelitian:

1. Pengumpulan data lebih efisien jika memiliki responden yang banyak.
2. Kuesioner lebih mengukur variabel yang bersifat subjektif, seperti sikap, persepsi, atau tingkat kepuasan.
3. Proses pengumpulan data terstandarisasi, sehingga hasil dapat dibandingkan antar responden.
4. Kuesioner dapat digunakan pada penelitian kuantitatif maupun kualitatif.

2.2.1 Jenis-Jenis Kuesioner

A. Berdasarkan Bentuk Pertanyaan

1. Kuesioner Tertutup

Jenis kuesioner ini memungkinkan responden memilih jawaban dari opsi yang disediakan. Kuesioner lebih mudah dianalisis karena jawaban terstandar, sehingga lebih hemat waktu dalam penelitian. Namun, penggunaan kuesioner tertutup tidak memungkinkan untuk mendapat jawaban di luar pilihan.

Contoh:

"Apakah Anda pernah melakukan pengobatan di Rumah Sakit ini sebelumnya? (Pernah/Tidak Pernah)"

"Jika pernah, berapa kali sudah melakukan pengobatan? (1 kali/ >1 kali)"

Kuesioner tertutup juga dapat berbentuk *checklist*, responden dapat menjawab dengan memberikan tanda pada jawaban. Contoh:

Tabel 2.1. Kuesioner Tertutup dalam Bentuk *Checklist*

Pernyataan	Sangat Tidak Setuju	Tidak Setuju	Setuju	Sangat Setuju
Pada usia 0-6 bulan bayi, saya hanya memberi ASI				√
Saya menyusui tergantung kemauan bayi, kurang lebih jarak 2 jam			√	

2. Kuesioner Terbuka

Kuesioner terbuka memungkinkan responden menuliskan jawaban bebas sesuai pandangan atau pengalamannya. Oleh karena itu, menggunakan jenis kuesioner ini akan mendapatkan data yang lebih banyak dan mendalam. Namun, karena banyaknya data tersebut pengolahan data menjadi lebih sulit, sehingga memerlukan waktu lebih lama.

Contoh:

"Bagaimana persepsi Anda mengenai pelayanan di Rumah Sakit ini?"

3. Kuesioner Campuran

Jenis kuesioner ini menggabungkan pertanyaan tertutup dan terbuka. Pengumpulan data menjadi lebih lengkap. Cocok untuk penelitian yang memerlukan data kuantitatif dan kualitatif.

B. Berdasarkan metode pengisian

1. *Self-administered Questionnaire*

Yaitu jenis kuesioner yang pengisiannya dilakukan sendiri oleh responden. Kuesionernya dapat berbentuk kertas (*paper-based*) atau elektronik (*online form*).

2. *Interviewer-administered Questionnaire*

Yaitu jenis kuesioner yang diisi oleh pewawancara berdasarkan jawaban responden. Biasanya digunakan pada responden dengan keterbatasan membaca/menulis, seperti anak-anak dan lansia.

2.2.2 Langkah Penyusunan Kuesioner

Menyusun kuesioner dengan baik dalam setiap langkahnya menjadikan kegiatan pengumpulan data dalam suatu penelitian menjadi terstruktur, sehingga kualitas penelitian terjaga. Berikut langkah dalam penyusunan kuesioner:

1. Menentukan tujuan dan variabel penelitian. Peneliti perlu memahami tujuan penelitiannya dan variabel yang diteliti, sehingga kuesioner yang dibuat sesuai dan dapat menjawab tujuan penelitian.

2. Memilih bentuk pertanyaan (tertutup, terbuka, atau campuran).
3. Menyusun indikator pengukuran untuk setiap variabel dan menentukan urutan pertanyaan yang logis.

Contoh: Penelitian mengenai Pengetahuan tentang ASI Eksklusif

Tabel 2.2. Menyusun Indikator Pengukuran Variabel

Variabel	Dimensi	Indikator	Nomor Pertanyaan
Pengetahuan	Prosedur pemberian	1. Usia pemberian ASI 2. Cara pemberian 3. Waktu pemberian	1 2, 3, 4 5, 6, 7
	Manfaat	1. Kesehatan bayi 2. Kesehatan ibu 3. Ekonomi	8, 9 10, 11, 12 13, 14, 15

4. Membuat pertanyaan sesuai indikator.
5. Menentukan skala pengukuran (nominal, ordinal, interval, rasio).
6. Melakukan *uji coba* (pre-test) untuk menguji validitas dan reliabilitas.
7. Melakukan revisi berdasarkan hasil uji coba.

Penyusunan kuesioner secara sistematis, menjadikan instrumen pengumpulan data **valid** yaitu mengukur variabel yang memang ingin diukur, **reliabel** yakni memberikan hasil yang konsisten jika digunakan berulang kali, **jelas dan singkat** sesuai kebutuhan, **relevan** sesuai dengan tujuan penelitian, dan **tersusun sistematis** berdasarkan urutan logis dari pertanyaan umum ke pertanyaan khusus.

2.3 Wawancara

Wawancara adalah metode pengumpulan data dengan cara tanya jawab langsung antara pewawancara (*interviewer*) dan responden (informan) untuk memperoleh informasi yang relevan dengan tujuan penelitian. Instrumen dalam wawancara adalah panduan wawancara (*interview guide*), yaitu daftar pertanyaan yang

disusun secara sistematis agar proses wawancara terarah dan konsisten.

Pengumpulan data menggunakan metode wawancara memudahkan peneliti dalam menggali informasi mendalam yang tidak bisa didapatkan dari kuesioner tertulis, jawaban dari responden lebih jelas karena peneliti akan mendapatkan data kualitatif, seperti pengalaman, persepsi, atau pendapat responden. Selain itu, metode wawancara dapat menjangkau responden dengan keterbatasan membaca dan atau menulis. Namun, metode ini kurang cocok untuk jumlah responden yang banyak jika sumber daya terbatas.

2.3.1 Jenis Wawancara

1. Wawancara Terstruktur

Yaitu metode wawancara menggunakan daftar pertanyaan tetap yang sama untuk semua responden. Jenis wawancara ini cocok untuk penelitian kuantitatif. Contoh penelitian menggunakan wawancara terstruktur seperti survei kesehatan masyarakat menggunakan formulir resmi.

2. Wawancara Semi-terstruktur

Yaitu jenis wawancara yang memiliki daftar pertanyaan inti, tetapi pewawancara boleh mengembangkan pertanyaan tambahan sesuai situasi. Jenis ini cocok untuk penelitian campuran (kuantitatif dan kualitatif).

3. Wawancara Tidak Terstruktur

Yaitu metode wawancara dengan pertanyaan penelitian sangat fleksibel, lebih seperti percakapan bebas dengan tujuan tertentu. Cocok untuk penelitian eksploratif atau studi kasus mendalam, seperti topik adat, budaya, seksual. Jenis wawancara ini lebih sulit untuk dianalisis dibandingkan jenis lainnya karena datanya yang luas.

2.3.2 Panduan Wawancara

Panduan atau pedoman wawancara perlu disiapkan, mencakup:

1. Identitas penelitian, seperti judul, tujuan, identitas pewawancara.

2. Identitas responden, seperti nama, umur, jenis kelamin, pendidikan, pekerjaan.
3. Pertanyaan pembuka untuk mencairkan suasana
4. Pertanyaan inti sesuai variabel penelitian
5. Pertanyaan penutup yang bersifat reflektif atau ringkasan
6. Catatan pewawancara untuk observasi non-verbal

Pelaksanaan wawancara, selain perlu mengikuti panduan wawancara juga perlu menerapkan prinsip validitas yakni pertanyaan yang diajukan sesuai dengan tujuan penelitian, reliabilitas yakni pertanyaan yang sama menghasilkan jawaban konsisten pada situasi serupa, netralitas yaitu pewawancara tidak mengarahkan atau memengaruhi jawaban responden, empati dengan menciptakan suasana nyaman agar responden terbuka, dan beretika seperti dengan meminta *informed consent* dan menjaga kerahasiaan data.

2.3.3 Teknik Wawancara

1. Wawancara Mendalam

In-Depth Interview atau wawancara mendalam adalah proses tanya jawab secara langsung antara pewawancara dan responden secara tatap muka, bersifat personal dan mendalam, dengan tujuan menggali informasi, pengalaman, atau pandangan secara detail. Bentuk pertanyaan pada teknik bersifat terbuka dan fleksibel, diajukan secara individual (satu responden per sesi)

Pengumpulan data dengan wawancara memerlukan kemampuan pewawancara dalam menggali jawaban lebih dalam melalui pertanyaan lanjutan (*probing*). Sehingga waktu pelaksanaan relatif lama (30–90 menit per responden). Tetapi, peneliti akan mendapatkan data yang detail, dan memungkinkan eksplorasi topik yang kompleks atau sensitif.

Pewawancara hendaknya menguasai panduan wawancara, ajukan pertanyaan dengan jelas, hindari mengajukan pertanyaan tertutup agar informan bisa memberikan jawaban yang lebih mendalam, juga ajukan pertanyaan satu per satu atau tunggal untuk memastikan semua pertanyaan terjawab penuh. Galilah jawaban tentang perilaku dan pengalaman sebelum pertanyaan

tentang pendapat dan persepsi, hal tersebut menghindari responden merubah jawaban tentang perilaku yang sebenarnya dilakukan dengan jawaban persepsi yang dianggap benar jika jawabannya berbeda.

2. Diskusi Kelompok Terfokus

Diskusi Kelompok Terfokus atau *Focus Group Discussion* (FGD) merupakan metode pengumpulan data melalui diskusi terarah dengan sekelompok kecil peserta yang memiliki karakteristik tertentu sesuai tujuan penelitian dengan pengarahan moderator/fasilitator dan notulen untuk mencatat hasil. Artinya FGD dilakukan dalam kelompok bukan individu seperti wawancara mendalam, dan pastikan responden memiliki latar belakang atau pengalaman yang relevan. Waktu pelaksanaan biasanya 1–2 jam bergantung keluasan topik penelitian, dipandu oleh moderator untuk jalannya diskusi sesuai panduan topik.

Kelebihan FGD dibandingkan wawancara mendalam diantaranya FGD menghasilkan berbagai perspektif dalam waktu singkat karena diajukan pada responden yang berkelompok, selain itu interaksi antar peserta dapat memunculkan ide baru, sehingga teknik ini cocok untuk menggali persepsi, pendapat, dan ide masyarakat. Namun, dalam kasusnya mungkin terdapat risiko dominasi pembicaraan oleh beberapa peserta. Kemudian data yang dihasilkan lebih bersifat kualitatif, dan tidak cocok untuk topik yang sangat sensitif karena diajukan berkelompok.

Karakteristik FGD lainnya adalah pastikan peserta homogen agar data yang digali relevan. Lakukan minimal 2 FGD untuk satu topik atau kelompok dengan jumlah peserta dalam satu kelompok 6-12 orang, dan pastikan memilih tempat dimana peserta merasa bebas dan tidak takut mengeluarkan pendapatnya.

2.4 Observasi

Observasi adalah metode pengumpulan data dengan cara mengamati secara langsung objek atau fenomena yang diteliti, baik secara sistematis maupun tidak, untuk mendapatkan informasi yang relevan. Dalam penelitian kesehatan, observasi sering digunakan untuk mencatat perilaku, kondisi lingkungan, prosedur pelayanan,

atau kejadian kesehatan tertentu tanpa selalu mengandalkan laporan responden. Sehingga metode observasi akan memberikan informasi tentang perilaku nyata/sesungguhnya. Namun, perlu sensitivitas terhadap ekspresi wajah dan bahasa tubuh subyek penelitian. Contohnya mengamati perilaku cuci tangan, atau mengamati prosedur pelayanan di poliklinik.

Pengumpulan data menggunakan metode observasi memungkinkan peneliti untuk dapat mengumpulkan data faktual sesuai kondisi nyata di lapangan, memperoleh informasi yang sulit diungkapkan responden melalui wawancara atau kuesioner, dan mampu memverifikasi atau melengkapi data dari metode lain, misalnya membandingkan laporan pasien dengan perilaku sebenarnya.

2.4.1 Jenis Observasi

1. Berdasarkan Peran Peneliti

a. Observasi Partisipatif

Yaitu peneliti terlibat langsung dalam kegiatan atau lingkungan yang diamati. Contoh: ikut serta dalam kegiatan posyandu untuk mengamati interaksi kader dengan ibu balita.

b. Observasi Non-Partisipatif

Yaitu peneliti hanya menjadi pengamat tanpa terlibat langsung.

2. Berdasarkan Struktur

a. Observasi Terstruktur

Yaitu menggunakan panduan atau checklist yang telah disiapkan sebelumnya, biasanya untuk penelitian kuantitatif.

b. Observasi Tidak Terstruktur

Yaitu jenis observasi yang lebih bebas, tanpa panduan baku, cocok untuk penelitian eksploratif atau kualitatif.

3. Berdasarkan Sifatnya

a. Observasi Terbuka

Yaitu objek mengetahui bahwa sedang diamati.

b. Observasi Tertutup

Yaitu objek tidak mengetahui bahwa sedang diamati, untuk menghindari perubahan perilaku akibat pengamatan (*Hawthorne effect*).

4. Berdasarkan Lamanya

a. *Single observation*

Yaitu waktu observasi yang terbatas, misalnya 1 jam.

b. *Several observation*

Yaitu observasi dilakukan beberapa kali selama waktu tertentu, misalnya 3 kali dalam 1 hari.

c. *Long-term observation*

Yaitu observasi dilakukan berkali-kali dalam waktu yang lama, misalnya 1 bulan, 1 tahun.

2.4.2 Instrumen Observasi

Instrumen yang digunakan dalam pengumpulan data melalui observasi meliputi:

1. *Checklist*

Yaitu daftar poin yang diamati, biasanya berisi indikator yang sudah ditetapkan.

2. *Rating Scale*

Yaitu skala penilaian (misalnya 1–5) untuk menilai intensitas atau kualitas perilaku/kejadian.

3. *Field Notes*

Yaitu catatan lapangan yang berisi deskripsi detail hasil pengamatan, misalnya peristiwa yang diamati, interaksi, keadaan lingkungan.

4. Kamera/Video

Untuk dokumentasi visual, dengan memperhatikan izin etis.

Pelaksanaan observasi, selain perlu menyiapkan instrumen observasi, juga perlu menerapkan prinsip objektivitas yakni mencatat fakta sesuai yang terlihat, bukan interpretasi pribadi, relevansi yaitu hanya mengamati aspek yang terkait dengan tujuan penelitian, konsistensi yaitu menggunakan metode pengamatan yang sama di setiap situasi, dan penerapan etika untuk menghormati privasi dan meminta izin jika diperlukan.

2.5 Pengukuran Klinis

Pengukuran klinis adalah metode pengumpulan data melalui pemeriksaan fisik atau pengukuran parameter kesehatan secara langsung pada subjek penelitian dengan menggunakan alat ukur medis standar. Tujuannya adalah memperoleh data objektif, kuantitatif, dan terukur tentang kondisi kesehatan responden. Contoh parameter yang diukur:

1. Tanda vital: tekanan darah, suhu tubuh, frekuensi nadi, frekuensi napas
2. Antropometri: berat badan, tinggi badan, lingkar perut, indeks massa tubuh (IMT)
3. Pemeriksaan fungsi tubuh: kapasitas paru-paru, ketajaman penglihatan, pendengaran
4. Pemeriksaan penunjang non-lab: EKG (elektrokardiografi), USG, spirometri

Jika dilihat berdasarkan contoh pengukuran klinis diatas, maka metode ini cocok untuk survei kesehatan masyarakat, pemeriksaan lapangan, atau studi epidemiologi. Peneliti yang hendak menggunakan metode pengukuran klinis dalam pengumpulan data perlu menggunakan alat ukur yang terstandar dan terkalibrasi, lakukan pengukuran oleh tenaga yang terlatih sesuai prosedur medis, catatlah hasil secara teliti dan objektif, serta tetap memperhatikan etika seperti *informed consent* dan privasi pasien.

2.6 Tes Laboratorium

Tes laboratorium adalah metode pengumpulan data melalui analisis sampel biologis, seperti darah, urin, feses, sputum, atau cairan tubuh lainnya di laboratorium untuk mengetahui kondisi kesehatan atau status biokimia seseorang. Contoh tes laboratorium:

1. Tes darah: kadar gula darah, kolesterol, hemoglobin, fungsi hati, fungsi ginjal
2. Tes urin: proteinuria, glukosuria, infeksi saluran kemih
3. Tes mikrobiologi: kultur bakteri, tes sensitivitas antibiotik
4. Tes imunologi: deteksi antibodi atau antigen (misalnya tes HIV, HBsAg)
5. Tes molekuler: PCR untuk deteksi DNA/RNA patogen

Tes laboratorium digunakan untuk diagnosis penyakit, pemantauan status kesehatan, dan validasi hasil pengukuran klinis. Maka peneliti yang menggunakan metode ini untuk pengumpulan data penelitian perlu memperhatikan:

1. Pengambilan sampel sesuai prosedur (SOP) untuk mencegah kontaminasi
2. Penyimpanan dan transportasi sampel sesuai standar suhu dan waktu
3. Analisis dilakukan oleh tenaga laboratorium terlatih
4. Menggunakan alat yang terkalibrasi dan reagen yang masih berlaku
5. Menjaga kerahasiaan dan keamanan data hasil tes

DAFTAR PUSTAKA

- Creswell, J.W. and Creswell, J.D., 2018. *Research design: Qualitative, quantitative, and mixed methods approaches*. 5th ed. Thousand Oaks, CA: Sage Publications.
- Gill, P., Stewart, K., Treasure, E. and Chadwick, B., 2008. Methods of data collection in qualitative research: interviews and focus groups. *British Dental Journal*, 204(6), pp.291–295.
- Setyawati, N.F., Yuliawuri, H., Raudah, S., Pristina, N., Kaisar, M.M.M., Sucipto, A., Mudhawaroh, M., Asriati, A., Ardiansyah, R.T., Achmad, B.K. and Maranata, M., 2023. Metodologi Riset Kesehatan.
- Ariyani, H., Hairuddin, H., Palilingan, R.A., Nugroho, H., Sarumi, R. and Aji, R., 2023. Metodologi Penelitian Kesehatan dan Statistika.
- Rochyani, Neny, et al. *BIOSTATISTIKA UNTUK PERGURUAN TINGGI*. Edited by Novita, Yuliatr, GET PRESS INDONESIA, 2024.
- Saryono, 2013. *Metodologi penelitian kesehatan*. Yogyakarta: Mitra Cendekia Press.

BAB 3

DISTRIBUSI DAN DESKRIPSI DATA

Oleh Nur Al-faida

3.1 Pendahuluan

Dalam dunia statistik, pemahaman tentang cara data tersebar sangat penting untuk mengambil kesimpulan yang akurat dan membuat keputusan yang tepat. Statistik erat kaitannya dengan proses pengolahan data. Metode statistik yang digunakan akan mempengaruhi proses dan hasil analisis data. Salah dalam memilih metode pengolahan data akan berakibat sangat fataal, misalnya hasil analisis yang tidak sesuai dengan data, sehingga kesimpulan hasil penelitian juga salah atau tidak sesuai dengan fakta. Dalam menentukan metode pengolahan data yang tepat perlu memperhatikan beberapa hal, salah satunya adalah mengetahui distribusi dan deskriptif dari data yang akan digunakan. Distribusi dan deskriptif data menunjukkan peluang yang mungkin terjadi dalam penelitian, sehingga mengetahui distribusi dan deskriptif yang dimiliki oleh data menjadi hal yang krusial.

Distribusi dan deskriptif data merujuk pada cara nilai-nilai dalam kumpulan data terdistribusi dan terdeskriptif disekitar berbagai nilai.

Statistik deskriptif pada dasarnya merupakan proses transformasi data penelitian dalam bentuk yang lebih mudah dipahami dan diinterpretasikan. Tabulasi menyajikan ringkasan, pengaturan, penyusunan data dalam bentuk numerik dan grafik. Statistik deskriptif umumnya digunakan oleh peneliti untuk memberikan informasi mengenai karakteristik variabel penelitian sekaligus *mensupport* variabel yang diteliti. Kegiatan yang berhubungan dengan statistik deskriptif seperti menghitung mean (rata-rata), median, modus, mencari standar deviasi dan melihat kemencengan distribusi data dan sebagainya.

3.2 Ditribusi Data

Distribusi data adalah cara data atau variabel tersebar atau terkelompok dalam berbagai nilai kategori. Hal ini membantu dalam memahami karakteristik dan pola data, yang penting untuk analisis dan pengambilan keputusan (Sugiyono, 2019).

Distribusi data adalah fungsi yang menentukan semua nilai yang mungkin untuk suatu variabel dan juga mengkuantifikasi frekuensi relatifnya (probabilitas seberapa sering nilai tersebut muncul). Distribusi dianggap sebagai populasi yang memiliki sebaran data. Penting untuk menentukan distribusi populasi agar kita dapat menerapkan metode statistik yang tepat saat menganalisisnya.

Distribusi data banyak digunakan dalam statistik. Bayangkan seorang insinyur mengumpulkan 500 titik data dilantai produksi. Hal ini tidak akan memberikan nilai apa pun bagi manajemen kecuali mereka mengkategorikan atau mengatur data tersebut dengan cara yang bermanfaat. Metode distribusi data mengorganisasikan data mentah ke dalam metode grafis (seperti histogram, diagram kotak, diagram run dll) dan memberikan informasi yang bermanfaat.

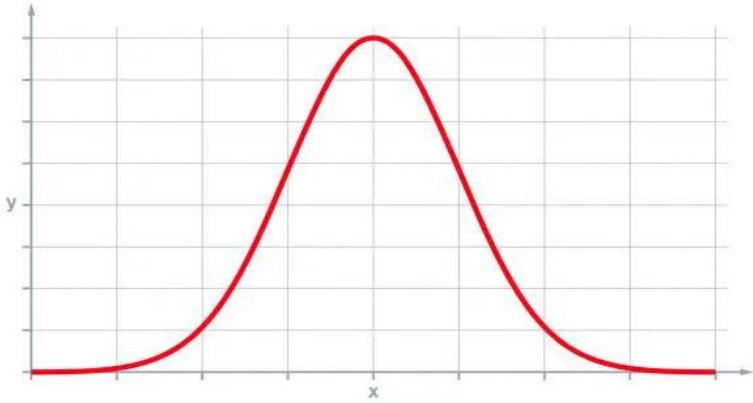
3.2.1 Jenis-jenis Distribusi

1. Distrubusi Normal

Distribusi normal adalah sebuah distribusi probabilitas yang berbentuk simetris dan berbentuk seperti lonceng (*bell-shaped curve*), di mana sebagian besar nilai data berada di sekitar nilai rata-rata (mean), dan semakin jauh dari mean, frekuensinya semakin kecil (Montgomery, 2014).

Distribusi normal ini termasuk ke dalam jenis variabel kontinu karena pendekatannya menggunakan kurva normal. Apabila kamu belum tahu apa itu variabel kontinu, jadi variabel kontinu adalah variabel yang mencakup bilangan riil dan jumlahnya tidak terbatas. Variabel kontinu ini juga mencakup bilangan pecahan dan bilangan desimal. Contohnya, ada seorang atlet lari yang berhasil lari dengan jarak 100 meter dalam waktu 9,52 detik. Nah, angka tersebut disebut dengan variabel kontinu. Bisa dikatakan variabel kontinu ini berkebalikan dengan angka bulat (variabel diskrit).

Jika kita sambungkan, variabel-variabel kontinu ini hanya bisa didapatkan dengan menggunakan distribusi peluang kontinu, salah satunya adalah distribusi normal.



Gambar 3.1. Kurva Distribusi Normal

Distribusi normal pertama kali diperkenalkan oleh Abraham DeMoivre (1733) sebagai pendekatan distribusi binomial untuk n besar. Selanjutnya dikembangkan oleh Pierre Simon de Laplace dan dikenal dengan Teorema Moivre-Laplace. Suatu data membentuk distribusi normal jika jumlah data di atas dan di bawah mean adalah sama. Grafiknya disebut kurva normal, oleh Jouffret (1872) disebut kurva lonceng/genta (*bell curve*). Alasan lain yang mendasari pentingnya Distribusi Normal adalah '*Central Limit Theorem*' yang menyebutkan bahwa, pengambilan sampel secara random dari sebuah distribusi (tanpa menghiraukan jenis distribusinya) maka distribusi dari sampel x akan mendekati distribusi normal, dan makin besar jumlah sampelnya distribusi akan makin mendekati distribusi normal.

Dalam suatu teori distribusi peluang atau probabilitas, distribusi normal menempati posisi penting di berbagai analisis statistika. Jenis distribusi ini juga kerap digunakan sebagai bahan perhitungan berbagai fenomena dalam kehidupan sehari-hari.

2. Parameter yang digunakan

Bentuk kurva serta nilai peluang distribusi normal didapatkan dari dua parameter, yaitu mean (rata-rata) dan standar deviasi (simpangan baku).

3. Nilai Rata-rata

Mean atau nilai rata-rata dalam distribusi normal akan menentukan di mana lokasi titik puncak (lengkungan dalam kurva lonceng), sedangkan nilai-nilai lainnya akan menyebar secara merata di dua sisi.

4. Standar Deviasi

Standar deviasi adalah perhitungan variabilitas yang mengukur seberapa lebar kurva distribusi normal. Fungsi standar ini adalah merepresentasikan seberapa jauh kecenderungan data dari posisi titik puncak data. Hal ini juga menunjukkan jarak (selisih) antara mean dengan data lain yang diobservasi. Apabila nilai standar deviasi semakin kecil, maka bentuk kurva akan semakin sempit dan runcing. Sebaliknya, apabila nilai standar deviasi semakin besar, maka kurva akan semakin lebar dan lebih tumpul.

3.2.2 Karakteristik Distribusi Normal

Distribusi normal yang simetris memiliki karakteristik sebagai berikut:

1. Penyebaran mean, median, dan modus memiliki nilai yang sama.
2. Nilai titik puncak (rata-rata) berada di tengah-tengah kurva kemudian data lainnya menyebar pada sekitar garis lurus yang ditarik ke bawah dari titik tengah.
3. Luas seluruh daerah di bawah kurva adalah 1, yaitu $\frac{1}{2}$ sisi kanan dan $\frac{1}{2}$ sisi kiri. Hal ini juga berlaku pada seluruh distribusi probabilitas kontinu.
4. Masing-masing ekor kurva pada kedua sisi (kanan dan kiri) memanjang tidak terbatas, bahkan di beberapa kasus, ekor kurva bisa memotong sumbu horizontal.

5. Distribusi Binomial

Distribusi binomial adalah jenis distribusi probabilitas diskrit yang menggambarkan jumlah keberhasilan dalam sejumlah percobaan yang bersifat independen, di mana setiap percobaan hanya memiliki dua kemungkinan hasil: sukses (berhasil) atau gagal (Walpole, 2012).

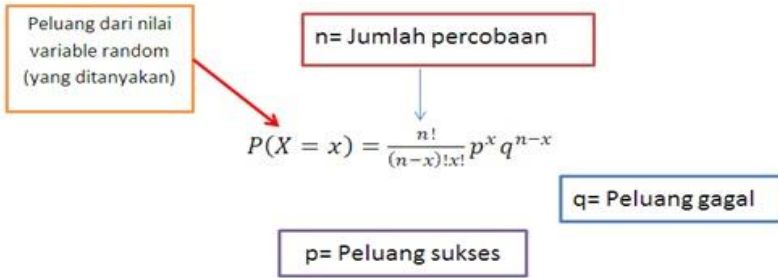
BI dalam kata BINOMIAL berarti dua. Hal ini merujuk ke setiap kali percobaan atau kesempatan, hasil yang MUNGKIN muncul hanya ada dua.

Pertanyaan yang biasanya muncul adalah, kapan kita menggunakan menghitung peluang dengan menggunakan persamaan distribusi peluang Binomial ? Jika kejadian tersebut memenuhi sifat-sifat di bawah ini maka ketika menghitung peluang kejadian tersebut terjadi maka persamaan yang digunakan adalah persamaan peluang dari distribusi binomial.

- a. Percobaan dilakukan sebanyak n kali.
- b. Setiap kali percobaan mempunyai dua kemungkinan hasil.
- c. Kemungkinan hasil dari masing-masing percobaan sama.
- d. Hasil yang diperoleh pada percobaan pertama tidak akan mempengaruhi hasil yang diperoleh pada percobaan-percobaan yang lain (saling independen).

Untuk persamaan hitung peluang dapat dilihat sebagai berikut:
Misalkan X adalah variabel random diskrit. Maka peluang dari X adalah:

$$P(X) = \frac{n!}{(n-x)!x!} p^x q^{n-x}$$



Gambar 3.2. Variabel Random Diskrit

Contoh:

Misalkan kita mempunyai satu buah koin yang terdiri atas dua sisi, depan dan belakang. Misalkan kita mengundi sebanyak 10 kali. Pada undian pertama, kemungkinan hasilnya hanya sisi depan atau sisi belakang. Pada undian kedua, kemungkinan hasilnya hanya sisi depan atau sisi belakang. Demikian seterusnya. Setiap kali mengundi, kemungkinan hasilnya sama, hanya dua yaitu sisi depan atau sisi belakang. Dari sepuluh kali percobaan, berapa peluang sisi depan muncul sebanyak dua kali ?

Diketahui:

Jumlah percobaan = $n = 10$.

Peluang sukses = peluang munculnya sisi depan dalam setiap percobaan = $p = 0.5$.

Peluang gagal = peluang tidak munculnya sisi depan dalam setiap percobaan = $q = 1-p = 0.5$.

Ditanyakan:

Dari sepuluh kali percobaan, berapa peluang sisi depan muncul sebanyak dua kali? Atau $P(X = 2)$ -> Yang ditanyakan

adalah peluang munculnya sisi depan maka kejadian yang dianggap sukses adalah jika sisi depan muncul ketika diundi.

Jawaban:

$$P(X = 2) = \frac{10!}{(10 - 2)! 2!} (0.5)^2 (1 - 0.5)^{10-2} = 0.0439$$

6. Distribusi Poisson

Distribusi Poisson adalah distribusi probabilitas diskrit yang menggambarkan probabilitas terjadinya sejumlah peristiwa dalam suatu interval waktu atau ruang tertentu, dengan asumsi bahwa peristiwa-peristiwa tersebut terjadi secara acak dan independen dengan laju rata-rata yang konstan.

Distribusi Poisson adalah distribusi probabilitas diskret yang menyatakan peluang jumlah peristiwa yang terjadi pada periode waktu tertentu apabila rata-rata kejadian tersebut diketahui dan dalam waktu yang saling bebas sejak kejadian terakhir. (distribusi Poisson juga dapat digunakan untuk jumlah kejadian pada interval tertentu seperti jarak, luas, atau volume) (Donald, 1969).

Ciri-ciri distribusi poisson:

- a. Banyaknya percobaan yang terjadi dalam suatu interval waktu atau suatu daerah tertentu tidak tergantung banyaknya hasil percobaan yang terjadi pada interval waktu atau daerah lain yang terpisah.
- b. Probabilitas hasil percobaan yang terjadi selama suatu interval waktu yang singkat atau daerah kecil, sebanding dengan panjangnya waktu atau besarnya daerah tersebut dan tidak tergantung pada banyaknya hasil percobaan yang terjadi diluar waktu atau daerah tersebut.
- c. Probabilitas lebih dari satu hasil percobaan yang terjadi dalam interval waktu yang singkat atau dalam daerah kecil yang dapat diabaikan (Feng, S. 2008).

Karakteristik poisson:

- a. Total seluruh kejadian (percobaan) yang sangat besar (lebih dari 50)
- b. Probabilitas hasil kejadian yang sangat kecil ($0,1 = 10$ persen atau lebih kecil)

Rumus Distribusi Poisson:

$$P(x) = (\mu^x \times e^{-\mu}) / x !$$

Dimana: $P(x)$ = probabilitas peristiwa x

μ = rata-rata

x = jumlah sukses

e = bilangan alam = 2,7182

Rata-rata distribusi poisson

$$\mu = n \times p$$

c. Penerapan Poisson:

- d. Proses Poisson merupakan proses mencacah yang menghasilkan bilangan N pada selang waktu tertentu. Bilangan N yang menyatakan banyaknya kejadian dalam suatu proses Poisson disebut peubah acak Poisson dan distribusi peluangnya disebut distribusi Poisson. Di dalam teori antrian, distribusi Poisson digunakan untuk menjabarkan jumlah kejadian pada selang waktu tertentu. Kejadian-kejadian tersebut dapat berupa kedatangan atau kepergian pelanggan.
- e. Asumsi yang diperlukan dalam pembahasan proses Poisson yaitu:

Kejadian terjadi secara acak Acak berarti bahwa:

Semua kejadian pada suatu interval waktu yang sangat pendek (Δt) mempunyai probabilitas yang sama. Apabila sebanyak n pelanggan berada dalam sistem antrian, maka probabilitas sebuah kedatangan terjadi antara waktu t dan $t + \Delta t$ adalah $t \ln D$ dan dinyatakan dengan $P \{ \text{sebuah kedatangan terjadi antara } t \text{ dan } t + \Delta t \} = t \ln D$, " $n^3 0$ " sehingga probabilitas nol kedatangan antara waktu t dan $t + \Delta t$ adalah $t 1 - \ln D$ dan dinyatakan dengan $P \{ \text{nol kedatangan terjadi antara } t \text{ dan } t + \Delta t \} = t 1 - \ln D$, " $n^3 0$ ". Sedangkan probabilitas sebuah kepergian terjadi antara waktu t dan $t + \Delta t$ adalah $t m n D$ dan dinyatakan dengan.

Probabilitas nol kepergian terjadi antara waktu t dan $t + Dt$ adalah $1 - mnD$ dan dinyatakan dengan $P\{ \text{nol kepergian terjadi antara } t \text{ dan } t + Dt \} = 1 - mnD, "n^3 0$.

Keterangan:

n : banyaknya pelanggan dalam sistem antrian

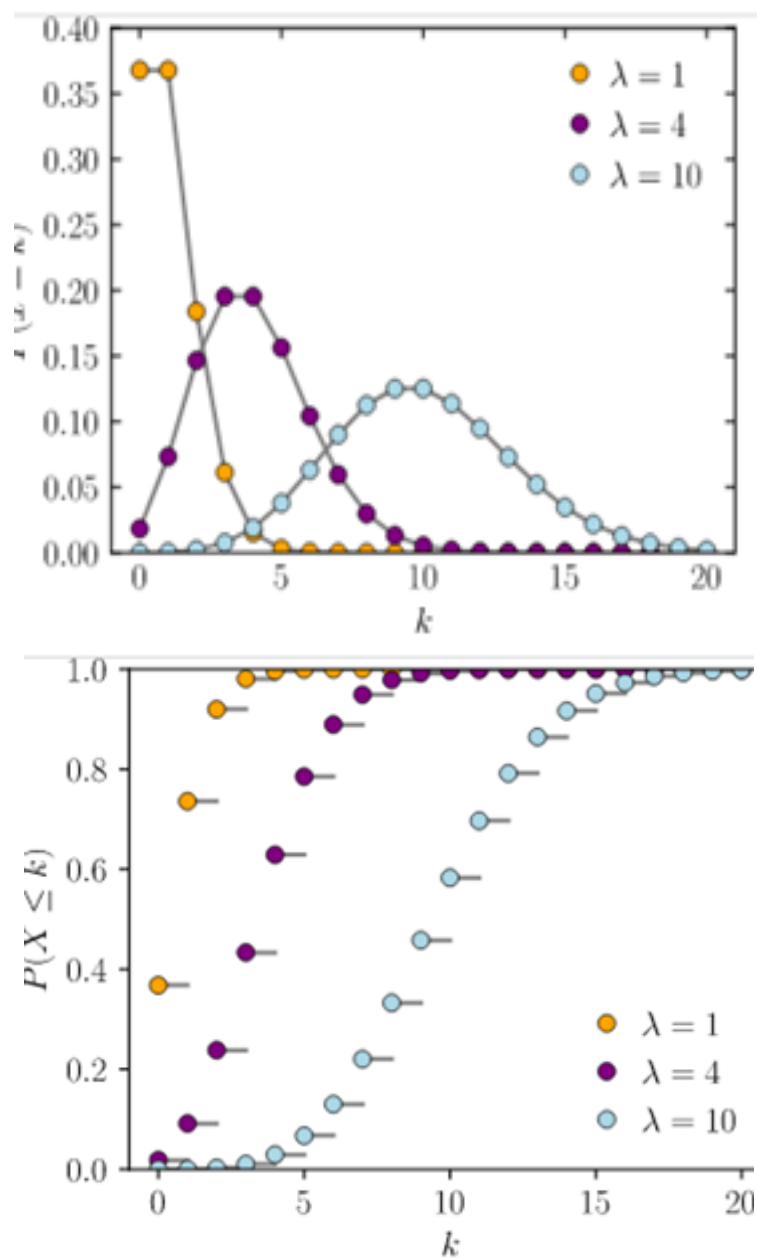
λ : laju kedatangan per satuan waktu jika sebanyak n pelanggan berada dalam sistem antrian.

μ : laju kepergian per satuan waktu jika sebanyak n pelanggan berada dalam sistem antrian.

P_n : probabilitas terdapat satu kedatangan pelanggan pada selang waktu Dt apabila terdapat n pelanggan pada sistem antrian.

$P_n D$: probabilitas terdapat satu kepergian pelanggan pada selang waktu Dt apabila terdapat n pelanggan pada sistem antrian.

Probabilitas terjadinya lebih dari satu kejadian pada selang waktu yang sangat pendek adalah sangat kecil sehingga dapat diabaikan. Persamaan tersebut dinyatakan dengan $P\{ \text{lebih dari satu kejadian antara } t \text{ dan } t + Dt \} = o(Dt)$.



Gambar 3.3. Distribusi Poisson

3.3 Deskriptif Data

Deskriptif data adalah metode penyajian data yang memberikan gambaran ringkas dan informatif mengenai data yang dikumpulkan, tanpa tujuan untuk membuat generalisasi atau inferensi ke populasi yang lebih luas. Ini mencakup berbagai cara untuk meringkas, mengorganisir, dan memvisualisasikan data agar mudah dipahami.

Sugiyono, (2022) mengartikan deskriptif data sebagai metode penelitian yang bertujuan untuk menggambarkan atau mendeskripsikan keadaan atau fenomena yang sedang diteliti, tanpa membuat perbandingan atau mencari hubungan dengan variabel lain. Deskriptif data tidak bertujuan untuk menguji hipotesis atau membuat generalisasi

Secara keseluruhan, deskriptif data membantu kita untuk memahami data dengan baik, mengidentifikasi pola dan tren dalam data, membuat keputusan berdasarkan informasi yang disajikan, menyampaikan informasi secara jelas dan ringkas. Dengan demikian, deskriptif data merupakan alat yang penting dalam analisis data, membantu kita untuk mendapatkan wawasan yang berharga dari data yang dikumpulkan

Tujuan Statistik Deskriptif:

1. Menyajikan data secara ringkas
2. Mempermudah pemahaman terhadap data
3. Menemukan pola atau tren awal

Jenis Data dalam Statistik Deskriptif

Data dalam statistik deskriptif dapat dikategorikan berdasarkan:

1. Skala Pengukuran: nominal, ordinal, interval, dan rasio
2. Sifat Data: kualitatif (kategori) dan kuantitatif (angka)
3. Ukuran Pemusatan (*Measures of Central Tendency*)

Ukuran yang menunjukkan titik pusat atau nilai tengah dari data.

Ukuran	Keterangan
Mean (Rata-rata)	Jumlah seluruh nilai dibagi banyaknya data
Median	Nilai tengah setelah data diurutkan
Modus	Nilai yang paling sering muncul

Ukuran Penyebaran (*Measures of Dispersion*)

Menunjukkan seberapa jauh data tersebar dari nilai pusatnya.

Ukuran	Keterangan
Range (Jangkauan)	Nilai maksimum - nilai minimum
Simpangan Rata-rata	Rata-rata dari selisih absolut tiap data terhadap mean
Variance (Ragam)	Rata-rata kuadrat dari selisih tiap data terhadap mean
Standard Deviation (Simpangan Baku)	Akar kuadrat dari variance

Ukuran Posisi (*Measures of Position*)

Menunjukkan posisi relatif suatu nilai dalam kumpulan data.

1. Kuartil: membagi data menjadi empat bagian yang sama
2. Desil: membagi data menjadi sepuluh bagian
3. Persentil: membagi data menjadi seratus bagian

Penyajian Data

Data deskriptif dapat disajikan dalam bentuk:

1. Tabel Frekuensi
2. Diagram Batang (Bar Chart)
3. Diagram Lingkaran (Pie Chart)
4. Histogram
5. Ogive (Kurva kumulatif)

Contoh Kasus Sederhana

Misal: Nilai ujian matematika 7 siswa: 70, 85, 90, 75, 80, 95, 100

1. Mean = $(70+85+90+75+80+95+100)/7 = 85$
2. Median = 85
3. Modus = Tidak ada (semua nilai muncul satu kali)
4. Range = $100 - 70 = 30$
5. Standard Deviation = Perlu perhitungan lebih lanjut, tapi menunjukkan variasi nilai siswa

Kegunaan Statistik Deskriptif

Statistik deskriptif digunakan dalam berbagai bidang seperti:

1. Pendidikan: untuk melihat nilai rata-rata siswa

2. Bisnis: menganalisis penjualan
3. Kesehatan: melihat data rata-rata pasien
4. Sosial: survei kepuasan masyarakat

Keterbatasan Statistik Deskriptif

1. Tidak dapat digunakan untuk menarik kesimpulan umum (tidak inferensial)
2. Hanya berlaku pada data yang dianalisis saja

DAFTAR PUSTAKA

- Donald E. Knuth (1969). *Seminumerical Algorithms*. The Art of Computer Programming, Volume 2. Addison Wesley.
- Feng, S. (2008). *The Poisson-Dirichlet Distribution and Related Topics: Models and Asymptotic*. Bandung: Ertama .
- Montgomery, D. C., & Runger, G. C. (2014). *Applied Statistics and Probability for Engineers* (6th Ed.). Wiley.
- Probability. (2009). *James*. New York: NY Publisher.
- Subagyo. (2010). *Metode Statistika untuk Kimia : Analisis Data dengan Excel*. Jakarta: Publisher corp.
- Sudjana, R. (2005). *Metode Statistika*. Bandung: Tarsito.
- Sugiyono. (2013). *Metode Penelitian Pendidikan (Pendekatan Kuantitatif, Kualitatif dan R&D)*. Bandung: Alfabeta.
- Sugiyono. (2022). *Metode Penelitian Kuantitatif, Kualitatif, dan R&D*. Bandung: Alfabeta.
- Sugiyono. (2019). *Statistik Untuk Penelitian*. Bandung: Alfabeta.
- Trihendradi, C. (2010). *Statistika* Trihendradi, C. (2010). *Statistika dengan SPSS*. Jakarta: Elex Media Komputindo.
- Walpole, R. E., Myers, R. H., Myers, S. L., & Ye, K. (2012). *Probability and Statistics for Engineers and Scientists* (9th Edition). Pearson.

BAB 4

PROBABILITAS DALAM BIOSTATISTIK

Oleh Wardiah Hamzah

4.1 Pendahuluan

Data memiliki fungsi penting dalam menjelaskan pola penyakit, mengevaluasi kemanjuran pengobatan, dan menginformasikan kebijakan kesehatan dalam domain kebidanan dan kedokteran(Machluf *et al.*, 2017). Biostatistik membekali peneliti dan profesional kesehatan dengan metodologi yang diperlukan untuk memanipulasi data, meneliti tren, dan merumuskan kesimpulan berbasis bukti. Biostatistik sebagai instrumen yang diperlukan dalam meningkatkan hasil kesehatan masyarakat dan memperkuat upaya investigasi di sektor medis dan farmasi(Pagano, 1993; Rosner, 2010).

Dalam bidang kesehatan, analisis statistik sangat diperlukan untuk memahami distribusi penyakit, faktor risiko, serta hubungan antara variabel yang memengaruhi kesehatan manusia. Misalnya, dalam studi epidemiologi, biostatistik digunakan untuk mengukur prevalensi dan insidensi suatu penyakit, membantu dalam prediksi wabah, serta mengidentifikasi kelompok yang berisiko tinggi terhadap penyakit tertentu(Newman, 2001). Tanpa pendekatan statistik yang tepat, interpretasi data medis dapat menjadi bias atau tidak akurat, yang berpotensi mengarah pada kesimpulan yang salah dan keputusan yang tidak efektif(Merrill, 2013).

Salah satu konsep utama dalam biostatistik adalah probabilitas(Rosner, 2010). Probabilitas membantu dalam memahami ketidakpastian dalam penelitian medis dan memberikan dasar untuk membuat inferensi statistik(Mikhail Nikulin, Daniel Commenges, 2006). Dalam dunia kesehatan, banyak fenomena yang tidak bisa diprediksi dengan kepastian mutlak, tetapi bisa diperkirakan menggunakan distribusi probabilitas tertentu. Misalnya, dalam uji klinis obat baru, probabilitas digunakan untuk

menilai apakah obat tersebut memiliki efek yang signifikan dibandingkan dengan plasebo.

Selain itu, konsep probabilitas juga diterapkan dalam pengujian diagnostik, seperti menentukan sensitivitas dan spesifisitas suatu tes medis (Weatherall, 2018). Probabilitas juga digunakan dalam perhitungan risiko suatu penyakit berdasarkan faktor genetik dan lingkungan (Sun *et al.*, 2024). Dengan memahami probabilitas dalam konteks biostatistik kesehatan, para profesional medis dan peneliti dapat mengambil keputusan yang lebih akurat dan berbasis data untuk meningkatkan layanan kesehatan dan keselamatan pasien.

Tulisan ini bertujuan untuk menjelaskan konsep dasar probabilitas dalam biostatistik kesehatan serta bagaimana konsep ini diterapkan dalam berbagai aspek penelitian medis. Dengan pemahaman yang baik mengenai probabilitas, tenaga kesehatan, peneliti, dan mahasiswa dapat lebih efektif dalam menganalisis data dan membuat prediksi terkait penyakit serta efektivitas intervensi medis. Selain itu, tulisan ini juga bertujuan untuk memberikan gambaran tentang bagaimana probabilitas digunakan dalam uji klinis, analisis epidemiologi, dan pengujian diagnostik.

4.2 Konsep Dasar Probabilitas dalam Biostatistik

Probabilitas adalah cabang dari matematika yang berfungsi untuk mengukur kemungkinan terjadinya suatu kejadian dalam suatu ruang sampel tertentu (Mikhail Nikulin, Daniel Commenges, 2006). Dalam biostatistik kesehatan, probabilitas digunakan untuk memahami dan menginterpretasikan pola serta ketidakpastian dalam data medis dan epidemiologi (Levy and Lemeshow, 2008). Misalnya, dalam studi klinis, probabilitas digunakan untuk menghitung kemungkinan seorang pasien merespons suatu pengobatan tertentu, atau dalam epidemiologi, digunakan untuk memperkirakan penyebaran suatu penyakit dalam populasi. Konsep probabilitas memungkinkan para peneliti untuk mengembangkan model prediktif, mengukur risiko, dan mengambil keputusan berbasis data untuk meningkatkan kesehatan masyarakat.

4.2.1 Defenisi Probabilitas

Probabilitas dapat didefinisikan sebagai ukuran kemungkinan terjadinya suatu kejadian dalam suatu ruang sampel (Levy and Lemeshow, 2008). Ruang sampel merupakan himpunan dari semua kemungkinan hasil yang dapat terjadi dalam suatu percobaan atau observasi. Sebagai contoh, dalam konteks kesehatan, jika dilakukan pengujian terhadap suatu kelompok pasien untuk menentukan apakah mereka memiliki suatu penyakit atau tidak, maka ruang sampelnya terdiri dari dua kemungkinan, yaitu "positif penyakit" dan "negatif penyakit".

Dalam statistik kesehatan, probabilitas sering digunakan untuk memperkirakan peluang seseorang terkena penyakit tertentu berdasarkan faktor risiko yang dimilikinya. Misalnya, jika diketahui bahwa probabilitas seseorang terkena diabetes meningkat jika memiliki riwayat keluarga dengan penyakit tersebut, maka probabilitas ini dapat dihitung dengan menggunakan pendekatan statistik. Dengan demikian, probabilitas memberikan alat yang kuat bagi para ilmuwan untuk membuat prediksi dan mengambil keputusan dalam praktik kesehatan dan kedokteran.

Selain itu, dalam uji klinis, probabilitas digunakan untuk menentukan tingkat efektivitas suatu obat baru dibandingkan dengan metode pengobatan yang sudah ada. Probabilitas juga menjadi dasar dalam penghitungan berbagai ukuran epidemiologi, seperti prevalensi, insidensi, dan angka kematian. Dengan memahami probabilitas, peneliti dapat menyusun model statistik yang dapat digunakan untuk menganalisis berbagai fenomena kesehatan (Cook and Moodie, 2024; Pérez-Guerrero *et al.*, 2024).

4.2.2 Jenis-Jenis Probabilitas

Berdasarkan jenisnya, probabilitas dapat dikelompokkan atas probabilitas klasik, empiris dan subjektif

1. Probabilitas Klasik

Probabilitas klasik adalah konsep probabilitas yang diperkenalkan berdasarkan prinsip dasar matematis, di mana semua kemungkinan hasil dianggap memiliki peluang yang sama. Rumus dasar probabilitas klasik dinyatakan sebagai berikut:

$$P(A) = \frac{\text{jumlah kejadian yang diinginkan}}{\text{jumlah total kemungkinan kejadian}}$$

Dalam konteks biostatistik kesehatan, probabilitas klasik sering digunakan dalam eksperimen yang memiliki hasil terbatas dan dapat dihitung secara pasti. Misalnya, dalam uji laboratorium, jika dilakukan pengujian terhadap 100 sampel darah dan diketahui bahwa 20 di antaranya mengandung virus tertentu, maka probabilitas ditemukannya virus dalam sampel secara acak adalah 20/100 atau 0,2 (20%).

Namun, dalam praktiknya, probabilitas klasik kurang fleksibel dalam situasi nyata di dunia medis, di mana banyak faktor dapat memengaruhi hasil yang diperoleh. Oleh karena itu, probabilitas klasik sering dikombinasikan dengan metode lain untuk memberikan estimasi yang lebih akurat terhadap risiko atau peluang kejadian kesehatan tertentu.

2. Probabilitas Empiris

Probabilitas empiris adalah probabilitas yang diperoleh berdasarkan hasil observasi atau eksperimen yang telah dilakukan. Probabilitas ini dihitung dengan membandingkan jumlah kejadian yang telah diamati dengan jumlah total percobaan atau observasi.

Sebagai contoh, jika sebuah penelitian dilakukan terhadap 500 pasien dengan tekanan darah tinggi dan ditemukan bahwa 150 pasien mengalami komplikasi jantung dalam 5 tahun ke depan, maka probabilitas empiris dari kejadian komplikasi jantung pada pasien dengan hipertensi adalah:

$$P(\text{Komplikasi Jantung}) = \frac{150}{500} = 0,3 \text{ atau } 30\%$$

Probabilitas empiris sangat berguna dalam epidemiologi dan studi klinis karena memberikan gambaran nyata mengenai pola penyakit dalam populasi berdasarkan data yang telah dikumpulkan. Probabilitas ini sering digunakan dalam studi kohort dan studi kasus-kontrol untuk menilai hubungan antara faktor risiko dan penyakit.

3. Probabilitas Subjektif

Probabilitas subjektif adalah probabilitas yang didasarkan pada keyakinan atau pendapat seseorang berdasarkan pengalaman atau informasi yang tersedia, bukan pada perhitungan matematis atau data empiris. Probabilitas ini sering digunakan dalam pengambilan keputusan klinis ketika data yang tersedia terbatas atau tidak dapat diukur dengan pasti.

Misalnya, seorang dokter yang berpengalaman mungkin memperkirakan bahwa pasien dengan riwayat merokok berat memiliki peluang lebih tinggi terkena kanker paru-paru dibandingkan pasien yang tidak merokok, meskipun ia tidak memiliki data statistik yang spesifik pada saat itu. Probabilitas subjektif juga sering digunakan dalam analisis risiko ketika tidak ada cukup data untuk menghitung probabilitas secara objektif.

4.2.3 Aturan Dasar Probabilitas

Aturan dasar probabilitas meliputi aturan penjumlahan dan perkalian serta probabilitas bersyarat dan independen

1. Aturan Penjumlahan dan Perkalian

Dalam probabilitas, aturan penjumlahan digunakan untuk menghitung peluang terjadinya salah satu dari beberapa kejadian, sedangkan aturan perkalian digunakan untuk menghitung peluang terjadinya kejadian secara bersamaan.

- a. **Aturan Penjumlahan:** Jika dua kejadian **A** dan **B** bersifat saling eksklusif (tidak bisa terjadi bersamaan), maka probabilitas salah satu terjadi adalah:

$$P(A \cup B) = P(A) + P(B)$$

Contohnya, jika probabilitas seseorang menderita diabetes adalah 0,1 dan probabilitas seseorang mengalami hipertensi adalah 0,2, serta kedua kondisi ini tidak saling berhubungan, maka probabilitas seseorang mengalami salah satu dari kedua kondisi ini adalah $0,1 + 0,2 = 0,3$ (30%).

- b. **Aturan Perkalian:** Jika dua kejadian **A** dan **B** bersifat independen (tidak saling memengaruhi), maka probabilitas kedua kejadian terjadi bersamaan adalah:

$$P(A \cap B) = P(A) \times P(B)$$

Sebagai contoh, jika probabilitas seorang pasien mengalami obesitas adalah 0,25 dan probabilitas pasien tersebut memiliki kadar kolesterol tinggi adalah 0,4, maka probabilitas pasien mengalami kedua kondisi tersebut adalah:

$$P(A \cap B) = 0,25 \times 0,4 = 0,10 \text{ atau } 10\%$$

2. Probabilitas Bersyarat dan Independensi

Probabilitas bersyarat adalah probabilitas suatu kejadian terjadi dengan asumsi bahwa kejadian lain telah terjadi. Secara matematis, probabilitas bersyarat dinyatakan sebagai:

$$P(A|B) = \frac{P(A \cap B)}{P(B)}$$

Sebagai contoh, dalam epidemiologi, jika diketahui bahwa probabilitas seseorang memiliki tekanan darah tinggi (B) adalah 0,3 dan probabilitas seseorang mengalami stroke (A) bersamaan dengan tekanan darah tinggi adalah 0,15, maka probabilitas seseorang mengalami stroke jika sudah memiliki tekanan darah tinggi adalah:

$$P(A|B) = \frac{P(A \cap B)}{P(B)} = \frac{0,15}{0,3} = 0,5 \text{ atau } 50\%$$

Probabilitas bersyarat sering digunakan dalam pengujian diagnostik dan penelitian epidemiologi untuk menilai hubungan antara faktor risiko dan penyakit

4.3 Distribusi Probabilitas dalam Biostatistik

Distribusi probabilitas adalah fungsi matematika yang menggambarkan kemungkinan terjadinya berbagai hasil dalam suatu percobaan atau fenomena acak (Mikhail Nikulin, Daniel Commenges, 2006). Dalam biostatistik kesehatan, distribusi probabilitas digunakan untuk memodelkan berbagai jenis data yang berkaitan dengan kesehatan dan penyakit, seperti jumlah pasien yang terinfeksi suatu penyakit, tingkat keberhasilan suatu terapi, atau waktu bertahan hidup pasien setelah mendapatkan perawatan tertentu. Dengan memahami distribusi probabilitas, peneliti dan tenaga kesehatan dapat menginterpretasikan data dengan lebih baik serta membuat keputusan berbasis bukti untuk meningkatkan kualitas layanan kesehatan.

Distribusi probabilitas dibagi menjadi dua kategori utama: distribusi probabilitas diskrit dan distribusi probabilitas kontinu (Rosner, 2010). Distribusi diskrit digunakan untuk variabel yang memiliki nilai terbatas atau dapat dihitung, seperti jumlah penderita suatu penyakit dalam kelompok tertentu. Sementara itu, distribusi kontinu digunakan untuk variabel yang dapat memiliki nilai dalam rentang tertentu, seperti berat badan pasien atau tekanan darah. Kedua jenis distribusi ini memiliki peran penting dalam analisis biostatistik dan diterapkan dalam berbagai metode penelitian kesehatan.

4.3.1 Distribusi Probabilitas Diskrit

Distribusi probabilitas diskrit digunakan untuk menggambarkan probabilitas kejadian dari variabel yang hanya memiliki nilai tertentu, biasanya dalam bentuk bilangan bulat. Dalam bidang kesehatan, distribusi ini sering digunakan untuk menghitung jumlah kasus penyakit dalam suatu populasi, jumlah kematian akibat suatu penyakit, atau jumlah pasien yang merespons suatu pengobatan. Dua distribusi diskrit yang paling umum digunakan dalam biostatistik kesehatan adalah distribusi binomial dan distribusi Poisson.

1. Distribusi Binomial

Distribusi binomial adalah jenis distribusi probabilitas diskrit yang menggambarkan jumlah keberhasilan dalam sejumlah percobaan independen yang memiliki dua kemungkinan hasil, yaitu "sukses" atau "gagal". Distribusi ini berguna dalam banyak aplikasi kesehatan, seperti menghitung probabilitas keberhasilan vaksinasi, efektivitas terapi, atau kemungkinan pasien sembuh setelah menjalani perawatan tertentu.

Secara matematis, probabilitas suatu kejadian dalam distribusi binomial diberikan oleh rumus:

$$P(X = k) = \binom{n}{k} p^k (1 - p)^{n-k}$$

dimana:

- a. $P(X = k)$ adalah probabilitas mendapatkan k keberhasilan dalam n percobaan

- b. p adalah keberhasilan dalam satu percobaan
- c. n adalah jumlah percobaan
- d. $\binom{n}{k}$ adalah kombinasi untuk memilih k keberhasilan dari n percobaan

Sebagai contoh, jika dalam uji klinis sebuah obat terdapat peluang keberhasilan penyembuhan sebesar 70% dan obat tersebut diberikan kepada 10 pasien, distribusi binomial dapat digunakan untuk menghitung kemungkinan berapa banyak pasien yang akan sembuh.

2. Distribusi Poisson

Distribusi Poisson digunakan untuk menghitung probabilitas jumlah kejadian dalam suatu interval waktu atau ruang tertentu ketika kejadian tersebut terjadi secara acak dengan tingkat rata-rata tertentu. Distribusi ini sering digunakan dalam epidemiologi untuk memodelkan jumlah kasus penyakit yang terjadi dalam jangka waktu tertentu atau jumlah pasien yang masuk ke unit gawat darurat dalam satu hari.

Distribusi Poisson dinyatakan dengan rumus:

$$P(X = k) = \frac{e^{-\lambda} \lambda^k}{k!}$$

dimana:

- a. $P(X = k)$ adalah probabilitas terjadi k kejadian dalam periode tertentu
- b. λ adalah rata-rata jumlah kejadian dalam periode tersebut
- c. e adalah jumlah bilangan Euler ($\approx 2,718$)

Sebagai contoh, jika rumah sakit mencatat bahwa rata-rata terdapat 5 pasien yang masuk ke IGD akibat stroke setiap malam, maka distribusi Poisson dapat digunakan untuk menghitung probabilitas bahwa pada suatu malam tertentu akan ada 7 pasien yang masuk akibat stroke.

4.3.2 Distribusi Probabilitas Kontinu

Distribusi probabilitas kontinu digunakan untuk menggambarkan probabilitas dari variabel yang dapat memiliki nilai

dalam rentang tertentu. Dalam bidang kesehatan, distribusi ini digunakan untuk menganalisis data seperti tinggi badan, tekanan darah, kadar gula darah, dan waktu bertahan hidup pasien setelah diagnosis suatu penyakit. Dua distribusi probabilitas kontinu yang paling umum dalam biostatistik kesehatan adalah distribusi normal dan distribusi eksponensial.

1. Distribusi Normal (*Gaussian*)

Distribusi normal, atau yang dikenal sebagai distribusi *Gaussian*, adalah distribusi probabilitas yang paling umum digunakan dalam analisis data kesehatan. Distribusi ini memiliki bentuk lonceng simetris dengan rata-rata (μ) sebagai titik tengah dan simpangan baku (σ) yang menentukan lebar kurva. Distribusi normal digunakan dalam berbagai aplikasi kesehatan, seperti mengukur tekanan darah rata-rata dalam populasi atau menguji apakah suatu variabel memiliki distribusi yang seimbang

Distribusi normal memiliki rumus sebagai berikut:

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(z-\mu)^2}{2\sigma^2}}$$

Sebagai contoh, jika tekanan darah sistolik rata-rata dalam populasi adalah 120 mmHg dengan standar deviasi 10 mmHg, maka distribusi normal dapat digunakan untuk menentukan probabilitas seseorang memiliki tekanan darah dalam rentang tertentu, misalnya antara 110 mmHg hingga 130 mmHg.

Distribusi normal juga menjadi dasar dalam banyak uji statistik, seperti uji t, ANOVA, dan regresi linier dalam penelitian medis. Oleh karena itu, pemahaman tentang distribusi normal sangat penting bagi para peneliti dan praktisi kesehatan dalam menganalisis data klinis dan epidemiologi.

2. Distribusi Eksponensial

Distribusi eksponensial digunakan untuk memodelkan waktu antara kejadian-kejadian yang terjadi secara acak, seperti waktu bertahan hidup pasien setelah diagnosis suatu penyakit atau waktu tunggu pasien sebelum mendapatkan perawatan di rumah sakit. Distribusi ini sering digunakan dalam analisis

kelangsungan hidup (*survival analysis*) dan dalam studi farmakokinetik.

Distribusi eksponensial memiliki fungsi probabilitas sebagai berikut:

$$f(x) = \lambda e^{-\lambda x}, \quad x \geq 0$$

dimana:

λ adalah rata-rata tingkat kejadian per unit waktu

Sebagai contoh, jika waktu rata-rata seorang pasien bertahan hidup setelah diagnosis HIV AIDS adalah 5 tahun, maka distribusi eksponensial dapat digunakan untuk menghitung probabilitas bahwa seorang pasien akan bertahan hidup lebih dari 3 tahun setelah diagnosis.

4.4 Penerapan Probabilitas dalam Biostatistik

Probabilitas memiliki peran penting dalam biostatistik kesehatan karena memungkinkan analisis data yang lebih akurat dan pengambilan keputusan berbasis bukti dalam berbagai aspek medis. Dengan memahami probabilitas, tenaga medis dan peneliti dapat memperkirakan risiko penyakit, mengukur efektivitas suatu pengobatan, serta menginterpretasikan hasil dari berbagai uji diagnostik. Dalam dunia medis, banyak kejadian tidak dapat diprediksi secara pasti, tetapi probabilitas memberikan alat untuk memahami ketidakpastian tersebut dengan lebih baik (Newman, 2001; Kestenbaum, 2009).

Beberapa penerapan utama probabilitas dalam biostatistik kesehatan meliputi analisis risiko penyakit, uji klinis, pengujian diagnostik, serta prediksi kejadian epidemiologi. Dengan pendekatan probabilistik, para profesional kesehatan dapat mengoptimalkan pengobatan, merancang kebijakan kesehatan yang lebih efektif, serta meningkatkan akurasi diagnosis dan pengobatan pasien. Berikut ini beberapa contoh utama penerapan probabilitas dalam biostatistik kesehatan.

4.4.1 Analisis Risiko Penyakit

Analisis risiko penyakit adalah penerapan probabilitas yang digunakan untuk menilai kemungkinan seseorang mengalami penyakit tertentu berdasarkan faktor risiko yang dimilikinya. Risiko

suatu penyakit dapat dihitung dengan metode probabilitas, baik secara absolut maupun relatif, untuk memahami hubungan antara variabel kesehatan dan kemungkinan kejadian penyakit

1. Risiko Absolut dan Risiko Relatif

Risiko absolut adalah kemungkinan seseorang terkena suatu penyakit dalam periode tertentu, sedangkan risiko relatif membandingkan kemungkinan suatu kelompok terkena penyakit dibandingkan dengan kelompok lain. Risiko relatif sering digunakan dalam studi epidemiologi untuk mengukur dampak faktor risiko terhadap suatu penyakit.

Sebagai contoh, dalam sebuah studi ditemukan bahwa 10% dari sopir angkutan umum mengalami kanker paru-paru akibat paparan polusi udara, sedangkan hanya 2% dari non sopir angkutan umum yang mengalami kondisi yang sama. Risiko relatif untuk kelompok perokok dibandingkan dengan non-perokok dapat dihitung sebagai berikut:

$$RR = \frac{10\%}{2\%} = 5$$

Ini berarti bahwa sopir angkutan umum memiliki peluang lima kali lebih besar untuk terkena kanker paru-paru dibandingkan dengan non sopir angkutan umum. Informasi ini sangat penting dalam pengambilan keputusan kesehatan masyarakat, seperti kampanye anti-rokok dan kebijakan pengendalian tembakau.

2. Probabilitas dalam Model Prediksi Penyakit

Probabilitas juga digunakan dalam model prediksi penyakit, seperti skor risiko stroke atau model prediksi diabetes. Model ini menggunakan data statistik dari berbagai faktor risiko, seperti tekanan darah, kadar kolesterol, dan indeks massa tubuh (IMT), untuk memprediksi kemungkinan seseorang mengalami penyakit tertentu dalam beberapa tahun ke depan.

Misalnya, dalam model **Framingham Risk Score**, probabilitas seseorang mengalami stoke dalam 10 tahun ke depan dihitung berdasarkan kombinasi faktor risiko individu (Dehghan *et al.*, 2023). Dengan memahami probabilitas ini, dokter dapat memberikan rekomendasi yang lebih akurat

untuk mencegah penyakit, seperti perubahan gaya hidup atau penggunaan terapi medis tertentu.

4.4.2 Uji Klinis dan Efektivitas Pengobatan

Uji klinis adalah metode ilmiah yang digunakan untuk mengevaluasi efektivitas dan keamanan suatu intervensi medis, seperti obat baru, terapi, atau vaksin. Probabilitas digunakan dalam berbagai aspek uji klinis, mulai dari perencanaan penelitian hingga analisis hasilnya.

1. Pengacakan dan Probabilitas dalam Uji Klinis

Dalam uji klinis, metode randomisasi digunakan untuk memastikan bahwa kelompok eksperimen dan kelompok kontrol memiliki karakteristik yang seimbang. Pengacakan ini mengurangi bias dan memungkinkan kesimpulan yang lebih valid tentang efek suatu intervensi.

Sebagai contoh, dalam sebuah uji klinis untuk mengevaluasi efektivitas vaksin, peserta dibagi secara acak ke dalam dua kelompok: kelompok yang menerima vaksin dan kelompok yang menerima plasebo. Probabilitas setiap individu untuk mendapatkan vaksin adalah 50%. Dengan cara ini, hasil yang diperoleh lebih dapat diandalkan karena tidak terpengaruh oleh faktor lain yang tidak terkontrol.

2. Analisis Statistik dan Probabilitas dalam Uji Klinis

Dalam analisis hasil uji klinis, probabilitas digunakan untuk menentukan apakah perbedaan antara kelompok eksperimen dan kontrol signifikan atau hanya terjadi karena kebetulan. Salah satu alat statistik yang digunakan adalah nilai *p-value*, yang menunjukkan probabilitas bahwa hasil yang diamati terjadi secara kebetulan jika tidak ada efek sebenarnya dari intervensi.

Sebagai contoh, jika sebuah penelitian menunjukkan bahwa pasien yang menerima obat baru memiliki tingkat kesembuhan 80%, sedangkan pasien yang menerima plasebo hanya memiliki tingkat kesembuhan 60%, analisis probabilitas digunakan untuk menentukan apakah perbedaan ini cukup signifikan untuk menyatakan bahwa obat tersebut efektif. Jika *p*-

value kurang dari 0,05, maka hasilnya dianggap signifikan dan bukan karena kebetulan.

4.4.3 Probabilitas dalam Pengujian Diagnostik

Pengujian diagnostik merupakan salah satu aspek penting dalam kedokteran yang menggunakan probabilitas untuk menentukan keakuratan suatu tes dalam mendeteksi penyakit. Dua ukuran probabilitas utama dalam pengujian diagnostik adalah sensitivitas dan spesifisitas.

1. Sensitivitas dan Spesifisitas

Sensitivitas adalah probabilitas bahwa tes akan memberikan hasil positif jika seseorang benar-benar memiliki penyakit. Spesifisitas adalah probabilitas bahwa tes akan memberikan hasil negatif jika seseorang tidak memiliki penyakit.

Sebagai contoh, jika suatu tes COVID-19 memiliki sensitivitas 95%, ini berarti bahwa dari 100 orang yang benar-benar terinfeksi, 95 akan mendapatkan hasil positif, sementara 5 lainnya akan mendapatkan hasil negatif palsu. Jika spesifisitasnya 90%, maka dari 100 orang yang tidak terinfeksi, 90 akan mendapatkan hasil negatif, tetapi 10 akan mendapatkan hasil positif palsu.

Probabilitas ini penting dalam mengevaluasi efektivitas tes diagnostik, terutama dalam menentukan kapan suatu tes dapat digunakan dalam skala besar.

2. Nilai Prediktif Positif dan Negatif

Selain sensitivitas dan spesifisitas, probabilitas dalam pengujian diagnostik juga melibatkan nilai prediktif positif (PPV) dan nilai prediktif negatif (NPV). PPV adalah probabilitas bahwa seseorang benar-benar memiliki penyakit jika hasil tesnya positif, sedangkan NPV adalah probabilitas bahwa seseorang benar-benar tidak memiliki penyakit jika hasil tesnya negatif.

Sebagai contoh, dalam suatu populasi dengan prevalensi rendah suatu penyakit, bahkan tes dengan sensitivitas dan spesifisitas tinggi mungkin memiliki nilai prediktif positif yang rendah. Hal ini karena ada banyak kasus positif palsu dibandingkan dengan jumlah kasus sebenarnya, sehingga

probabilitas bahwa seseorang yang mendapatkan hasil positif benar-benar memiliki penyakit menjadi lebih kecil.

4.4.3 Prediksi Kejadian Epidemiologi

Probabilitas juga digunakan dalam epidemiologi untuk memprediksi penyebaran penyakit dalam suatu populasi. Model matematika probabilistik, seperti model SIR (*Susceptible-Infected-Recovered*), digunakan untuk memahami bagaimana suatu penyakit menyebar dan untuk merancang strategi intervensi yang efektif.

Sebagai contoh, dalam model probabilistik penyebaran COVID-19, digunakan nilai reproduksi dasar (R_0), yang menunjukkan rata-rata jumlah individu yang dapat tertular oleh satu orang yang terinfeksi. Jika $R_0 > 1$, maka penyakit akan terus menyebar, tetapi jika $R_0 < 1$, maka wabah akan mereda.

Dengan menggunakan probabilitas dalam epidemiologi, kebijakan kesehatan masyarakat dapat dirancang dengan lebih efektif, seperti menentukan kapan harus menerapkan *lockdown* atau kapan harus meningkatkan kampanye vaksinasi.

4.5 Konsep Studi Kasus: Probabilitas Biostatistik

Penerapan probabilitas dalam biostatistik kesehatan tidak hanya bersifat teoritis, tetapi juga digunakan dalam berbagai studi kasus nyata untuk membantu pengambilan keputusan berbasis data. Studi kasus ini mencakup berbagai aspek, mulai dari prediksi risiko penyakit, analisis efektivitas pengobatan, hingga pemodelan penyebaran penyakit dalam populasi (Tsokani *et al.*, 2022; Rella *et al.*, 2023; Sun *et al.*, 2024).

Dengan menggunakan probabilitas, para peneliti dapat menilai pola kejadian suatu penyakit, mengevaluasi keberhasilan intervensi medis, serta mengembangkan kebijakan kesehatan yang lebih efektif. Berikut adalah beberapa studi kasus dan contoh aplikasi probabilitas dalam biostatistik kesehatan yang menunjukkan pentingnya konsep ini dalam dunia medis dan epidemiologi

4.5.1 Prediksi Risiko Penyakit Jantung

Penyakit jantung merupakan salah satu penyebab utama kematian di seluruh dunia. Oleh karena itu, pemodelan risiko

penyakit ini menggunakan probabilitas sangat penting untuk mendeteksi individu dengan risiko tinggi(Sun *et al.*, 2024) dan menerapkan langkah-langkah pencegahan lebih awal.

1. Model *Framingham Risk Score*

Salah satu metode probabilistik yang banyak digunakan dalam analisis risiko penyakit jantung adalah ***Framingham Risk Score***. Model ini menggunakan variabel-variabel seperti usia, jenis kelamin, tekanan darah, kadar kolesterol, status merokok, dan riwayat diabetes untuk memperkirakan kemungkinan seseorang mengalami penyakit jantung dalam 10 tahun ke depan(Dehghan *et al.*, 2023).

Sebagai contoh, seorang pria berusia 55 tahun dengan tekanan darah tinggi, kadar kolesterol LDL yang tinggi, dan riwayat merokok memiliki probabilitas lebih besar terkena penyakit jantung dibandingkan dengan seseorang yang tidak memiliki faktor risiko tersebut. Dengan menggunakan metode probabilistik, dokter dapat memberikan rekomendasi pencegahan yang lebih spesifik, seperti perubahan gaya hidup atau penggunaan obat statin untuk menurunkan kadar kolesterol.

2. Pengambilan Keputusan Klinis

Dalam praktik klinis, probabilitas juga digunakan untuk membantu dokter membuat keputusan mengenai intervensi medis. Sebagai contoh, jika model prediksi menunjukkan bahwa seseorang memiliki peluang lebih dari 20% mengalami serangan jantung dalam 10 tahun ke depan, dokter dapat merekomendasikan terapi pengobatan lebih agresif, seperti penggunaan aspirin atau obat penurun tekanan darah.

Dengan menggunakan pendekatan berbasis probabilitas, pengambilan keputusan medis dapat lebih berbasis data, mengurangi ketidakpastian, serta meningkatkan efektivitas perawatan bagi pasien berisiko tinggi.

4.5.2 Efektivitas Vaksin

Vaksinasi merupakan salah satu strategi kesehatan masyarakat yang paling efektif dalam mencegah penyakit menular(Rella *et al.*,

2023). Untuk mengukur efektivitas suatu vaksin, probabilitas digunakan dalam analisis uji klinis dan studi epidemiologi.

1. Uji Klinis Efektivitas Vaksin COVID-19

Sebelum vaksin COVID-19 diberikan kepada masyarakat umum, uji klinis dilakukan untuk menentukan efektivitasnya. Dalam uji klinis, sekelompok relawan secara acak dibagi menjadi dua kelompok: satu kelompok menerima vaksin, dan kelompok lainnya menerima plasebo. Setelah periode tertentu, jumlah individu yang terinfeksi di setiap kelompok dibandingkan untuk menghitung efektivitas vaksin (Raji *et al.*, 2024).

Sebagai contoh, dalam uji klinis vaksin *Pfizer-BioNTech*, ditemukan bahwa dari 100 individu yang terinfeksi COVID-19, hanya 5 berasal dari kelompok yang divaksinasi, sementara 95 berasal dari kelompok plasebo. Efektivitas vaksin dihitung menggunakan rumus probabilitas berikut:

$$E = \left(\frac{R_u - R_v}{R_u} \right) \times 100\%$$

dimana:

- R_u adalah risiko infeksi pada kelompok tanpa vaksin
- R_v adalah risiko infeksi pada kelompok yang divaksinasi

Jika risiko infeksi pada kelompok tanpa vaksin adalah 10%, dan risiko pada kelompok vaksin adalah 1%, maka efektivitasnya adalah:

$$E = \left(\frac{10\% - 1\%}{10\%} \right) \times 100\% = 90\%$$

Hasil ini menunjukkan bahwa vaksin memiliki probabilitas tinggi dalam mencegah infeksi, dan data probabilitistik ini digunakan untuk mendukung keputusan kebijakan vaksinasi global.

2. Analisis *Herd Immunity*

Selain efektivitas individu, probabilitas juga digunakan dalam analisis *herd immunity* atau kekebalan kelompok (Gamboa and Lopez-Herrero, 2022). Jika proporsi individu yang divaksinasi dalam suatu populasi mencapai ambang batas tertentu, maka penyebaran penyakit akan berkurang secara signifikan.

Misalnya, dalam kasus campak, ambang herd immunity diperkirakan sekitar 95%. Artinya, jika minimal 95% populasi divaksinasi, maka probabilitas penyebaran penyakit di komunitas akan sangat kecil, bahkan bagi mereka yang tidak bisa divaksinasi karena alasan medis.

4.5.3 Pengujian Diagnostik dan Nilai Prediktif

Pengujian diagnostik adalah aspek penting dalam kedokteran yang menggunakan probabilitas untuk menentukan keakuratan suatu tes dalam mendeteksi penyakit (Tsokani *et al.*, 2022).

1. Tes COVID-19: Sensitivitas dan Spesifisitas

Ketika seseorang menjalani tes PCR COVID-19, ada dua kemungkinan utama dalam hasilnya: positif atau negatif. Namun, karena tidak ada tes yang 100% akurat, probabilitas digunakan untuk menginterpretasikan hasil tes dengan lebih baik.

Misalnya, jika sensitivitas tes PCR adalah 95%, ini berarti bahwa 95 dari 100 individu yang benar-benar terinfeksi akan mendapatkan hasil positif (positif benar), tetapi 5 individu akan mendapatkan hasil negatif palsu. Jika spesifisitasnya adalah 98%, maka 98 dari 100 individu yang tidak terinfeksi akan mendapatkan hasil negatif (negatif benar), tetapi 2 individu akan mendapatkan hasil positif palsu.

2. Nilai Prediktif Positif dan Nilai Prediktif Negatif

Selain sensitivitas dan spesifisitas, probabilitas juga digunakan untuk menghitung nilai prediktif positif (PPV) dan nilai prediktif negatif (NPV). PPV menunjukkan seberapa besar kemungkinan seseorang benar-benar sakit jika hasil tesnya positif, sementara NPV menunjukkan seberapa besar kemungkinan seseorang benar-benar sehat jika hasil tesnya negatif.

Sebagai contoh, dalam populasi dengan tingkat infeksi rendah, meskipun tes memiliki spesifisitas tinggi, masih mungkin terdapat banyak hasil positif palsu, sehingga nilai PPV menjadi lebih rendah. Oleh karena itu, dokter menggunakan pendekatan probabilistik untuk mempertimbangkan hasil tes bersama dengan faktor risiko lain sebelum membuat diagnosis pasti

4.6 Kesimpulan dan Saran

4.6.1 Kesimpulan

Probabilitas dalam biostatistik kesehatan memiliki peran yang sangat penting dalam memahami dan menganalisis berbagai fenomena dalam bidang kesehatan dan kedokteran. Dengan menggunakan konsep probabilitas, para peneliti dan tenaga medis dapat memperkirakan risiko penyakit, mengevaluasi efektivitas pengobatan, serta meningkatkan akurasi pengujian diagnostik. Probabilitas membantu dalam memahami ketidakpastian yang melekat dalam data medis, sehingga pengambilan keputusan dapat dilakukan secara lebih objektif dan berdasarkan bukti ilmiah.

Studi kasus yang telah dibahas menunjukkan bagaimana probabilitas digunakan dalam berbagai aspek kesehatan, seperti prediksi risiko penyakit jantung, efektivitas vaksin, dan pengujian diagnostik COVID-19. Dalam prediksi risiko penyakit, model probabilistik membantu dokter dalam mengidentifikasi individu dengan risiko tinggi, sehingga tindakan pencegahan dapat dilakukan lebih dini. Dalam uji klinis vaksin, probabilitas digunakan untuk mengukur efektivitas suatu vaksin dalam melindungi populasi dari penyakit menular. Sementara itu, dalam pengujian diagnostik, probabilitas digunakan untuk mengevaluasi sensitivitas, spesifisitas, serta nilai prediktif suatu tes medis.

Secara keseluruhan, penerapan probabilitas dalam biostatistik kesehatan memungkinkan pengambilan keputusan yang lebih akurat dan berbasis data. Dengan memahami probabilitas, tenaga medis, kesehatan dan peneliti dapat mengurangi kesalahan diagnosis, meningkatkan efektivitas pengobatan, serta merancang strategi kesehatan masyarakat yang lebih optimal. Oleh karena itu, pemahaman yang baik tentang probabilitas dalam biostatistik menjadi suatu keharusan dalam dunia medis, kesehatan dan penelitian kesehatan.

4.6.2 Saran

Agar pemanfaatan probabilitas dalam biostatistik kesehatan semakin optimal, diperlukan peningkatan pemahaman dan keterampilan tenaga medis serta peneliti dalam mengaplikasikan konsep probabilitas dalam analisis data kesehatan. Salah satu

langkah yang dapat dilakukan adalah dengan mengintegrasikan pembelajaran probabilitas dalam kurikulum pendidikan kedokteran dan kesehatan masyarakat. Dengan pemahaman yang lebih baik, tenaga kesehatan dapat lebih percaya diri dalam menginterpretasikan hasil statistik dan menerapkan informasi tersebut dalam pengambilan keputusan klinis dan kebijakan kesehatan.

Selain itu, diperlukan peningkatan penggunaan teknologi dalam analisis probabilistik di bidang kesehatan. Penggunaan perangkat lunak statistik dan kecerdasan buatan (AI) dapat membantu dalam pengolahan data yang lebih cepat dan akurat. Misalnya, algoritma machine learning yang berbasis probabilitas dapat digunakan untuk memprediksi penyebaran penyakit menular atau menentukan efektivitas suatu terapi berdasarkan data klinis pasien. Dengan pemanfaatan teknologi yang lebih luas, hasil analisis probabilitas dalam biostatistik kesehatan dapat lebih akurat dan aplikatif dalam dunia medis.

Terakhir, kolaborasi antara tenaga kesehatan, peneliti, dan pembuat kebijakan sangat diperlukan untuk memastikan bahwa hasil analisis probabilitas dapat digunakan secara efektif dalam praktik kesehatan. Kebijakan berbasis data harus menjadi dasar dalam perencanaan program kesehatan masyarakat, termasuk strategi pencegahan penyakit dan intervensi medis. Dengan pendekatan berbasis probabilitas yang lebih luas, diharapkan kualitas pelayanan kesehatan dapat meningkat dan memberikan manfaat yang lebih besar bagi masyarakat secara keseluruhan.

DAFTAR PUSTAKA

- Cook, R. J. and Moodie, E. E. M. (2024) 'A retrospective and prospective study of biostatistics in Canada', *Canadian Journal of Public Health*. doi: 10.17269/s41997-024-00866-w.
- Dehghan, A. *et al.* (2023) 'A comparison of laboratory-based and office-based Framingham risk scores to predict 10-year risk of cardiovascular diseases: a population-based study', *Journal of Translational Medicine*, 21(1). doi: 10.1186/s12967-023-04568-8.
- Gamboa, M. and Lopez-Herrero, M. (2022) 'Measures to assess a warning vaccination level in a stochastic SIV model with imperfect vaccine', *Studies in Applied Mathematics*, 148(4). doi: 10.1111/sapm.12479.
- Kestenbaum, B. (2009) *Epidemiology and Biostatistics: An Introduction to Clinical Research*. United States of America.: Springer Science+Business Media.
- Levy, P. S. and Lemeshow, S. (2008) *Sampling of Populations: Methods and Applications*. Fourth Edi. United States of America.: A John Wiley & Sons, Inc., Publication.
- Machluf, Y. *et al.* (2017) 'From Population Databases to Research and Informed Health Decisions and Policy', *Frontiers in Public Health*, 5. doi: 10.3389/fpubh.2017.00230.
- Merrill, R. M. (2013) *Fundamentals of Epidemiology and Biostatistics: Combining the Basics*. United States of America.: Jones & Bartlett Learning, LLC.
- Mikhail Nikulin, Daniel Commenges, C. H. (2006) *Probability, Statistics And Modelling In Public Health*. United States of America.: Springer Science+Business Media, Inc.
- Newman, S. C. (2001) *Biostatistical Methods in Epidemiology*. A John Wiley & Sons, Inc., Publication.
- Pagano, M. (1993) *Principles of biostatistics*. 2nd ed. CRC Press.
- Pérez-Guerrero, E. E. *et al.* (2024) 'Methodological and statistical considerations for cross-sectional, case-control, and cohort studies', *Journal of Clinical Medicine*, 13(14), p. 4005.
- Raji, T. *et al.* (2024) 'Efficacy and effectiveness of COVID-19 vaccines in Africa: A systematic review', *Plos one*, 19(6), p. e0306309.

- Rella, S. A. *et al.* (2023) 'Complex vaccination strategies prevent the emergence of vaccine resistance', *European Journal of Public Health*, 33(Supplement_2), pp. ckad160-597.
- Rosner, B. (2010) *Fundamentals of Biostatistics*. Seventh Ed. United States of America.: Cengage Learning, Inc.
- Sun, C. *et al.* (2024) 'A review of disease risk prediction methods and applications in the omics era', *Proteomics*, p. 2300359.
- Tsokani, S. *et al.* (2022) 'Assessing the performance of diagnostic test accuracy measures', *American Journal of Orthodontics and Dentofacial Orthopedics*. doi: 10.1016/j.ajodo.2021.12.007.
- Weatherall, M. (2018) 'Information provided by diagnostic and screening tests: Improving probabilities', *Postgraduate Medical Journal*. doi: 10.1136/postgradmedj-2017-135273.

BAB 5

INFERENSI STATISTIK ESTIMASI DAN PENGUJIAN HIPOTESIS

Oleh Nasruddin Syam

5.1 Pendahuluan

Penelitian medis, menggunakan biostatistik untuk mengolah, menganalisis, dan menyimpulkan data yang diperoleh dari studi klinis, epidemiologi, dan eksperimen laboratorium (Pagano, 1993; Cook and Moodie, 2024; Rutkowski *et al.*, 2024). Salah satu aspek utama dalam biostatistik adalah inferensi statistik, yang memungkinkan para peneliti membuat kesimpulan tentang populasi berdasarkan data sampel. Inferensi statistik terdiri dari dua komponen utama, yaitu estimasi parameter dan pengujian hipotesis, yang keduanya berperan penting dalam pengambilan keputusan berbasis data dalam dunia kesehatan (Levy and Lemeshow, 2008; Rosner, 2010).

Estimasi parameter digunakan untuk memperkirakan nilai karakteristik suatu populasi berdasarkan data sampel. Dalam konteks kesehatan, estimasi ini sering digunakan dalam studi prevalensi penyakit, efektivitas pengobatan, serta analisis faktor risiko. Estimasi dapat dilakukan melalui estimasi titik, yang memberikan satu nilai sebagai perkiraan terbaik suatu parameter, atau estimasi interval, yang memberikan rentang nilai yang mencakup parameter populasi dengan tingkat kepercayaan tertentu. Teknik estimasi ini sangat penting dalam penelitian medis, karena keputusan klinis sering kali bergantung pada hasil perkiraan yang akurat mengenai berbagai parameter kesehatan (Mielke, Denwood and Christiansen, 2023).

Sementara itu, pengujian hipotesis berfungsi untuk menguji kebenaran suatu asumsi atau klaim berdasarkan data yang tersedia. Dalam dunia kesehatan, pengujian hipotesis digunakan untuk menentukan efektivitas suatu terapi, membandingkan hasil antara

dua kelompok pasien, atau menilai hubungan antara faktor risiko dan kejadian penyakit. Melalui metode statistik seperti uji-t, uji ANOVA, regresi linier, dan uji *Chi-square*, para peneliti dapat mengevaluasi signifikansi hubungan dan menentukan apakah hasil penelitian mereka memiliki dasar statistik yang kuat atau hanya terjadi secara kebetulan.

Dalam praktiknya, penerapan estimasi dan pengujian hipotesis dalam biostatistik kesehatan tidak terlepas dari berbagai tantangan, seperti kesalahan pengambilan sampel, bias dalam penelitian, serta kesalahan interpretasi hasil statistik. Oleh karena itu, pemahaman yang mendalam mengenai metode inferensi statistik menjadi sangat penting agar hasil penelitian dapat digunakan untuk mendukung keputusan klinis yang berbasis bukti (*evidence-based medicine*). Dengan pemahaman yang tepat, tenaga medis, peneliti, dan pembuat kebijakan dapat menggunakan data statistik untuk meningkatkan kualitas layanan kesehatan dan kebijakan publik di bidang kesehatan.

Seiring dengan perkembangan teknologi dan metode statistik, penggunaan inferensi statistik dalam biostatistik kesehatan semakin luas dan kompleks. Penggunaan perangkat lunak statistik seperti SPSS, R, atau Python memungkinkan analisis data yang lebih cepat dan akurat. Oleh karena itu, memahami konsep dasar estimasi dan pengujian hipotesis menjadi langkah awal yang krusial bagi setiap profesional di bidang kesehatan yang ingin menerapkan metode statistik dalam penelitian dan praktik mereka.

Tulisan ini bertujuan untuk menjelaskan konsep dasar inferensi statistik dalam biostatistik kesehatan, dengan fokus pada estimasi parameter dan pengujian hipotesis. Pembahasan mencakup metode utama yang digunakan dalam kedua proses tersebut, serta contoh penerapannya dalam penelitian kesehatan. Diharapkan, pembaca dapat memahami pentingnya inferensi statistik dalam analisis data kesehatan dan mampu mengaplikasikannya dalam penelitian atau praktik profesional mereka.

5.2 Konsep Dasar Inferensi Biostatistik Kesehatan

Inferensi statistik dalam biostatistik kesehatan didefinisikan sebagai proses pengambilan keputusan atau penarikan kesimpulan mengenai populasi berdasarkan analisis

data dari sampel. Karena mustahil untuk mengamati seluruh individu dalam suatu populasi, metode statistik digunakan untuk memberikan estimasi yang akurat terhadap parameter populasi berdasarkan data yang terbatas. Inferensi statistik melibatkan penggunaan berbagai alat analisis, seperti estimasi interval kepercayaan, pengujian hipotesis, serta model statistik prediktif(Ihemelandu and Ekstrand, 2021).

Dalam konteks kesehatan, inferensi statistik membantu dalam memahami pola penyakit, menilai efektivitas intervensi medis, serta mengembangkan kebijakan kesehatan berbasis bukti. Misalnya, dalam studi epidemiologi, inferensi statistik dapat digunakan untuk menentukan apakah paparan terhadap zat tertentu meningkatkan risiko suatu penyakit. Dengan kata lain, metode ini memungkinkan peneliti untuk membuat generalisasi yang valid tentang populasi berdasarkan hasil penelitian yang diperoleh dari sampel yang representatif.

5.2.1 Definisi dan Ruang Lingkup

1. Pengertian Inferensi Biostatistik

Inferensi statistik dalam biostatistik kesehatan didefinisikan sebagai proses pengambilan keputusan atau penarikan kesimpulan mengenai populasi berdasarkan analisis data dari sampel. Karena mustahil untuk mengamati seluruh individu dalam suatu populasi, metode statistik digunakan untuk memberikan estimasi yang akurat terhadap parameter populasi berdasarkan data yang terbatas. Inferensi statistik melibatkan penggunaan berbagai alat analisis, seperti estimasi interval kepercayaan, pengujian hipotesis, serta model statistik prediktif(Ihemelandu and Ekstrand, 2021).

Dalam konteks kesehatan, inferensi statistik membantu dalam memahami pola penyakit, menilai efektivitas intervensi medis, serta mengembangkan kebijakan kesehatan berbasis bukti. Misalnya, dalam studi epidemiologi, inferensi statistik dapat digunakan untuk menentukan apakah paparan terhadap zat tertentu meningkatkan risiko suatu penyakit. Dengan kata lain,

metode ini memungkinkan peneliti untuk membuat generalisasi yang valid tentang populasi berdasarkan hasil penelitian yang diperoleh dari sampel yang representatif.

2. Peran dalam Penelitian Epidemiologi dan Klinis

Dalam penelitian epidemiologi, inferensi statistik digunakan untuk mengidentifikasi hubungan antara faktor risiko dan kejadian penyakit. Misalnya, dalam studi kohort atau studi kasus-kontrol, metode ini membantu menentukan apakah suatu faktor tertentu, seperti merokok atau pola makan, berkontribusi terhadap perkembangan penyakit tertentu. Selain itu, inferensi statistik memungkinkan peneliti untuk mengukur probabilitas terjadinya suatu kejadian kesehatan dalam populasi yang lebih luas berdasarkan hasil dari sampel terbatas (Rothman, 2012).

Dalam penelitian klinis, inferensi statistik sangat penting dalam mengevaluasi efektivitas pengobatan dan intervensi medis. Uji klinis acak terkontrol (*randomized controlled trial* – RCT), misalnya, menggunakan inferensi statistik untuk menentukan apakah suatu terapi baru lebih efektif dibandingkan dengan terapi standar. Dengan menggunakan teknik seperti uji-t, ANOVA, dan analisis regresi, peneliti dapat membandingkan hasil antar kelompok dan menentukan apakah perbedaan yang diamati bersifat signifikan secara statistik atau hanya terjadi secara kebetulan.

5.2.2 Jenis Data dalam Biostatistik

1. Data kategorik dan numerik

Dalam biostatistik kesehatan, data dikategorikan ke dalam dua jenis utama: data kategorik dan data numerik. Data kategorik mencakup variabel yang dikelompokkan ke dalam kategori tanpa nilai numerik yang jelas, seperti jenis kelamin (laki-laki atau perempuan) dan status vaksinasi (sudah divaksin atau belum divaksin). Data ini sering digunakan dalam studi epidemiologi dan survei kesehatan untuk menggambarkan karakteristik populasi yang diteliti.

Sebaliknya, data numerik terdiri dari variabel yang memiliki nilai kuantitatif dan dapat diukur dalam satuan tertentu, seperti tekanan darah, kadar gula darah, atau usia seseorang. Data numerik memungkinkan analisis statistik yang lebih kompleks, seperti analisis regresi atau uji perbedaan rata-rata. Dengan memahami perbedaan antara data kategorik dan numerik, peneliti dapat memilih metode analisis yang paling sesuai untuk menjawab pertanyaan penelitian mereka.

2. Skala pengukuran

Berdasarkan karakteristiknya, data dalam biostatistik kesehatan juga diklasifikasikan menurut skala pengukurannya, yaitu nominal, ordinal, interval, dan rasio.

- a. **Data Nominal** adalah jenis data kategorik yang hanya menunjukkan kategori tanpa urutan tertentu, seperti golongan darah (A, B, AB, O) atau status perokok (perokok atau bukan perokok). Data ini sering dianalisis menggunakan metode statistik non-parametrik, seperti uji *Chi-square*.
- b. **Data Ordinal** memiliki tingkatan atau urutan tertentu tetapi tidak memiliki perbedaan nilai yang tetap antar kategori, seperti tingkat keparahan penyakit (ringan, sedang, berat). Meskipun ada urutan, perbedaan antara kategori tidak dapat diukur secara pasti.
- c. **Data Interval** merupakan data numerik yang memiliki jarak tetap antara nilai, tetapi tanpa nol absolut. Contohnya adalah suhu tubuh dalam skala Celsius, di mana perbedaan antara 30°C dan 40°C sama dengan perbedaan antara 20°C dan 30°C, tetapi nol pada skala ini bukan berarti "tidak ada suhu".
- d. **Data Rasio** memiliki semua karakteristik data interval, tetapi dengan nol absolut yang berarti. Contohnya termasuk berat badan, tekanan darah, dan kadar glukosa dalam darah. Karena memiliki nol absolut, data rasio memungkinkan perhitungan proporsi dan perbandingan.

Pemahaman tentang skala pengukuran ini sangat penting karena menentukan metode analisis yang dapat diterapkan dalam penelitian kesehatan. Kesalahan dalam mengidentifikasi jenis data dapat menyebabkan penggunaan metode statistik yang tidak sesuai, yang pada akhirnya berpengaruh terhadap validitas hasil penelitian. Oleh karena itu, memilih metode analisis yang tepat berdasarkan jenis data merupakan langkah penting dalam proses inferensi statistik dalam biostatistik kesehatan.

5.3 Estimasi dalam Inferensi Biostatistik

Estimasi dalam inferensi statistik adalah metode yang digunakan untuk memperkirakan nilai suatu parameter populasi berdasarkan data sampel. Karena sulit atau bahkan tidak mungkin untuk mengamati seluruh populasi, estimasi statistik memungkinkan para peneliti untuk menafsirkan data dengan lebih efisien dan akurat. Dalam biostatistik kesehatan, estimasi digunakan untuk memahami berbagai aspek kesehatan masyarakat, seperti prevalensi penyakit, rerata tekanan darah dalam suatu populasi, atau efektivitas suatu terapi.

Estimasi dalam inferensi statistik terdiri dari dua jenis utama, yaitu estimasi titik dan estimasi interval. Estimasi titik memberikan satu nilai terbaik sebagai perkiraan suatu parameter populasi, sedangkan estimasi interval memberikan rentang nilai yang kemungkinan besar mencakup parameter populasi dengan tingkat kepercayaan tertentu. Kedua metode ini berperan penting dalam penelitian kesehatan, terutama dalam pengambilan keputusan berbasis bukti untuk perawatan pasien dan kebijakan kesehatan masyarakat.

5.3.1 Pengertian dan Jenis-Jenis Estimasi

1. Pengertian Estimasi dalam Inferensi Biostatistik

Estimasi adalah teknik statistik yang digunakan untuk memperkirakan nilai suatu parameter populasi berdasarkan informasi dari sampel. Karena data yang dikumpulkan dalam penelitian kesehatan umumnya berasal dari sampel, estimasi menjadi alat utama dalam membuat kesimpulan tentang populasi yang lebih luas. Sebagai contoh, dalam penelitian

tentang tingkat kejadian hipertensi di suatu wilayah, peneliti tidak dapat mengukur tekanan darah seluruh penduduk, sehingga mereka menggunakan sampel untuk memperkirakan rerata tekanan darah populasi tersebut.

Estimasi dalam inferensi statistik bertujuan untuk memastikan bahwa hasil penelitian dapat digunakan untuk membuat keputusan yang tepat dan akurat. Dalam konteks biostatistik kesehatan, estimasi digunakan dalam berbagai aplikasi, termasuk studi epidemiologi, uji klinis, dan analisis efektivitas terapi. Dengan menggunakan teknik estimasi yang tepat, para peneliti dapat memastikan bahwa hasil penelitian mereka memiliki relevansi yang tinggi dan dapat digunakan dalam pengambilan keputusan kesehatan masyarakat maupun klinis.

2. Jenis-Jenis Estimasi dalam Biostatistik

Secara umum, terdapat dua jenis utama estimasi dalam inferensi statistik, yaitu estimasi titik dan estimasi interval.

- a. **Estimasi Titik** adalah metode yang menghasilkan satu nilai tunggal sebagai perkiraan suatu parameter populasi. Misalnya, jika sampel memiliki rerata tekanan darah sebesar 120 mmHg, maka estimasi titik untuk rerata tekanan darah populasi adalah 120 mmHg. Estimasi titik bersifat sederhana dan mudah dipahami, tetapi memiliki kelemahan utama, yaitu tidak memberikan informasi tentang ketidakpastian estimasi tersebut.
- b. **Estimasi Interval** adalah metode yang memberikan rentang nilai yang kemungkinan besar mencakup parameter populasi dengan tingkat kepercayaan tertentu. Interval kepercayaan (*confidence interval*) biasanya dinyatakan dalam bentuk “nilai perkiraan $\pm$ *margin of error*.” Misalnya, jika estimasi rerata tekanan darah adalah 120 mmHg dengan interval kepercayaan 95% sebesar 118-122 mmHg, maka kita dapat menyatakan bahwa ada kemungkinan 95% bahwa rerata tekanan darah populasi berada dalam rentang tersebut.

Estimasi interval lebih disukai dalam analisis biostatistik karena memberikan informasi tambahan tentang tingkat ketidakpastian dalam estimasi. Dengan demikian, metode ini lebih berguna dalam pengambilan keputusan berbasis data dalam penelitian kesehatan.

5.3.1 Estimasi Parameter dalam Biostatistik

1. Estimasi Parameter Rerata dan Proporsi

Dalam biostatistik kesehatan, dua parameter utama yang sering diestimasi adalah rerata (mean) dan proporsi.

- a. **Estimasi Rerata Populasi** digunakan ketika data bersifat numerik dan peneliti ingin mengetahui nilai rata-rata suatu variabel dalam populasi. Misalnya, dalam studi mengenai kadar hemoglobin pada pasien anemia, peneliti dapat menggunakan sampel untuk memperkirakan rerata kadar hemoglobin seluruh pasien dengan teknik estimasi interval kepercayaan.
- b. **Estimasi Proporsi Populasi** digunakan ketika data bersifat kategorik, seperti dalam studi mengenai prevalensi penyakit. Misalnya, peneliti ingin memperkirakan proporsi populasi yang mengalami obesitas di suatu wilayah berdasarkan sampel survei kesehatan. Teknik estimasi proporsi sering digunakan dalam studi epidemiologi untuk memahami distribusi penyakit dalam populasi.

Kedua metode estimasi ini memiliki peran penting dalam perencanaan intervensi kesehatan dan pengambilan keputusan kebijakan kesehatan. Dengan memahami cara mengestimasi rerata dan proporsi, peneliti dapat

2. Pengaruh Ukuran Sampel terhadap Estimasi

Ukuran sampel memiliki dampak besar terhadap keakuratan estimasi dalam inferensi statistik. Semakin besar ukuran sampel, semakin kecil *margin of error* dalam estimasi dan semakin tinggi tingkat kepercayaan terhadap hasil penelitian.

Dalam penelitian biostatistik, pemilihan ukuran sampel yang optimal sangat penting untuk memastikan hasil yang dapat digeneralisasikan ke populasi. Misalnya, dalam penelitian

tentang efektivitas vaksin, sampel yang terlalu kecil dapat menyebabkan estimasi yang tidak akurat dan hasil yang tidak dapat diandalkan. Sebaliknya, sampel yang terlalu besar dapat meningkatkan biaya dan waktu penelitian tanpa memberikan manfaat tambahan yang signifikan. Oleh karena itu, metode statistik seperti *power analysis* sering digunakan untuk menentukan ukuran sampel yang optimal dalam studi kesehatan.

Dengan pemahaman yang baik tentang pengaruh ukuran sampel terhadap estimasi, peneliti dapat meningkatkan kualitas dan validitas hasil penelitian mereka, yang pada akhirnya berdampak pada efektivitas kebijakan dan intervensi kesehatan.

5.4 Uji Hipotesis dalam Biostatistik

Pengujian hipotesis adalah salah satu teknik utama dalam inferensi statistik yang digunakan untuk menilai validitas klaim atau asumsi mengenai populasi berdasarkan data sampel. Dalam biostatistik kesehatan, pengujian hipotesis memiliki peran penting dalam menentukan hubungan antar variabel, mengevaluasi efektivitas intervensi medis, serta membuat keputusan berbasis bukti dalam penelitian epidemiologi dan klinis.

Secara umum, proses pengujian hipotesis melibatkan beberapa langkah utama, termasuk perumusan hipotesis nol (null hypothesis, H_0) dan hipotesis alternatif (*alternative hypothesis*, H_1), pemilihan tingkat signifikansi, pemilihan uji statistik yang sesuai, serta interpretasi hasil berdasarkan nilai p (p -value) atau interval kepercayaan. Dengan metode ini, peneliti dapat menghindari kesalahan dalam pengambilan keputusan dan memastikan bahwa hasil penelitian memiliki validitas yang tinggi.

5.4.1 Konsep Dasar Uji Hipotesis

1. Definisi dan Tujuan Uji Hipotesis

Pengujian hipotesis adalah prosedur statistik yang digunakan untuk mengevaluasi klaim mengenai parameter populasi berdasarkan data sampel. Dalam konteks biostatistik

kesehatan, metode ini sering diterapkan dalam studi klinis dan epidemiologi untuk menentukan apakah suatu perbedaan atau hubungan yang diamati dalam sampel cukup kuat untuk digeneralisasikan ke populasi yang lebih luas.

Tujuan utama dari pengujian hipotesis adalah untuk membedakan apakah suatu efek yang diamati dalam penelitian terjadi karena faktor acak atau memang mencerminkan kondisi nyata dalam populasi. Misalnya, dalam uji klinis obat baru, pengujian hipotesis digunakan untuk menentukan apakah obat tersebut memberikan efek yang signifikan dibandingkan dengan plasebo atau pengobatan standar.

2. Langkah-Langkah dalam Uji Hipotesis

Proses pengujian hipotesis dalam biostatistik kesehatan melibatkan beberapa langkah sistematis:

a. Merumuskan Hipotesis Nol (H_0) dan Hipotesis Alternatif (H_1)

- 1) Hipotesis nol (*null hypothesis*, H_0) menyatakan bahwa tidak ada perbedaan atau hubungan yang signifikan dalam populasi.
- 2) Hipotesis alternatif (*alternative hypothesis*, H_1) menyatakan bahwa ada perbedaan atau hubungan yang signifikan.

b. Menentukan Tingkat Signifikansi (α)

Tingkat signifikansi biasanya ditetapkan sebesar 5% ($\alpha = 0,05$), yang berarti ada kemungkinan 5% bahwa kesimpulan yang diambil berdasarkan sampel adalah kesalahan.

c. Memilih Uji Statistik yang Sesuai

Pemilihan uji statistik tergantung pada jenis data dan desain penelitian, misalnya uji-t, ANOVA, uji *Chi-square*, atau regresi logistik.

d. Menghitung Statistik Uji dan Nilai p (p -value)

- 1) Nilai p menunjukkan probabilitas mendapatkan hasil yang sama atau lebih ekstrem jika H_0 benar.
- 2) Jika $p < \alpha$, maka H_0 ditolak, dan dapat disimpulkan bahwa ada hubungan atau perbedaan yang signifikan.

e. Membuat Kesimpulan

- 1) Jika H_0 ditolak, maka hasil penelitian mendukung H_1 dan menunjukkan adanya efek atau hubungan yang signifikan.
- 2) Jika H_0 gagal ditolak, maka tidak cukup bukti untuk menyatakan bahwa ada efek atau hubungan yang nyata dalam populasi.

5.4.2 Jenis Uji Hipotesis dalam Biostatistik

1. Uji Hipotesis Parametrik dan Nonparametrik

Proses pengujian hipotesis dalam biostatistik kesehatan melibatkan beberapa langkah sistematis:

- a. **Uji Parametrik** digunakan ketika asumsi tentang distribusi data dapat dipenuhi, seperti distribusi normal. Contohnya adalah uji-t untuk membandingkan dua rata-rata dan ANOVA untuk membandingkan lebih dari dua kelompok.
- b. **Uji Nonparametrik** digunakan ketika data tidak memenuhi asumsi distribusi tertentu, seperti pada data ordinal atau data dengan distribusi yang tidak normal. Contohnya adalah uji *Mann-Whitney*, uji *Wilcoxon*, dan uji *Chi Square*.

Penggunaan uji parametrik atau nonparametrik bergantung pada karakteristik data yang dianalisis. Jika data tidak berdistribusi normal atau memiliki ukuran sampel kecil, maka uji nonparametrik lebih direkomendasikan untuk mendapatkan hasil yang lebih akurat

2. Penggunaan Uji Hipotesis dalam Kesehatan

Dalam penelitian kesehatan, pengujian hipotesis digunakan untuk berbagai tujuan, seperti:

a. Menguji efektivitas pengobatan

Misalnya, dalam uji klinis terhadap obat hipertensi baru, peneliti menggunakan uji-t berpasangan untuk membandingkan tekanan darah pasien sebelum dan setelah pengobatan. Jika hasil menunjukkan perbedaan signifikan, maka obat tersebut dapat dianggap efektif.

b. Menentukan hubungan faktor risiko dengan penyakit

Studi epidemiologi sering menggunakan uji *Chi Square* untuk mengevaluasi apakah terdapat hubungan signifikan antara faktor risiko tertentu (misalnya kebiasaan merokok) dengan kejadian penyakit tertentu (misalnya kanker paru-paru).

c. Membandingkan kelompok populasi

Misalnya, dalam studi tentang kadar gula darah antara pasien diabetes tipe 1 dan tipe 2, peneliti menggunakan **ANOVA** untuk menguji apakah ada perbedaan signifikan di antara kedua kelompok.

5.4.3 Kesalahan Uji Hipotesis dan Implikasinya

1. Kesalahan Tipe I dan Tipe II dalam Uji Hipotesis

Dalam pengujian hipotesis, terdapat dua jenis kesalahan yang perlu diwaspadai:

a. Kesalahan Tipe I (α): Menolak H_0 yang sebenarnya benar

- 1) Ini berarti kesimpulan penelitian menyatakan ada perbedaan atau hubungan yang signifikan padahal sebenarnya tidak ada dalam populasi.
- 2) Dalam penelitian kesehatan, kesalahan tipe I dapat menyebabkan kesimpulan yang salah tentang efektivitas suatu pengobatan, sehingga pasien dapat menerima terapi yang tidak efektif atau berisiko.

b. Kesalahan Tipe II (β): Gagal menolak H_0 yang sebenarnya salah

- 1) Ini berarti penelitian gagal mendeteksi perbedaan atau hubungan yang sebenarnya ada dalam populasi.
- 2) Kesalahan tipe II dapat menyebabkan intervensi kesehatan yang efektif tidak diadopsi, sehingga berpotensi merugikan pasien dan kebijakan kesehatan masyarakat.

2. Uji Hipotesis Parametrik dan Nonparametrik

Kesalahan dalam pengujian hipotesis dapat berdampak luas pada keputusan medis dan kebijakan kesehatan. Oleh karena itu, sangat penting untuk:

- a. Memastikan ukuran sampel yang cukup agar analisis memiliki kekuatan statistik (*statistical power*) yang tinggi.
- b. Menggunakan metode analisis yang tepat sesuai dengan karakteristik data.
- c. Menetapkan tingkat signifikansi yang sesuai untuk mengontrol risiko kesalahan.

Dalam penelitian medis dan kesehatan masyarakat, penggunaan pengujian hipotesis yang tepat dapat membantu meningkatkan akurasi dalam menganalisis data, sehingga menghasilkan kebijakan dan intervensi kesehatan yang lebih efektif.

5.5 Tantangan dan Keterbatasan Inferensi Biostatistik

Inferensi statistik dalam biostatistik kesehatan memiliki peran yang sangat penting dalam penelitian medis, epidemiologi, dan kebijakan kesehatan masyarakat. Dengan metode ini, para peneliti dapat menarik kesimpulan tentang populasi berdasarkan data sampel, memungkinkan pengambilan keputusan yang lebih akurat dan berbasis bukti. Namun, meskipun memiliki banyak manfaat, inferensi statistik juga memiliki tantangan dan keterbatasan yang dapat mempengaruhi validitas dan reliabilitas hasil penelitian.

Tantangan utama dalam inferensi statistik meliputi keterbatasan dalam desain penelitian, kesalahan dalam pengumpulan dan analisis data, serta asumsi statistik yang sering kali sulit dipenuhi dalam penelitian kesehatan (Gelman and Brown, 2024; Kaufman, 2024). Selain itu, keterbatasan sumber daya, bias dalam sampel, serta interpretasi hasil yang tidak tepat dapat menyebabkan kesimpulan yang salah dan berdampak pada pengambilan keputusan klinis atau kebijakan kesehatan.

masyarakat(Rutkowski *et al.*, 2024). Oleh karena itu, memahami tantangan dan keterbatasan inferensi statistik menjadi sangat penting untuk meningkatkan kualitas penelitian dan mengurangi risiko kesalahan dalam analisis data biostatistik kesehatan.

5.5.1 Tantangan dalam Inferensi Biostatistik

1. Kesulitan dalam Pemenuhan Asumsi Statistik

Banyak metode statistik yang digunakan dalam biostatistik kesehatan didasarkan pada asumsi tertentu, seperti distribusi normal data, independensi antar observasi, serta homogenitas varians. Namun, dalam praktiknya, data yang diperoleh dari studi kesehatan sering kali tidak memenuhi asumsi ini.

Misalnya, dalam penelitian klinis, data pasien sering kali tidak berdistribusi normal karena variasi biologis yang luas atau efek samping yang tidak terprediksi dari suatu terapi. Jika asumsi tidak terpenuhi, hasil pengujian statistik dapat menjadi bias dan tidak dapat diandalkan. Oleh karena itu, peneliti harus memastikan bahwa metode analisis yang digunakan sesuai dengan karakteristik data, misalnya dengan menggunakan uji nonparametrik jika distribusi data tidak normal.

Selain itu, banyak studi kesehatan menggunakan data longitudinal atau data yang dikumpulkan secara berulang dari individu yang sama. Hal ini dapat menyebabkan adanya ketergantungan antar observasi yang melanggar asumsi independensi dalam inferensi statistik. Dalam situasi seperti ini, metode statistik khusus seperti model efek campuran (*mixed-effects models*) atau regresi logistik terstratifikasi perlu digunakan untuk mengakomodasi struktur data yang lebih kompleks.

2. Tantangan Pengambilan Sampel yang Representatif

Inferensi statistik mengandalkan prinsip bahwa sampel yang digunakan harus mewakili populasi yang lebih luas. Namun, dalam biostatistik kesehatan, sering kali sulit untuk mendapatkan sampel yang benar-benar representatif. Beberapa faktor yang dapat menyebabkan bias dalam pemilihan sampel meliputi:

a. **Non-respons atau drop-out dalam studi longitudinal:**

Dalam penelitian kesehatan, tidak semua partisipan akan menyelesaikan studi karena berbagai alasan, seperti efek

sampling terapi atau kesulitan dalam mengakses layanan kesehatan. Ini dapat menyebabkan bias dalam hasil penelitian.

- b. **Kesulitan menjangkau populasi tertentu:** Beberapa kelompok populasi, seperti masyarakat pedalaman atau kelompok dengan akses terbatas ke layanan kesehatan, sering kali kurang terwakili dalam penelitian. Akibatnya, hasil studi dapat menjadi kurang relevan untuk populasi tersebut.

5.5.2 Keterbatasan dalam Inferensi Biostatistik

1. Potensi Bias dan Kesalahan dalam Analisis Data

Inferensi statistik dalam biostatistik kesehatan sangat rentan terhadap bias dan kesalahan dalam analisis data. Beberapa bentuk bias yang sering terjadi dalam penelitian kesehatan meliputi:

- a. **Bias seleksi:** Terjadi ketika sampel yang digunakan dalam penelitian tidak benar-benar acak dan tidak mewakili populasi yang lebih luas. Misalnya, jika suatu penelitian tentang diabetes hanya melibatkan pasien yang datang ke rumah sakit, hasilnya mungkin tidak dapat digeneralisasikan ke populasi yang tidak memiliki akses ke layanan kesehatan.
- b. **Bias informasi:** Terjadi ketika terdapat kesalahan dalam pengukuran variabel penelitian, seperti penggunaan alat diagnostik yang kurang akurat atau pencatatan data yang tidak konsisten.
- c. **Confounding variables:** Faktor luar yang tidak dikontrol dalam penelitian dapat menyebabkan hubungan yang salah antara variabel yang diteliti. Misalnya, penelitian yang menghubungkan konsumsi kopi dengan risiko penyakit jantung mungkin tidak mempertimbangkan variabel perancu seperti kebiasaan merokok.

Untuk mengurangi risiko bias, peneliti harus menggunakan desain studi yang ketat, seperti uji klinis terkontrol secara acak (*randomized controlled trials*), serta menerapkan metode analisis statistik yang dapat menyesuaikan faktor perancu, seperti regresi multivariat atau propensity score matching.

2. Keterbatasan dalam Interpretasi dan Generalisasi

Meskipun inferensi statistik memungkinkan kita untuk menarik kesimpulan dari data sampel, hasil penelitian tidak selalu dapat digeneralisasikan secara langsung ke populasi yang lebih luas. Beberapa keterbatasan yang dapat mempengaruhi interpretasi hasil dalam biostatistik kesehatan meliputi:

- a. **Keterbatasan ukuran sampel:** Studi dengan ukuran sampel kecil sering kali memiliki tingkat ketidakpastian yang lebih tinggi dan dapat menghasilkan estimasi yang kurang akurat. Oleh karena itu, penting untuk menggunakan metode *power analysis* sebelum penelitian untuk memastikan ukuran sampel yang cukup.
- b. **Hasil yang signifikan secara statistik belum tentu signifikan secara klinis:** Meskipun suatu hasil dapat menunjukkan perbedaan yang signifikan secara statistik ($p\text{-value} < 0,05$), hal ini belum tentu memiliki dampak yang bermakna dalam praktik klinis. Sebagai contoh, penurunan tekanan darah sebesar 1 mmHg mungkin signifikan secara statistik tetapi tidak memiliki dampak klinis yang besar bagi pasien hipertensi.
- c. **Masalah ekstrapolasi hasil ke populasi yang berbeda:** Banyak penelitian dilakukan pada populasi tertentu dengan karakteristik khusus, seperti pasien di rumah sakit atau kelompok usia tertentu. Oleh karena itu, hasil penelitian belum tentu dapat diterapkan ke populasi yang lebih luas dengan latar belakang yang berbeda.

Untuk meningkatkan validitas eksternal hasil penelitian, peneliti perlu melakukan replikasi studi di berbagai populasi dan menggunakan pendekatan analisis yang mempertimbangkan faktor sosial, ekonomi, dan demografi yang lebih luas.

5.6 Kesimpulan dan Rekomendasi

1. Kesimpulan

Inferensi statistik merupakan elemen fundamental dalam biostatistik kesehatan yang memungkinkan peneliti menarik kesimpulan berdasarkan data sampel. Dengan menggunakan teknik estimasi dan pengujian hipotesis, metode ini membantu

dalam memahami berbagai fenomena kesehatan, termasuk efektivitas pengobatan, faktor risiko penyakit, serta tren epidemiologi. Meskipun demikian, keakuratan dan validitas hasil sangat tergantung pada kualitas desain penelitian dan analisis data yang digunakan.

Salah satu tantangan terbesar dalam inferensi statistik adalah memastikan bahwa asumsi statistik yang digunakan sesuai dengan karakteristik data. Ketidakesesuaian dalam asumsi, seperti distribusi data yang tidak normal atau ketergantungan antar observasi, dapat menyebabkan hasil yang bias atau tidak dapat diandalkan. Oleh karena itu, pemilihan metode analisis yang tepat, seperti penggunaan uji nonparametrik atau model statistik yang lebih kompleks, menjadi sangat penting dalam meningkatkan validitas penelitian.

Selain itu, keterbatasan dalam pengambilan sampel yang representatif sering kali menjadi kendala dalam generalisasi hasil penelitian. Bias seleksi, bias informasi, serta keberadaan faktor perancu dapat mengurangi validitas eksternal suatu studi. Untuk mengatasi tantangan ini, peneliti harus menerapkan strategi sampling yang lebih inklusif serta menggunakan teknik analisis statistik yang mampu mengontrol bias, seperti regresi multivariat atau *propensity score matching*.

2. Rekomendasi

Untuk meningkatkan kualitas dan keandalan inferensi statistik dalam biostatistik kesehatan, penting bagi peneliti untuk merancang studi dengan metodologi yang ketat dan memperhatikan prinsip validitas serta reliabilitas. Penggunaan teknik sampling yang representatif, pemilihan metode analisis yang sesuai dengan karakteristik data, serta penerapan strategi kontrol bias harus menjadi prioritas utama dalam setiap penelitian biostatistik kesehatan.

Selain itu, pelatihan dan peningkatan kapasitas dalam analisis statistik bagi tenaga kesehatan dan peneliti menjadi hal yang sangat penting. Banyak kesalahan dalam interpretasi data yang terjadi akibat kurangnya pemahaman terhadap prinsip-prinsip inferensi statistik. Oleh karena itu, penyelenggaraan pelatihan statistik medis, penggunaan perangkat lunak analisis

yang canggih, serta bimbingan dalam penerapan metode statistik yang tepat akan membantu meningkatkan akurasi dalam pengambilan keputusan berbasis data.

Terakhir, kolaborasi antara ahli biostatistik, tenaga medis, dan pembuat kebijakan perlu diperkuat agar hasil penelitian dapat diterapkan secara optimal dalam kebijakan kesehatan masyarakat. Dengan mengintegrasikan hasil penelitian berbasis inferensi statistik ke dalam strategi kesehatan nasional, diharapkan pengelolaan penyakit, alokasi sumber daya, serta kebijakan intervensi kesehatan dapat dilakukan secara lebih efektif dan efisien.

DAFTAR PUSTAKA

- Cook, R. J. and Moodie, E. E. M. (2024) 'A retrospective and prospective study of biostatistics in Canada', *Canadian Journal of Public Health*. doi: 10.17269/s41997-024-00866-w.
- Gelman, A. and Brown, N. J. L. (2024) 'How statistical challenges and misreadings of the literature combine to produce unreplicable science: An example from psychology', *Advances in Methods and Practices in Psychological Science*, 7(4), p. 25152459241276400.
- Ihemelandu, N. and Ekstrand, M. D. (2021) 'Statistical inference: The missing piece of RecSys experiment reliability discourse', in *CEUR Workshop Proceedings*.
- Kaufman, J. S. (2024) 'Causal Inference Challenges in the Relationship Between Social Determinants and Cardiovascular Outcomes', *Canadian Journal of Cardiology*. doi: 10.1016/j.cjca.2024.02.005.
- Levy, P. S. and Lemeshow, S. (2008) *Sampling of Populations: Methods and Applications*. Fourth Edi. United States of America.: A John Wiley & Sons, Inc., Publication.
- Mielke, A., Denwood, M. and Christiansen, L. E. (2023) 'Estimating true prevalence through questionnaire data', *Journal of Medical Virology*, 95(7). doi: 10.1002/jmv.28908.
- Pagano, M. (1993) *Principles of biostatistics*. 2nd ed. CRC Press.
- Rosner, B. (2010) *Fundamentals of Biostatistics*. Seventh Ed. United States of America.: Cengage Learning, Inc.
- Rothman, K. J. (2012) 'Epidemiology in Clinical Settings', *Epidemiology. An Introduction*, 72.
- Rutkowski, D. et al. (2024) 'The limits of inference: reassessing causality in international assessments', *Large-scale Assessments in Education*, 12(1), p. 9.

BAB 6

ANALISIS REGRESI DALAM KESEHATAN

Oleh Islamiyati

6.1 Pendahuluan

Analisis regresi adalah teknik statistik yang digunakan dalam bidang kesehatan untuk mempelajari pola, mengukur bagaimana variabel berinteraksi satu sama lain, dan memprediksi hasil yang berkaitan dengan kesehatan. Teknik ini membantu mengidentifikasi faktor-faktor risiko, mengevaluasi intervensi, dan membuat perkiraan. Analisis regresi dapat digunakan untuk berbagai jenis variabel, termasuk variabel biner (misalnya, sakit atau tidak sakit) dan variabel kontinu (misalnya, tekanan darah). Variabel yang akan diprediksi disebut variabel terikat (dependen), sedang variabel yang menjelaskan perubahan variabel terikat tadi disebut variabel bebas (independen) (Carolina, 2021).

6.2 Manfaat Analisis Regresi dalam Kesehatan

6.2.1 Mengidentifikasi Faktor Risiko

Analisis regresi membantu mengidentifikasi faktor-faktor yang berhubungan dengan hasil kesehatan, seperti penyakit atau kematian. Misalnya, dapat digunakan untuk menentukan apakah merokok, obesitas, atau kurang olahraga berhubungan dengan peningkatan risiko penyakit jantung.

6.2.2 Mengevaluasi Intervensi

Analisis regresi bisa dimanfaatkan untuk mengevaluasi keefektifitasan suatu program kesehatan, seperti program imunisasi atau kampanye keluarga berencana. Dengan menganalisis dampak suatu program kesehatan terhadap kesehatan masyarakat penerima program, peneliti dapat

menentukan apakah program tersebut efektif dan jika belum efektif bagaimana cara memperbaikinya.

6.2.3 Memprediksi Hasil Kesehatan

Hasil kesehatan di masa depan, seperti risiko terkena penyakit atau tingkat kematian, dapat diprediksi dengan menggunakan model regresi. Prediksi ini dapat membantu dalam merencanakan sumber daya dan pembagian dana anggaran.

6.2.4 Memahami Hubungan Antar Variabel

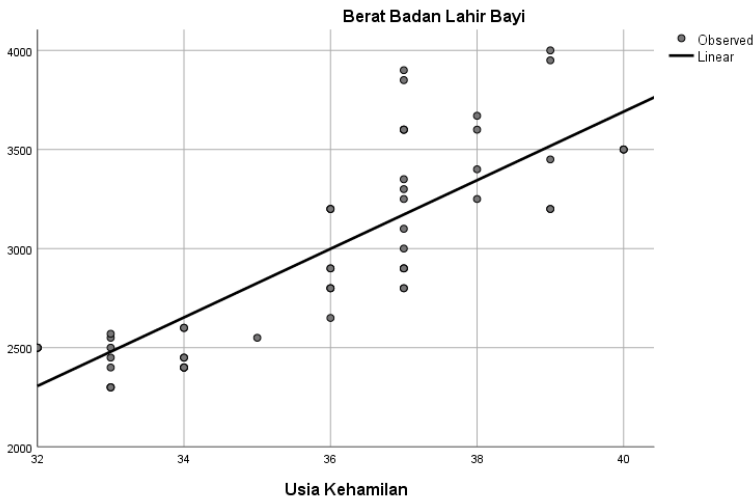
Analisis regresi memungkinkan peneliti untuk memahami bagaimana berbagai faktor kesehatan saling terkait. Misalnya, dapat membantu memahami bagaimana tingkat pendidikan, status sosial ekonomi, dan akses ke layanan kesehatan mempengaruhi kesehatan masyarakat.

6.3 Jenis-jenis Analisis Regresi

6.3.1 Regresi Linear

Regresi linier adalah salah satu regresi yang paling mendasar dalam pembelajaran otomatis. Ketika variabel dependen (Y) adalah variabel kontinu, maka model regresi linier digunakan. Ada keterkaitan secara linier antara variabel yang dipengaruhi dan variabel yang mempengaruhi dalam model ini. Misalnya, memprediksi berat badan lahir bayi berdasarkan usia kehamilan, kenaikan BB ibu selama hamil, dan frekuensi ANC.

Regresi linier bertujuan menyesuaikan hyperplan (garis untuk titik data 2D) dengan meminimalkan jumlah kuadrat rata-rata untuk setiap titik data. Berikut gambaran kurve regresi linier yang dihasilkan dari analisis regresi linier antar 2 variabel Berat Badan Lahir Bayi dan usia kehamilan.



Gambar 6.1. Contoh Kurva estimasi Regresi Linier
(Sumber : Dokumentasi pribadi penulis)

Persamaan Umum Rumus Regresi Linier;

$$Y = a + bX$$

Dimana :

Y = Subyek dalam variabel terikat yg diprediksi

a = harga y bila x

b = angka arah atau koefisien regresi bila b(+) maka ada kenaikan, bila (-) maka ada penurunan

X = subyek pada variabel bebas yang memiliki nilai tertentu

Penghitungan nilai a dan b menggunakan rumus :

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

$$a = \frac{\sum Y}{n} - b \cdot \frac{\sum X}{n}$$

Regresi Linear Sederhana

Sejauh mana hubungan sebab akibat antara variabel X , yang menjadi faktor penyebab, dan variabel Y, sebagai akibat dapat ditentukan dengan menggunakan regresi linier sederhana. Model regresi ini mengikutsertakan sebuah variabel bebas dan

sebuah variabel terikat, dengan hubungan yang diasumsikan linier (Suyono, 2015).

Beberapa asumsi dan persyaratan yang harus dipenuhi sebelum pengujian data dengan analisis regresi sederhana dapat dilakukan, antara lain (Setyawan, dkk, 2021) :

1. Variabel independen tidak berhubungan dengan *disturbance term* (Error). Nilai *disturbance term* 0 atau dengan simbol $E(U/X) = 0$,
2. Jika variabel independen (yang mempengaruhi) lebih dari satu, maka tidak ada hubungan linier yang nyata antara variabel independen (*explanatory*)
3. Model regresi dianggap sesuai jika angka signifikansi pada ANOVA sebesar < 0.05 dan variabel penyebab yang dipakai sebagai variabel independen dianggap sesuai. Kesesuaian ini diketahui bila nilai *Standard Error of Estimate* $< Standard Deviation$,
4. Ada signifikansi Koefisien regresi. Uji T digunakan untuk menguji koefisien regresi. Nilai T hitung $>$ dari nilai T tabel, atau nilai kritis, maka koefisien regresi adalah signifikan.
5. Nilai koefisien determinasi, yang didefinisikan sebagai $KD = R \text{ Square} \times 100\%$, digunakan untuk menjelaskan model regresi. Jika nilai koefisien determinasi semakin besar, model regresi semakin baik; jika nilai koefisien determinasi mendekati 1, maka semakin baik model regresi
6. Residual wajib berdistribusi normal
7. Skala data yaitu interval atau rasio
8. Satu variabel dianggap sebagai variabel prediktor atau variabel bebas, sedangkan yang lain dianggap sebagai variabel respons. Dengan demikian, kedua variabel memiliki sifat dependen.

Regresi Linear Berganda

Hubungan sebab akibat antara lebih dari satu variabel X (yang menjadi faktor penyebab) dan Y (yang menjadi akibat) diuji dengan regresi linier berganda. Dalam regresi ini, lebih dari satu variabel independen digunakan untuk memprediksi variabel dependen.

Iba & Wardhana (2023) menyatakan bahwa Regresi Linear Berganda merupakan salah satu metode analisis statistik yang dipakai guna mengevaluasi keterkaitan antara satu variabel terikat (Y) dan lebih dari satu variabel bebas (X1, X2, X3, dst.). Tujuan dari analisis ini adalah untuk memahami dan memprediksi bagaimana perubahan pada salah satu atau lebih variabel bebas X mempengaruhi perubahan pada variabel terikat Y. Hubungan antara variabel dalam regresi linear berganda disajikan dalam bentuk persamaan garis lurus (seperti pada gambar 1). Persamaan regresi linear berganda sangat membantu dalam menentukan pengaruh relatif dari masing-masing variabel bebas terhadap variabel terikat. Selain itu, persamaan ini dapat menggunakan nilai variabel bebas untuk memprediksi nilai variabel terikat.

Bentuk persamaan umum dari Regresi Linear Berganda adalah:

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_n X_n + \varepsilon$$

Y	= variabel yang dipengaruhi
X ₁ , X ₂ , X ₃ , ..., X _n	= variabel yang mempengaruhi
β ₀	= konstanta (intercept)
β ₁ , β ₂ , β ₃ , ..., β _n	= koefisien regresi yang menunjukkan seberapa besar perubahan yang diharapkan dalam Y terjadi ketika masing-masing satu unit dipengaruhi oleh variabel bebas
ε	= kesalahan atau residu yang menjadi perbedaan antara nilai Y yang sebenarnya dan nilai yang diperkirakan oleh model.

Proses analisis regresi linear berganda biasanya adalah sebagai berikut:

1. Pengumpulan Data.

Mengumpulkan data yang berisi nilai variabel dependen dan variabel independen akan digunakan dalam analisis.

2. Estimasi Model.

Hitung model regresi linear berganda dengan menemukan koefisien regresi (β₀, β₁, β₂, dll.) yang paling sesuai dengan

data. Estimasi ini biasanya dilakukan dengan metode kuadrat terkecil untuk menemukan garis atau hiperbidur terbaik yang sesuai dengan data.

3. Evaluasi Model.

Evaluasi model dilakukan guna mengetahui apakah model menjelaskan dengan baik hubungan antara variabel terikat dan bebas. Ini melibatkan penggunaan statistik seperti koefisien determinasi (*Rsquared*) untuk menentukan seberapa besar variasi yang dapat dijelaskan oleh model untuk variabel terikat tertentu.

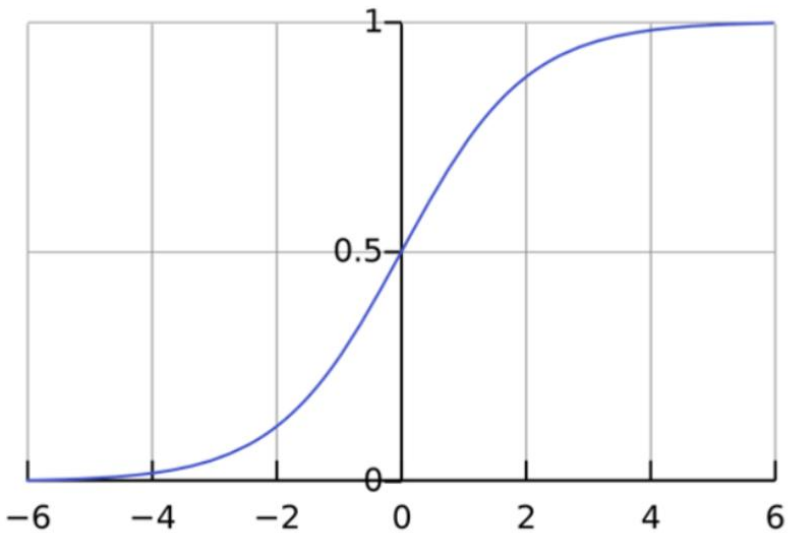
4. Prediksi dan Interpretasi.

Nilai variabel bebas digunakan sebagai dasar untuk membuat prediksi tentang variabel terikat. Model regresi linear berganda juga menginterpretasikan koefisien regresi untuk mengetahui bagaimana masing-masing variabel bebas memengaruhi variabel terikat.

6.3.2 Regresi Logistik

Menurut Romadhoni, Fania Ananta (2022) Analisis regresi logistik ganda adalah teknik statistik yang dipakai untuk menemukan model terbaik dari suatu masalah dan menganalisis pengaruh antar variabel. Pengaruh beberapa variabel bebas terhadap variabel terikat ditentukan melalui penggunaan metode ini. Misalnya, memprediksi kemungkinan seorang pasien terkena kanker payudara berdasarkan faktor risiko seperti usia, riwayat penggunaan kontrasepsi, dan riwayat menyusui

Menurut Murti (1997) analisis regresi logistik ganda bertujuan untuk menemukan model regresi yang paling cocok, paling efisien, dan paling masuk akal secara biologis untuk menunjukkan keterkaitan antara variabel respons (Y) dan satu set variabel prediktor (X) dalam populasi.



Gambar 6.2. Kurva Persamaan Regresi Logistik

(Sumber : <https://aws.amazon.com/id/what-is/logistic-regression>)

Regresi Logistik Biner

Regresi logistik biner adalah model statistik yang dipakai guna memprediksi peluang suatu kejadian/variabel prediktor (X) dengan dua kategori (biner), seperti ya/tidak, benar/salah, atau 0/1 yang dihubungkan dengan variabel respon (Y). Ada perbedaan model statistik ini dengan regresi linier yaitu variabel dependennya bersifat kategorikal, sedang regresi linier variabel dependennya bersifat kontinu. Contoh kategori variabel dependen biner adalah lulus dan tidak lulus, kompeten dan tidak kompeten, sakit dan tidak sakit.

Ada 2 jenis regresi logistik biner, yaitu:

1. Regresi logistik biner tunggal : apabila regresi logistik biner hanya mempunyai 1 variabel bebas
2. Regresi Logistik biner ganda : apabila regresi logistik biner mempunyai lebih dari 1 variabel bebas

Berikut merupakan persamaan dasar Regresi Logistik Biner:

$$\ln(p / (1 - p)) = b_0 + b_1 \cdot x_1 + b_2 \cdot x_2 + \dots + b_n \cdot x_n$$

Dimana:

- **P** : adalah probabilitas kejadian sukses (misalnya, 1).
- **1 - p** : adalah probabilitas kejadian gagal (misalnya, 0).
- **$\ln(p/(1 - p))$** : adalah [logit\(p\)](#), yaitu logaritma dari odds (peluang) untuk kejadian sukses.
- **b0** : adalah konstanta (intersep).
- **b1, b2, ..., bn** : adalah koefisien regresi untuk masing-masing variabel prediktor (x1, x2, ..., xn).
- **x1, x2, ..., xn** : adalah nilai-nilai variabel prediktor

Interpretasi Koefisien:

- **b0 (intersep)** : Adalah nilai logit ketika semua variabel prediktor bernilai nol.
- **b1, b2, ..., bn** : Menunjukkan perubahan dalam logit (dan dalam log-odds) untuk setiap peningkatan satu unit pada variabel prediktor yang bersangkutan, dengan asumsi variabel prediktor lainnya tetap.

Regresi logistik biner dapat dimanfaatkan dalam berbagai bidang. Sebagai contoh:

1. **Kesehatan:** Memprediksi risiko penyakit, kemungkinan keberhasilan pengobatan.
2. **Pemasaran:** Memprediksi kemungkinan pelanggan melakukan pembelian, probabilitas churn.
3. **Keuangan:** Memprediksi kemungkinan gagal bayar pinjaman, kemungkinan fraud.
4. **Kriminalogi:** Memprediksi kemungkinan seseorang melakukan tindak kriminal.
5. **Pendidikan:** Memprediksi kelulusan seseorang mahasiswa

Regresi Logistik Multinomial

Regresi logistik multinomial dapat dilakukan saat variabel dependen memiliki skala lebih dari dua kategori. Contoh regresi logistik dengan skala lebih dari 2 kategori, model regresi dengan

variabel dependen yang memiliki skala nominal dalam tiga kategori dapat dikodekan hasil Y dengan kategori 0, 1, dan 2. Sebelum menggunakan variabel Y sebagai parameter yang memiliki dua fungsi logistik, harus ditentukan kategori variabel hasil mana yang akan dimanfaatkan sebagai pembanding. Yang biasa digunakan sebagai pembanding digunakan Y = 0. Dimana dalam membuat fungsi logit, akan membandingkan Y = 0 terhadap Y = 1 dan Y = 2.

Berikut ini merupakan rumus umum persamaan regresi logistik multinomial

$$P(Y=k|X) = \exp(\beta_k X) / \sum_{j=1}^K \exp(\beta_j X)$$

Dimana:

P(Y=k|X) : adalah probabilitas bahwa variabel dependen (Y) termasuk dalam kategori k, given nilai variabel independen (X).

K : adalah jumlah total kategori dalam variabel dependen.

X : adalah vektor variabel independen.

β_k : adalah vektor parameter (koefisien regresi) untuk kategori k.

exp() : adalah fungsi eksponensial.

Σ : adalah notasi sigma yang menandakan penjumlahan.

Rumus berikut menunjukkan bentuk model regresi logistik dengan p variabel prediktor:

$$\pi(x) = \frac{e^{\beta_0 + \beta_1 x_1 + \dots + \beta_p x_p}}{1 + e^{\beta_0 + \beta_1 x_1 + \dots + \beta_p x_p}}$$

Ada 2 jenis regresi logistik multinomial, yaitu:

1. Regresi logistik multinomial tunggal : apabila regresi logistik multinomial hanya mempunyai 1 variabel bebas
2. Regresi Logistik multinomial ganda : apabila regresi logistik multinomial mempunyai lebih dari 1 variabel bebas

Regresi Logistik Ordinal

Alat statistik yang disebut regresi ordinal, juga dikenal sebagai regresi logistik ordinal, digunakan untuk membuat model hubungan antara lebih dari dua kategori yang berurutan

dari variabel dependen yang berskala ordinal. Misalnya, regresi ordinal mengukur variabel dependen kepuasan pasien dengan tingkatan yang menggunakan jenis kelamin, usia, dan jenis ruang perawatan. Variabel dependen kepuasan pasien diukur dalam analisis regresi ordinal dengan empat skala Likert, yaitu: sangat puas, puas, tidak puas dan sangat tidak puas. Variabel demografi misalnya jenis kelamin dengan menggunakan 0 dan 1 untuk membedakan pria dan wanita, usia responden serta membedakan juga jenis ruang perawatan.

Dalam membuat model regresi ordinal, harus memperhatikan dua hal yaitu pilih fungsi hubungan serta variabel mana yang harus dimasukkan ke dalam model dalam rangka menunjukkan kesesuaian model. Jika skala interval atau rasio maka variabel independen atau bebas berupa kovariat dan jika skala nominal atau ordinal maka bias dianggap sebagai factor.

Model regresi logistik ordinal memodelkan logit dari peluang kumulatif sebagai fungsi linear dari variabel independen. Jika ada k kategori pada variabel dependen, maka akan ada $(k-1)$ persamaan logit. Persamaan umum untuk setiap persamaan logit adalah:

$$\text{logit}(P(Y \leq j|X)) = \alpha_j + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_p x_p$$

Dimana:

1. $P(Y \leq j|X)$ adalah peluang kumulatif kategori j atau lebih rendah, given nilai variabel independen X .
2. α_j adalah intercept untuk kategori j .
3. $\beta_1, \beta_2, \dots, \beta_p$ adalah koefisien regresi untuk variabel independen $X_1, X_2, \dots, X_p$.
4. $x_1, x_2, \dots, x_p$ adalah nilai variabel independen

Dimana Interpretasi Koefisien:

1. Koefisien regresi (β) menunjukkan perubahan dalam logit peluang untuk setiap peningkatan satu unit pada variabel independen, dengan asumsi variabel independen lainnya konstan.

2. Untuk menginterpretasikan efek variabel independen, biasanya digunakan odds ratio (OR), yang merupakan eksponensial dari koefisien regresi: $OR = \exp(\beta)$.
3. Odds ratio menunjukkan berapa kali lipat odds (peluang) suatu kategori atau kategori yang lebih rendah akan terjadi, untuk setiap peningkatan satu unit pada variabel independen.

DAFTAR PUSTAKA

- Amazon Web Service. 2024. *Apa Itu Regresi Logistik*.
<https://aws.amazon.com/id/what-is/logistic-regression/>
- Carolina I, Komarudin R dan Ramanda K, 2021. *Statistika*.
Yogyakarta: Graha Ilmu
- Hiola R, Yusuf H & Otok B.W. 2014. *Metode Regresi Terapan untuk Ilmu Kesehatan*. Yogyakarta: Deepublish
- Iba, B. R., & Wardhana, A. 2023. *Statistika Terapan untuk Mahasiswa Kesehatan*. Yogyakarta: Deepublish.
- Pezzullo J.C. 2013. *Biostatistic for Dummies*. Washington D.C: Georgetown University
- Reza W. 2024. *Buku Ajar Analisis Regresi : Pendekatan Praktis dan Sistematis*. Purbalingga: Eureka Media Aksara
- Roflin E, Riana F, Munarsih E, Pariyana & Liberty I.A. 2023. *Regresi Logistik Biner dan Multinomial*. Pekalongan: PT. Nasya Expanding Management (Penerbit NEM)
- Setyawan. D.A., Devriyani A, Huda N, Rahmadiliyani N, Patriyani R.E.H & Sulustyowati E.C. 2021. *Buku Ajar Statistika*. Indramayu: Penerbit Adab
- Susilo W.H, Aima H dan Suprpti F. 2021. *Biostatistik Lanjut dan Aplikasi Riset*. Jakarta: CV. Trans Info Media
- Suyono. 2015. *Analisis Regresi untuk Penelitian*. Yogyakarta: Deepublish
- Yani R.W.E, Permatasari E & Armiyanti Y. 2023. *Buku Ajar Biostatistik*. Jember: Repository Universitas Jember
- Yuantari M.G.C dan Handayani S. 2017. *Buku Ajar Biostatistik Deskriptif dan Inferensial*. Semarang: Badan Penerbit Universitas Dian Nuswantoro

BAB 7

METODE META ANALISIS DALAM KESEHATAN

Oleh Gede Wirabuana Putra

7.1 Konsep Dasar dan Prinsip Meta Analisis dalam Penelitian Kesehatan

Meta analisis merupakan metodologi penelitian kuantitatif yang telah menjadi instrumen fundamental dalam kedokteran berbasis bukti, memungkinkan peneliti untuk menggabungkan hasil dari berbagai studi independen guna menghasilkan estimasi efek yang lebih presisi dengan kekuatan statistik yang superior (Borenstein et al., 2021). Dalam konteks penelitian kesehatan, meta analisis berfungsi sebagai alat sintesis bukti ilmiah yang mengatasi keterbatasan ukuran sampel individual dan menghasilkan kesimpulan yang lebih dapat digeneralisasi terhadap populasi yang lebih luas. Pendekatan ini secara sistematis mengintegrasikan temuan dari penelitian-penelitian terpisah menggunakan metode statistik formal untuk mengurangi ketidakpastian dan meningkatkan presisi estimasi efek (Higgins et al., 2023).

Perbedaan mendasar antara meta analisis dengan tinjauan naratif terletak pada pendekatan sistematis dan kuantitatif yang digunakan (Page et al., 2021). Sementara tinjauan naratif mengandalkan interpretasi subjektif penulis dalam mensintesis temuan, meta analisis menggunakan metode statistik yang objektif dan terstandarisasi untuk menggabungkan hasil dari studi-studi yang memenuhi kriteria inklusi yang telah ditetapkan secara ketat. Karakteristik ini memberikan tingkat objektivitas yang jauh lebih tinggi dan secara signifikan mengurangi bias interpretasi yang sering terjadi dalam tinjauan tradisional.

Sebagai ilustrasi aplikasi praktis, dapat dipertimbangkan penelitian tentang efektivitas terapi musik untuk mengurangi kecemasan pada pasien bedah. Studi individual menunjukkan hasil

yang bervariasi dengan ukuran efek berkisar dari kecil hingga besar, menciptakan dilema bagi praktisi klinis dalam pengambilan keputusan. Meta analisis dari 23 penelitian dengan total 1.856 partisipan menghasilkan ukuran efek gabungan sebesar -0,84 dengan interval kepercayaan 95% (-1,12 hingga -0,56), memberikan bukti yang kuat tentang efektivitas intervensi musik dalam mengurangi kecemasan pra bedah. Temuan ini memiliki implikasi klinis yang signifikan karena memberikan dasar empiris yang solid untuk implementasi terapi musik sebagai intervensi non farmakologis dalam setting perioperatif.

Tabel 7.1. Perbandingan Metode Sintesis Bukti Ilmiah

Karakteristik	Tinjauan Naratif	Tinjauan Sistematis	Meta Analisis
Pendekatan pencarian	Tidak sistematis	Sistematis komprehensif	Sistematis komprehensif
Kriteria seleksi	Subjektif	Objektif terstruktur	Objektif terstruktur
Sintesis data	Deskriptif kualitatif	Kualitatif sistematis	Kuantitatif statistik
Kontrol bias	Minimal	Moderat	Tinggi
Kekuatan statistik	Rendah	Sedang	Tinggi
Presisi estimasi	Rendah	Sedang	Tinggi
Reproduksibilitas	Rendah	Tinggi	Sangat tinggi

Dalam hierarki piramida bukti ilmiah, meta analisis menempati posisi puncak bersama dengan uji klinis acak terkontrol skala besar. Posisi strategis ini mencerminkan kemampuan meta analisis untuk menyediakan bukti dengan tingkat kepercayaan tertinggi yang diperlukan untuk pengambilan keputusan klinis yang berbasis bukti. Sebagai ilustrasi konkret, meta analisis Cochrane tentang efektivitas vaksinasi influenza pada populasi lansia berhasil mengintegrasikan 75 studi dengan total 5,6 juta partisipan, menghasilkan estimasi efektivitas yang jauh lebih presisi dan dapat diandalkan dibandingkan dengan studi individual manapun.

Prinsip fundamental meta analisis didasarkan pada asumsi bahwa studi-studi yang dianalisis mengukur konstruk

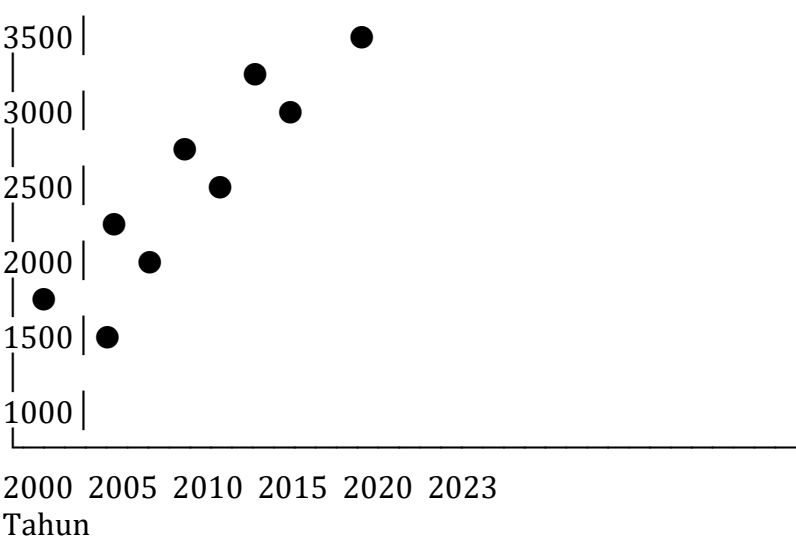
yang sama dalam populasi yang secara klinis dapat dibandingkan. Asumsi homogenitas ini memungkinkan penggabungan hasil statistik yang bermakna secara metodologis. Namun demikian, dalam praktik penelitian sesungguhnya, heterogenitas antar studi sering terjadi dan harus dikelola dengan tepat melalui teknik statistik yang sesuai untuk memastikan validitas kesimpulan yang dihasilkan.

Kasus yang menarik untuk dipahami adalah meta analisis tentang efektivitas suplementasi vitamin D untuk pencegahan infeksi saluran pernapasan. Penelitian individual menunjukkan hasil yang sangat beragam, mulai dari tidak ada efek hingga pengurangan risiko yang substansial. Heterogenitas ini terjadi karena perbedaan dosis, durasi suplementasi, karakteristik populasi, dan tingkat defisiensi vitamin D baseline (Simanjourang et al., 2023). Meta analisis dari 25 uji klinis acak terkontrol dengan 11.321 partisipan menggunakan pendekatan model efek acak untuk mengakomodasi heterogenitas, menghasilkan rasio odds 0,88 dengan interval kepercayaan 95% (0,81-0,96), menunjukkan pengurangan risiko infeksi sebesar 12%.

Konsep validitas internal dan eksternal menjadi pertimbangan krusial dalam interpretasi hasil meta analisis. Validitas internal berkaitan dengan kualitas metodologis studi-studi individual yang dimasukkan, termasuk randomisasi yang tepat, penyamaran yang memadai, dan minimisasi bias seleksi. Validitas eksternal atau generalisabilitas berkaitan dengan sejauh mana temuan meta analisis dapat diterapkan pada populasi dan setting klinis yang berbeda dari studi-studi yang dianalisis.

Perkembangan teknologi informasi telah merevolusi cara meta analisis dilakukan, dengan tersedianya perangkat lunak khusus seperti Review Manager, Comprehensive Meta-Analysis, dan paket statistik R yang memungkinkan analisis yang lebih canggih dan visualisasi yang lebih informatif. Otomatisasi proses pencarian dan ekstraksi data melalui pembelajaran mesin dan pemrosesan bahasa alami mulai diterapkan untuk meningkatkan efisiensi dan mengurangi bias manusia dalam tahapan tinjauan sistematis.

Jumlah Publikasi Meta Analisis



Gambar 7.1. Grafik Evolusi Jumlah Meta Analisis dalam Bidang Kesehatan (2000-2023)

Meta analisis data pasien individual merupakan pendekatan yang semakin berkembang dimana data mentah dari studi-studi individual dikumpulkan dan dianalisis ulang secara terpusat. Pendekatan ini memungkinkan analisis subkelompok yang lebih presisi dan identifikasi faktor moderator yang tidak dapat dideteksi melalui meta analisis konvensional yang hanya menggunakan data agregat. Sebagai contoh, meta analisis data pasien individual tentang efektivitas statin untuk pencegahan kardiovaskular memungkinkan identifikasi ambang batas kolesterol LDL yang optimal untuk berbagai kelompok usia dan jenis kelamin (Higgins et al., 2023)..

Tantangan metodologis dalam meta analisis modern mencakup penanganan data yang hilang, bias publikasi, dan heterogenitas yang kompleks. Teknik imputasi ganda untuk data yang hilang, analisis sensitivitas yang ekstensif, dan penggunaan model Bayesian untuk mengakomodasi ketidakpastian

parameter menjadi pendekatan yang semakin diadopsi untuk meningkatkan ketahanan temuan meta analisis.

7.2 Metodologi Pencarian dan Seleksi Studi untuk Meta Analisis

Strategi pencarian literatur dalam meta analisis merupakan fondasi metodologis yang menentukan validitas dan komprehensivitas tinjauan sistematis. Pendekatan yang sistematis dan dapat direproduksi dalam pencarian literatur tidak hanya memastikan identifikasi semua studi yang relevan, tetapi juga meminimalkan bias seleksi yang dapat mengkompromikan validitas kesimpulan akhir. Kerangka kerja PICO (*Population, Intervention, Comparison, Outcome*) menyediakan struktur yang terorganisir untuk mengembangkan strategi pencarian yang fokus dan efektif dalam mengidentifikasi studi-studi yang memenuhi kriteria inklusi yang telah ditetapkan.

Formulasi pertanyaan penelitian menggunakan kerangka kerja PICO memungkinkan operasionalisasi yang presisi terhadap komponen-komponen penelitian yang akan dicari (Liberati et al., 2009). Sebagai contoh konkret, pertanyaan penelitian "Apakah terapi kognitif perilaku efektif untuk mengurangi gejala depresi pada remaja?" dapat diuraikan menjadi komponen-komponen yang spesifik. *Population* mengacu pada remaja usia 12-18 tahun dengan diagnosis depresi, *Intervention* berupa terapi kognitif perilaku, *Comparison* mencakup terapi konvensional, plasebo, atau daftar tunggu, dan *Outcome* berupa skor depresi yang diukur dengan instrumen yang tervalidasi. Formulasi yang tepat ini akan mengarahkan pemilihan kata kunci pencarian dan basis data yang sesuai untuk memaksimalkan sensitivitas dan spesifisitas pencarian.

Pengembangan strategi pencarian memerlukan keseimbangan yang cermat antara sensitivitas dan spesifisitas. Sensitivitas tinggi memastikan identifikasi semua studi yang potensial relevan namun menghasilkan banyak hasil yang tidak relevan dan meningkatkan beban kerja penyaringan. Sebaliknya, spesifisitas tinggi mengurangi noise namun berisiko melewatkan

studi penting yang dapat mempengaruhi kesimpulan akhir. Pendekatan iteratif dengan pengujian awal menggunakan set studi referensi yang diketahui relevan membantu mengoptimalkan strategi pencarian untuk mencapai keseimbangan optimal antara sensitivitas dan spesifisitas.

Tabel 7.2. Karakteristik Basis Data Literatur Utama untuk Meta Analisis

Basis Data	Cakupan Temporal	Keunggulan Utama	Kosakata Terkontrol	Estimasi Artikel
PubMed/MEDLINE	1946-sekarang	Literatur biomedis umum	MeSH (Medical Subject Headings)	>35 juta
EMBASE	1947-sekarang	Farmakologi, toksikologi	EMTREE	>42 juta
Cochrane CENTRAL	1996-sekarang	Uji klinis terkurasi	MeSH + manual	>1,8 juta
Web of Science	1900-sekarang	Multidisiplin, sitasi	Tidak ada	>180 juta
CINAHL	1981-sekarang	Keperawatan, kesehatan sekutu	CINAHL Subject Headings	>6 juta

Pencarian literatur telah menjadi komponen yang semakin penting dalam meta analisis kontemporer untuk mengurangi bias publikasi yang dapat mempengaruhi validitas kesimpulan. Repositori institusi, register uji klinis seperti ClinicalTrials.gov dan Platform Register Uji Klinis Internasional WHO, serta basis data disertasi dan tesis menyediakan akses terhadap studi yang tidak dipublikasikan dalam jurnal yang melalui tinjauan sejawat. Sebagai ilustrasi dampaknya, analisis komprehensif register uji klinis untuk meta analisis tentang efektivitas obat antihipertensi baru mengidentifikasi 23% studi tambahan yang tidak dipublikasikan, yang secara signifikan mengubah estimasi efek gabungan dari 0,73 menjadi 0,81.

Proses seleksi studi melibatkan penyaringan bertahap yang dimulai dari evaluasi judul, kemudian abstrak, dan akhirnya teks lengkap artikel. Penggunaan dua peninjau independen

dengan perhitungan statistik kappa Cohen untuk mengukur keandalan antar penilai telah menjadi standar metodologis yang diterima secara universal (Page et al., 2021). Nilai kappa di atas 0,7 menunjukkan kesepakatan yang baik antar peninjau, sementara nilai di bawah 0,4 mengindikasikan perlunya klarifikasi kriteria seleksi atau pelatihan tambahan untuk meningkatkan konsistensi penilaian.

Kasus yang ilustratif untuk dipahami adalah meta analisis tentang efektivitas intervensi berbasis mindfulness untuk mengurangi stres pada tenaga kesehatan. Pencarian awal di lima basis data menghasilkan 3.247 artikel potensial. Setelah duplikasi dihilangkan, tersisa 2.891 artikel yang disaring berdasarkan judul dan abstrak. Dari proses ini, 156 artikel dinilai layak untuk evaluasi teks lengkap. Akhirnya, 28 studi memenuhi semua kriteria inklusi dan dimasukkan dalam meta analisis. Kappa agreement antar peninjau untuk tahap penyaringan abstrak adalah 0,82, menunjukkan konsistensi yang sangat baik dalam penerapan kriteria seleksi (Higgins et al., 2023)..

Ekstraksi data merupakan tahapan kritis yang memerlukan standarisasi format dan prosedur untuk memastikan konsistensi dan kelengkapan informasi yang dikumpulkan. Formulir ekstraksi data harus mencakup karakteristik studi (desain, setting, durasi), karakteristik partisipan (demografi, kriteria inklusi/eksklusi), detail intervensi, pengukuran luaran, dan hasil numerik yang diperlukan untuk analisis statistik. Ekstraksi ganda oleh peninjau independen dengan resolusi diskrepansi melalui diskusi atau konsultasi peninjau ketiga memastikan akurasi dan reliabilitas data yang digunakan untuk analisis (Page et al., 2021).

Penilaian kualitas dan risiko bias studi individual menggunakan instrumen yang tervalidasi seperti alat Risk of Bias Cochrane untuk uji klinis atau Newcastle-Ottawa Scale untuk studi observasional. Penilaian ini tidak hanya menginformasikan interpretasi hasil tetapi juga dapat digunakan dalam analisis sensitivitas untuk mengevaluasi ketahanan temuan terhadap variasi kualitas studi yang dimasukkan (Higgins et al., 2023).

Tabel 7.3. Instrumen Penilaian Kualitas Studi Berdasarkan Desain Penelitian

Desain Studi	Instrumen Penilaian	Domain Utama	Skala Penilaian
RCT	Cochrane RoB 2.0	Randomisasi, penyamaran, luaran, pelaporan selektif	Rendah/Tinggi/Tidak jelas
Kohort	Newcastle-Ottawa	Seleksi, komparabilitas, luaran	0-9 bintang
Kasus-Kontrol	Newcastle-Ottawa	Seleksi, komparabilitas, eksposur	0-9 bintang
Potong lintang	Alat AXIS	Metode, bias, perancu	Ya/Tidak/Tidak jelas

Teknologi kecerdasan buatan dan pembelajaran mesin mulai diintegrasikan dalam proses penyaringan dan ekstraksi data untuk meningkatkan efisiensi dan mengurangi beban kerja manual. Alat seperti Rayyan, Covidence, dan ASReview menggunakan algoritma pembelajaran mesin untuk memprioritaskan artikel yang kemungkinan besar relevan, secara signifikan mengurangi waktu yang diperlukan untuk penyaringan. Namun demikian, supervisi manusia tetap diperlukan untuk memastikan akurasi dan menangani kasus-kasus yang ambigu yang memerlukan pertimbangan kontekstual.

Dokumentasi proses pencarian dan seleksi harus mengikuti standar PRISMA (*Preferred Reporting Items for Systematic Reviews and Meta-Analyses*) untuk memastikan transparansi dan reproduksibilitas. Diagram alur PRISMA memberikan visualisasi yang jelas tentang jumlah studi pada setiap tahap proses seleksi, alasan eksklusi, dan hasil akhir yang dimasukkan dalam analisis. Dokumentasi yang komprehensif ini memungkinkan pembaca untuk mengevaluasi kualitas metodologis tinjauan dan memfasilitasi pemutakhiran tinjauan

masa depan yang menggunakan strategi pencarian yang sama (Higgins et al., 2023).

Tantangan khusus dalam pencarian literatur untuk topik kesehatan global mencakup akses terhadap publikasi dalam bahasa non-Inggris dan literatur dari negara berkembang yang mungkin tidak terindeks dalam basis data mainstream. Strategi pencarian yang inklusif harus mempertimbangkan basis data regional seperti LILACS untuk literatur Amerika Latin, WPRIM untuk Asia Pasifik, dan AIM untuk Afrika. Kolaborasi dengan peneliti lokal dan institusi dapat memfasilitasi akses terhadap literatur yang tidak terpublikasi secara internasional namun relevan untuk konteks global.

7.3 Analisis Statistik dan Interpretasi Hasil Meta Analisis

Analisis statistik dalam meta analisis menyediakan kerangka kerja matematika yang canggih untuk menggabungkan ukuran efek dari studi-studi independen dengan cara yang secara metodologis valid dan statistik dapat diandalkan. Pemilihan model statistik yang tepat merupakan keputusan kritis yang bergantung pada asumsi tentang sifat heterogenitas antar studi dan karakteristik distribusi data yang akan dianalisis. Dua paradigma utama yang mendominasi praktik meta analisis adalah model efek tetap dan model efek acak, masing-masing dengan implikasi teoritis dan praktis yang berbeda terhadap interpretasi dan generalisabilitas hasil.

Model efek tetap beroperasi berdasarkan asumsi fundamental bahwa semua studi yang dianalisis mengukur ukuran efek populasi yang identik, dan variasi yang diamati hanya disebabkan oleh kesalahan sampling acak yang melekat dalam setiap studi (Riley et al., 2011). Model ini memberikan bobot yang proporsional dengan presisi studi, yang secara praktis berarti studi dengan sampel besar dan galat baku kecil akan mendominasi kontribusi terhadap estimasi gabungan. Sebagai contoh ilustratif, dalam meta analisis efektivitas statin untuk pencegahan primer kardiovaskular, model efek tetap memberikan bobot sebesar 45% kepada studi JUPITER dengan

17.802 partisipan, sementara studi yang lebih kecil dengan 500 partisipan hanya mendapat bobot 2%.

Sebaliknya, model efek acak mengakui bahwa ukuran efek sesungguhnya dapat bervariasi antar studi sebagai konsekuensi dari perbedaan populasi, variasi implementasi intervensi, atau heterogenitas dalam kondisi studi. Model ini menambahkan komponen varians antar studi (tau kuadrat) dalam perhitungan bobot, menghasilkan interval kepercayaan yang lebih lebar dan konservatif yang lebih tepat mencerminkan ketidakpastian yang melekat dalam heterogenitas antar studi. Dalam contoh meta analisis yang sama, model efek acak menghasilkan estimasi risiko relatif 0,76 dengan interval kepercayaan 95% (0,69-0,84) dibandingkan 0,73 dengan interval kepercayaan 95% (0,70-0,77) dari model efek tetap.

Tabel 7.4. Perbandingan Komprehensif Model Efek Tetap vs Efek Acak

Aspek	Model Efek Tetap	Model Efek Acak
Asumsi dasar	Satu ukuran efek populasi	Distribusi ukuran efek
Perhitungan bobot	1/variens dalam studi	1/(variens dalam + antar studi)
Interval kepercayaan	Lebih sempit	Lebih lebar
Sensitivitas terhadap studi besar	Tinggi	Sedang
Generalisabilitas	Terbatas pada studi yang dianalisis	Lebih luas ke populasi studi
Interpretasi	Ukuran efek rata-rata	Ukuran efek rata-rata populasi
Kesesuaian	Heterogenitas rendah	Heterogenitas tinggi

Perhitungan ukuran efek memerlukan standarisasi metrik untuk memungkinkan penggabungan hasil dari studi dengan skala pengukuran yang potensial berbeda. Untuk data kontinu, Cohen's d dan Hedges' g merupakan ukuran efek terstandar yang paling banyak diadopsi dalam literatur meta

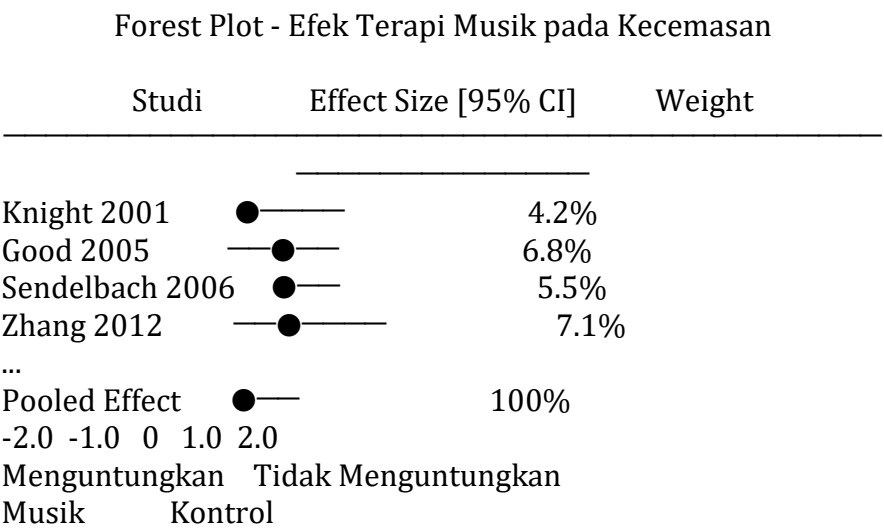
analisis. Hedges' g memberikan koreksi bias sampel kecil yang secara statistik penting untuk studi dengan ukuran sampel terbatas, menghasilkan estimasi yang kurang bias terutama ketika $n < 20$ per kelompok. Formula Hedges' g adalah $g = d \times (1 - 3/(4df-1))$, dimana d adalah Cohen's d dan df adalah derajat kebebasan (Page et al., 2021).

Untuk data dikotomis, rasio odds, rasio risiko, dan perbedaan risiko masing-masing memiliki interpretasi dan asumsi yang berbeda yang harus dipertimbangkan dalam konteks klinis spesifik. Rasio odds cenderung memberikan estimasi efek yang lebih besar terutama untuk luaran yang umum terjadi, sementara rasio risiko memberikan interpretasi yang lebih intuitif untuk praktisi klinis. Perbedaan risiko memberikan informasi tentang manfaat absolut yang langsung relevan untuk perhitungan jumlah yang perlu diobati.

Sistem pembobotan menggunakan kebalikan varians merupakan pendekatan optimal yang memastikan bahwa studi dengan presisi lebih tinggi berkontribusi proporsional lebih besar terhadap estimasi gabungan. Dalam model efek tetap, bobot dihitung sebagai $w_i = 1/SE_i^2$, dimana SE_i adalah galat baku studi ke- i (DerSimonian & Laird, 1986). Model efek acak memodifikasi perhitungan ini menjadi $w_i = 1/(SE_i^2 + \tau^2)$, dimana τ^2 adalah varians antar studi yang diestimasi menggunakan metode seperti DerSimonian-Laird, restricted maximum likelihood (REML), atau metode Paule-Mandel yang lebih mutakhir.

Sebagai contoh aplikasi praktis yang komprehensif, meta analisis tentang efektivitas terapi musik untuk mengurangi kecemasan pada pasien yang akan menjalani pembedahan mengombinasikan 23 studi dengan ukuran sampel yang bervariasi dari 30 hingga 372 partisipan. Heterogenitas yang signifikan dengan $I^2 = 68\%$ mengindikasikan perlunya penggunaan model efek acak. Ukuran efek gabungan menunjukkan perbedaan rerata terstandar sebesar -0,84 dengan interval kepercayaan 95% (-1,12 hingga -0,56), mengindikasikan efek besar yang secara klinis bermakna. Analisis sensitivitas dengan secara sistematis mengeluarkan studi pencilan

mengurangi heterogenitas menjadi $I^2 = 42\%$ tanpa mengubah signifikansi temuan secara substansial, menunjukkan ketahanan hasil terhadap variasi metodologis (Page et al., 2021).



Gambar 7.2. Grafik Distribusi Ukuran Efek dalam Meta Analisis Terapi Musik

Penilaian heterogenitas menggunakan statistik I^2 dan Q Cochran memberikan kerangka kerja kuantitatif untuk mengevaluasi kesesuaian penggabungan dan memilih model analisis yang tepat (Higgins & Thompson, 2002). Interpretasi konvensional mengkategorikan $I^2 < 25\%$ sebagai heterogenitas rendah, 25-50% sebagai sedang, 50-75% sebagai substansial, dan $> 75\%$ sebagai sangat tinggi. Namun, interpretasi ini harus selalu mempertimbangkan konteks klinis dan signifikansi praktis dari perbedaan yang diamati.

Meta regresi memungkinkan eksplorasi sumber heterogenitas dengan examine hubungan antara kovariat tingkat studi dan ukuran efek yang diamati. Teknik ini sangat berharga untuk mengidentifikasi variabel moderator yang dapat menjelaskan variasi dalam efek pengobatan antar studi. Sebagai contoh, meta regresi dalam studi efektivitas intervensi olahraga

untuk depresi menunjukkan bahwa durasi intervensi secara signifikan memprediksi magnitudo ukuran efek ($\beta = -0,12$ per minggu, $p = 0,03$), dengan intervensi yang lebih lama menghasilkan manfaat yang lebih besar (Page et al., 2021).

Analisis sensitivitas merupakan komponen kritis untuk mengevaluasi ketahanan temuan terhadap keputusan metodologis dan variasi kualitas studi. Analisis leave-one-out, analisis pembatasan berdasarkan kualitas studi, dan model statistik alternatif dapat mengungkapkan kerapuhan atau stabilitas kesimpulan. Meta analisis yang tahan uji harus menunjukkan temuan yang konsisten di berbagai rentang analisis sensitivitas yang wajar. Sebagai ilustrasi, meta analisis efektivitas terapi kognitif perilaku untuk gangguan kecemasan menunjukkan hasil yang konsisten ketika analisis dibatasi pada RCT berkualitas tinggi (ukuran efek = 0,68) dibandingkan dengan semua studi yang dimasukkan (ukuran efek = 0,71), mengindikasikan ketahanan temuan.

Penilaian bias publikasi menggunakan funnel plot dan uji statistik seperti uji Egger dan uji Begg memberikan wawasan tentang potensi pelaporan selektif yang dapat membiaskan hasil meta analisis (Egger et al., 1997; Sterne et al., 2011). Metode trim-and-fill dapat memberikan estimasi yang disesuaikan ketika bias publikasi terdeteksi, meskipun interpretasi harus berhati-hati karena asumsi yang mendasari metode ini tidak selalu terpenuhi dalam praktik.

7.4 Aplikasi Meta Analisis dalam Evidence-Based Medicine dan Kebijakan Kesehatan

Meta analisis telah menjadi instrumen yang tidak terpisahkan dalam kedokteran berbasis bukti dan formulasi kebijakan kesehatan, menyediakan fondasi empiris yang kuat untuk pengambilan keputusan klinis dan administratif. Dalam konteks kedokteran berbasis bukti, meta analisis berperan sebagai sintesis bukti tingkat tertinggi yang memungkinkan klinisi untuk membuat keputusan yang diinformasikan berdasarkan agregasi bukti terbaik yang tersedia. Pendekatan ini tidak hanya meningkatkan kualitas perawatan pasien tetapi juga

berkontribusi pada pengembangan pedoman praktik klinis yang berbasis bukti dan standardisasi protokol perawatan di berbagai setting pelayanan kesehatan (Page et al., 2021).

Implementasi meta analisis dalam pengembangan pedoman praktik klinis memerlukan integrasi yang hati-hati antara bukti kuantitatif dengan pertimbangan klinis, nilai pasien, dan konteks lokal. Organisasi profesional seperti American Heart Association, European Society of Cardiology, dan World Health Organization secara rutin menggunakan temuan meta analisis sebagai dasar untuk rekomendasi mereka. Sebagai contoh konkret, pedoman tatalaksana hipertensi yang dikeluarkan oleh Joint National Committee (JNC) mengintegrasikan temuan dari meta analisis yang melibatkan lebih dari 600.000 pasien untuk menetapkan target tekanan darah optimal dan strategi pengobatan lini pertama.

Dalam domain kebijakan kesehatan masyarakat, meta analisis menyediakan evidence base yang diperlukan untuk alokasi sumber daya yang efisien dan pengembangan program intervensi yang efektif. Pembuat kebijakan menggunakan temuan meta analisis untuk menentukan prioritas kesehatan masyarakat, merancang program pencegahan, dan mengevaluasi efektivitas biaya dari berbagai intervensi kesehatan (Alkafi, 2022). Meta analisis tentang efektivitas program vaksinasi HPV, misalnya, telah menginformasikan keputusan kebijakan di lebih dari 80 negara mengenai implementasi program vaksinasi nasional dan strategi targeting populasi.

Tabel 7.5. Aplikasi Meta Analisis dalam Berbagai Domain Kesehatan

Domain Aplikasi	Contoh Intervensi	Dampak Kebijakan	Ukuran Efek Tipikal
Farmakologi Klinis	Efektivitas antidepresan	Formularium nasional	0,3-0,8 (SMD)
Kesehatan Masyarakat	Program vaksinasi	Kebijakan imunisasi wajib	0,7-0,9 (RR)
Pencegahan Penyakit	Skrining kanker	Pedoman skrining nasional	0,8-0,95 (HR)
Rehabilitasi	Terapi fisik stroke	Standar perawatan rehabilitasi	0,4-0,7 (SMD)
Kesehatan Mental	Intervensi psikososial	Program kesehatan mental komunitas	0,2-0,6 (SMD)

Evaluasi ekonomi kesehatan berdasarkan meta analisis telah menjadi komponen krusial dalam pengambilan keputusan alokasi sumber daya di sektor kesehatan. Meta analisis ekonomi mengintegrasikan data efektivitas dengan informasi biaya untuk menghasilkan estimasi rasio efektivitas biaya yang dapat digunakan untuk membandingkan intervensi yang berbeda. *Health Technology Assessment* (HTA) agencies di berbagai negara seperti NICE di Inggris, CADTH di Kanada, dan PBAC di Australia secara sistematis menggunakan meta analisis ekonomi untuk menginformasikan keputusan cakupan asuransi kesehatan dan penetapan harga obat.

Kasus yang ilustratif adalah meta analisis ekonomi tentang program skrining mammografi untuk kanker payudara yang mengintegrasikan data dari 11 uji klinis acak terkontrol dengan total 600.000 wanita. Analisis ini menunjukkan bahwa skrining mammografi pada wanita usia 50-69 tahun menghasilkan *Incremental Cost-Effectiveness Ratio* (ICER) sebesar \$35.000 per *Quality-Adjusted Life Year* (QALY), yang berada di bawah ambang batas yang dapat diterima di sebagian

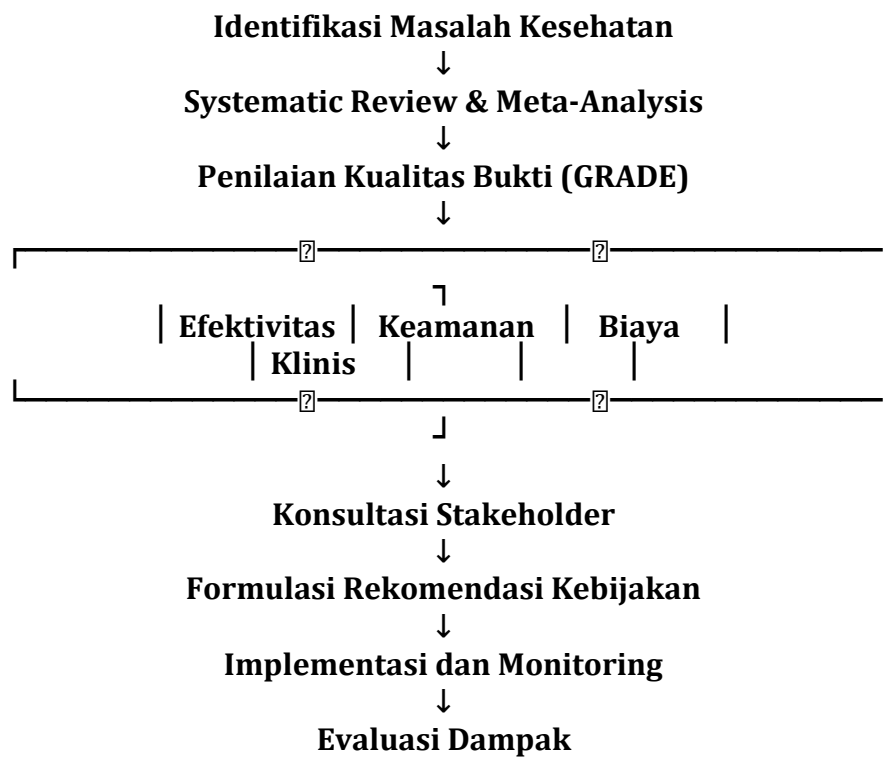
besar negara maju. Temuan ini telah menginformasikan kebijakan skrining kanker payudara di lebih dari 50 negara dan berkontribusi pada standarisasi protocol skrining internasional.

Tantangan dalam aplikasi meta analisis untuk kebijakan kesehatan mencakup generalisabilitas temuan antar populasi yang berbeda, variasi dalam sistem pelayanan kesehatan, dan pertimbangan konteks lokal yang mungkin tidak tercermin dalam studi-studi yang dianalisis (Azizah & Fadillah, 2022). Adaptasi temuan meta analisis untuk konteks lokal memerlukan pertimbangan faktor-faktor seperti prevalensi penyakit, karakteristik demografis populasi, infrastruktur kesehatan yang tersedia, dan preferensi budaya. Sebagai contoh, meta analisis tentang efektivitas intervensi pencegahan HIV yang dikembangkan berdasarkan studi di negara maju mungkin memerlukan adaptasi signifikan untuk implementasi di negara berkembang dengan prevalensi HIV yang tinggi dan sumber daya terbatas.

Peran meta analisis dalam pengembangan guideline evidence-based telah berkembang menjadi proses yang semakin sistematis dan transparan. Framework GRADE (*Grading of Recommendations Assessment, Development and Evaluation*) menyediakan metodologi terstruktur untuk menerjemahkan temuan meta analisis menjadi rekomendasi klinis dengan tingkat kepercayaan yang explicit. Sistem ini mengevaluasi kualitas bukti berdasarkan lima domain: risiko bias, inkonsistensi, ketidaklangsungan bukti, ketidakpresisian, dan bias publikasi, kemudian mengintegrasikan informasi ini dengan pertimbangan nilai, preferensi, dan sumber daya untuk menghasilkan rekomendasi yang dapat ditindaklanjuti (Liberati et al., 2009).

Meta analisis jaringan (*network meta-analysis*) merupakan pengembangan metodologis yang memungkinkan perbandingan simultan multiple intervensi bahkan ketika tidak semua intervensi dibandingkan secara langsung dalam studi individual. Pendekatan ini sangat berharga dalam konteks kebijakan kesehatan dimana pembuat keputusan perlu membandingkan efektivitas relatif dari beberapa alternatif

intervensi. Sebagai contoh, meta analisis jaringan dari 42 studi tentang pengobatan depresi mayor mengintegrasikan 21 antidepresan berbeda, memungkinkan ranking efektivitas dan tolerabilitas yang komprehensif yang tidak mungkin dicapai melalui meta analisis pairwise tradisional.



Gambar 7.3. Grafik Proses Integrasi Meta Analisis dalam Kebijakan Kesehatan

Implementasi temuan meta analisis dalam praktik klinis seringkali menghadapi hambatan yang signifikan, termasuk resistensi terhadap perubahan praktik, keterbatasan sumber daya, dan kompleksitas translasi bukti agregat ke pengambilan keputusan individual pasien. Penelitian implementasi telah menunjukkan bahwa strategi passive seperti publikasi guideline saja tidak efektif, dan diperlukan pendekatan aktif yang

melibatkan opinion leaders lokal dan integrasi rekomendasi ke dalam sistem informasi klinis.

Aspek etis dalam aplikasi meta analisis untuk kebijakan kesehatan mencakup pertimbangan tentang keadilan distribusi manfaat kesehatan, transparansi dalam proses pengambilan keputusan, dan akuntabilitas terhadap bukti ilmiah. Meta analisis dapat mengungkapkan disparitas dalam efektivitas intervensi antar subpopulasi yang berbeda, menimbulkan pertanyaan etis tentang bagaimana alokasi sumber daya harus mempertimbangkan perbedaan ini. Analisis subkelompok dalam meta analisis tentang efektivitas terapi kardiovaskular, misalnya, telah mengungkapkan perbedaan respons yang signifikan berdasarkan ras, jenis kelamin, dan status sosial ekonomi, menginformasikan kebijakan yang lebih equitable (Liberati et al., 2009).

Masa depan aplikasi meta analisis dalam kebijakan kesehatan akan semakin didorong oleh perkembangan dalam big data analytics, real-world evidence, dan artificial intelligence. Living systematic reviews dan living meta-analyses yang diperbarui secara kontinyu dengan bukti baru akan memungkinkan kebijakan kesehatan yang lebih responsif terhadap perkembangan ilmiah terkini. Platform digital yang mengintegrasikan meta analisis dengan data surveillance kesehatan real-time akan memfasilitasi evidence-informed policy making yang lebih dinamis dan adaptive.

Tantangan metodologis yang berkelanjutan dalam aplikasi meta analisis mencakup penanganan heterogenitas yang kompleks, integrasi bukti dari berbagai jenis studi, dan pengembangan metode untuk mengakomodasi ketidakpastian dalam proyeksi jangka panjang. Pendekatan Bayesian yang semakin sophisticated memungkinkan integrasi prior knowledge dengan bukti empirical, sementara machine learning techniques dapat membantu identifikasi pattern yang kompleks dalam heterogenitas antar studi (Page et al., 2021).

Kolaborasi internasional dalam meta analisis kesehatan global telah memfasilitasi pengembangan bukti yang relevan untuk berbagai konteks. Cochrane Collaboration, Campbell

Collaboration, dan initiative seperti Evidence for Policy and Practice Information merupakan contoh upaya coordinated untuk mensintesis bukti berkualitas tinggi yang dapat menginformasikan kebijakan di berbagai setting (Putra & Pradnyani, 2022; Suryadi et al., 2023). Standardisasi metodologi dan pelaporan melalui guidelines seperti PRISMA dan CONSORT telah meningkatkan kualitas dan comparability meta analisis antar peneliti dan institusi.

DAFTAR PUSTAKA

- Alkafi, M. (2022). Meta-analisis risiko hipertensi dengan penyakit stroke di Asia. *JIK Jurnal Ilmu Kesehatan*, 6(2), 245-252. DOI: 10.35816/jik.v6i2.539
- Azizah, R., & Fadillah, G. (2022). Analisis faktor risiko perilaku dengan kasus malaria pada masyarakat di Indonesia - meta analysis 2016-2021: Literature review. *Media Publikasi Promosi Kesehatan Indonesia (MPPKI)*, 5(11), 1336-1345. DOI: 10.56338/mppki.v5i11.2733
- Borenstein, M., Hedges, L. V., Higgins, J. P., & Rothstein, H. R. (2021). *Introduction to meta-analysis* (2nd ed.). Chichester: John Wiley & Sons. DOI: 10.1002/9781119558378
- DerSimonian, R., & Laird, N. (1986). Meta-analysis in clinical trials. *Controlled Clinical Trials*, 7(3), 177-188. DOI: 10.1016/0197-2456(86)90046-2
- Egger, M., Smith, G. D., Schneider, M., & Minder, C. (1997). Bias in meta-analysis detected by a simple, graphical test. *BMJ*, 315(7109), 629-634. DOI: 10.1136/bmj.315.7109.629
- Higgins, J. P., & Thompson, S. G. (2002). Quantifying heterogeneity in a meta-analysis. *Statistics in Medicine*, 21(11), 1539-1558. DOI: 10.1002/sim.1186
- Higgins, J. P., Thomas, J., Chandler, J., Cumpston, M., Li, T., Page, M. J., & Welch, V. A. (Eds.). (2023). *Cochrane handbook for systematic reviews of interventions* (Version 6.4). London: Cochrane. <https://training.cochrane.org/handbook>
- Liberati, A., Altman, D. G., Tetzlaff, J., Mulrow, C., Gøtzsche, P. C., Ioannidis, J. P., Clarke, M., Devereaux, P. J., Kleijnen, J., & Moher, D. (2009). The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate health care interventions: explanation and elaboration. *PLoS Medicine*, 6(7), e1000100. DOI: 10.1371/journal.pmed.1000100
- Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., Shamseer, L., Tetzlaff, J. M., Akl, E. A., Brennan, S. E., Chou, R., Glanville, J., Grimshaw, J. M.,

- Hróbjartsson, A., Lalu, M. M., Li, T., Loder, E. W., Mayo-Wilson, E., McDonald, S., McGuinness, L. A., Stewart, L. A., Thomas, J., Tricco, A. C., Welch, V. A., Whiting, P., & Moher, D. (2021). The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. *BMJ*, 372, n71. DOI: 10.1136/bmj.n71
- Putra, G. W., & Pradnyani, P. E. (2022). Determinan keberhasilan pengobatan pasien tuberculosis di Kota Denpasar tahun 2021. *Indonesian Journal of Health Information Management (INOHIM)*, 10(2), 125-132. DOI: 10.47007/inohim.v10i2.456
- Putra, G. W., Pradnyani, P. E., Paramita, I. A. P. F., Putri, P. C. S., & Romansyah, D. E. (2023). Analisis keakuratan kode diagnosis pada kasus tuberculosis berdasarkan ICD 10 di Rumah Sakit Prima Medika. *MAINTTEKKES: The Journal of Management Information and Health Technology*, 7(1), 89-96. DOI: 10.53579/maintekkes.v7i1.234
- Simanjorang, C., Tarigan, K. M., Sudrajat, D. A. C., Wangsawinangun, R. Z. R., & Zahra, A. S. A. (2023). Vaccination as a protective factor against hospitalization for SARS-CoV-2 infections: an observational study using surveillance data. *Jurnal Biostatistik, Kependudukan, dan Informatika Kesehatan (BIKFOKES)*, 5(2), 89-98. DOI: 10.7454/bikfokes.v5i2.1084
- Suryadi, I., Besral, B., & Anhari, A. (2023). Studi meta-analisis: faktor risiko kejadian pneumonia pada balita di Indonesia. *Journal of Community Mental Health and Public Policy*, 5(2), 134-145. DOI: 10.25311/jcmhp.v5i2.1456

BAB 8

ANALISIS DATA SPASIAL DALAM KESEHATAN MASYARAKAT

Oleh M. Habibullah Arief

8.1 Pendahuluan

Pengolahan data spasial semakin berkembang pesat dalam beberapa tahun terakhir, terutama dengan adanya teknologi yang semakin maju dan terjangkau seperti Sistem Informasi Geografis (SIG). SIG memungkinkan kita untuk mengumpulkan, menyimpan, memproses, dan menganalisis data geospasial, yang kemudian dapat digunakan untuk berbagai keperluan seperti pemetaan, perencanaan, dan pengambilan keputusan.

Model data spasial adalah salah satu komponen penting dalam SIG, yang digunakan untuk merepresentasikan objek-objek geografis dalam bentuk digital (Longley *et al.*, 2015). Model data spasial ini terdiri dari dua kategori utama, yaitu data raster dan data vektor. Data raster merepresentasikan objek-objek geografis dalam bentuk grid, sedangkan data vektor merepresentasikan objek-objek geografis dalam bentuk titik, garis, atau poligon.

Praktik penggunaan model data spasial dalam SIG memiliki tantangan tersendiri, seperti kompleksitas data, kesulitan dalam integrasi data dari berbagai sumber, dan kesulitan dalam interpretasi data spasial. Oleh karena itu buku ini hadir sebagai upaya untuk memberikan pemahaman yang lebih baik tentang model data spasial dalam bidang kesehatan masyarakat, sehingga dapat membantu pengguna dalam melakukan analisis spasial dan pengambilan keputusan yang lebih baik.

Beberapa tahun terakhir data spasial telah dimanfaatkan di berbagai bidang, termasuk perencanaan tata ruang, mitigasi bencana, pengelolaan sumber daya alam, transportasi, kesehatan, dan lingkungan. Dalam tata ruang data spasial membantu merancang kawasan perkotaan yang efisien dan berkelanjutan (Arief *et al.*,

2021). Di bidang kesehatan, data spasial digunakan untuk memetakan penyebaran penyakit (Kartiko *et al.*, 2025), memetakan aksesibilitas fasilitas kesehatan (Arief *et al.*, 2023) dan merancang strategi intervensi yang efektif (Longley *et al.*, 2015). Selain itu SIG juga memiliki peran penting dalam mitigasi bencana dengan memberikan informasi lokasi rawan bencana dan memandu proses evakuasi (Wardana *et al.*, 2018).

Saat ini data spasial menjadi salah satu komponen terpenting dalam pengambilan keputusan di berbagai sektor. ESRI menyatakan bahwa sekitar 80% data di dunia memiliki unsur keruangan, artinya data tersebut dapat dipetakan dan dianalisis berdasarkan lokasi geografisnya esri (ESRI, 2018). Keunggulan data spasial terletak pada kemampuannya menghubungkan informasi dengan konteks wilayah, sehingga memudahkan analisis pola, tren, dan hubungan antarvariabel di permukaan bumi. Data ini tidak hanya dimanfaatkan dalam bidang perencanaan wilayah, transportasi, dan manajemen sumber daya alam, tetapi juga memiliki peran strategis dalam bidang kesehatan masyarakat, misalnya untuk memantau persebaran penyakit, mengidentifikasi wilayah rawan wabah, dan merancang intervensi berbasis lokasi yang lebih tepat sasaran.

8.2 Representasi Data Spasial

Model data dalam data spasial memiliki hubungan erat dengan *representing geography* (representasi geografi). Representasi geografi adalah proses merepresentasikan dunia nyata ke dalam bentuk yang dapat dipahami oleh manusia atau komputer (Longley *et al.*, 2015). Representasi ini dapat dilakukan dengan berbagai cara, seperti dalam bentuk peta, citra satelit, atau model data spasial dalam SIG.

Model data spasial dalam SIG merepresentasikan objek-objek geografis dalam bentuk digital dengan menggunakan kategori data raster atau data vektor. Model ini digunakan untuk merepresentasikan fitur-fitur geografis seperti jalan, bangunan, sungai, dan lain-lain dalam bentuk yang dapat diproses oleh komputer dan dapat digunakan untuk berbagai analisis spasial.

Dalam representasi fenomena spasial, model data spasial dalam SIG dapat digunakan sebagai salah satu cara

merepresentasikan fitur-fitur geografis yang ada di dunia nyata. Model data spasial dalam SIG juga dapat dikombinasikan dengan data lain, seperti data sosial dan ekonomi, untuk membentuk sebuah representasi yang lebih lengkap dan dapat digunakan untuk berbagai keperluan seperti perencanaan, pengambilan keputusan, dan pengembangan sistem informasi geografis yang lebih canggih.

Dewasa ini model data spasial dalam SIG memiliki peran yang penting dalam representasi fenomena geografi, karena dapat membantu merepresentasikan fitur-fitur geografis dengan lebih akurat dan efisien, serta memungkinkan dilakukannya analisis spasial yang lebih kompleks dan informatif.

Representasi geografi memiliki peran yang sangat penting dalam berbagai bidang, seperti ilmu geografi, penginderaan jauh, sistem informasi geografis, perencanaan kota, pengembangan wilayah, dan lain sebagainya. Berikut beberapa alasan mengapa data spasial dalam fenomena geografi penting:

1. Membantu pemahaman tentang geografi

Data spasial dalam representasi geografi seperti peta dan citra satelit, dapat membantu orang memahami lebih baik tentang geografi dan topografi suatu wilayah. Representasi ini juga dapat digunakan untuk memperlihatkan hubungan antara suatu wilayah dengan wilayah lainnya, sehingga dapat membantu dalam pengambilan keputusan yang lebih baik.

2. Membantu perencanaan layanan kesehatan masyarakat

Data spasial dapat digunakan untuk memetakan distribusi penduduk, fasilitas kesehatan, serta persebaran penyakit di suatu wilayah. Informasi ini membantu perencanaan dan pengembangan layanan kesehatan masyarakat yang lebih efektif dan efisien, seperti menentukan lokasi rumah sakit, puskesmas, atau posko kesehatan agar dapat menjangkau masyarakat secara optimal.

3. Mendukung pengambilan keputusan

Data spasial dapat digunakan sebagai alat bantu dalam pengambilan keputusan, seperti dalam penentuan lokasi fasilitas kesehatan, pemantauan perkembangan kesehatan wilayah, dan lain sebagainya.

4. Menunjang pengembangan sistem informasi geografis
Data spasial merupakan salah satu komponen penting dalam sistem informasi geografis, yang dapat digunakan untuk merepresentasikan data geospasial dalam bentuk digital dan melakukan analisis spasial yang lebih kompleks.
5. Mendorong penelitian dan pengembangan ilmu geografi kesehatan
Data spasial dapat digunakan sebagai bahan penelitian dan pengembangan ilmu geografi kesehatan, seperti dalam pemetaan penyakit, analisis spasial penyakit menular, dan pemantauan perubahan distribusi wilayah rawan pandemi.

Data spasial dalam representasi geografi memiliki peran yang sangat penting dalam berbagai bidang dan dapat membantu dalam pengambilan keputusan yang lebih baik dan efektif, serta meningkatkan pemahaman dan pengembangan ilmu geografi kesehatan.

Dalam penerapan representasi geografi tidak lepas dari permasalahan, menurut Longley *et al.*, (2015) representasi geografi memiliki hubungan erat dengan *uncertainty* atau ketidakpastian. Ketidakpastian ini bisa terjadi dalam berbagai aspek dalam representasi geografi, seperti ketidakpastian dalam pengukuran, ketidakpastian dalam model atau algoritma yang digunakan, dan ketidakpastian dalam keterbatasan data yang tersedia.

Dalam representasi geografi ketidakpastian ini dapat mempengaruhi akurasi dan kualitas representasi yang dihasilkan. Perlu dilakukan manajemen ketidakpastian untuk meminimalkan dampak dari ketidakpastian tersebut. Hal ini dapat dilakukan dengan berbagai cara, seperti melakukan analisis sensitivitas, penggunaan teknik-teknik statistik untuk memperkirakan tingkat ketidakpastian, dan penggunaan model-model yang dapat menangani ketidakpastian.

Dalam penelitiannya yang berjudul "*A Conceptual Framework for Uncertainty Analysis in Map-Based Urban Spatial Planning*", Arief *et al.*, (2021) membahas tentang kerangka konseptual untuk menganalisis ketidakpastian dalam implementasi data spasial. Artikel ini mencakup tinjauan singkat tentang ketidakpastian yang

terkait dengan data spasial dan metode analisis spasial, dan membahas secara mendalam konsep ketidakpastian dalam perencanaan tata ruang perkotaan.

Penulis menjelaskan bahwa ketidakpastian dalam representasi geografi dapat terjadi karena berbagai faktor, seperti ketidakpastian dalam data spasial, ketidakpastian dalam metode analisis spasial, dan ketidakpastian dalam asumsi dan model yang digunakan dalam perencanaan. Oleh karena itu penulis mengusulkan kerangka konseptual yang terdiri dari tiga tahap utama: identifikasi ketidakpastian, penilaian ketidakpastian, dan pengelolaan ketidakpastian.

Pada tahap pengelolaan ketidakpastian, penulis memberikan beberapa saran dan strategi untuk mengurangi ketidakpastian dalam representasi geografi, seperti menggunakan metode analisis spasial yang lebih akurat, memperbaiki kualitas data spasial, menggunakan teknologi informasi geografis yang lebih baik, dan meningkatkan keterampilan analisis spasial para perencana.

Penerapan manajemen ketidakpastian dalam Representasi geografi dapat membantu meningkatkan kualitas dan akurasi representasi geografi yang dihasilkan. Hal ini sangat penting dalam berbagai aplikasi, seperti dalam pengambilan keputusan, perencanaan, dan manajemen sumber daya. Oleh karena itu peran Representasi geografi dalam manajemen ketidakpastian sangatlah penting. SIG tidak dapat merepresentasikan dunia nyata secara sempurna (100%) (Longley *et al.*, 2015), akan tetapi setiap melakukan analisis spasial harus mengelola *uncertainty* yang berpotensi muncul (Fisher, 1999; Leyk *et al.*, 2005), dan analisis *uncertainty* merupakan bagian dari tahapan pengujian perangkat lunak (Vliet, 2007; Sommerville, 2013; Pressman, 2014; Pressman and Maxim, 2014; Muttaqin *et al.*, 2023).

8.2.1 Representasi Data Diskrit

Representasi data diskrit dalam konteks analisis data spasial merupakan salah satu bentuk representasi data geografis yang paling umum digunakan. Data diskrit didefinisikan sebagai kumpulan objek geografis terpisah dan dapat diidentifikasi secara individual, seperti bangunan, jalan, atau wilayah administratif (Longley *et al.*, 2015).

Data diskrit merupakan salah satu teknik representasi data spasial dengan cara memetakan data menjadi titik-titik atau piksel-piksel diskrit yang tersebar pada suatu wilayah. Data ini dikelompokkan menjadi beberapa kelas dan setiap piksel atau titik diberi nilai yang merepresentasikan kelas tersebut. Contohnya data curah hujan dapat direpresentasikan dengan menggunakan pola titik-titik diskrit pada peta dengan setiap titik mewakili nilai curah hujan di lokasi tersebut.

Pada representasi data diskrit, data biasanya direpresentasikan dalam bentuk raster, yaitu kumpulan sel atau piksel yang membentuk *grid* atau *matrix* dengan ukuran dan bentuk yang sama. Setiap piksel pada grid merepresentasikan nilai dari data yang dipetakan pada suatu lokasi tertentu. Contoh data yang dapat direpresentasikan dalam bentuk raster antara lain citra satelit, data elevasi, data curah hujan, dan data suhu.

Sedangkan untuk data diskrit yang direpresentasikan dalam bentuk data vektor, yaitu dengan menyimpan kumpulan koordinat yang membentuk titik, garis, atau poligon. Pada data vektor, titik merepresentasikan lokasi suatu objek, garis merepresentasikan jalan atau sungai, sedangkan poligon merepresentasikan wilayah atau bangunan. Selain lokasi, data diskrit juga menyimpan informasi atribut pada setiap objek, seperti nama jalan, jenis bangunan, atau luas tanah. Informasi ini dapat ditambahkan dalam bentuk metadata pada data vektor.

Kelebihan dari representasi data diskrit adalah kemampuannya untuk merepresentasikan objek geografis dengan detail yang tinggi. Data vektor juga memungkinkan untuk melakukan operasi analisis spasial, seperti *overlay* dan *buffer analysis*, serta memungkinkan untuk membuat peta dengan detail yang tinggi dan tampilan yang menarik. Sedangkan untuk representasi data diskrit dalam bentuk raster adalah mudah dalam penggunaannya dan memberikan representasi yang jelas mengenai nilai data spasial pada setiap lokasi. Data dalam bentuk raster juga memudahkan user dalam melakukan analisis spasial, seperti *overlay*, *buffer*, dan analisis jarak.

Kelemahan dari representasi data diskrit dalam bentuk data vektor adalah besarnya ukuran *file* yang dihasilkan. Kumpulan data

diskrit dengan detail yang tinggi dapat menghasilkan *file* yang sangat besar dan memerlukan penyimpanan yang besar pula. Selain itu representasi data diskrit juga kurang cocok untuk merepresentasikan fenomena geografis yang meluas dalam suatu wilayah, seperti elevasi atau suhu udara. Representasi data diskrit dalam bentuk raster juga kurang mampu merepresentasikan objek spasial yang memiliki batas yang kompleks (Ramdani, 2017). Pada beberapa kasus representasi data diskrit dalam bentuk raster menghasilkan data yang terlalu besar sehingga sulit dalam penyimpanan dan pemrosesan. Contoh visualisasi dari data diskrit dapat dilihat pada Gambar 8.1.



Gambar 8.1. Beruang
(Sumber : Longley *et al.*, 2015)

Beruang adalah objek spasial yang terbatas, dimana wilayah keberadaannya dibatasi oleh faktor geografi dan lingkungan tertentu. Dalam representasi data spasial beruang umumnya direpresentasikan sebagai data diskrit, yaitu dengan menggunakan model data vektor. Model data vektor merepresentasikan objek spasial sebagai titik, garis, atau poligon dengan nilai atribut yang terkait dengan objek tersebut.

Pada representasi beruang, model data vektor akan merepresentasikan wilayah keberadaan beruang sebagai poligon

dengan nilai atribut yang mewakili karakteristik beruang seperti jenis, umur, dan berat badan. Karakteristik dan wilayah keberadaannya dapat diidentifikasi dengan jelas dan terbatas, sehingga cocok direpresentasikan sebagai data diskrit dengan menggunakan model data vektor. Hal tersebut dapat bervariasi tergantung pada konteks dan tujuan representasi data spasial yang diinginkan.

8.2.2 Representasi Data Kontinu

Menurut Longley *et al.*, (2015), representasi data kontinu dalam data spasial melibatkan pemetaan nilai data pada suatu wilayah yang terus berubah secara kontinu atau halus. Dalam representasi data kontinu, wilayah dipandang sebagai suatu kontinum yang terdiri dari nilai data yang berubah secara halus, sehingga diperlukan suatu fungsi matematis untuk merepresentasikan perubahan nilai data tersebut.

Salah satu cara untuk merepresentasikan data kontinu adalah dengan menggunakan model data vektor. Pada model data vektor, wilayah direpresentasikan dalam bentuk poligon atau garis yang membatasi wilayah dengan nilai data yang berbeda di setiap titik pada wilayah tersebut. Setiap titik pada wilayah tersebut diberi koordinat dan nilai data yang sesuai. Model data vektor ini lebih cocok digunakan untuk merepresentasikan data seperti peta topografi, peta iklim, peta tanah, dan peta hidrologi.

Selain model data vektor, representasi data kontinu juga dapat dilakukan dengan menggunakan model data raster. Pada model data raster, wilayah direpresentasikan sebagai kumpulan sel yang diberi nilai data yang mewakili nilai data pada suatu wilayah tertentu. Setiap sel memiliki ukuran yang sama, dan wilayah direpresentasikan sebagai suatu matriks atau tabel dua dimensi. Model data raster lebih cocok digunakan untuk merepresentasikan data seperti citra satelit dan foto udara, serta data yang kontinu dan memerlukan resolusi tinggi seperti data elevasi.

Kelebihan dari representasi data kontinu adalah mampu merepresentasikan perubahan nilai data secara halus dan akurat pada suatu wilayah. Representasi data kontinu dalam bentuk vektor juga mampu merepresentasikan objek spasial yang memiliki batas

yang kompleks dengan lebih baik dibandingkan dengan representasi data diskrit dalam bentuk raster. Representasi data kontinu dapat digunakan untuk melakukan analisis spasial yang lebih kompleks seperti interpolasi, *overlay*, dan analisis keterkaitan spasial (Longley *et al.*, 2015). Data dalam bentuk vektor juga lebih efisien dalam penyimpanan dan pemrosesan karena hanya menyimpan informasi tentang batas dan nilai data pada setiap titik penting (Ramdani, 2017).

Kelemahan dari representasi data kontinu adalah memerlukan teknik pemrosesan yang lebih kompleks dan memerlukan ruang penyimpanan yang lebih besar dibandingkan dengan representasi data diskrit. Representasi data kontinu juga kurang efektif dalam merepresentasikan objek spasial yang memiliki batas yang kompleks. Sedangkan kelemahan dari representasi data kontinu dalam bentuk vektor adalah kurang mudah dalam penggunaannya dan memerlukan teknik pemrosesan data yang lebih kompleks untuk melakukan analisis spasial. Representasi data kontinu dalam bentuk vektor juga kurang memberikan representasi yang jelas mengenai nilai data spasial pada setiap lokasi dibandingkan dengan representasi data diskrit dalam bentuk raster. Contoh visualisasi dari data kontinu dapat dilihat pada Gambar 8.2.

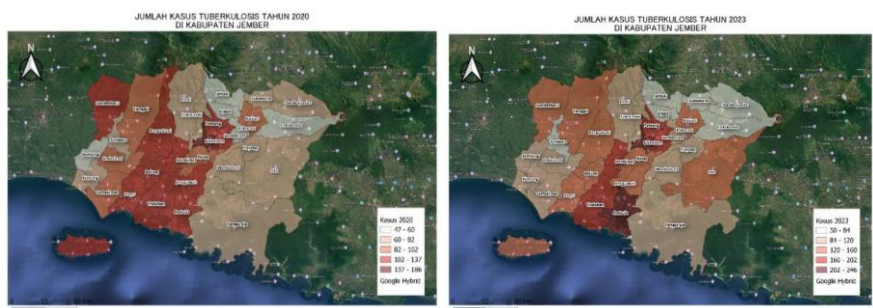


Gambar 8.2. Danau di Helsinki, Finlandia
(Sumber : Longley *et al.*, 2015)

Danau merupakan objek spasial yang luas dan kontinu, dimana perubahannya terjadi secara halus dan terus-menerus pada suatu wilayah yang luas. Dalam representasi data spasial, danau umumnya direpresentasikan sebagai data kontinu, yaitu dengan menggunakan model data raster atau model data vektor. Model data raster akan merepresentasikan danau sebagai kumpulan sel dengan nilai data yang mewakili nilai elevasi atau kedalaman pada suatu wilayah tertentu. Sedangkan pada model data vektor, danau direpresentasikan sebagai poligon dengan nilai atribut yang mewakili nilai elevasi atau kedalaman pada suatu titik pada danau.

8.3 Implementasi Analisis Data Spasial dalam Bidang Kesehatan Masyarakat

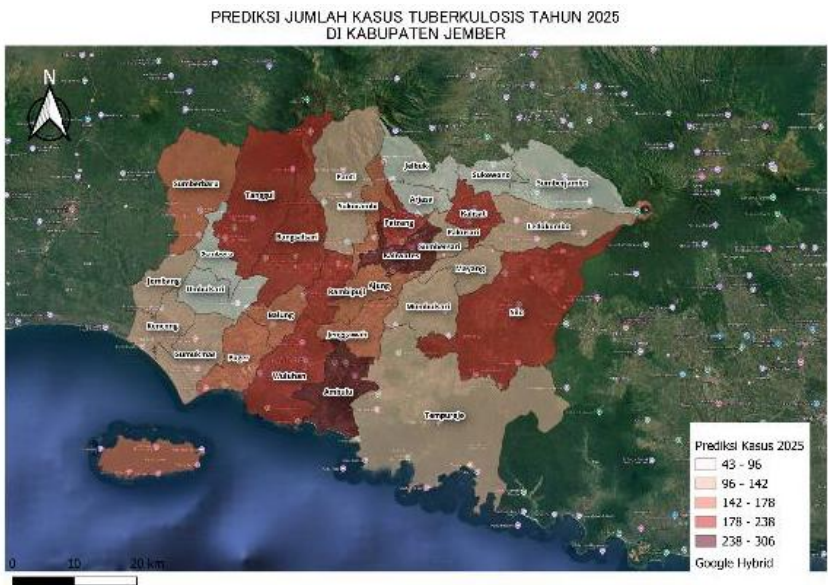
Implementasi analisis data spasial telah menjadi fondasi strategis dalam memperkuat sistem kesehatan masyarakat di Indonesia, dengan berbagai aplikasi mulai dari pemetaan penyakit hingga perencanaan dan evaluasi layanan. Di Kabupaten Jember misalnya, penelitian seperti Pemetaan Potensi Persebaran Kasus Tuberkulosis menggunakan pendekatan analisis spasial untuk mengidentifikasi wilayah dengan risiko penularan tinggi, sebuah analisis penting untuk memprioritaskan intervensi medis dan alokasi sumber daya.



Gambar 8.3. Sebaran Kasus TBC di Kabupaten Jember 2020-2023
(Sumber : Kartiko *et al.*, 2025)

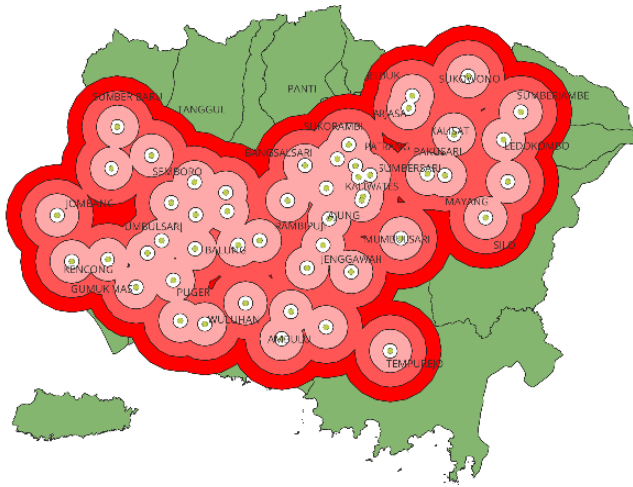
Tuberkulosis (TBC) merupakan penyakit menular yang masih menjadi tantangan serius di Indonesia termasuk Kabupaten Jember, karena berdampak signifikan terhadap

kesehatan dan kualitas hidup masyarakat. Gambar 8.3 merupakan hasil analisis spasial untuk memetakan persebaran kasus TBC di Jember berdasarkan data kesehatan aktual tahun 2020–2023 dari seluruh fasilitas kesehatan. Analisis mencakup pola persebaran, proyeksi kasus untuk tahun 2025 seperti pada Gambar 12.4, serta identifikasi faktor risiko di setiap kecamatan. Hasilnya diharapkan menjadi dasar perumusan strategi pencegahan dan penanggulangan TBC yang lebih tepat sasaran, membantu pemerintah daerah dalam merancang program kesehatan terintegrasi guna meningkatkan efektivitas pengendalian penyakit menular secara berkelanjutan.



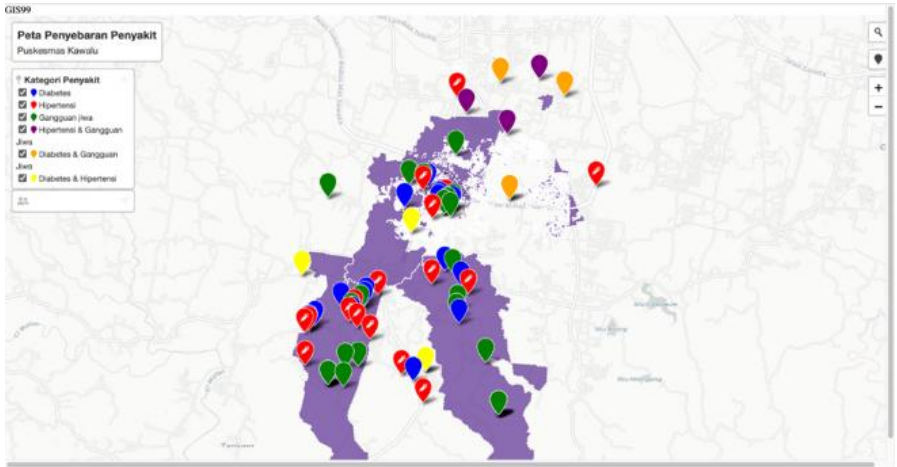
Gambar 8.4. Pemetaan Prediksi Sebaran Kasus TBC di Kabupaten Jember Tahun 2025
(Sumber : Kartiko *et al*, 2025)

Studi mengenai Analisis Spasial Aksesibilitas Fasilitas Kesehatan di Kabupaten Jember memanfaatkan analisis jarak dan waktu tempuh untuk mengevaluasi keterjangkauan fasilitas kesehatan seperti pada Gambar 8.5, penting dalam merancang lokasi fasilitas yang strategis dan inklusif.



Gambar 8.5. Peta Keterjangkauan Puskesmas di Kab. Jember
(Sumber : Arief *et al.*, 2023)

Di tingkat praktis pelayanan primer, Puskesmas Kawalu di Tasikmalaya menjadi contoh implementasi GIS yang inovatif. Melalui pengabdian masyarakat dari Universitas Telkom, Puskesmas ini menggunakan aplikasi GIS berbasis Leaflet dan Google Sheet untuk memvisualisasikan distribusi penyakit, cakupan layanan, dan data spasial secara real-time, sehingga mempermudah tenaga kesehatan dalam pengambilan keputusan. Model tersebut juga diuji ulang dalam bentuk aplikasi GIS di Puskesmas Urug, Tasikmalaya. Membantu pemetaan penyebaran penyakit tidak menular, meningkatkan efisiensi manajemen data dan respons kesehatan lokal.



Gambar 8.6. Peta Penyebaran Penyakit
(Sumber : Hakim *et al.*, 2025)

Puskesmas Merdeka Bogor mengintegrasikan GIS dalam sistem informasi kesehatan di wilayah kerjanya. Teknologi GIS digunakan untuk memetakan infrastruktur kesehatan, kondisi lingkungan, serta distribusi penyakit di wilayah operasional, sebagai bagian dari upaya memperkuat sistem informasi kesehatan melalui survei lapangan dan analisis spasial (Dinata, Cahyadi and Muharam, 2024). Sementara itu pengalaman Jakarta Smart City dalam menghadapi pandemi COVID-19 menunjukkan betapa ampuhnya visualisasi dan analisis spasial saat proses vaksinasi. Analisis spasial digunakan untuk memetakan daerah prioritas berdasarkan tingkat cakupan vaksinasi, membantu koordinasi antara dinas kesehatan, statistik, dan perencanaan wilayah untuk memperkuat distribusi vaksin secara tepat sasaran.



Gambar 8.7. *Vaccine Landing Page of Jakarta COVID-19 official website*

DAFTAR PUSTAKA

- Arief, M.H. *et al.* (2023) 'Analisis Spasial Aksesibilitas Fasilitas Kesehatan di Kabupaten Jember', *Jurnal Minfo Polgan*, 12(September), pp. 1764–1771. Available at: <https://doi.org/https://doi.org/10.33395/jmp.v12i2.12984>.
- Arief, M.H., Ramdani, F. and Bachtiar, F.A. (2021) 'A Conceptual Framework for Uncertainty Analysis in Map-Based Urban Spatial Planning', in *ACM International Conference Proceeding Series*. Association for Computing Machinery, pp. 197–202. Available at: <https://doi.org/10.1145/3479645.3479683>.
- Dinata, S.A., Cahyadi, W.A. and Muharam, D. (2024) 'Penerapan Sistem Informasi Geografis (GIS) Berbasis Komputer Sebagai Upaya Penguatan Sistem Informasi Kesehatan Di Wilayah Kerja Puskesmas Merdeka Kota Bogor', *Abdi Laksana: Jurnal Pengabdian Kepada Masyarakat*, 5(Volume 5, Nomor 1), pp. 240–249. Available at: <http://openjournal.unpam.ac.id/index.php/JAL>.
- ESRI (2018) *Spatially represent your data, ArcGIS Python Libraries*. Available at: <https://www.esri.com/en-us/arcgis/products/arcgis-python-libraries/overview>.
- Fisher, P.F. (1999) 'Models of uncertainty in spatial data', in *Geographical Information Systems: Principles, Techniques, Management and Applications*, pp. 191–205.
- Hakim, B.P. *et al.* (2025) 'Konvergensi Geographic Information System dalam Manajemen Kesehatan Unit Pelaksana Teknis Daerah', *Jurnal Abdimas Madani dan Lestari (JAMALI)*, 07, pp. 50–58. Available at: <https://doi.org/10.20885/jamali.vol7.iss1.art6>.
- Kartiko, E.Y. *et al.* (2025) 'Pemetaan Potensi Persebaran Kasus Tuberkulosis di Kabupaten Jember dengan Pendekatan Analisis Spasial', 8(4), pp. 1182–1195.
- Leyk, S., Boesch, R. and Weibel, R. (2005) 'A conceptual framework for uncertainty investigation in map-based land cover change modelling', *Transactions in GIS*, 9(3), pp. 291–322. Available at: <https://doi.org/10.1111/j.1467-9671.2005.00220.x>.
- Longley, P.A. *et al.* (2015) *Geographic Information Sciences and*

- Systems*. Fourth Edi, Wiley. Fourth Edi. Edited by R. Flahive and J. Fiorillo. London: Wiley. Available at: https://doi.org/10.1007/springerreference_62122.
- Muttaqin *et al.* (2023) *Audit Sistem Informasi*. Cetakan 1. Edited by R. Watrianthos. Medan: Yayasan Kita Menulis.
- Pressman, R.S. (2014) *Software Engineering: A Practitioner's Approach*. Seventh Ed, Mc Graw-Hill. Seventh Ed. Edited by F.M. Schilling. New York: MC Graw Hill. Available at: <https://doi.org/10.1002/9781118830208>.
- Pressman, R.S. and Maxim, B. (2014) *Software Engineering: A Practitioner's Approach*. 8th edn. McGraw Hill.
- Ramdani, F. (2017) *Pengantar Ilmu Geoinformatika*. Pertama. Malang: UB Press.
- Sommerville, I. (2013) *Software Engineering*. Ninth Edit, *Clinical Engineering: A Handbook for Clinical and Biomedical Engineers*. Ninth Edit. Edited by M. Horton *et al.* Boston: Pearson Education. Available at: <https://doi.org/10.1016/B978-0-12-396961-3.00009-3>.
- Vliet, H. van (2007) *Software Engineering: Principles and Practice*. 3rd edn, *Software Engineering Journal*. 3rd edn. Amsterdam: Wiley. Available at: <https://doi.org/10.1049/sej.1994.0029>.
- Wardana, D.P.T., Ramdani, F. and Pradana, F. (2018) 'Real-Time Landslide Monitoring System Based on 3D WebGIS and Embedded System', in *4th International Symposium on Geoinformatics, ISyG 2018*. IEEE, pp. 1–7. Available at: <https://doi.org/10.1109/ISYG.2018.8612028>.

BIODATA PENULIS



**Dr.Arif Rachman,drg.,SH.,MH.,MM.,MTr.Hanla.,
Sp.Pros.,CIQnR.,CIQaR.,CSCBM.,CHRMp.,CRMS.,CAIA**
Dosen Pendidikan Sarjana Kedokteran
Fakultas Kedokteran Militer Universitas Pertahanan RI

Penulis lahir pada tanggal 07 Februari 1971 di Surabaya – Jawa Timur, dari ayah yang bernama Dulmanan dan Ibu Romlah. Alamat di Jl. Cipinang Cempedak IV No. 60B Jakarta Timur 13340, e-mail ar604299@gmail.com / arifrachmanlecturer@gmail.com / arif.rachman@idu.ac.id

Pendidikan yang pernah ditempuh :Program Sarjana Jurusan Sarjana Pendidikan dokter gigi di FKG Unair Lulus Tahun 1995, Program Profesi dokter gigi di FKG Unair Lulus Tahun 1997, Program Spesialis Prosthodontia FKG Unair Lulus Tahun 2007, Program Magister Manajemen di Ubhara Surabaya Lulus Tahun 2008, Program Sarjana Ilmu Hukum di Ubhara Surabaya Lulus Tahun 2020, Program Magister Hukum di Ubhara Surabaya Lulus Tahun 2019, Program Doktorat Ilmu Kedokteran Dasar di Fakultas Kedokteran Unair Lulus Tahun 2019, Sertifikasi Peneliti Kuantitatif dan kualitataif di Quantum Lulus Tahun 2021.

Pekerjaan/Jabatan yang di pegang sekarang adalah: Dosen Tetap Fakultas Kedokteran Militer Universitas Pertahanan RI di Bogor, Kepala Laboratorium Fakultas Kedokteran Militer Universitas Pertahanan RI di Bogor, Dosen Tidak Tetap Universitas Adhirajasa Reswara Sanjaya di Bandung.

Organisasi Dosen: PDGI (anggota), IPROSI (anggota), MHKI (anggota), APEPINDO (anggota)

Karya Tulis dalam bentuk buku dan jurnal penelitian adalah:
Buku Referensi dan Bahan Ajar :

1. Reduction of Osx expression in the osseointegration process of dental implants with human adipose-derived mesenchymal stem cell intervention: an experimental study. Padjadjaran Journal of Dentistry, Volume 37, Number 2, July 2025; 142-154.
<https://doi.org/10.24198/pjd.vol37no2.62084>
2. Dasar-Dasar Biologi Sel Dan Molekuler. Get Press Indonesia, ISBN: 978-623-125-848-9, Juni 2025
3. Sistem Informasi E-Hospital. Lakeisha, ISBN: 978-623-119-614-9, Maret 2025
4. Etika, Hukum, Dan Regulasi Dental Implan Untuk Drg Umum. Get Press Indonesia, ISBN: 978-623-125-659-1, Maret 2025
5. Pengantar Metodologi Kuantitatif Dan Kualitatif. CV. Saba Jaya Publisher ISBN 978-623-8730-33-9, Maret 2025
6. Prevention of Military Stress Fractures using the hadMCS Strategy Approach in Health Resilience Projections in National Defence: Preliminary Research. PJJSS 2024 Posted: Oct 19, 2024. Volume 22(2). 2024; 8791-8798.
<https://doi.org/10.57239/PJJSS-2024-22.2.00663> (Q4).
7. The Correlation of Exclusive and Non-Exclusive Asi on Stunting In The Context of State Defense: Initial Research on PCBM In Bogor District. PJJSS 2024 Posted: Oct 20, 2024. Volume 22(2). 2024; 8976-8985.
<https://doi.org/10.57239/PJJSS-2024-22.2.00679> (Q4).
8. Acceleration of Healing of TB Patients through Dietary Supplements on Anti-Tuberculosis Drugs. PJJSS 2024 Posted: Oct 19, 2024. Volume 22(2). 2024; 8799-8807.
<https://doi.org/10.57239/PJJSS-2024-22.2.00664> (Q4).
9. Prevention Of Stunting In The First 450 Days Of Life (FDL) Through Fortification Of Tuna Fish Cake And Formula Programming In Supporting Indonesia's National Defense In

- The Health Sector. PJLSS 2024 Posted: Oct 20, 2024. Volume 22(2). 2024; 8986-8996. <https://doi.org/10.57239/PJLSS-2024-22.2.00680>. Q4)
10. Increased Runx-2 Expression In The Y-TZP Biomaterial Osteointegration Process With hADMSC Seeding. SEEJPH 2024 Posted: 02-08-2024. Volume XXV 2024; 216-224.<https://doi.org/10.70135/seejph.vi.1470> (Q4).
 11. Manajemen Epidemiologi, Get Press Indonesia, ISBN: 978-623-125-387-3, September, 2024.
 12. Pengantar Hukum Indonesia, Get Press Indonesia, ISBN: 978-623-125-239-5, Juni 2024.
 13. Komunikasi Dalam Praktik Keperawatan (Komunikasi Efektif Layanan Kesehatan), Agdosi Makassar, ISBN: 978-623-10-0651-6, Mei 2024
 14. Manajemen Risiko dan Asuransi, PT. MAFY Literasi Indonesia, ISBN: 978-623-863-871-0, Mei 2024.
 15. Metode Penelitian dalam Pendidikan_2, Get Press Indonesia, ISBN: 978-623-125-155-8, April 2024.
 16. Dasar Biostatistika untuk Peneliti, Get Press Indonesia, ISBN: 978-623-125-137-4, April 2024.
 17. Pengantar Manajemen, Get Press Indonesia, ISBN: 978-623-125-138-1, April 2024.
 18. Konsep Autoimun, PT. MAFY Literasi Indonesia, ISBN: 978-623-8575-21-3, Maret 2024.
 19. An Innovative Approach to Dental Implantology Through DPMSC Therapy in Military Dentistry, Tahta Media Group, ISBN: 978-623-147-325-7, Februari 2024.
 20. Metodologi Penelitian Kuantitatif, Kualitatif, dan R&D, CV. Saba Jaya Publisher ISSN: 978-623-09-7582-0, Januari 2024
 21. Fakultas Kedokteran Militer Universitas Pertahanan Cita dan Jejak Sejarah, Unhan Press, Januari Tahun 2021.
 22. Bioetika Multidisplin, Multidisiplinary Prespectives on Bioethics, Unhan Press, Januari Tahun 2022.
 23. Aesthetic and masticatory rehabilitation on post mandibular resection with combination of hollow obturator

- and hybrid prosthesis, Dent. J. (Maj. Ked. Gigi), Vol. 42. No. 2 April-June 2009: 76-81
24. Biocompatibility of Yttria-Tetragonal Zirconia Polycrystal Seeded with Human Adipose Derived Mesenchymal Stem Cell, ACTA INFORM MED. 2018 DEC; 26(4): 249-253 (Q3)
 25. Addition Of Adipose Derived Stem Cell To Beta Tricalcium Phosphate and Human Cancellous Bone for Craniofacial Bone Tissue Engineering: An In Vitro Study, Malaysian Journal of Medicine and Health Sciences. 2020; 16(4): 83-86 (Q3)
 26. Comparison of Bone regeneration in hADMSC Versus hUCBMSC with hBMMSC as a Reference: A Literature Review of Potential Bone Regeneration, Research J. Pharm. and Tech. 14(4): April 2021 (Q3)
 27. CBCT: Tracing Dominant Region of Periodontitis on Psychological Stress, Sys Rev Pharm. 2020;11(11):1567-1572 (Q2)
 28. Ideal Bone Defects Distance on Orthodontic Tooth Movement for Preparation of hADMSC-Scaffold Chitosan Intervention, Sys Rev Pharm. 2020;11(8):571-575 (Q2)
 29. Accuracy of Volume of Interest Determination with Cone-Beam Computed Tomography in Periodontitis due to Psychological Stress., Journal of International Dental and Medical Research ISSN 1309-100X <http://www.jidmr.com> (Q3).
 30. Antitumor Effects of Myrmecodia pendans: a Scoping Review, Journal of International Dental and Medical Research, Volume 16 Number 2 , 2023;1-5, ISSN 1309-100X (Q3).
 31. Evaluation of the Use of Antibiotics on Therapy Results of Sepsis Patients in the Intensive Care Unit (ICU) of Fatmawati Hospital, Jakarta, J Med Chem Sci. 2023;7(1):262-74. (Q3).
 32. Enhancing Denture Care Efficiency: Mobile Prosto Open-Source Software for Indonesian National Army Soldiers, J Int Dent Med Res 2024; 17(1): 209-214; ISSN 1309-100X. (Q3).
 33. The Impact of Lifestyle Factors on the Human Microbiome, Medical Studies and Health Journal (SEHAT) Vol 1 (1) 2024 : 113-122.

34. Exploring the Therapeutic Potential of Ant's Nest Plant Extract on Epithelial Dysplasia in an Oral Squamous Cell Carcinoma Model: A Study on Apoptosis Markers in Sprague Dawley Mice, *Journal of International Dental and Medical Research*, Volume 17 Number 2, 2024;543–548, ISSN 1309-100X (Q3).
35. Pendidikan Kesehatan, Pemeriksaan dan Deteksi Dini Pencegahan Penyakit Tidak Menular melalui Kegiatan Olah Raga Fisik di Wilayah Kerja Puskesmas Tamalanrea Makassar, *Sahabat Sosial Jurnal Pengabdian Masyarakat*, e-ISSN: 2964-9196, Vol.2 No.4 September 2024

BIODATA PENULIS



Iha Nursolihah, S.S.T., M.K.M.

Dosen Program Studi Administrasi Rumah Sakit
Fakultas Ilmu Kesehatan Universitas Singaperbangsa Karawang

Penulis lahir di Karawang tanggal 23 Oktober 1991. Pada tahun 2019, penulis menyelesaikan Strata2 (S2) pada peminatan Biostatistika dan Kependudukan, Fakultas Kesehatan Masyarakat Universitas Indonesia. Setelah menamatkan S2, penulis bekerja di Rumah Sakit Anak dan Bunda Harapan Kita, sebagai Analis Data dan Informasi Rawat Inap pada Substansi Pelayanan Medik. Sejak tahun 2022, Penulis menjadi salah satu dosen tetap PNS pada Fakultas Ilmu Kesehatan, Universitas Singaperbangsa Karawang. Penulis merupakan dosen pada program studi Administrasi Rumah Sakit, beberapa mata kuliah yang diampu diantaranya Biostatistik, Statistika Kesehatan, Epidemiologi, Manajemen Pengelolaan Analisis Data Rumah Sakit, dan mata kuliah lainnya.

BIODATA PENULIS



Nur Al-faida, SKM., M. Kes
Dosen Program Studi Sarjana Gizi
STIKes Persada Nabire Papua Tengah

Penulis lahir di Sumulluk tanggal 25 September 1993. Penulis adalah dosen tetap pada Program Studi Sarjana Gizi STIKes Persada Nabire sejak tahun 2020. Tahun 2020-2025 penulis mendapat amanah sebagai Ketua Lembaga Penelitian dan Pengabdian kepada Masyarakat (LPPM) di Sekolah Tinggi Ilmu Kesehatan Persada Nabire (STIKPEN).

Penulis menyelesaikan pendidikan S1 pada Fakultas Kesehatan dengan Jurusan Kesehatan Masyarakat Konsentrasi Epidemiologi di Universitas Muhammadiyah Parepare pada tahun 2016 dan melanjutkan S2 pada Jurusan Kesehatan Masyarakat Konsentrasi Epidemiologi lulus tahun 2018 di Universitas Muslim Indonesia.

Penulis menjadi pengampu pada mata kuliah: Matematika, Dasar Sosiologi, Fisiologi, Ilmu Bahan Makanan, Biostatistik, Gizi Ibu dan Anak 1, Antropologi Sosial Gizi, Epidemiologi Gizi, Teknologi Pangan Gizi, Metodologi Penelitian Gizi, dan Inovasi Ilmu Pangan Lokal

BIODATA PENULIS



Dr. Wardiah Hamzah, SKM, M.Kes

Dosen Program Studi Kesehatan Masyarakat
Fakultas Kesehatan Masyarakat Universitas Muslim Indonesia

Dr. Wardiah Hamzah, SKM, M.Kes dilahirkan di Kota Makassar pada tanggal 7 Juni 1975. Setelah nyantri selama 6 tahun di Pesantren IMMIM Putri, penulis melanjutkan pendidikan S1 dan S2 di FKM Unhas, yang selanjutnya meraih gelar doktor di FKM Unair Surabaya. Penulis aktif mengajar sebagai dosen tetap di FKM UMI sejak tahun 2002, dengan fokus pada Sumber Daya Manusia. Aktif melakukan riset dan pengabdian masyarakat. Sebelumnya penulis aktif sebagai supervisor pada proyek *Nutrition Surveillance System* Sulawesi Selatan (*Hellen Keller International*) selama 6 tahun.

BIODATA PENULIS



Nasruddin Syam

Dosen Program Studi Kesehatan Masyarakat
Fakultas Kesehatan Masyarakat Universitas Muslim Indonesia

Penulis adalah dosen pada Program Studi Kesehatan Masyarakat, Fakultas Kesehatan Masyarakat, Universitas Muslim Indonesia. Sebelumnya, pernah bekerja sebagai ASN di Balai Teknik Kesehatan Lingkungan Kemenkes di Makassar tahun 2002-2008 dan Balai Besar Pelatihan Kesehatan Kemenkes di Makassar 2008-2011. Penulis menyelesaikan pendidikan S1 dan S2 di Fakultas Kesehatan Masyarakat, Universitas Hasanuddin. Saat ini, penulis terlibat aktif pada riset dan pengabdian masyarakat serta konsultasi analisis data hasil riset khususnya mengenai dampak lingkungan khususnya logam berat terhadap anak dengan *Neurodevelopmental Disorder*.

BIODATA PENULIS



Islamiyati, AK, S,Keb, Bdn., MKM.
Dosen Program Studi Kebidanan Metro
Poltekkes Kemenkes Tanjungkarang

Penulis lahir di Baturaja tanggal 03 April 1972. Penulis adalah dosen tetap pada Program Studi Kebidanan Metro Poltekkes Kemenkes Tanjungkarang. Menyelesaikan pendidikan D4 Keperawatan pada Fakultas Kedokteran USU Medan tahun 1999 dan S1 serta Profesi Kebidanan pada STIKES Abdi Nusantara Jakarta tahun 2025 dan menyelesaikan S2 pada Program Studi Ilmu Kesehatan Masyarakat Universitas Indonesia peminatan Biostatistik pada tahun 2005. Penulis menekuni bidang Menulis buku Kebidanan meliputi Etika Kebidanan, Kebidanan Komunitas, Asuhan Kebidanan Bayi Balita dan Evidence Based Kebidanan, Biostatistik, serta Psikologi Perkembangan.

BIODATA PENULIS



Gede Wirabuana Putra, S.KM, M.K.M

*Dosen Program Studi Sarjana Terapan
Manajemen Informasi Kesehatan
Politeknik Kesehatan Kartini Bali*

Gede Wirabuana Putra adalah akademisi dan praktisi kesehatan masyarakat yang berdedikasi dalam pengembangan pembiayaan kesehatan dan manajemen data informasi kesehatan. Lahir di Denpasar pada 4 September 1994, menyelesaikan pendidikan S1 Kesehatan Masyarakat di Fakultas Kedokteran Universitas Udayana (2016) dan meraih gelar Magister Kesehatan Masyarakat dari Universitas Indonesia (2020).

Pengalaman Profesional: Perjalanan karier dimulai sebagai Staff Kepesertaan BPJS Kesehatan Cabang Denpasar (2015), dilanjutkan sebagai Provider Relation Officer di Mandiri Inhealth Denpasar (2017-2018). Pengalaman internasional diperoleh sebagai Technical Officer Consultant USAID TBPS pada Project Belanja Kesehatan Strategis Program Tuberkulosis (2021-2023).

Kontribusi Akademik: Saat ini menjabat sebagai Wakil Direktur 2 Politeknik Kesehatan Kartini Bali, mengampu mata kuliah strategis seperti Statistik Fasilitas Kesehatan, Sistem Pembiayaan Kesehatan, dan Pembiayaan Kesehatan Berbasis

Diagnosis. Aktif dalam Tridharma Perguruan Tinggi melalui penelitian, publikasi, dan pengabdian masyarakat.

Publikasi dan Buku Unggulan:

1. "Analysis of The Relationship Between Risk Management, Economic Situation, Customer Behavior And Insurance Business Growth"
2. "Determinan Keberhasilan Pengobatan Pasien Tuberkulosis di Kota Denpasar Tahun 2021"
3. "Gambaran Penerapan Konsolidasi Data Sistem Informasi Tuberkulosis (SITB) Dengan Data Jaminan Kesehatan Nasional (JKN) Terhadap Pencatatan Pasien Terduga Tuberkulosis Di Klinik Swasta Kota Denpasar".
4. Manajemen Rekam Medis Hybrid, (CV Globa Aksara Pers Tahun 2025, ISBN : 9786234628067)
5. Pembiayaan Kesehatan Berbasis Diagnosis (Filosofis Indonesia Press Tahun 2025, ISBN : 9786347174437)

Kontak: buanawira09@gmail.com

BIODATA PENULIS



M. Habibullah Arief, S.Kom., M.Kom.

Dosen Program Studi Sistem Informasi
Fakultas Ilmu Komputer

M. Habibullah Arief, S.Kom., M.Kom. lahir di Kabupaten Jember pada tanggal 11 Februari 1992. Ia memulai pendidikan tinggi pada Program Studi Manajemen Informatika di Universitas Telkom Bandung dan meraih gelar Ahli Madya Komputer. Melanjutkan pendidikannya di Universitas Brawijaya Malang, ia berhasil menyelesaikan Program Sarjana Komputer pada tahun 2017 dan Program Magister Komputer pada tahun 2021.

Saat ini aktif sebagai dosen tetap pada Program Studi Sistem Informasi, Fakultas Ilmu Komputer, Universitas Jember, sekaligus menjabat sebagai Project Manager di CV. Sinergi Spasial Indonesia, sebuah perusahaan yang bergerak di bidang pengembangan sistem informasi geografis dan solusi spasial. Sebagai pengajar menguasai berbagai mata kuliah antara lain Pengembangan Perangkat Lunak untuk Agroindustri Modern, Pengantar Sistem Informasi Geografi, Pengantar Ilmu Geospasial, Spatial Basis Data, dan Penginderaan Jauh. Selain itu, juga mengajarkan mata kuliah Manajemen Proyek Teknologi Informasi, E-Government, serta Ekonomi Regional, yang semakin memperluas cakupan keilmuan dan pendekatan multidisipliner dalam proses belajar-mengajar.

Dalam bidang penelitian, ia memiliki minat yang kuat pada tema-tema terkait analisis spasial, Sistem Informasi Geografis (SIG),

transformasi digital, dan bisnis digital. Ketertarikan ini ia wujudkan dalam bentuk riset, kegiatan pengabdian masyarakat, serta kolaborasi lintas sektor yang mendorong pemanfaatan teknologi spasial dan digital dalam pemerintahan maupun sektor privat.