



DASAR BIOSTATISTIKA UNTUK PENELITI

PENULIS :

Reza Fahlevi
Patrich Phill Edrich Papilaya
Wara Alfa Syukrilla
Putu Erma Pradnyani
Nur Al-faida
Ardiana Fatma Dewi
Ayatullah Harun
Arif Rachman
Asruria Sani Fajriah

DASAR BIOSTATISTIKA UNTUK PENELITI

**Reza Fahlevi
Patrich Phill Edrich Papilaya
Wara Alfa Syukrilla
Putu Erma Pradnyani
Nur Al-faida
Ardiana Fatma Dewi
Ayatullah Harun
Arif Rachman
Asruria Sani Fajriah**



GETPRESS INDONESIA

DASAR BIOSTATISTIKA UNTUK PENELITI

Penulis : Reza Fahlevi

Patrich Phill Edrich Papilaya

Wara Alfa Syukrilla

Putu Erma Pradnyani

Nur Al-faida

Ardiana Fatma Dewi

Ayatullah Harun

Arif Rachman

Asruria Sani Fajriah

ISBN : 978-623-125-137-4

Editor : Dr. Oktavianis, M.Biomed.

Penyunting: Mila Sari, M.Si

Desain Sampul dan Tata Letak : Tri Putri Wahyuni, S.Pd.

Penerbit: CV GETPRESS INDONESIA

Anggota IKAPI No. 033/SBA/2022

Redaksi:

Jl. Palarik RT 01 RW 06, Kelurahan Air Pacah
Kecamatan Koto Tengah, Padang, Sumatera Barat

website: www.getpress.co.id

email: adm.getpress@gmail.com

Cetakan pertama, April 2024

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

KATA PENGANTAR

Dengan mengucapkan puji syukur kehadiran Allah SWT, atas limpahan rahmat dan hidayahNya, maka Penulisan Buku dengan judul Dasar Biostatistika Untuk Peneliti dapat diselesaikan dengan kerjasama tim penulis. Buku ini merupakan pengantar singkat tentang Biostatistika, sebuah bidang penting dalam ilmu kedokteran dan kesehatan. Buku ini dirancang untuk memberikan pemahaman dasar tentang konsep-konsep statistik yang diperlukan dalam menganalisis data medis.

Dalam buku ini, Anda akan diajak untuk mempelajari berbagai hal, mulai dari definisi Biostatistika, perkembangannya, hingga penerapan metode statistik baik secara deskriptif maupun inferensial. Anda juga akan belajar tentang macam-macam data, skala ukuran, populasi, sampel, dan metode pengumpulan data.

Buku ini ditujukan untuk mahasiswa dan praktisi kesehatan yang ingin memahami dasar-dasar Biostatistika. Semoga buku ini dapat membantu Anda dalam memahami dan mengaplikasikan konsep-konsep statistik dalam dunia kedokteran dan kesehatan.

Buku ini masih banyak kekurangan dalam penyusunannya. Oleh karena itu, kami sangat mengharapkan kritik dan saran demi perbaikan dan kesempurnaan buku ini selanjutnya. Kami mengucapkan terima kasih kepada berbagai pihak yang telah membantu dalam proses penyelesaian Buku ini. Semoga Buku ini dapat menjadi sumber referensi dan literatur yang mudah dipahami dan dapat digunakan diberbagai Perguruan Tinggi.

Padang, April 2024
Penulis

DAFTAR ISI

KATA PENGANTAR.....	i
DAFTAR ISI.....	ii
DAFTAR GAMBAR	v
BAB 1 PENGERTIAN BIOSTATISTIK.....	1
1.1 Definisi Biostatistik.....	1
1.2 Ruang Lingkup Biostatistik.....	4
1.3 Data Komponen Utama Sebuah Studi.....	11
DAFTAR PUSTAKA	16
BAB 2 PERKEMBANGAN BIOSTATISTIKA	17
2.1 Sejarah Perkembangan.....	17
2.2 Pembelajaran Mesin dan Kecerdasan Buatan	47
DAFTAR PUSTAKA	58
BAB 3 STATISTIKA DESKRIPTIF	71
3.1 Pendahuluan.....	71
3.2 Ukuran Pemusatan Data	71
3.3 Ukuran Variabilitas Data	74
3.4 Ukuran Distribusi Frekuensi.....	76
3.5 Statistika Deskriptif <i>Univariate</i> dan <i>Bivariate</i>	80
DAFTAR PUSTAKA	82
BAB 4 STATISTIKA INFERENSIAL.....	83
4.1 Pendahuluan.....	83
4.2 Kegunaan Statistika Inferensial.....	85
4.3 Syarat Statistika Inferensial	86
4.4 Jenis Statistika Inferensial.....	87
4.5 Implementasi Statistika Inferensial yang sering digunakan dalam Penelitian	89
DAFTAR PUSTAKA	102
BAB 5 MACAM-MACAM DATA.....	103
5.1 Pendahuluan.....	103
5.2 Jenis Data Menurut Cara Memperolehnya.....	104
5.3 Klasifikasi Data Berdasarkan Sifatnya	110

5.3.1 Data Kuantitatif.....	111
5.3.2 Data Kualitatif.....	112
5.4 Jenis Data Menurut Waktu Pengumpulannya	112
5.4.1 Data Cross Section	113
5.4.2. Data <i>Time Series</i> / Berkala.....	114
DAFTAR PUSTAKA	117
BAB 6 SKALA UKURAN.....	119
6.1 Skala Pengukuran.....	119
6.2 Skala Nominal	119
6.3 Skala Ordinal	120
6.4 Skala Interval.....	121
6.4 Skala Rasio.....	122
DAFTAR PUSTAKA	124
BAB 7 POPULASI.....	125
7.1 Pengertian Populasi.....	128
7.2 Jenis Populasi	130
7.3 Kriteria Populasi	137
7.4 Populasi dalam Penelitian Kualitatif.....	140
7.5 Hubungan Populasi dan Sampel	143
DAFTAR PUSTAKA	145
BAB 8 SAMPEL.....	147
8.1 Pendahuluan.....	147
8.2 Sampel dalam Konteks Penelitian.....	148
8.3 Proses Pemilihan Sampel	151
8.4 Ukuran Sampel	156
8.5 Teknik Pengumpulan Data dari Sampel	160
8.6 Analisis Statistik pada Sampel.....	162
8.7 Pentingnya Konsep Sampel dalam Penelitian dan Biostatistika.....	163
DAFTAR PUSTAKA	164

BAB 9 METODE PENGUMPULAN DATA.....	167
9.1 Pendahuluan.....	167
9.2 Jenis Data	169
9.3 Metode Pengumpulan Data	171
DAFTAR PUSTAKA	181
BIODATA PENULIS	

DAFTAR GAMBAR

Gambar 1. 1. Angka kematian bayi per 1000 kelahiran hidup menurut ras di California dan Amerika Serikat pada tahun 1967	13
Gambar 1. 2. Angka kematian bayi per 1000 kelahiran hidup dengan metode konvensional dan kohort berdasarkan ras di California, 1965–1967	14
Gambar 2. 1. Ronald A. Fisher, ahli statistik dan genetika.....	22
Gambar 3. 1. Histogram data.....	79
Gambar 3. 2. Diagram Batang	79
Gambar 7. 1. Skema hubungan antara populasi Target dan Populasi Terjangkau	134

BAB 1

PENGERTIAN BIOSTATISTIK

Oleh Reza Fahlevi

1.1 Definisi Biostatistik

Biostatistik didefinisikan sebagai penerapan prinsip statistik dalam kedokteran, kesehatan masyarakat, atau biologi. Prinsip statistik didasarkan pada matematika terapan dan mencakup alat dan teknik untuk mengumpulkan informasi atau data dan kemudian merangkum, menganalisis, dan menafsirkan hasil tersebut. Prinsip-prinsip ini mencakup pembuatan kesimpulan dan penarikan kesimpulan yang memperhitungkan ketidakpastian secara tepat. Biostatistik adalah penerapan metode statistik pada ilmu biologi dan kehidupan. Metode statistik mencakup prosedur untuk (1) merancang penelitian, (2) mengumpulkan data, (3) menyajikan dan merangkum data, dan (4) menarik kesimpulan dari data sampel ke suatu populasi. Metode-metode ini sangat berguna dalam penelitian yang melibatkan manusia karena proses yang diteliti seringkali sangat kompleks. Karena kerumitan ini, biasanya dilakukan pengukuran dalam jumlah besar pada subjek penelitian untuk membantu proses penemuan; namun, kompleksitas dan banyaknya data ini sering kali menutupi proses yang mendasarinya. Dalam situasi inilah metode sistematis yang ditemukan dalam statistik membantu menciptakan keteraturan dari kekacauan yang tampak. Ini adalah beberapa bidang penerapannya:

1. Pengumpulan statistik penting misalnya, angka kematian yang digunakan untuk memberikan *informasi* dan *memantau* status kesehatan masyarakat.

2. Analisis catatan kecelakaan untuk *mengetahui* waktu-waktu dalam setahun ketika jumlah kecelakaan terbesar terjadi di suatu instalasi dan *memutuskan* kapan kebutuhan akan instruksi keselamatan paling tinggi.
3. Uji klinis untuk *menentukan* apakah pengobatan hipertensi baru memiliki kinerja yang lebih baik dibandingkan *pengobatan* standar untuk hipertensi esensial ringan hingga sedang.
4. Survei untuk memperkirakan proporsi perempuan usia subur berpenghasilan rendah yang menderita anemia defisiensi besi.
5. Studi untuk *mengetahui* apakah paparan medan elektromagnetik merupakan faktor risiko leukemia atau tidak.

Teknik biostatistik dapat digunakan untuk menjawab setiap pertanyaan di atas. Dalam biostatistik terapan, tujuannya biasanya untuk membuat kesimpulan tentang populasi tertentu. Menurut definisinya, populasi ini adalah kumpulan semua individu yang ingin kita nyatakan pernyataannya. Populasi yang menjadi perhatian mungkin adalah semua orang dewasa yang tinggal di Amerika Serikat atau semua orang dewasa yang tinggal di kota Boston. Definisi populasi bergantung pada pertanyaan penelitian peneliti, yang merupakan tujuan analisis. Misalkan populasi yang diteliti adalah semua orang dewasa yang tinggal di Amerika Serikat dan kita ingin memperkirakan proporsi semua orang dewasa yang menderita penyakit kardiovaskular. Untuk menjawab pertanyaan ini secara lengkap, kami akan memeriksa setiap orang dewasa di Amerika Serikat dan menilai apakah mereka menderita penyakit kardiovaskular. Pilihan yang lebih baik dan realistis adalah dengan menggunakan analisis statistik untuk memperkirakan proporsi yang diinginkan.

Biostatistik membantu administrator, legislator, manajer pabrik, dan peneliti dalam menjawab pertanyaan. Pertanyaan-pertanyaan yang menarik secara eksplisit terlihat pada contoh 2, 3 dan 5 di atas, apakah pola kecelakaan musiman memberikan petunjuk untuk mengurangi kejadiannya?; apakah obat baru tersebut lebih efektif dibandingkan obat standar?; dan apakah paparan medan elektromagnetik merupakan faktor risiko? Dalam contoh 1 dan 4, nilai atau perkiraan yang diperoleh merupakan pengukuran pada suatu titik waktu yang dapat digunakan dengan pengukuran pada titik waktu lain untuk menentukan apakah suatu kebijakan berubah atau tidak, misalnya peningkatan pendanaan Medicaid sebesar 10 persen di setiap negara bagian, berpengaruh. Kata-kata yang dicetak miring pada daftar sebelumnya menunjukkan bahwa statistik bukanlah kumpulan pengetahuan substantif, melainkan kumpulan metode untuk memperoleh, mengatur, merangkum, dan menyajikan informasi serta menarik kesimpulan. Namun, setiap kali kita mengambil kesimpulan, ada kemungkinan kita salah. Untungnya, metode statistik menggabungkan gagasan probabilitas yang memungkinkan kita menentukan peluang membuat kesimpulan yang salah. Seperti yang dikemukakan oleh Profesor C.R. Rao (1989), "Statistik memberikan peluang untuk berhasil."

Dalam biostatistik mempelajari sampel atau subkumpulan populasi yang diteliti. Dalam contoh ini, misalnya sampel orang dewasa yang tinggal di Amerika Serikat dan menilai apakah masing-masing orang tersebut menderita penyakit kardiovaskular atau tidak. Jika sampel mewakili populasi, maka proporsi orang dewasa dalam sampel yang menderita penyakit kardiovaskular harus merupakan perkiraan yang baik mengenai proporsi orang dewasa dalam populasi yang menderita penyakit kardiovaskular. Dalam biostatistik menganalisis sampel dan kemudian membuat kesimpulan tentang populasi berdasarkan analisis sampel tersebut.

Kesimpulan ini merupakan suatu lompatan besar, terutama jika populasinya besar (misalnya populasi Amerika Serikat yang berjumlah 300 juta jiwa) dan sampelnya relatif kecil (misalnya 5.000 orang). Ketika kita mendengarkan laporan berita atau membaca tentang penelitian, kita sering berpikir tentang bagaimana hasilnya dapat diterapkan pada diri kita secara pribadi. Sebagian besar dari kita tidak pernah terlibat dalam studi penelitian. Kita sering bertanya-tanya apakah kita harus mempercayai hasil penelitian padahal kita, atau siapa pun yang kita kenal, tidak pernah berpartisipasi dalam penelitian tersebut.

1.2 Ruang Lingkup Biostatistik

Melakukan dan menafsirkan aplikasi biostatistik dengan tepat memerlukan perhatian pada sejumlah isu penting. Hal ini termasuk, namun tidak terbatas pada, hal-hal berikut:

1. Mendefinisikan dengan jelas tujuan atau pertanyaan penelitian
2. Memilih desain penelitian yang sesuai (yaitu cara pengumpulan data)
3. Memilih sampel yang representatif dan memastikan bahwa sampel mempunyai ukuran yang cukup
4. Mengumpulkan dan menganalisis data dengan cermat
5. Menghasilkan ringkasan pengukuran atau statistik yang sesuai
6. Menghasilkan ukuran dampak atau keterkaitan yang tepat
7. Mengukur ketidakpastian
8. Akuntansi yang tepat untuk hubungan antar karakteristik
9. Membatasi kesimpulan pada populasi yang sesuai

Bagaimana Tingkat Penyakit di Suatu Kelompok atau Wilayah Diukur?

Idealnya, sampel individu dalam kelompok atau wilayah yang diminati dipilih. Sampel tersebut harus cukup besar sehingga hasil analisis sampel cukup tepat. Secara umum, sampel yang lebih besar untuk analisis lebih disukai; namun, kami tidak ingin mengambil sampel peserta lebih banyak dari yang dibutuhkan, baik karena alasan finansial maupun etika. Sampel juga harus mewakili populasi. Misalnya, jika populasinya 60% perempuan, idealnya kita ingin sampelnya berjumlah sekitar 60% perempuan. Setelah sampel dipilih, setiap peserta dinilai berdasarkan status penyakitnya. Proporsi sampel yang menderita penyakit dihitung dengan mengambil rasio jumlah penderita penyakit terhadap jumlah sampel total. Proporsi ini merupakan perkiraan proporsi penduduk yang menderita penyakit. Misalkan proporsi sampel dihitung sebesar 0,17 (yaitu, 17% dari sampel menderita penyakit tersebut). Kami memperkirakan proporsi penduduk yang mengidap penyakit sekitar 0,17 (atau 17%). Karena ini berdasarkan perkiraan pada satu sampel, kita harus memperhitungkan ketidakpastian, dan hal ini tercermin dalam apa yang disebut margin kesalahan. Hal ini mungkin mengakibatkan perkiraan kami bahwa proporsi penduduk yang mengidap penyakit berkisar antara 0,13 hingga 0,21 (atau 13% hingga 21%).

Penelitian ini kemungkinan besar akan dilakukan pada satu waktu; jenis penelitian ini biasa disebut dengan studi *cross sectional*. Perkiraan kami mengenai luasnya penyakit hanya mengacu pada periode penelitian. Tidaklah tepat untuk membuat kesimpulan tentang tingkat penyakit di masa depan berdasarkan penelitian ini. Jika kita memilih orang dewasa yang tinggal di Boston sebagai populasi kita, maka tidak tepat untuk menyimpulkan bahwa tingkat penyakit di kota-kota lain atau di

bagian lain Massachusetts akan sama dengan yang diamati pada sampel warga Boston.

Tugas memperkirakan tingkat penyakit di suatu wilayah atau kelompok tampak mudah di permukaan. Namun, ada banyak permasalahan yang membuat segalanya menjadi rumit. Misalnya, dari mana kita mendapatkan daftar populasi, bagaimana kita memutuskan siapa yang akan dijadikan sampel, bagaimana kita memastikan bahwa kelompok tertentu terwakili (misalnya perempuan) dalam sampel, dan bagaimana kita menemukan orang-orang yang kita identifikasi. Untuk sampel dan meyakinkan mereka untuk berpartisipasi? Semua pertanyaan ini harus dijawab dengan benar untuk menghasilkan data yang valid dan kesimpulan yang benar.

Bagaimana Tingkat Perkembangan Penyakit Baru Diperkirakan?

Untuk memperkirakan laju perkembangan suatu penyakit baru misalnya penyakit kardiovaskular kita memerlukan strategi pengambilan sampel yang spesifik. Untuk analisis ini, kami hanya akan mengambil sampel orang-orang yang bebas dari penyakit kardiovaskular dan memantau mereka secara prospektif (di masa mendatang) untuk menilai perkembangan penyakit tersebut. Masalah utama dalam jenis studi ini adalah periode tindak lanjut; penyelidik harus memutuskan apakah akan mengikuti peserta selama 1, 5, atau 10 tahun, atau periode lain, untuk mengetahui perkembangan penyakitnya. Jika kita ingin memperkirakan perkembangan penyakit dalam kurun waktu 10 tahun, maka diperlukan pengamatan terhadap setiap partisipan dalam sampel selama 10 tahun untuk menentukan status penyakit mereka. Rasio jumlah kasus penyakit baru terhadap jumlah sampel mencerminkan proporsi atau kejadian kumulatif penyakit baru selama periode tindak lanjut yang telah ditentukan. Misalkan kita

mengikuti setiap peserta dalam sampel kita selama 5 tahun dan menemukan bahwa 2,4% mengidap penyakit. Sekali lagi, secara umum penting untuk memberikan serangkaian nilai yang masuk akal untuk proporsi kasus penyakit baru; hal ini dicapai dengan memasukkan margin kesalahan untuk mencerminkan ketepatan perkiraan kami. Memasukkan margin kesalahan dapat menghasilkan perkiraan kejadian penyakit kumulatif antara 1,2% hingga 3,6% selama 5 tahun. Epidemiologi adalah bidang studi yang berfokus pada studi tentang kesehatan dan penyakit pada populasi manusia, pola kesehatan atau penyakit, dan faktor-faktor yang mempengaruhi pola tersebut. Penelitian yang dijelaskan di sini adalah contoh penelitian epidemiologi.

Bagaimana Faktor Risiko atau Karakteristik yang Mungkin Terkait dengan Perkembangan atau Perkembangan Penyakit Diidentifikasi?

Misalkan kita berhipotesis bahwa faktor risiko atau paparan tertentu berhubungan dengan perkembangan suatu penyakit. Ada beberapa desain penelitian atau cara berbeda yang dapat digunakan untuk mengumpulkan informasi guna menilai hubungan antara faktor risiko potensial dan timbulnya penyakit. Desain studi yang paling tepat bergantung, antara lain, pada distribusi faktor risiko dan hasil pada populasi yang diteliti (misalnya, berapa banyak peserta yang kemungkinan memiliki faktor risiko tertentu atau tidak). Terlepas dari desain spesifik yang digunakan, baik faktor risiko maupun hasil harus diukur pada setiap anggota sampel. Jika tertarik pada hubungan antara faktor risiko dan perkembangan penyakit, akan kembali melibatkan peserta yang bebas penyakit pada awal penelitian dan memantau perkembangan penyakit semua peserta. Untuk menilai apakah ada hubungan antara faktor risiko dan hasil, memperkirakan proporsi (atau persentase) peserta dengan faktor risiko yang kemudian

mengembangkan penyakit dan membandingkannya dengan proporsi (atau persentase) peserta yang tidak memiliki faktor risiko. memiliki faktor risiko dan terus mengembangkan penyakit. Ada beberapa cara untuk membuat perbandingan ini; hal ini dapat didasarkan pada perbedaan proporsi atau rasio proporsi.

Misalkan di antara mereka yang mempunyai faktor risiko, 12% mengembangkan penyakit selama masa tindak lanjut, dan di antara mereka yang bebas dari faktor risiko, 6% mengembangkan penyakit. Rasio proporsinya disebut risiko relatif dan di sini sama dengan $0,12 / 0,06 = 2,0$. Interpretasinya adalah orang dengan faktor risiko terkena penyakit dua kali lebih banyak dibandingkan orang tanpa faktor risiko. Persoalannya kemudian adalah menentukan apakah perkiraan ini, yang diamati dalam satu sampel penelitian, mencerminkan peningkatan risiko dalam populasi. Memperhitungkan ketidakpastian dapat menghasilkan perkiraan risiko relatif antara 1,1 hingga 3,2 kali lebih tinggi bagi orang-orang dengan faktor risiko tersebut. Karena rentang tersebut mengandung nilai risiko yang lebih besar dari 1, maka data mencerminkan peningkatan risiko (karena nilai 1 menunjukkan tidak ada peningkatan risiko).

Masalah lain dalam menilai hubungan antara faktor risiko tertentu dan status penyakit melibatkan pemahaman hubungan yang kompleks di antara faktor-faktor risiko. Orang yang mempunyai faktor risiko mungkin berbeda dengan orang yang bebas dari faktor risiko; misalnya, mereka mungkin lebih tua dan lebih mungkin memiliki faktor risiko lain. Terdapat metode yang dapat digunakan untuk menilai hubungan antara faktor risiko yang dihipotesiskan dan status penyakit sambil mempertimbangkan dampak faktor risiko lainnya. Teknik-teknik ini melibatkan pemodelan statistik.

Bagaimana Efektivitas Obat Baru Ditentukan?

Desain penelitian yang ideal dari sudut pandang statistik adalah uji coba terkontrol secara acak atau uji klinis. (Istilah klinis berarti penelitian ini melibatkan manusia.) Misalnya, kita ingin menilai efektivitas obat baru yang dirancang untuk menurunkan kolesterol. Kebanyakan uji klinis melibatkan kriteria inklusi dan eksklusi spesifik. Misalnya, kita mungkin ingin memasukkan hanya orang-orang dengan kadar kolesterol total melebihi 200 atau 220, karena pengobatan baru ini kemungkinan besar mempunyai peluang terbaik untuk menunjukkan efek pada orang-orang dengan kadar kolesterol tinggi. Kami mungkin juga mengecualikan orang-orang dengan riwayat penyakit kardiovaskular. Setelah kriteria inklusi dan eksklusi ditentukan, kami merekrut peserta. Setiap peserta secara acak ditugaskan untuk menerima obat percobaan baru atau obat kontrol. Komponen pengacakan adalah fitur utama dalam penelitian ini.

Pengacakan secara teoritis mendorong keseimbangan antara kelompok pembanding. Obat kontrol dapat berupa plasebo (zat *inert*) atau obat penurun kolesterol yang dianggap sebagai standar perawatan saat ini. Pemilihan pembanding yang tepat bergantung pada sifat penyakitnya. Misalnya, pada penyakit yang mengancam jiwa, tidak etis jika kita menunda pengobatan; jadi pembanding plasebo tidak akan pernah tepat. Dalam contoh ini, plasebo mungkin cocok selama kadar kolesterol partisipan tidak terlalu tinggi sehingga memerlukan pengobatan. Ketika peserta didaftarkan dan diacak untuk menerima perlakuan eksperimental atau pembanding, mereka tidak diberitahu perlakuan mana yang diberikan kepada mereka. Ini disebut membutakan atau menutupi. Peserta kemudian diinstruksikan tentang dosis yang tepat dan setelah waktu yang telah ditentukan, kadar kolesterol diukur dan dibandingkan antar kelompok. Karena peserta secara acak ditugaskan ke kelompok perlakuan, kelompok tersebut harus

sebanding dalam semua karakteristik kecuali perlakuan yang diterima. Jika menemukan bahwa kadar kolesterol berbeda antar kelompok, perbedaan tersebut kemungkinan besar disebabkan oleh pengobatan.

Sekali lagi, kita harus menafsirkan perbedaan yang diamati setelah memperhitungkan peluang atau ketidakpastian. Jika kami mengamati perbedaan besar dalam kadar kolesterol antara peserta yang menerima obat percobaan dan pembanding, kami dapat menyimpulkan bahwa obat percobaan tersebut efektif. Namun, kesimpulan mengenai efek obat hanya dapat digeneralisasikan pada populasi yang menjadi asal partisipan khususnya, pada populasi yang ditentukan berdasarkan kriteria inklusi dan eksklusi. Uji klinis harus dirancang dan dianalisis dengan cermat. Terdapat sejumlah isu khusus mengenai uji klinis. Uji klinis dibahas secara luas dalam berita, khususnya baru-baru ini. Obat-obatan ini diatur secara ketat di Amerika Serikat oleh Badan Pengawas Obat dan Makanan (FDA). Baru-baru ini laporan berita membahas penelitian yang melibatkan obat-obatan yang diberikan persetujuan untuk indikasi tertentu dan kemudian dikeluarkan dari pasar karena masalah keamanan. Kami meninjau studi-studi ini dan menilai bagaimana studi-studi tersebut dilakukan dan, yang lebih penting, mengapa studi-studi tersebut dievaluasi ulang. Untuk mengevaluasi obat, uji coba terkontrol secara acak dianggap sebagai standar emas. Namun, hal tersebut dapat menimbulkan kontroversi. Penelitian selain uji klinis kurang ideal dan seringkali lebih kontroversial.

Apa yang Dapat Menjelaskan Hasil Kontradiksi Antara Berbagai Penelitian pada Penyakit yang Sama?

Semua studi statistik didasarkan pada analisis sampel dari populasi yang diminati. Terkadang, penelitian tidak dirancang dengan tepat sehingga hasilnya mungkin dipertanyakan.

Terkadang, terlalu sedikit peserta yang terdaftar, sehingga dapat menyebabkan hasil yang tidak tepat dan bahkan tidak akurat. Ada juga contoh di mana penelitian dirancang dengan tepat, namun dua replikasi yang berbeda memberikan hasil yang berbeda. Sepanjang buku ini, kita akan membahas bagaimana dan kapan hal ini mungkin terjadi.

Dalam buku ini, masing-masing poin sebelumnya dibahas secara bergantian. Pada buku ini akan menjelaskan cara mengumpulkan dan meringkas data dan cara membuat kesimpulan yang tepat. Untuk mencapai hal ini, menggunakan prinsip biostatistik yang didasarkan pada teori matematika dan probabilitas. Tujuan utamanya adalah untuk memahami dan menafsirkan analisis biostatistik. Sekarang mari kita meninjau kembali pertanyaan awal kita dan memikirkan beberapa permasalahan yang telah diidentifikasi sebelumnya.

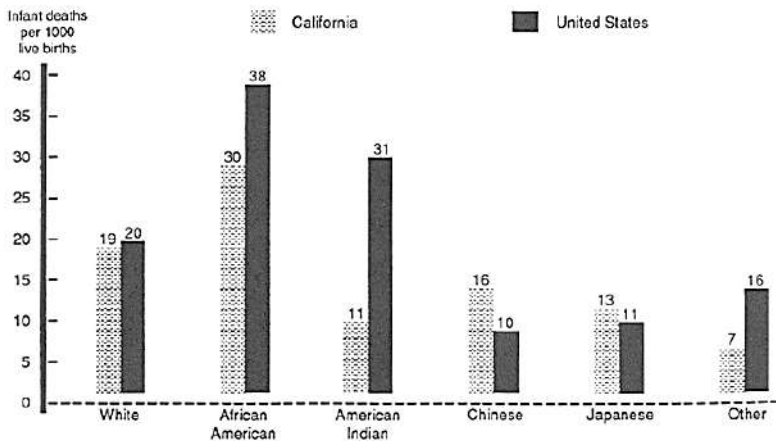
1.3 Data Komponen Utama Sebuah Studi

Penting untuk memahami metode-metode dan penggunaan metode biostatistik karena ini akan membantu Anda memahami laporan penelitian, merancang studi, dan melakukan penelitian. Namun, pembaca tidak perlu merasa terbebani dengan banyaknya metode analisis dan perhitungan terkait yang disajikan. Yang lebih penting daripada metode yang digunakan dalam analisis adalah penggunaan desain penelitian yang tepat dan definisi serta pengukuran variabel penelitian yang tepat. Anda tidak dapat melakukan penelitian yang baik tanpa data yang baik. Terkadang, karena pemahaman data yang tidak lengkap atau kemungkinan adanya masalah pada data, kesimpulan dari suatu penelitian mungkin bermasalah. Misalnya, pertimbangkan sebuah penelitian untuk menentukan apakah sunat dikaitkan dengan kanker serviks. Salah satu permasalahan yang harus ditetapkan oleh peneliti

adalah bagaimana menentukan status sunat. Cara termudah adalah dengan bertanya kepada laki-laki apakah dia sudah disunat; namun, Lilienfeld dan Graham (1958) menemukan bahwa 34 persen dari 192 pasien laki-laki yang mereka teliti secara berturut-turut memberikan jawaban yang salah tentang status sunat mereka. Sebagian besar tanggapan yang salah disebabkan oleh para pria yang tidak mengetahui bahwa mereka telah disunat. Oleh karena itu, penggunaan pertanyaan langsung dibandingkan pemeriksaan dapat menyebabkan kesimpulan yang salah mengenai hubungan antara sunat dan kanker serviks.

Dalam contoh ketergantungan pada ingatan atau pengetahuan subjek penelitian bisa saja merupakan sebuah kesalahan. Yaffe dan Shapiro (1979) memberikan contoh lain mengenai potensi masalah ketika tanggapan subjek penelitian digunakan. Mereka memeriksa keakuratan laporan subjek mengenai pemanfaatan dan pengeluaran layanan kesehatan selama tujuh bulan dibandingkan dengan yang ditunjukkan dalam catatan medis dan asuransi mereka untuk dua wilayah geografis. Di wilayah Baltimore yang menyediakan data dari sekitar 375 rumah tangga, subjek melaporkan hanya 73 persen dari kunjungan ke kantor dokter yang teridentifikasi dan hanya 54 persen dari kunjungan ke klinik. Hasil untuk Washington County, Maryland yang didasarkan pada sekitar 315 rumah tangga, menunjukkan akurasi 84 persen untuk kunjungan ke dokter tetapi hanya 39 persen akurasi untuk kunjungan klinik. Oleh karena itu, laporan pemanfaatan layanan kesehatan menurut subjek mungkin terlalu meremehkan pemanfaatan sebenarnya, dan, mungkin yang lebih penting, keakuratannya dapat berbeda-beda menurut jenis pemanfaatan dan subkelompok populasi. Perlu dicatat bahwa kesimpulan ini didasarkan pada asumsi bahwa catatan medis/asuransi lebih akurat dibandingkan ingatan subjek.

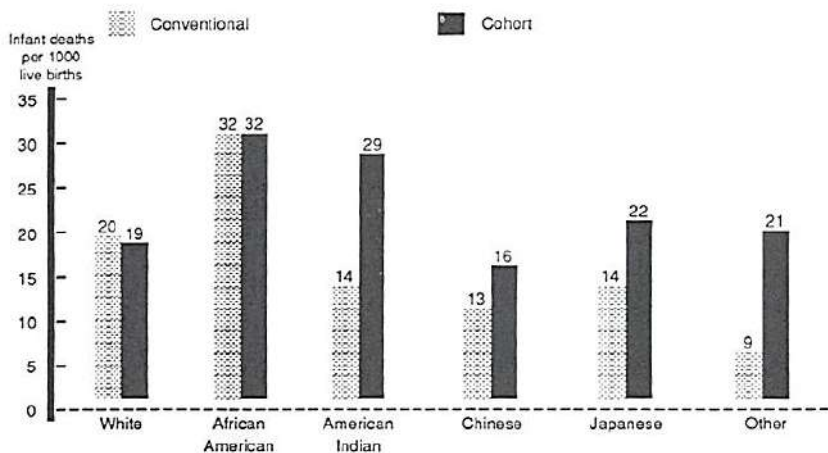
Salah satu contoh bagaimana kesimpulan yang salah dapat dicapai karena kegagalan memahami cara pengumpulan data berasal dari Norris dan Shipley (1971). Gambar 1 menunjukkan angka kematian bayi yang dihitung secara konvensional sebagai rasio jumlah kematian bayi terhadap jumlah kelahiran hidup selama periode yang sama dikalikan 1000 untuk kelompok ras berbeda di Kalifornia dan Amerika Serikat pada tahun 1967.



Gambar 1. 1. Angka kematian bayi per 1000 kelahiran hidup menurut ras di California dan Amerika Serikat pada tahun 1967 (Sumber: Forthofer, Lee, & Hernandez, 2007)

Norris dan Shipley mempertanyakan keakuratan angka tersebut untuk orang Indian Amerika karena angka tersebut jauh lebih rendah daripada angka orang Indian Amerika di AS dan bahkan lebih rendah daripada angka untuk orang Amerika keturunan Tionghoa dan Jepang di Kalifornia. Oleh karena itu, mereka menggunakan metode *kohort* untuk menghitung ulang angka kematian bayi. Tingkat *kohort* didasarkan pada pelacakan semua anak yang lahir di California selama satu tahun dan pengamatan berapa banyak bayi yang meninggal sebelum mereka mencapai usia satu tahun. Beberapa kematian tidak tercakup

misalnya, bayi yang meninggal di luar Kalifornia namun diperkirakan bahwa hampir 97 persen kematian bayi pada kelompok tersebut dicatat dalam catatan kematian di Kalifornia. Norris dan Shipley menggunakan data selama tiga tahun dalam pemeriksaan ulang mereka. angka kematian bayi untuk memberikan stabilitas angka yang lebih baik. Gambar 2 menunjukkan angka konvensional dan angka *kohort* pada periode 1965–1967 berdasarkan ras. Penggunaan data dari tiga tahun tidak banyak mengubah tarif konvensional. Tarif konvensional untuk orang Indian Amerika di Kalifornia masih jauh lebih rendah dibandingkan tarif di AS untuk orang Indian Amerika, meskipun sekarang tarifnya sedikit di atas tarif di Tiongkok dan Jepang. Namun, angka *kohort* untuk orang Indian Amerika kini mendekati angka serupa yang ditemukan di Amerika Serikat. Angka kematian pada ras Cina, Jepang, dan Amerika serta ras lainnya juga meningkat secara signifikan ketika metode penghitungan *kohort* digunakan. Apa penjelasan atas perbedaan hasil antara metode penghitungan angka kematian bayi ini?



Gambar 1. 2. Angka kematian bayi per 1000 kelahiran hidup dengan metode konvensional dan kohort berdasarkan ras di California, 1965–1967 (Sumber: Forthofer, Lee, & Hernandez, 2007)

Norris dan Shipley mengaitkan perbedaan besar ini dengan cara pembuatan akta kelahiran dan kematian, yang digunakan dalam metode konvensional. Mereka menemukan bahwa akta kelahiran biasanya diisi oleh staf rumah sakit yang terutama menangani ibu, sehingga akta kelahiran biasanya mencerminkan ras dari ibu. Direktur pemakaman bertanggung jawab untuk melengkapi catatan kematian dan biasanya berhubungan dengan ayah, yang mungkin berasal dari kelompok ras yang berbeda dari ras pada ibu. Oleh karena itu, identifikasi ras bayi dapat bervariasi antara catatan kelahiran dan kematian ketidaksesuaian antara pembilang (kematian) dan penyebut (kelahiran) dalam penghitungan angka kematian bayi. Metode kohort tidak terpengaruh oleh kemungkinan perbedaan ini karena metode ini hanya menggunakan ras anak yang tercatat dalam akta kelahiran.

Sejak data tahun 1989, Pusat Statistik Kesehatan Nasional (NCHS) terutama menggunakan ras pada ibu yang diambil dari akta kelahiran dalam membuat tabulasi data kelahiran. Perubahan ini harus menghilangkan masalah yang disebabkan oleh adanya orang tua dari dua kelompok ras yang menggunakan metode konvensional dalam menghitung angka kematian bayi. Jadi kita dapat melihat bahwa data jarang berbicara dengan jelas dan biasanya memerlukan penerjemah. Penerjemah seseorang seperti Norris dan Shipley akrab dengan pokok bahasan, memahami apa yang seharusnya diwakili oleh data, dan mengetahui bagaimana data dikumpulkan.

DAFTAR PUSTAKA

- Forthofer, R. N., Lee, E. S., & Hernandez, M. (2006). *Biostatistics: a guide to design, analysis and discovery*. Elsevier.
- Kaikkonen, K. S., Kortelainen, M. L., Linna, E., & Huikuri, H. V. (2006). Family history and the risk of sudden cardiac death as a manifestation of an acute coronary event. *Circulation*, 114(14), 1462-1467.
- Lilienfeld, A. M., & Graham, S. (1958). Validity of determining circumcision status by questionnaire as related to epidemiological studies of cancer of the cervix. *Journal of the National Cancer Institute*, 21(4), 713-720.
- Magnus, M. (2007). *Essentials of Infectious Disease Epidemiology*. Sudbury, MA: Jones and Bartlett.
- Norris, F. D., & Shipley, P. W. (1971). A closer look at race differentials in California's infant mortality, 1965-7. *HSMHA Health Reports*, 86(9), 810.
- Panel II, A. T. (1993). Expert Panel on Detection, Evaluation, and Treatment of High Blood Cholesterol in Adults. Summary of the second report of The National Cholesterol Education Program (NCEP) expert panel on detection, evaluation, and treatment of high blood cholesterol in adults. *J Am Med Assoc*, 269, 3015-23.
- Sytkowski, P. A., D'Agostino, R. B., Belanger, A., & Kannel, W. B. (1996). Sex and time trends in cardiovascular disease incidence and mortality: the Framingham Heart Study, 1950-1989. *American journal of epidemiology*, 143(4), 338-350.
- Wilson, P. W., D'Agostino, R. B., Levy, D., Belanger, A. M., Silbershatz, H., & Kannel, W. B. (1998). Prediction of coronary heart disease using risk factor categories. *Circulation*, 97(18), 1837-1847.

BAB 2

PERKEMBANGAN BIOSTATISTIKA

Oleh Patrich Phill Edrich Papilaya

2.1 Sejarah Perkembangan

Metode statistika berkembang dari waktu ke waktu. Metode statistika sendiri telah dikenal dari zaman dahulu/kuno dan berkembang sampai abad modern sekarang ini. Perkembangan metode statistika dibagi atas beberapa tahapan yang mana setiap tahapan perkembangan menampilkan temua-temua tersendiri yang mendasari kemajuan dan perkembangan ilmu statistika itu sendiri.

2.1.1 Perkembangan Metode Statistika Abad Permulaan- abad 17

Awal perkembangan statistika dari zaman kuno hingga abad ke-16 mencakup beberapa tahapan penting:

a) Zaman Kuno dan Kekaisaran Awal

Sejak awal peradaban, bentuk dasar statistika telah digunakan. Kekaisaran awal seperti Han dan Romawi mengumpulkan data sensus untuk mengetahui jumlah populasi, luas geografis, dan kekayaan.

b) Metode Statistika di Yunani Kuno

Sejarah statistika dapat dilacak kembali hingga abad ke-5 SM. Sejarawan Thucydides dalam "History of the Peloponnesian War" menjelaskan bagaimana orang-orang Athena menghitung tinggi dinding Platea dengan menghitung jumlah bata pada bagian dinding yang tidak dipilester. Mereka

mengulangi penghitungan beberapa kali dan nilai yang paling sering muncul (mode) digunakan untuk menentukan tinggi tangga yang diperlukan.

c) Penggunaan Metode Sampling Statistika (Abad ke-12)

"Trial of the Pyx" adalah tes kemurnian mata uang yang dilakukan oleh Royal Mint sejak abad ke-12. Tes ini menggunakan metode sampling statistika dengan memilih sampel koin dari kotak untuk diuji kemurniannya.

d) Statistika di Abad Pertengahan

"Nuova Cronica," sejarah Florence abad ke-14 oleh Giovanni Villani, berisi informasi statistik tentang populasi, perundang-undangan, perdagangan, pendidikan, dan fasilitas keagamaan. Ini dianggap sebagai pengenalan pertama statistika sebagai elemen positif dalam sejarah.

e) Perkembangan Konsep Statistika (Abad ke-16)

Konsep mean aritmetika, meskipun dikenal oleh orang Yunani, tidak digeneralisasi hingga abad ke-16. Simon Stevin memperkenalkan sistem desimal pada tahun 1585, yang memudahkan perhitungan ini. Konsep median diperkenalkan oleh Edward Wright dalam bukunya tentang navigasi pada tahun 1599. Istilah 'statistik' diperkenalkan oleh Girolamo Ghilini pada tahun 1589 (Stigler, 1990).

Pada tahun 1657, Christian Huygens, seorang ilmuwan Jerman, menerbitkan buku "*De Ratiocinilis in Ludo Aleae*" yang membahas tentang perjudian. Buku ini secara signifikan mempopulerkan teori peluang. Huygens dianggap sebagai guru dari Leibniz, salah satu tokoh penting dalam sejarah matematika dan filosofi. Selain itu, Huygens juga dikenal sebagai ilmuwan yang

merumuskan hukum untuk menggambarkan tumbukan elastis antara dua tubuh dan menulis persamaan kuadrat untuk apa yang akan menjadi hukum gerak Newton yang kedua. Perkembangan teori peluang yang dipopulerkan oleh Huygens kemudian menjadi bagian integral dari perkembangan ilmu statistika (Cornelis, 2005).

Pada abad ke-17, statistika deskriptif mulai berkembang sebagai cabang statistika yang berfokus pada pengumpulan, penyusunan, dan penyajian data sehingga dapat dengan mudah dipahami dan diinterpretasikan. Sementara itu, ilmu peluang juga mulai muncul, awalnya berasal dari meja judi. Perkembangan ini menandai awal dari penerapan konsep-konsep statistika dalam berbagai aspek kehidupan, termasuk dalam analisis data, pengambilan keputusan, dan pemodelan fenomena acak (Gani, 2012).

Beberapa tokoh-tokoh penting yang muncul pada abad ke-17, yaitu dalam ilmu statistika deskriptif dan ilmu peluang berkontribusi signifikan. Jacob Bernoulli (1654-1705) dan Abraham de Moivre (1667-1754) adalah dua dari mereka (Bernoulli dan Syeilla, 2006).

Jacob Bernoulli:

- a. Anggota keluarga Bernoulli yang memberikan kontribusi besar untuk matematika dan ilmu statistik.
- b. Menyusun "Ars Conjectandi", buku yang mengandung karya kombinasi dan kelas-kelas matematika yang diri oleh Jacob Bernoulli dan tertua setelah kematiannya pada tahun 1713.
- c. Menemukan ketiga-ketiga kasus utama dalam ilmu peluang: kasus penuh syarat, kasus tidak syarat, dan kasus akibat syarat.
- d. Merupakan salah satu dari pencipta kasus Bernoulli, yang merupakan salah satu dari kasus terkenal dalam ilmu peluang.

Abraham de Moivre:

- a. Matematikawan Prancis yang terkenal karena penemuan kurva Normal dan persamaannya pada tahun 1733, yang menjadi dasar penting dalam teori statistik.
- b. Mengembangkan sekuen asimptotik untuk fungsi faktorial, yang digunakan untuk mengkaitkan distribusi binomial dengan distribusi Normal.
- c. Berkontribusi dalam pengembangan ilmu peluang bersama dengan Pascal dan Fermat, yang menjadi bahan pertukaran pikiran antara Pascal dan Fermat.

Kontribusi mereka ini menandai awal perkembangan ilmu statistika dan ilmu peluang pada abad ke-17.

2.1.2 Perkembangan Biostatistika pada Abad ke-18 dan 19

Pada abad ke-18, perkembangan awal biostatistika dimulai dengan penelitian tentang kesehatan populasi dan epidemiologi. Pada tahun 1760-an, British Government mengumpulkan data mortalitas dan penyebab utama kematian di seluruh Inggris. Hal ini dianggap sebagai evaluasi epidemiologi modern pertama. Selanjutnya, pada tahun 1799, Lancelletto menerbitkan buku tentang statistik medis dan merupakan salah satu buku pertama tentang statistik medis/kesehatan (Elandt-Johnson dan Johnson, 1980).

Pada abad ke-19, statistika menjadi alat untuk mempelajari variasi dalam data medis dan kesehatan populasi. William Farr, seorang dokter Inggris, menerbitkan laporan berkala tentang statistik medis Inggris sejak tahun 1838 (Pearson, 1978). Ia membuat statistik kematian lebih sistematis dan membangun dasar untuk data vital saat ini.

Pada tahun 1854, John Snow menggunakan analisis spasial untuk menghubungkan sumber air dengan wabah kolera di London, yang merupakan salah satu studi epidemiologi deskriptif pertama (Johnson, 1983). Selanjutnya pada akhir abad ke-19, statistika menjadi alat penting dalam bidang kesehatan masyarakat. Contohnya adalah karya Adolphe Quetelet tentang manusia dan pengembangan indeks massa tubuh (Cole, 2000). Pada masa itu terjadi peningkatan pemahaman tentang hubungan antara faktor sosial, lingkungan dan kesehatan populasi, sehingga merupakan awal dari biostatistika modern.

Beberapa aplikasi praktis perkembangan biostatistika pada abad ke-18 dan 19 antara lain:

- a. Pengumpulan dan analisis data vital/kematian secara sistematis untuk mempelajari epidemiologi penyakit dan meningkatkan kesehatan populasi (misalnya laporan statistik medis William Farr di Inggris).
- b. Penggunaan analisis statistik untuk menghubungkan faktor lingkungan dan sosial dengan kejadian penyakit, seperti yang dilakukan John Snow dalam menghubungkan sumber air dengan wabah kolera di London.
- c. Pengembangan indeks atau standar untuk mempelajari variasi antar individu dan kelompok, seperti indeks massa tubuh yang dikembangkan Adolphe Quetelet.
- d. Evaluasi dampak program kesehatan masyarakat dan intervensi, misalnya penilaian langkah-langkah sipil untuk menurunkan angka kematian.
- e. Pembuatan proyeksi penduduk berdasarkan data vital yang tersedia untuk perencanaan kebijakan kesehatan dan SDM.
- f. Identifikasi faktor risiko penyakit melalui studi asosiasi dengan menggunakan analisis statistik.

2.1.3 Perkembangan Biostatistika pada Awal Abad ke-20

Perkembangan Biostatistika pada Abad 20 melalui beberapa tahapan, antara lain :

1. Statistik dalam Biologi dan Genetika

Pada awal abad ke-20, biostatistik mulai dikenal terutama melalui pekerjaan **Ronald A. Fisher**, seorang ahli statistik dan genetika. Fisher mengembangkan metode statistik untuk analisis data biologi, terutama dalam studi genetika dan evolusi.

Statistik dalam biologi dan genetika merupakan area yang luas dan kompleks, yang mencakup berbagai aspek penting. Fokus utamanya adalah pada penggunaan metode statistik untuk memahami dan menganalisis data biologis dan genetika.



Gambar 2. 1. Ronald A. Fisher, ahli statistik dan genetika

Berikut adalah beberapa aspek kunci yang mendasari perkembangan Biostatistik :

- a) Analisis Variasi Genetik: Menggunakan statistik untuk memahami bagaimana variasi genetik mempengaruhi ciri-ciri organisme. Ini termasuk studi tentang pola warisan, mutasi, dan rekombinasi.

- b) Pemetaan Genetik dan Asosiasi Genetik: Pemetaan genetik mengidentifikasi lokasi gen-gen spesifik dalam kromosom. Studi asosiasi genetik, terutama genom-wide association studies (GWAS), menggunakan statistik untuk menghubungkan variasi genetik dengan penyakit atau ciri-ciri tertentu.
- c) Analisis Populasi Genetik: Memahami struktur populasi dan dinamika evolusi melalui analisis frekuensi alel dan genotip, migrasi, seleksi alam, dan efek genetik drift.
- d) Analisis Ekspresi Gen dan Transcriptomics: Menggunakan teknik statistik untuk menganalisis data ekspresi gen, seperti data microarray dan RNA-seq, untuk memahami bagaimana gen diaktifkan atau dinonaktifkan dalam sel atau jaringan tertentu.
- e) Metode Statistik dalam Genetika Kuantitatif: Memahami bagaimana gen-gen tertentu mempengaruhi ciri-ciri kuantitatif (seperti tinggi badan atau tekanan darah) melalui analisis statistik.
- f) Biostatistik dalam Genomik Fungsional: Menerapkan statistik untuk memahami fungsi gen dan interaksi dalam jaringan biologis. Ini termasuk studi tentang proteomik, metabolomik, dan interaksi gen-gen.
- g) Statistik dalam Epigenetik: Analisis statistik data epigenetik, termasuk metilasi DNA dan modifikasi histon, untuk memahami bagaimana faktor lingkungan mempengaruhi ekspresi gen.
- h) Bioinformatika: Menggunakan metode statistik dan komputasi untuk menganalisis data biologis besar. Bioinformatika sangat penting dalam era genomik untuk mengelola dan menganalisis data yang sangat besar dan kompleks.

- i) Desain Eksperimental dan Analisis Data: Pengembangan desain eksperimental yang kuat dan metode analisis data untuk studi genetik dan biologis, termasuk kontrol terhadap bias dan kesalahan.
- j) Pemodelan Statistik dan Komputasi: Mengembangkan model statistik dan algoritma komputasi untuk menyederhanakan dan menafsirkan data biologi dan genetik yang kompleks.

Statistik dalam biologi dan genetika terus berkembang dengan kemajuan teknologi dan pengetahuan. Hal ini memungkinkan peneliti untuk menjawab pertanyaan yang lebih kompleks dan memberikan wawasan baru dalam biologi dan kedokteran.

Dalam bidang statistik dalam biologi dan genetika, beberapa tokoh terkemuka telah memberikan kontribusi signifikan. Mereka tidak hanya mengembangkan konsep-konsep statistik penting tetapi juga menerapkannya dalam biologi dan genetika. Berikut adalah beberapa tokoh berpengaruh beserta karya-karya penting mereka:

- a) Ronald A. Fisher
 - Fisher dikenal sebagai salah satu bapak biostatistik modern. Ia mengembangkan banyak konsep statistik dasar yang masih digunakan hingga hari ini.
 - Buku penting: "The Genetical Theory of Natural Selection" (1999) – Buku ini menggabungkan prinsip genetika Mendelian dengan teori seleksi alam Charles Darwin dan memperkenalkan metode statistik dalam genetika. Tulisan penting: "*Statistical Methods for Research Workers*" (2017) Buku ini merupakan salah satu karya paling berpengaruh dalam statistik,

memperkenalkan teknik-teknik seperti analisis varians (ANOVA).

b) Sewall Wright

- Wright adalah seorang genetis Amerika yang membuat kontribusi besar dalam bidang genetika populasi.
- Karyanya tentang "jalur koefisien inbreeding" dan "landscapes adaptif" sangat berpengaruh dalam evolusi teoritis dan genetika populasi.
- Tulisan penting: "*Evolution in Mendelian Populations*" (1931) – Paper ini membahas tentang drift genetik, seleksi, dan struktur populasi.

c) J.B.S. Haldane

- Haldane adalah seorang ilmuwan dan genetikus asal Inggris yang juga memberikan kontribusi besar pada biostatistik.
- Ia dikenal karena pekerjaannya dalam genetika populasi dan evolusi.
- Tulisan penting: "*The Causes of Evolution*" (1932) – Buku ini merangkum kerja Haldane tentang teori evolusi, termasuk aspek statistiknya.

d) Karl Pearson

- Pearson adalah seorang pendiri statistik matematika dan biometrika.
- Ia mengembangkan "koefisien korelasi produk-momen Pearson" yang sangat penting dalam statistik.
- Tulisan penting: "*On the Theory of Contingency and Its Relation to Association and Normal Correlation*" (2018)
- Di sini, ia mengembangkan beberapa konsep dasar dalam statistik.

e) Francis Galton

- Galton, sepupu Darwin, merupakan pelopor dalam eugenika dan biometrika.
- Ia memperkenalkan konsep regresi dan korelasi.
- Tulisan penting: *"Natural Inheritance"* (1889) – Buku ini mengeksplorasi ide warisan sifat dan distribusi statistik dari ciri-ciri keturunan.

Karya-karya ini dan kontribusi para ilmuwan tersebut membentuk dasar dari statistik modern dalam biologi dan genetika, dan terus mempengaruhi cara kita memahami dan menganalisis data dalam ilmu hayati.

2. Epidemiologi dan Kesehatan Masyarakat

Di bidang epidemiologi, biostatistik memainkan peran penting dalam mengidentifikasi pola dan penyebab penyakit, serta dalam desain dan analisis studi epidemiologis. Ini termasuk pengembangan teknik sampel, percobaan terkontrol, dan survei kesehatan.

Perkembangan biostatistik dalam konteks epidemiologi dan kesehatan masyarakat telah menjadi kunci penting dalam memahami dan mengatasi masalah kesehatan pada populasi.

Pada awal abad ke-20, pengumpulan dan analisis data vital seperti kelahiran, kematian, dan penyakit menjadi lebih sistematis. Ini menjadi fondasi surveilans kesehatan masyarakat, yang merupakan titik awal penting dalam sejarah epidemiologi. Selanjutnya, metodologi studi epidemiologi seperti kasus-kontrol dan kohort dikembangkan. Sebagai contoh, Studi Framingham Heart, yang dimulai pada tahun 1948, adalah studi kohort yang memberikan pemahaman mendalam tentang faktor risiko penyakit jantung.

Revolusi lain terjadi pada pertengahan abad ke-20 dengan pengembangan *Randomized Controlled Trials* (RCT). Ini mengubah cara kita mengevaluasi efektivitas intervensi kesehatan, dan menjadi standar emas dalam penelitian klinis. Sementara itu, kemajuan dalam teknologi komputasi memungkinkan analisis data kesehatan skala besar, memungkinkan lahirnya bidang baru, yaitu bioinformatika, yang berfokus pada data kesehatan dan penyakit.

Pada saat yang sama, model matematika semakin digunakan dalam epidemiologi. Pengembangan model seperti SIR (*Susceptible, Infected, Recovered*) membantu kita memahami dinamika penyebaran penyakit dan dampak intervensi. Selain itu, integrasi genetika molekuler dalam epidemiologi, yang dikenal sebagai epidemiologi molekuler, memperkaya pemahaman kita tentang bagaimana faktor genetik mempengaruhi risiko penyakit.

Era digital juga membawa perubahan signifikan. Transisi ke catatan kesehatan elektronik dan penggunaan data digital, seperti data dari wearable devices, memperluas kemungkinan dalam pengumpulan dan analisis data kesehatan. Kemajuan dalam teknologi data dan machine learning memungkinkan analisis data kesehatan yang lebih canggih, termasuk prediksi risiko penyakit dan respons terhadap perawatan.

Pandemi global seperti H1N1, Ebola, dan COVID-19, memperlihatkan peran penting biostatistik dalam pemodelan penyebaran penyakit, evaluasi intervensi, dan pengembangan strategi vaksinasi. Selain itu, semakin banyak kolaborasi interdisipliner antara biostatistisi, epidemiolog, profesional kesehatan, dan pakar lainnya dalam proyek-proyek penelitian besar untuk mengatasi masalah kesehatan global.

Secara keseluruhan, semua perkembangan ini telah membantu membentuk bagaimana kita memahami dan mengatasi masalah kesehatan pada populasi saat ini, serta dalam pengembangan kebijakan kesehatan dan intervensi klinis.

3. Era Komputer dan Metode Statistik Lanjutan (Pertengahan hingga Akhir Abad ke-20)

Dengan munculnya komputer pada pertengahan abad ke-20, biostatistik mengalami kemajuan signifikan. Komputer memungkinkan pemrosesan data besar dan analisis statistik yang lebih kompleks, yang penting untuk studi klinis dan penelitian biomedis. Perkembangan biostatistik selama era komputer dan penggunaan metode statistik lanjutan telah mengubah cara penelitian dilakukan dalam bidang kesehatan dan biologi secara dramatis. Perubahan ini dimulai sejak pertengahan hingga akhir abad ke-20 dan masih berlanjut hingga kini.

Di era komputer, kemajuan dalam teknologi komputer telah memungkinkan otomatisasi dan pemrosesan data statistik secara cepat. Ini mengubah cara kita mengumpulkan, menganalisis, dan menginterpretasikan data dalam penelitian biomedis. Selain itu, kemampuan untuk mengolah dataset skala besar, atau yang dikenal dengan 'big data', telah membawa kemajuan signifikan dalam biostatistik. Hal ini memungkinkan analisis data yang lebih kompleks dan detail.

Pengembangan perangkat lunak statistik canggih seperti SAS, SPSS, R, dan Python juga telah memudahkan peneliti dalam menerapkan metode statistik yang kompleks tanpa harus memiliki keahlian pemrograman mendalam. Di sisi lain, penggunaan metode statistik lanjutan juga telah berperan penting dalam perkembangan ini. Metode seperti

analisis komponen utama (PCA), analisis regresi logistik, dan model efek campuran telah menjadi alat standar dalam analisis data biostatistik. Hal ini memungkinkan pemahaman yang lebih baik tentang hubungan yang kompleks dalam data. Penggunaan model prediktif dan teknik pembelajaran mesin dalam biostatistik, seperti 'random forest' dan 'neural networks', telah meningkatkan kemampuan dalam membuat prediksi berdasarkan data kesehatan.

Era komputer juga memungkinkan pengembangan bioinformatika, khususnya dalam analisis data genomik dan proteomik, yang sangat penting dalam penelitian genetik dan penemuan obat.

Pengembangan metode statistik lebih lanjut dalam desain dan analisis uji klinis, termasuk metode adaptif dan Bayesian, telah memperbaiki efisiensi dan keefektifan uji klinis.

Kemajuan dalam analisis data longitudinal dan time-series sangat penting untuk studi epidemiologi dan klinis yang memantau perubahan seiring waktu.

Kemajuan dalam visualisasi data memungkinkan penampilan data yang lebih kompleks menjadi lebih mudah dipahami dan diinterpretasikan (Cleveland, 1993 ; Hastie et al, 2001).

Akhirnya, kemampuan untuk mengintegrasikan data dari berbagai sumber dan disiplin ilmu, seperti data klinis, genetik, dan lingkungan, telah menjadi mungkin. Hal ini memberikan wawasan yang lebih komprehensif dalam penelitian kesehatan.

Kemajuan ini tidak hanya meningkatkan kemampuan kita dalam analisis statistik tetapi juga membuka pintu untuk penemuan dan inovasi baru dalam biologi dan kesehatan masyarakat. Era komputer dan metode statistik lanjutan telah mengubah wajah biostatistik, menjadikannya lebih dinamis, interaktif, dan terintegrasi dengan bidang lain.

4. Pengembangan Uji Klinis dan Metodologi Statistik

Biostatistik menjadi kunci dalam desain dan analisis uji klinis. Metodologi statistik, termasuk randomisasi, blinding, dan analisis survival, menjadi standar dalam penelitian klinis.

Perkembangan Biostatika pada abad ke-20, khususnya dalam pengembangan uji klinis dan metodologi statistik, mencakup berbagai aspek penting:

a. Pertumbuhan Metodologi Statistik

Terjadi pertumbuhan yang pesat dalam pengembangan metodologi statistik selama beberapa dekade terakhir abad ke-20. Penelitian di bidang kedokteran dan kesehatan masyarakat, yang melibatkan biostatistisi secara dramatis, telah mengambil manfaat dari metodologi baru ini sekaligus menjadi sumber masalah baru. Kontribusi metodologi statistik dalam desain, pelaksanaan, dan analisis sangat penting dalam penelitian medis, termasuk penerimaan bahwa konsep biostatistik merupakan bagian integral dari penelitian medis yang solid (Ellenberg et al, 1990)

b. Pengaruh pada Evaluasi Obat

Biostatistika memiliki peran penting dalam evaluasi obat. Terdapat persamaan dan perbedaan dalam metodologi evaluasi obat di Jepang, Amerika Serikat, dan Eropa. Isu-isu statistik saat ini dan masa lalu termasuk

dalam pembahasan, terutama berkaitan dengan pengembangan obat dan uji klinis (Sasaki, 1995)

c. Pengembangan Uji Klinis Acak

Uji klinis acak adalah salah satu alat metodologis paling penting dalam biostatistika. Metode ini dianggap sebagai kemajuan signifikan dalam pengendalian bias seleksi dalam studi medis dan menjadi pusat dalam pengobatan berbasis bukti. Randomisasi dalam desain eksperimental tidak terbatas pada studi subjek manusia, tetapi juga digunakan dalam eksperimen pertanian (Harrington, 2000)

d. Naiknya Uji Klinis Terandomisasi

Sebelum berkembangnya uji klinis terandomisasi, probabilitas dan statistik harus disatukan. Sir RA Fisher memperkenalkan randomisasi pada tahun 1920-an, dan mulai tahun 1930-an dan 1940-an, uji klinis terandomisasi pada manusia telah dilakukan menggunakan paradigma pengujian hipotesis statistik (Fisher, 1999)

e. Pengembangan Metodologis dalam Uji Klinis

Sejak pertengahan abad ke-20, uji klinis dan metodologi statistik telah berkembang signifikan, dengan kontribusi penting dari individu seperti Sir A.B. Hill di Inggris dan J. Cornfield di Amerika Serikat. Pengembangan uji klinis modern dari sudut pandang statistik telah banyak ditinjau (White, 1990).

Kesimpulannya, perkembangan Biostatika pada abad ke-20 sangat berdampak pada pengembangan uji klinis dan metodologi statistik, dengan penekanan pada peran penting biostatistika dalam penelitian medis, evaluasi obat, dan penerapan uji klinis terandomisasi. Biostatistika telah

menjadi kunci dalam membuat keputusan yang didasarkan pada bukti ilmiah, terutama dalam bidang kedokteran dan kesehatan masyarakat.

5. Kolaborasi Interdisipliner

Era ini juga menandai peningkatan kolaborasi antara statistikawan, biolog, dan profesional kesehatan, memperkaya penelitian di bidang genetika, farmakologi, dan epidemiologi.

Perkembangan biostatistika pada abad ke-20, khususnya dalam konteks kolaborasi interdisipliner, menyoroti beberapa aspek penting:

a. Kolaborasi dalam Penelitian Ilmiah

Sebuah studi yang mengeksplorasi perkembangan kolaborasi dalam penelitian ilmiah dari tahun 1901 hingga 1991 menemukan bahwa kolaborasi antara peneliti dari berbagai institusi dan negara meningkat secara signifikan pada paruh kedua abad ke-20. Tren menuju interdisiplineritas menjadi lebih menonjol sejak tahun 1960-an, yang berhubungan dengan jumlah rata-rata penulis per makalah interdisipliner dan jenis kolaborasi yang terlibat (Qin, 1994)

b. Teori Kolaborasi Ilmiah

Dari perspektif historis dan sosiologis, penelitian tentang kolaborasi ilmiah mengembangkan teori komprehensif pertama tentang kolaborasi penelitian ilmiah: kolaborasi penelitian ilmiah, yang secara formal diakui melalui ko-penulisan makalah ilmiah, berasal, berkembang, dan terus dipraktikkan sebagai respons terhadap profesionalisasi ilmu pengetahuan (Beaver & Rosen, 1978)

- c. Mengestimasi Nilai Layanan Konsultasi Biostatistika
Konsultasi biostatistika adalah bisnis layanan. Tujuan jangka panjang dari biostatistisi konsultan adalah hubungan kolaboratif dengan para peneliti. Penting untuk menunjukkan bahwa manfaat kelompok konsultasi biostatistika melebihi biaya aktualnya bagi institusi (Parker, 2000)
- d. Kolaborasi dalam Riset Kesehatan dan Ilmu Sosial
Terdapat banyak proyek yang menggabungkan ilmuwan sosial dan kesehatan untuk mempelajari dan merekomendasikan solusi untuk berbagai masalah kesehatan. Proses ini, yang biasanya disebut sebagai penelitian multidisiplin atau interdisiplin, memiliki preseden historis yang penting dan menyoroti berbagai jenis masalah yang ditangani (Rosenfield, 1992)
- e. Mitos dan Salah Paham Tentang Kolaborasi Profesional
Penelitian yang berfokus pada kolaborasi interdisipliner menyoroti pentingnya kolaborasi dalam pendidikan dan pendidikan khusus, serta menyajikan wawasan tentang bagaimana kolaborasi dapat memengaruhi produktivitas dan kualitas kerja interdisipliner (Friend, 2000)

Kesimpulannya, biostatistika pada abad ke-20 menunjukkan pentingnya kolaborasi interdisipliner dalam penelitian ilmiah, khususnya dalam bidang kesehatan dan ilmu sosial. Kolaborasi ini telah berkontribusi pada pengembangan metodologi statistik, peningkatan efektivitas layanan konsultasi, dan pemahaman yang lebih baik tentang bagaimana kolaborasi interdisipliner dapat memengaruhi penelitian dan praktik ilmiah.

6. Bioinformatika dan Genomik (Tahun 2000-an)

Awal abad ke-21 melihat pertumbuhan pesat dalam bioinformatika dan genomik. Biostatistik menjadi sangat penting dalam analisis data genomik dan proteomik, termasuk dalam proyek-proyek seperti Human Genome Project.

Perkembangan biostatistika pada tahun 2000-an, terutama dalam konteks bioinformatika dan genomik, menghadirkan berbagai terobosan signifikan:

a. Pengaruh Bioinformatika dalam Genomik

Sejak penyelesaian proyek genom manusia pada tahun 2000, bioinformatika telah memainkan peran kunci dalam interpretasi data genomik, yang menciptakan peluang luar biasa untuk identifikasi target obat potensial. Pendekatan bioinformatika menjadi alat yang sangat penting untuk mendukung kegiatan penemuan obat target secara sistematis (Sanseau, 2001)

b. Perkembangan dan Sejarah Bioinformatika

Bioinformatika berkembang pesat pada tahun 1990-an dan 2000-an, didorong oleh penyelesaian proyek genom manusia dan kemajuan teknologi komputasi. Bioinformatika berkembang sebagai disiplin ilmu yang memadukan biologi dan informatika, dan menghadapi tantangan dalam integrasi, representasi, dan visualisasi data biomedis yang heterogen (Lewis & Bartlett, 2013)

c. Bioinformatika Pasca-Sequencing

Bioinformatika telah menjadi bagian integral dari penelitian dan pengembangan di bidang ilmu biomedis. Disiplin ini memiliki peran penting dalam menguraikan data genetik, transkriptomik, dan proteomik yang

dihasilkan oleh teknologi eksperimental berkecepatan tinggi dan dalam mengatur informasi yang diperoleh dari biologi tradisional (Kanehisa & Bork, 2003)

d. Interaksi Protein dan Penyakit

Bioinformatika memainkan peran kunci dalam memahami interaksi kompleks antara gen dan protein, termasuk peran mereka dalam penyakit. Alat dan metodologi bioinformatika telah dikembangkan untuk merumuskan prediksi tentang peran fungsional gen dan protein, termasuk keterlibatan mereka dalam penyakit (Kann, 2007)

e. Biostatistika dan Bioinformatika dalam Uji Klinis

Baik biostatistika maupun bioinformatika berhubungan dengan akuisisi dan interpretasi informasi kuantitatif dalam penelitian medis. Kedua disiplin ini terlibat dalam analisis data dan desain eksperimen, dengan bioinformatika cenderung berurusan dengan data dalam banyak dimensi, seperti dalam genomik (Berry & Coombes, 2020)

Kesimpulannya, biostatistika pada tahun 2000-an menunjukkan integrasi yang kuat dengan bioinformatika dan genomik, memberikan kontribusi penting dalam pemahaman mendalam tentang sistem biologis pada level genomik dan postgenomik. Pendekatan bioinformatika telah berperan penting dalam revolusi genetika dan genomik, memberikan wawasan baru dalam penelitian biomedis dan pengembangan obat.

7. Analisis Data Skala Besar

Biostatistik beradaptasi dengan tantangan analisis data skala besar, menggunakan teknik statistik canggih dan pembelajaran mesin untuk mengatasi kompleksitas data biomedis dan kesehatan masyarakat.

Perkembangan biostatistika pada tahun 2000-an, terutama dalam hal analisis data skala besar, menyoroti beberapa aspek kunci:

- a. Kontribusi Biostatistika dalam Pengembangan Ilmu Biomedis Biostatistika telah berkontribusi secara signifikan dalam pengembangan ilmu biomedis dalam banyak area, termasuk kedokteran laboratorium, uji klinis teracak, pengambilan keputusan klinis, dan pengembangan obat baru. Dengan ledakan penelitian ilmiah dalam beberapa dekade terakhir, penggunaan metode statistik yang tidak tepat telah meningkat. Oleh karena itu, akal sehat selalu harus menjadi prioritas utama dalam analisis data skala besar [(Albert, 2004).

- b. Biostatistika dalam Eksplorasi, Validasi, dan Interpretasi Data Klinis

Biostatistika memainkan peran kunci dalam semua fase penelitian klinis, mulai dari desain hingga pemantauan, pengumpulan data, analisis data, dan interpretasi hasil. Pemahaman yang jelas tentang kerangka kerja statistik sangat penting untuk integritas ilmiah studi dan penerimaannya di komunitas medis umum [(Mandrekar & Mandrekar, 2009)

- c. Pengujian Hipotesis dan Analisis Statistik dalam Mikrobiom

Sejak inisiasi Proyek Mikrobiom Manusia pada tahun 2008, berbagai alat biostatistik dan bioinformatika

telah dikembangkan dan diterapkan dalam studi mikrobiom. Ulasan ini membahas hipotesis penelitian dan statistik dalam studi mikrobiom, dengan fokus pada konsep mekanistik yang mendasari hubungan kompleks antara inang, mikrobiom, dan lingkungan [(Xia & Sun, 2017)]

d. Biostatistika dan Bioinformatika dalam Uji Klinis

Baik biostatistika maupun bioinformatika berkaitan dengan akuisisi dan interpretasi informasi kuantitatif dalam penelitian medis. Kedua disiplin ini terlibat dalam analisis data dan desain eksperimen, dengan bioinformatika cenderung berurusan dengan data dalam banyak dimensi, seperti dalam genomik. Keduanya berfokus pada inferensi dan pengukuran bukti berdasarkan data yang diamati [(Berry & Coombes, 2020)]

Kesimpulannya, biostatistika pada tahun 2000-an telah berkontribusi secara signifikan dalam pengembangan ilmu biomedis, khususnya dalam analisis data skala besar. Peran biostatistika sangat penting dalam merancang, mengumpulkan, menganalisis, dan menginterpretasikan data dalam penelitian klinis dan medis, serta dalam menghadapi tantangan yang timbul dari kemajuan teknologi biologi dan medis.

8. Pandemi dan Kesehatan Global

Dengan munculnya pandemi seperti HIV/AIDS dan kemudian COVID-19, biostatistik menjadi sangat penting dalam pemodelan epidemi, pengembangan vaksin, dan strategi kesehatan masyarakat.

Perkembangan biostatistika pada tahun 2000-an, terutama dalam konteks pandemi dan kesehatan global, menyoroti beberapa aspek penting:

a. Pemodelan Statistik Epidemiologi untuk Biosurveilans

Dua dekade terakhir telah melihat munculnya penyakit infeksi baru dan peningkatan kasusnya secara global. Kekhawatiran tentang bioterorisme, influenza pandemik, dan ancaman kesehatan masyarakat lainnya, serta peningkatan volume data epidemiologi, telah mendorong pengembangan sistem biosurveilans yang canggih. Pemodelan regresi statistik telah dikembangkan untuk deteksi dini dan akurat wabah penyakit dalam sistem surveilans epidemiologi (Parpoula, Karagrigoriou, & Lambrou, 2017)

b. Memperkuat Sumber Daya Biostatistika di Afrika Sub-Sahara.

Pada tahun 2009, National Institute of Allergy and Infectious Diseases (NIAID) mengadakan workshop tentang penguatan sumber daya biostatistika di Afrika Sub-Sahara. Inisiatif ini didorong oleh peningkatan pengeluaran global untuk penelitian kesehatan dan munculnya HIV/AIDS, malaria, tuberkulosis, dan agen infeksius lainnya sebagai pendorong utama peningkatan penelitian biomedis dan program perawatan klinis di Afrika Sub-Sahara (Gezmu et al., 2011)

c. Pandemi Influenza dan Sejarahnya

Studi tentang pandemi influenza telah menunjukkan pentingnya memahami asal-usul virus pandemik dan bagaimana mereka menjadi patogen yang sukses. Analisis filogenetik virus influenza 1918, misalnya, telah digunakan untuk mengidentifikasi gen

virus influenza yang berkaitan secara evolusioner sebagai petunjuk sumber pandemi 1918. Metode molekuler juga telah digunakan untuk menyelidiki virus influenza avian H5N1 dan H9N2 yang baru-baru ini menginfeksi manusia di Hong Kong (Cox & Subbarao, 2000)

d. Konsep One Health

10 Tahun Berjalan dan Jalan Panjang ke Depan. Dekade terakhir telah menunjukkan peningkatan signifikan dalam sirkulasi agen infeksius. Penyebaran dan munculnya epizootik, zoonosis, dan epidemi, serta ancaman pandemi, telah menjadi lebih kritis. Kesehatan manusia dan hewan juga terancam oleh resistensi antimikroba, polusi lingkungan, dan perkembangan penyakit kronis dan multifaktorial. Hal ini menyoroti pentingnya perspektif global dan lintas disiplin dalam memahami evolusi dan munculnya patogen (Destoumieux-Garzon et al., 2018)

Kesimpulannya, biostatistika pada tahun 2000-an telah memainkan peran penting dalam mengatasi tantangan pandemi dan kesehatan global. Dari pengembangan sistem biosurveilans yang canggih hingga konsep One Health yang holistik, biostatistika telah membantu dalam pemahaman, deteksi dini, dan pengendalian pandemi, serta meningkatkan kerja sama internasional dalam bidang kesehatan global.

9. Personalisasi Pengobatan dan Biostatistik

Munculnya konsep pengobatan personalisasi, yang menggunakan informasi genetik dan biologis pasien untuk mengarahkan terapi, juga mengandalkan metode statistik lanjutan untuk penelitian dan aplikasinya.

Perkembangan biostatistika pada tahun 2000-an, khususnya dalam konteks personalisasi pengobatan, menyoroti beberapa aspek penting:

a. Kontribusi Biostatistika dalam Pengembangan Ilmu Biomedis

Biostatistika telah berkontribusi secara signifikan dalam pengembangan ilmu biomedis di berbagai area, termasuk kedokteran laboratorium, uji klinis teracak (RCT), pengambilan keputusan klinis, dan pengembangan obat baru. Terjadi ledakan penelitian ilmiah dalam beberapa dekade terakhir, namun hal ini juga menyebabkan penyalahgunaan metode statistik. Oleh karena itu, akal sehat harus selalu menjadi prioritas utama dalam personalisasi pengobatan (Albert, 2004)

b. Pengobatan Personalisasi: Kembalinya Kunjungan Rumah.

Pengobatan personalisasi, yang memanfaatkan informasi klinis, genetik, genomik, dan lingkungan setiap individu, telah lama dijadikan metode utama. Namun, tantangannya adalah bagaimana rata-rata klinisi atau peneliti memahami informasi yang terus berkembang, termasuk manfaat bagi subkelompok tertentu. Kunci untuk mencapai hal ini adalah mempercayai pelatihan, pengalaman, dan naluri, dipadukan dengan pemahaman biostatistika (Cutter & Liu, 2012)

c. Penelitian Efektivitas Komparatif, Pengobatan Personalisasi Genomik, dan Perawatan Kesehatan Belajar Cepat

Translasi biomarker ke dalam praktik klinis sering terhambat karena kurangnya bukti tentang kegunaan klinis mereka. Metode biostatistika seperti penelitian

efektivitas komparatif (CER) yang menggunakan ulasan sistematis dan penelitian hasil kesehatan dapat menyediakan kerangka kerja metodologi untuk menilai subkelompok biologis pasien (Ginsburg & Kuderer, 2012)

d. Estimasi Rejimen Pengobatan Dinamis melalui Teknik Berbasis Regresi

Pengobatan personalisasi, di mana perawatan disesuaikan dengan pasien tertentu, merupakan area yang menarik di bidang biostatistika dan epidemiologi. Rejimen pengobatan dinamis (DTRs) adalah bagian penting dari kerangka kerja ini, memungkinkan pengobatan personalisasi pasien dengan kondisi jangka panjang, dengan mempertimbangkan kondisi dan riwayat medis mereka saat ini. Mengidentifikasi DTR optimal dalam konteks tertentu adalah masalah yang tidak sederhana, sehingga metodologi khusus telah dikembangkan untuk tujuan tersebut (Wallace, Moodie, & Stephens, 2017)

10. Peningkatan Akses dan Kolaborasi Internasional

Era digital juga membawa peningkatan akses ke data dan kolaborasi internasional dalam penelitian biostatistik, memperluas kemungkinan penemuan dan inovasi.

Perkembangan biostatistika pada tahun 2000-an, terutama dalam hal peningkatan akses dan kolaborasi internasional, menyoroti beberapa aspek penting:

a. Jaringan Kolaborasi Ilmiah Global Antara Kota

Sebuah penelitian menunjukkan bahwa kolaborasi ilmiah antar kota di dalam negara meningkat bersamaan dengan hubungan internasional. Di sebagian besar negara, kolaborasi domestik meningkat lebih cepat daripada kolaborasi internasional. Hal ini menunjukkan

bahwa, meskipun ada proses globalisasi, sistem riset nasional telah menguat selama tahun 2000-an (Maisonobe et al., 2016).

b. Mengubah Kesehatan Global: Peran Sains, Teknologi, dan Inovasi

Adopsi Tujuan Pembangunan Milenium pada tahun 2000 berbarengan dengan pengakuan atas peran sains dan teknologi dalam menyelesaikan masalah manusia, termasuk munculnya penyakit menular baru. Perkembangan ini telah membantu mendefinisikan riset biomedis sebagai salah satu isu kebijakan publik paling kritis yang dihadapi komunitas global (Juma & Lee Yee-Cheong, 2005).

c. Strategi Pengembangan Sumber Daya Biostatistika dalam Pusat Kesehatan Akademik

Biostatistika memainkan peran penting dalam riset kesehatan, namun sering kali sumber dayanya terfragmentasi atau berlebihan dalam pusat kesehatan akademik (AHC). Strategi untuk pengembangan sumber daya biostatistika dalam AHC termasuk merekrut dan mempertahankan biostatistisi, menggunakan sumber daya biostatistika secara efisien, dan meningkatkan kontribusi biostatistika terhadap ilmu pengetahuan (Welty et al., 2013)

d. Biostatistika dan Manajemen Data dalam Pengembangan Obat Global

Pengembangan obat global yang melibatkan studi klinis di berbagai negara menuntut tim internasional dan penanganan biostatistika dan manajemen data oleh tim biometrik proyek yang terdiri dari biostatistisi dan manajer data proyek dan studi. Standardisasi dan

manajemen alur kerja sangat penting dalam konteks ini (Esser, 2001)

Biostatistik telah berkembang dari menjadi sekadar alat dalam penelitian biologi dan kesehatan, menjadi disiplin ilmu yang penting dan dinamis sendiri, memengaruhi banyak aspek penelitian dan praktik dalam bidang biomedis dan kesehatan masyarakat. Biostatistik saat ini tidak hanya penting dalam analisis data, tetapi juga dalam merumuskan kebijakan kesehatan, pengembangan obat, dan dalam banyak aspek pengambilan keputusan klinis dan kesehatan publik.

2.1.4 Perkembangan biostatistik di abad ke-21

Dalam era abad ke-21, revolusi yang terjadi di bidang ***Big Data dan Analisis Data Skala Besar*** telah memberikan dampak signifikan pada biostatistik, terutama dalam sektor kesehatan dan kedokteran. Inovasi dalam bidang ini telah memfasilitasi para peneliti dan praktisi kesehatan dalam mengelola data yang berukuran besar, kompleks, dan beragam secara lebih efektif dan informatif. Berikut adalah narasi ilmiah yang mendeskripsikan aspek-aspek utama dari perkembangan ini:

a. Kemajuan Teknologi Penyimpanan dan Pengolahan Data

Peningkatan yang signifikan dalam kapasitas penyimpanan dan kecepatan pengolahan data telah memungkinkan akumulasi serta analisis dataset yang sangat luas, yang sering kali dikategorikan sebagai "big data". Hal ini memungkinkan para peneliti untuk menyimpan dan menganalisis data dalam skala yang lebih besar daripada sebelumnya, mencakup berbagai jenis data seperti genomik, data klinis, dan data dari sumber-sumber lain yang beragam (Berisha et al. 2002)

b. Integrasi dan Analisis Data Multidimensi

Kemampuan untuk mengintegrasikan dan menganalisis data dari berbagai sumber dan jenis, termasuk data genetik, klinis, lingkungan, dan perilaku, telah menjadi kunci dalam penanganan big data. Analisis data multidimensi ini memberikan wawasan yang lebih luas mengenai kesehatan dan penyakit, meliputi faktor risiko, jalannya penyakit, dan respons terhadap pengobatan (Keith, 2019)

c. Metode Statistik Lanjutan dan Pembelajaran Mesin

Pengaplikasian metode statistik lanjutan dan algoritma pembelajaran mesin telah memungkinkan identifikasi pola dan hubungan dalam dataset besar yang tidak dapat terungkap melalui metode tradisional. Hal ini membuka jalan bagi prediksi yang lebih akurat mengenai risiko penyakit, hasil pengobatan, serta pengembangan pendekatan pengobatan yang lebih terpersonalisasi (Rajula et al. 2020)

d. Visualisasi Data

Pengembangan teknik visualisasi data yang canggih memudahkan penyajian data kompleks secara lebih intuitif dan informatif. Ini membantu para peneliti dan praktisi kesehatan untuk lebih mudah memahami tren, pola, dan anomali yang ada di dalam data besar.

e. Keamanan dan Privasi Data

Seiring dengan meningkatnya penggunaan data skala besar, isu keamanan dan privasi data menjadi semakin krusial. Hal ini mendorong pengembangan protokol keamanan data yang lebih ketat dan metode yang efektif untuk mengelola data yang sensitif, seperti data kesehatan pribadi (Dietrich et al, 2015).

Kemajuan ini menandai era baru dalam biostatistik dan penelitian kesehatan, di mana analisis data yang besar dan kompleks menjadi kunci dalam memperoleh wawasan yang mendalam dan membuat keputusan berbasis bukti di bidang medis dan kedokteran.

Perkembangan terbaru di bidang **Bioinformatika** dan **Genomik** pada abad ke-21 telah memberikan kemajuan signifikan dalam teknologi sekuen genetik dan bioinformatika, yang pada gilirannya telah membuka pintu bagi penelitian genetik dalam skala yang sebelumnya tidak pernah terbayangkan. Dampaknya yang paling mencolok adalah dalam pemahaman genetika penyakit, yang telah membawa kita ke era baru dalam pengobatan personalisasi dan terapi yang ditargetkan (David W. Mount, 2004).

Perkembangan terkini di bidang bioinformatika dan genomik pada abad ke-21 telah menghadirkan perubahan mendasar dalam penelitian ilmiah, kedokteran, dan industri farmasi. Kemajuan dalam teknologi sekuensing genetik dan bioinformatika telah memungkinkan pengkajian genetik pada skala yang belum pernah terjadi sebelumnya. Dampak dari pemahaman genetika penyakit telah membawa era baru dalam pengobatan personalisasi dan terapi yang ditargetkan.

Salah satu tonggak penting dalam genomik pada abad ini adalah penyelesaian Proyek Genom Manusia pada tahun 2003, yang dipimpin oleh Francis S. Collins dan Craig Venter. Proyek ini berhasil mengidentifikasi dan merinci urutan lengkap gen manusia, membuka pintu bagi pemahaman yang lebih dalam tentang dasar genetik manusia. International Human Genome Sequencing Consortium, Nature (2001).

Teknologi CRISPR-Cas9 yang dikembangkan oleh Jennifer Doudna dan Emmanuelle Charpentier pada tahun 2012 telah mengubah paradigma dalam memahami dan memanipulasi genom.

Teknologi ini memungkinkan rekayasa gen yang lebih presisi dan telah digunakan dalam berbagai penelitian dan aplikasi, termasuk dalam terapi gen. Jennifer A. Doudna & Emmanuelle Charpentier, Science (2012).

Penggunaan teknologi sekuensing generasi berikutnya (*Next-Generation Sequencing*, NGS) telah mengubah cara kita menganalisis genom dengan cepat dan efisien. Tokoh seperti Eric Lander dan David Haussler berperan dalam mengembangkan metode analisis genom yang lebih canggih, yang telah mendukung banyak penemuan penting di bidang ini. Mouse Genome Sequencing Consortium, Nature (2002).

Genomik kedokteran, yang menggabungkan informasi genom individu dengan pemahaman tentang penyakit untuk personalisasi pengobatan, telah menjadi area penelitian yang signifikan. George Church adalah salah satu tokoh yang berperan dalam mengembangkan konsep ini. - George M. Church, Science (2014).

Dalam bidang bioinformatika, para ilmuwan seperti Ewan Birney telah mengembangkan alat analisis dan basis data untuk memahami genom secara lebih mendalam, memungkinkan penelitian yang lebih mendalam dan komprehensif. "An overview of Ensembl" - Ewan Birney et al., Genome Research (2004).

Selain itu, pemahaman genom tumbuhan dan mikroba juga telah berkembang pesat. Sebagai contoh, The Arabidopsis Initiative berhasil mengidentifikasi genom *Arabidopsis thaliana*, tanaman model yang memiliki aplikasi luas dalam penelitian biologi tumbuhan.

2.2 Pembelajaran Mesin dan Kecerdasan Buatan

Pengembangan terkini dalam bidang Biostatistika, khususnya dalam konteks Pembelajaran Mesin (Machine Learning) dan Kecerdasan Buatan (Artificial Intelligence) pada abad ke-21, telah menghadirkan kemajuan yang signifikan dalam analisis data medis dan prediksi di berbagai aspek kesehatan. Teknik-teknik AI dan pembelajaran mesin telah diterapkan dengan sukses dalam biostatistik, memberikan kontribusi besar terhadap penelitian dan praktik medis. Ton J. Cleophas dan Aeilko H. Zwinderman (2015).

Beberapa perkembangan dan referensi terkait adalah sebagai berikut:

- a. "Machine Learning Methods in Municipal Formation" (Ong et al, 2021)

Penelitian ini memfokuskan pada penerapan pembelajaran mesin untuk memecahkan masalah perkotaan dan dalam pengambilan keputusan manajerial dalam konteks pengelolaan pembangunan sosial dan ekonomi wilayah. Penelitian ini juga memiliki relevansi dalam meramalkan indikator penting pembangunan sosial-ekonomi.

- b. "Panduan AI dan Pembelajaran Mesin bagi Spesialis Retina" (María Margoth Bonilla-Jiménez et al, 2023)

Artikel ini memberikan panduan tentang cara mengakses dan memanfaatkan kecerdasan buatan dan pembelajaran mesin dalam praktik spesialis retina di abad ke-21. Hal ini menyoroti potensi besar AI dalam bidang kedokteran mata.

- c. "Strategi Pengajaran melawan Kecerdasan Buatan dalam Pembelajaran Abad ke-21" (2023)

Penelitian ini menekankan pentingnya mengembangkan strategi pengajaran yang inovatif, fleksibel, kritis, prospektif, dan panduan untuk menghadapi gangguan yang mungkin disebabkan oleh perkembangan kecerdasan buatan. Ini

menekankan pentingnya mempersiapkan siswa untuk menghadapi perubahan teknologi yang cepat.

- d. Kecerdasan Buatan, Kompetensi Abad ke-21, dan Pembelajaran Sosio-emosional dalam Pendidikan (2022)

Artikel ini mungkin membahas hubungan antara kecerdasan buatan, perkembangan kompetensi abad ke-21, dan pentingnya pembelajaran sosio-emosional dalam konteks pendidikan. Ini menggambarkan kompleksitas peran AI dalam pengembangan pendidikan modern.

- e. Analisis Awal dan Penentu Kecerdasan Buatan dan Pembelajaran di Abad ke-21 (2022)

Penelitian ini menyajikan analisis kritis tentang perkembangan AI dalam konteks pembelajaran. Penelitian ini meramalkan peningkatan yang signifikan dalam penggunaan kecerdasan buatan dalam pendidikan setelah periode stabilitas relatif, mengikuti perkembangan pesat AI dalam pembelajaran.

Perkembangan ini mencerminkan peran penting yang dimainkan oleh Pembelajaran Mesin dan Kecerdasan Buatan dalam transformasi praktik medis dan pendidikan di abad ke-21. Referensi utama dalam mempelajari konteks ini lebih dalam dapat merujuk pada buku berjudul "Machine Learning in Medicine - a Complete Overview" yang ditulis oleh Ton J. Cleophas dan Aeilko H. Zwinderman pada tahun 2015, yang menyediakan pandangan komprehensif tentang peran pembelajaran mesin dalam kedokteran.

Studi Genom-Wide Association (GWAS)

Studi *Genom-Wide Association* (GWAS) adalah pendekatan yang telah memberikan kontribusi penting dalam pemahaman tentang variabilitas genetik yang terkait dengan penyakit pada populasi besar di abad ke-21. Hal ini telah menghasilkan wawasan

yang berharga mengenai akar penyebab genetik dari banyak penyakit dan kondisi kesehatan. Berikut adalah beberapa perkembangan terkini dalam bidang GWAS dan beberapa referensi yang relevan:

1. Genome-Wide Association Studies and Genomic Selection in Pearl Millet: Advances and Prospects" (2020)

Penelitian ini membahas penerapan seleksi genomik pada tanaman millet mutiara dengan menggunakan pendekatan WGRS, RAD, dan tGBS. Hal ini penting dalam konteks program pemuliaan tanaman (Srivastava et al. 2020)

2. Genome-wide association studies revealed complex genetic architecture and breeding perspective of maize ear traits" (2022)

Studi ini memberikan wawasan tentang arsitektur genetik kompleks yang terlibat dalam sifat-sifat telinga jagung. Informasi ini memiliki relevansi penting dalam pemuliaan tanaman jagung (Khatun et al, 2022)

3. From single-SNP to wide-locus: genome-wide association studies identifying functionally related genes and intragenic regions in small sample studies." (2013)

Penelitian ini menunjukkan bagaimana GWAS dapat digunakan untuk mengidentifikasi kluster gen yang berhubungan dengan jalur biologis tertentu, bahkan dalam sampel yang relatif kecil (Wittkowski et al. 2013)

4. Genome-wide association study of production and stability traits in barley cultivated under future climate scenarios" (2015)

Studi ini menggunakan GWAS untuk mengidentifikasi penanda peningkatan produksi pada barley dalam kondisi perubahan iklim. Informasi ini memiliki dampak penting dalam pertanian di masa depan (Angueira, 2015).

5. New Insights Into Gestational Glucose Metabolism: Lessons Learned From 21st Century Approaches" (2015)

Artikel ini membahas perkembangan dalam pemahaman tentang metabolisme glukosa selama kehamilan, dengan menggunakan pendekatan GWAS dan metabolomik.

6. GWAS on family history of Alzheimer's disease" (2018)

Studi ini menyoroti validitas penggunaan laporan diri tentang riwayat keluarga Alzheimer sebagai proxy dalam GWAS. Hal ini memiliki implikasi dalam pemahaman genetik penyakit Alzheimer (Marioni et al, 2018)

7. The genetics of asthma and allergic disease: a 21st century perspective" (2011)

Artikel ini membahas variasi genetik yang terkait dengan penyakit asma dan alergi, dengan mengacu pada gen interleukin-33 (IL-33) dan thymic stromal lymphopoietin (TSLP) (Gondro et al. 2013)

Uji Klinis Adaptif dan Metode Bayes

Di era modern biostatistika, khususnya sepanjang abad ke-21, telah terjadi revolusi dalam desain dan analisis uji klinis, yang ditandai oleh pengembangan dan implementasi Uji Klinis Adaptif serta Metode Bayes. Perubahan ini telah mengakibatkan peningkatan fleksibilitas dan efisiensi dalam mengevaluasi keamanan dan efikasi obat-obatan.

Dalam konteks ini, literatur biostatistika terkini telah mengungkapkan beberapa temuan penting. Salah satunya adalah penelitian oleh Berry (2006), yang menyoroti keunggulan pendekatan Bayesian dalam uji klinis, terutama dalam pengembangan terapi kanker yang dipersonalisasi. Pendekatan ini, menurut Berry, lebih selaras dengan cara kerja ilmu pengetahuan dibandingkan dengan metode frekuensial, karena memungkinkan

pembaruan berkelanjutan informasi, yang mendukung desain uji klinis yang adaptif.

Selanjutnya, Chevret (2012) meneliti penggunaan metode Bayesian dalam desain Uji Klinis Adaptif. Temuannya menunjukkan bahwa metode ini telah diusulkan untuk segala fase uji klinis, dari Fase I hingga Fase III, namun penggunaannya dalam praktik masih terbatas. Studi bibliometrik yang dilakukan oleh Chevret mengungkapkan bahwa sebagian besar penelitian tentang metode Bayesian berasal dari Amerika Serikat, terutama dari MD Anderson Cancer Center.

Simon dan Simon (2018) juga memberikan kontribusi penting dengan mengusulkan kerangka kerja untuk Uji Klinis Adaptif yang memanfaatkan model Bayesian dalam pengambilan keputusan selama uji klinis. Meski begitu, mereka tetap menggunakan tes hipotesis frekuensial di akhir uji klinis. Pendekatan ini menunjukkan bagaimana integrasi antara metode Bayesian dan frekuensial dapat meningkatkan efektivitas uji klinis.

Pada tahun 2020, Lin dan Lee membahas tentang desain Uji Klinis Adaptif Bayesian terbaru dan aplikasinya dalam uji klinis kanker. Mereka memberikan gambaran umum tentang berbagai aplikasi desain adaptif Bayesian, termasuk penemuan dosis adaptif, kalkulasi probabilitas posterior dan probabilitas prediktif, serta randomisasi adaptif berdasarkan hasil.

Terakhir, Gonen (2009) menekankan peningkatan penggunaan metode Bayesian dalam penelitian medis. Metode ini dianggap lebih adaptif terhadap informasi yang terakumulasi selama uji klinis, memungkinkan uji klinis yang lebih kecil namun lebih informatif.

Secara keseluruhan, perkembangan biostatistika dalam Uji Klinis Adaptif dan Metode Bayes menandai sebuah era baru dalam penelitian klinis. Metode Bayesian, dengan kemampuannya untuk menyesuaikan desain berdasarkan informasi yang berkembang, membawa perubahan signifikan dalam cara uji klinis dirancang dan dianalisis, memungkinkan penelitian yang lebih efisien dan adaptif. Kajian lebih lanjut dan aplikasi praktis dari metode ini diharapkan terus berkembang, membuka jalan bagi inovasi lebih lanjut dalam bidang ini.

Epidemiologi Digital dan Penggunaan Data Real-Time

Di era ke-21, terjadi perkembangan signifikan dalam biostatistika, khususnya dalam bidang epidemiologi digital dan penggunaan data real-time. Hal ini telah membuka horison baru dalam pengawasan dan respons terhadap wabah penyakit. Akses ke data real-time dan digital memungkinkan percepatan respons terhadap krisis kesehatan publik, seperti yang terlihat selama pandemi COVID-19. Marcel Salathé dalam karyanya "Digital Epidemiology" (2018) menggarisbawahi pentingnya perkembangan ini.

Sejumlah penelitian telah memberikan wawasan yang mendalam tentang bidang ini. Pertama, Goldstein, Levasseur, dan McClure (2020) menyoroti konvergensi antara epidemiologi, biostatistika, dan ilmu data. Mereka menekankan bahwa ketiga bidang ini berfokus pada pendekatan kuantitatif untuk menangani isu-isu kesehatan yang kompleks, terutama dengan munculnya data kesehatan elektronik.

Kemudian, Salathé et al. (2012) mendeskripsikan revolusi epidemiologi digital yang terjadi sebagai akibat dari kemajuan dalam komunikasi dan data digital. Sumber data digital, termasuk media sosial dan perangkat mobile, memberikan informasi cepat dan lokal tentang dinamika penyakit pada populasi global.

Selanjutnya, Al-Imam et al. (2019) menggambarkan penggunaan analisis statistik real-time dalam epidemiologi digital, khususnya melalui database Google Trends. Ini menunjukkan kemajuan signifikan dalam pengumpulan dan analisis data epidemiologi secara real-time.

Shortreed et al. (2019) menyoroti tantangan dan peluang yang muncul dari penggunaan data rekam medis elektronik (EHR) dalam penelitian medis. Meskipun metodologi baru telah meningkatkan kemampuan ini, terdapat beberapa hambatan seperti kualitas data dan integrasi sumber data tambahan yang perlu diatasi.

Terakhir, Pastor-Escuredo dan Olmos (2022) memberikan ulasan tentang penelitian terkini dalam epidemiologi digital, mencakup model transmisi penyakit, analisis mobilitas, dan studi penyebaran informasi terkait penyakit.

Kesimpulannya, biostatistika di abad ke-21 telah berkembang pesat dengan integrasi konsep epidemiologi digital dan analisis data real-time. Kolaborasi ini dengan ilmu data telah meningkatkan kemampuan dalam memahami dan menangani masalah kesehatan melalui pendekatan yang lebih luas dan analisis yang lebih canggih.

Personalisasi Pengobatan dan Farmakogenomik

Dalam abad ke-21, perkembangan biostatistika telah menjadi fokus utama dalam upaya untuk mewujudkan personalisasi pengobatan dan farmakogenomik. Perkembangan ini telah menghasilkan peningkatan signifikan dalam pemahaman dan implementasi pengobatan yang disesuaikan berdasarkan profil genetik individu. Dalam konteks ini, beberapa temuan utama dari literatur yang relevan adalah sebagai berikut:

Pendekatan Baru dalam Farmakogenomik

Farmakogenomik memiliki dampak besar pada sistem kesehatan dan penemuan obat. Fokusnya adalah pada konsep

"obat yang tepat untuk pasien yang tepat" yang mencakup seluruh spektrum gen yang mempengaruhi perilaku dan sensitivitas terhadap obat (Fujita, 2000).

Model Matematika dalam Farmakogenomik

Penggunaan model matematika dalam sistem farmakogenomik untuk personalisasi pengobatan telah menjadi topik penelitian yang penting. Penelitian ini mengintegrasikan parameter klinis-patologis dengan profil genetik dan molekuler untuk menciptakan produk dan pendekatan yang bermanfaat secara klinis (Wu & Wang, 2013).

Bioetika dalam Farmakogenomik:

Pertimbangan bioetika sangat penting dalam penelitian farmakogenomik dan pengobatan personalisasi, terutama mengingat perkembangan data yang intensif dari ilmu pengetahuan generasi keempat (Ozdemir et al., 2013).

Genomik dan Pengobatan Personalisasi

Kemajuan dalam genetika dan praktik pengobatan personalisasi memiliki fokus pada penggunaan informasi genomik untuk mengarahkan pengobatan berbasis biomarker (Katsios & Roukos, 2010).

Tantangan dan Peluang dalam Farmakogenomik

Tantangan dan peluang yang muncul dari farmakogenomik dapat mempengaruhi perkembangan obat dan praktek klinis. Oleh karena itu, perlu mempertimbangkan pendekatan yang sesuai (Lee, 2005).

Pendekatan Berbasis Partisi Rekursif dalam Genetik

Metode partisi rekursif, seperti pohon klasifikasi dan regresi, digunakan dalam farmakogenomik untuk mengidentifikasi faktor genetik yang mempengaruhi respons terhadap pengobatan (Vanness, 2011).

Integrasi Data

Perkembangan Biostatistik abad ke-21, dengan fokus pada Integrasi Data, telah membawa perubahan signifikan dalam cara data dianalisis dan diinterpretasi dalam ilmu biomedis. Berikut adalah beberapa temuan penting dari literatur yang relevan:

- a. Integrasi Bioinformatika untuk Genomika dan Proteomika.

Wu (2011) menyoroti pentingnya integrasi sistem dalam biologi abad ke-21. Penelitian ini menggabungkan penambangan data dengan penambangan teks dan analisis jaringan dalam konteks biologi sistem untuk penemuan jaringan biologis (Wu, 2011).

- b. Genom, Transkriptom, dan Proteom

Manzoni et al. (2016) menjelaskan bagaimana data omik diintegrasikan untuk mendukung pemahaman yang lebih baik tentang sistem biologis dan pengembangan obat yang dipersonalisasi (Manzoni et al., 2016).

- c. OBO Foundry dan Integrasi Data Biomedis

Smith et al. (2007) membahas inisiatif OBO Foundry, yang bertujuan untuk mengkoordinasikan reformasi ontologi yang ada dan menciptakan ontologi baru untuk mendukung integrasi data biomedis (Smith et al., 2007).

- d. Pelajaran dari Toksikologi untuk Penelitian Medis

Langley et al. (2015) membahas pentingnya pendekatan sistem dalam biomedis untuk memahami penyakit dan patofisiologi, yang sangat bergantung pada integrasi data tingkat lanjut (Langley et al., 2015).

- e. Kontribusi Biostatistik dalam Pengembangan Ilmu Biomedis

Albert (2004) menekankan peran penting biostatistik dalam mengubah informasi menjadi pengetahuan, terutama dalam pengembangan obat dan pengambilan keputusan klinis (Albert, 2004).

f. **BioDW: Sistem Integrasi Data Bioinformatika**

Zhongping (2005) mengembangkan sistem integrasi data bioinformatika, BioDW, yang berfokus pada interoperabilitas sumber data, penyelesaian heterogenitas semantik, dan pembaruan data (Zhongping, 2005).

Perkembangan biostatistik dalam konteks integrasi data telah memberikan kontribusi yang signifikan dalam penggabungan dan analisis berbagai jenis data biomedis. Pendekatan ini telah memungkinkan para peneliti untuk mengajukan pertanyaan biologis yang lebih kompleks dan memperoleh pemahaman yang lebih baik tentang proses biologis dan penyakit.

Integrasi Data Multi-Sumber

Perkembangan Integrasi data dari berbagai sumber termasuk data klinis, genomik, dan lingkungan, telah secara signifikan memperkaya pemahaman kita tentang kesehatan dan penyakit. Pengaruhnya yaitu pendekatan ini memfasilitasi penelitian medis yang lebih komprehensif dan holistik.

Evolusi biostatistika pada abad ke-21, khususnya dalam integrasi data, telah merevolusi analisis dan interpretasi data dalam bidang biomedis. Temuan penting dari literatur terkait meliputi:

a. **Integrasi Bioinformatika untuk Genomik dan Proteomik**

Wu (2011) menekankan pentingnya integrasi sistem dalam biologi modern, menggabungkan penambangan data, penambangan teks, dan analisis jaringan dalam biologi sistem untuk identifikasi jaringan biologis.

b. **Genom, Transkriptom, dan Proteom**

Manzoni et al. (2016) membahas integrasi data omik untuk meningkatkan pemahaman sistem biologis dan pengembangan obat personalisasi.

c. OBO Foundry dan Integrasi Data Biomedis

Smith et al. (2007) menjelaskan inisiatif OBO Foundry untuk menyelaraskan ontologi yang ada dan menciptakan ontologi baru guna mendukung integrasi data biomedis.

d. Pelajaran dari Toksikologi untuk Riset Medis

Langley et al. (2015) menyoroti pentingnya pendekatan sistem dalam biomedis untuk memahami penyakit dan patofisiologinya, bergantung pada integrasi data tingkat lanjut.

e. Kontribusi Biostatistika dalam Pengembangan Ilmu Biomedis

Albert (2004) menyoroti peran kunci biostatistika dalam mengubah informasi menjadi pengetahuan, terutama dalam pengembangan obat dan keputusan klinis.

f. BioDW: Sistem Integrasi Data Bioinformatika

Zhongping (2005) mengembangkan BioDW, sebuah sistem integrasi data bioinformatika yang fokus pada interoperabilitas data, penyelesaian heterogenitas semantik, dan pembaruan data.

Kesimpulannya, kemajuan biostatistika dalam integrasi data telah berperan penting dalam menggabungkan dan menganalisis berbagai jenis data biomedis. Pendekatan ini memungkinkan peneliti untuk menjawab pertanyaan biologis yang lebih kompleks dan mendapatkan pemahaman yang lebih mendalam tentang proses biologis dan penyakit.

DAFTAR PUSTAKA

- Al-Imam, Ahmed, et al. "Real-Time Inferential Analytics Based on Online Databases of Trends: A Breakthrough within the Discipline of Digital Epidemiology of Dentistry and Oral-Maxillofacial Surgery." *Modern Applied Science*, vol. 13, no. 2, 3 Jan. 2019, pp. 81–81, <https://doi.org/10.5539/mas.v13n2p81>. Accessed 26 Apr. 2023.
- Albert, A. "[Biostatistics Contribution to the Development of Biomedical Sciences]." *Bulletin et Memoires de l'Academie Royale de Medecine de Belgique*, vol. 159 5-6, 2004, pp. 317–24324, consensus.app/papers/biostatistics-contribution-development-sciences-albert/1e64cd52e89850d38f83433c806cf11e/?utm_source=chatgpt. Accessed 26 Jan. 2024.
- Angueira, Anthony R., et al. "New Insights into Gestational Glucose Metabolism: Lessons Learned from 21st Century Approaches." *Diabetes*, vol. 64, no. 2, 22 Jan. 2015, pp. 327–334, [diabetes.diabetesjournals.org/content/64/2/327](https://doi.org/10.2337/db14-0877), <https://doi.org/10.2337/db14-0877>. Accessed 21 Sept. 2019.
- Aylmer, Ronald. *The Genetical Theory of Natural Selection*. Oxford, Oxford University Press, 1999.
- Beaver, D. deB, and R. Rosen. "Studies in Scientific Collaboration." *Scientometrics*, vol. 1, no. 1, Sept. 1978, pp. 65–84, <https://doi.org/10.1007/bf02016840>. Accessed 2 Feb. 2020.
- Berisha, Blend, et al. "Big Data Analytics in Cloud Computing: An Overview." *Journal of Cloud Computing*, vol. 11, no. 1, 6 Aug. 2022, journalofcloudcomputing.springeropen.com/articles/10.1186/s13677-022-00301-w, <https://doi.org/10.1186/s13677-022-00301-w>.

- Berry, D., and K. Coombes. "Biostatistics and Bioinformatics in Clinical Trials." *Abeloff's Clinical Oncology*, 2020, consensus.app/papers/biostatistics-bioinformatics-clinical-trials-berry/4b12f9eb8c55524997e564ad2ac59851/?utm_source=chatgpt, <https://doi.org/10.1016/B978-1-4557-2865-7.00019-9>. Accessed 26 Jan. 2024.
- CAMERON, DONALD, and IAN G JONES. "John Snow, the Broad Street Pump and Modern Epidemiology." *International Journal of Epidemiology*, vol. 12, no. 4, 1983, pp. 393–396, <https://doi.org/10.1093/ije/12.4.393>.
- Chevret, Sylvie. "Bayesian Adaptive Clinical Trials: A Dream for Statisticians Only?" *Statistics in Medicine*, vol. 31, no. 11-12, 9 Sept. 2011, pp. 1002–1013, <https://doi.org/10.1002/sim.4363>. Accessed 23 Apr. 2022.
- Cleophas, Ton J., and Aeilko H. Zwinderman. *Machine Learning in Medicine – a Complete Overview*. Amazon, 2nd ed. 2020 edition ed., Cham, Springer, 4 Mar. 2020, www.amazon.com/Machine-Learning-Medicine-Complete-Overview/dp/3030339696. Accessed 26 Jan. 2024.
- Cleveland, William S. *Visualizing Data*. Hobart Press, 1993.
- Cole, T. J. "Secular Trends in Growth." *Proceedings of the Nutrition Society*, vol. 59, no. 2, May 2000, pp. 317–324, www.cambridge.org/core/journals/proceedings-of-the-nutrition-society/article/div-classtitlesecular-trends-in-growthdiv/73F62046A2FCE9E274153AE20E0F926A, <https://doi.org/10.1017/s0029665100000355>.
- Cornelis Dirk Andriessse. *Huygens : The Man behind the Principle*. Cambridge ; New York, Cambridge University Press, 2005.

- Cox, N. J., and K. Subbarao. "Global Epidemiology of Influenza: Past and Present." *Annual Review of Medicine*, vol. 51, no. 1, Feb. 2000, pp. 407–421, <https://doi.org/10.1146/annurev.med.51.1.407>.
- Cutter, Gary R., and Yuliang Liu. "Personalized Medicine." *Neurology: Clinical Practice*, vol. 2, no. 4, 1 Dec. 2012, pp. 343–351, www.ncbi.nlm.nih.gov/pmc/articles/PMC3613206/, <https://doi.org/10.1212/CPJ.0b013e318278c328>.
- Destoumieux-Garzón, Delphine, et al. "The One Health Concept: 10 Years Old and a Long Road Ahead." *Frontiers in Veterinary Science*, vol. 5, no. 14, 12 Feb. 2018, <https://doi.org/10.3389/fvets.2018.00014>.
- Dietrich, David, et al. *Data Science & Big Data Analytics Discovering, Analyzing, Visualizing and Presenting Data*. Indianapolis Wiley, 2015.
- Donald , Berry. "The Bayesian Approach to Clinical Trials and Individualized Cancer Therapy." *Cancer Biology & Therapy*, vol. 5, no. 2, Feb. 2006, pp. 136–136, <https://doi.org/10.4161/cbt.5.2.2502>. Accessed 25 Oct. 2019.
- Elandt-Johnson, Regina C, and Norman Lloyd Johnson. *Survival Models and Data Analysis*. New York, Wiley, 1980.
- Ellenberg, Jonas H., et al. "Biostatistical Collaboration in Medical Research." *Biometrics*, vol. 46, no. 1, 1990, pp. 1–32, www.jstor.org/stable/2531627, <https://doi.org/10.2307/2531627>. Accessed 9 Feb. 2022.
- Esser, Regina. "Biostatistics and Data Management in Global Drug Development." *Drug Information Journal*, vol. 35, no. 3, 1 July 2001, pp. 643–653, <https://doi.org/10.1177/009286150103500302>. Accessed 26 Jan. 2024.

- Fisher, Lloyd D. "Advances in Clinical Trials in the Twentieth Century." *Annual Review of Public Health*, vol. 20, no. 1, May 1999, pp. 109–124, <https://doi.org/10.1146/annurev.publhealth.20.1.109>. Accessed 13 Dec. 2019.
- Fisher, R a. *Statistical Methods for Research Workers*. Gyan Books, 2017.
- Foote, Keith D. "Big Data Integration 101: The What, Why, and How." *DATAVERSITY*, 27 Feb. 2019, www.dataversity.net/big-data-integration-101-the-what-why-and-how/.
- Friend, Marilyn. "Myths and Misunderstandings about Professional Collaboration." *Remedial and Special Education*, vol. 21, no. 3, May 2000, pp. 130–160, <https://doi.org/10.1177/074193250002100301>. Accessed 10 Sept. 2019.
- Fujita, Y. "[a New Approach to Pharmacogenomics]." *Nihon Yakurigaku Zasshi. Folia Pharmacologica Japonica*, vol. 116 3, 2000, pp. 149–57, consensus.app/papers/approach-pharmacogenomics-fujita/7fd5d53b0827571aacb11fc7fd29de6a/?utm_source=chatgpt, <https://doi.org/10.1254/FPJ.116.149>. Accessed 26 Jan. 2024.
- Galton, Francis. *Natural Inheritance*. Google Books, Macmillan, 1889, books.google.co.id/books/about/Natural_Inheritance.html?id=pGIVAAAYAAJ&redir_esc=y. Accessed 25 Jan. 2024.
- Gani, J. *The Making of Statisticians*. Springer Science & Business Media, 6 Dec. 2012.
- Gezmu, Misrak, et al. "Strengthening Biostatistics Resources in Sub-Saharan Africa: Research Collaborations through U.S. Partnerships." *Statistics in Medicine*, vol. 30, no. 7, 12 Jan. 2011, pp. 695–708, <https://doi.org/10.1002/sim.4144>. Accessed 21 Nov. 2022.

- Ginsburg, Geoffrey S., and Nicole M. Kuderer. "Comparative Effectiveness Research, Genomics-Enabled Personalized Medicine, and Rapid Learning Health Care: A Common Bond." *Journal of Clinical Oncology*, vol. 30, no. 34, 1 Dec. 2012, pp. 4233–4242, <https://doi.org/10.1200/jco.2012.42.6114>. Accessed 10 Mar. 2020.
- Gondro, Cedric, et al. *Genome-Wide Association Studies and Genomic Prediction*. 2013.
- Gönen, Mithat. "Bayesian Clinical Trials: No More Excuses." *Clinical Trials*, vol. 6, no. 3, June 2009, pp. 203–204, <https://doi.org/10.1177/1740774509105374>. Accessed 11 Nov. 2021.
- Haldane, J. B. S. "APA PsycNet." *Psycnet.apa.org*, 1932, psycnet.apa.org/record/1934-03536-000. Accessed 25 Jan. 2024.
- Harrington, David P. "The Randomized Clinical Trial." *Journal of the American Statistical Association*, vol. 95, no. 449, 1 Mar. 2000, pp. 312–315, <https://doi.org/10.1080/01621459.2000.10473931>. Accessed 9 Nov. 2023.
- Hastie, Trevor, et al. *The Elements of Statistical Learning : Data Mining, Inference, and Prediction*. New York, Ny Springer New York Imprint, Springer, 2001.
- Jakob Bernoulli, and Edith Dudley Sylla. *The Art of Conjecturing, Together with Letter to a Friend on Sets in Court Tennis*. Baltimore, Md., Johns Hopkins University Press, 2006.
- Juma, Calestous, and Lee Yee-Cheong. "Reinventing Global Health: The Role of Science, Technology, and Innovation." *The Lancet*, vol. 365, no. 9464, 19 Mar. 2005, pp. 1105–1107, [www.thelancet.com/journals/lancet/article/PIIS0140-6736\(05\)71147-8/fulltext](https://www.thelancet.com/journals/lancet/article/PIIS0140-6736(05)71147-8/fulltext), [https://doi.org/10.1016/S0140-6736\(05\)71147-8](https://doi.org/10.1016/S0140-6736(05)71147-8). Accessed 19 July 2020.

- Kanehisa, Minoru, and Peer Bork. "Bioinformatics in the Post-Sequence Era." *Nature Genetics*, vol. 33, no. S3, Mar. 2003, pp. 305–310, <https://doi.org/10.1038/ng1109>. Accessed 3 Sept. 2021.
- Kann, M. G. "Protein Interactions and Disease: Computational Approaches to Uncover the Etiology of Diseases." *Briefings in Bioinformatics*, vol. 8, no. 5, 20 June 2007, pp. 333–346, <https://doi.org/10.1093/bib/bbm031>. Accessed 6 June 2022.
- Karl 1857-1936 Pearson. *On the Theory of Contingency and Its Relation to Association and Normal Correlation*. Franklin Classics Trade Press, 31 Oct. 2018.
- Katsios, Christos, and Dimitrios H Roukos. "Individual Genomes and Personalized Medicine: Life Diversity and Complexity." *Personalized Medicine*, vol. 7, no. 4, July 2010, pp. 347–350, <https://doi.org/10.2217/pme.10.30>. Accessed 8 Dec. 2019.
- Khatun, Mita, et al. "Genome-Wide Association Studies Revealed Complex Genetic Architecture and Breeding Perspective of Maize Ear Traits." *BMC Plant Biology*, vol. 22, no. 1, 18 Nov. 2022, p. 537, pubmed.ncbi.nlm.nih.gov/36397013, <https://doi.org/10.1186/s12870-022-03913-1>. Accessed 26 Jan. 2024.
- Langley, Gill, et al. "Lessons from Toxicology: Developing a 21st-Century Paradigm for Medical Research." *Environmental Health Perspectives*, vol. 123, no. 11, 1 Nov. 2015, pp. A268–272, pubmed.ncbi.nlm.nih.gov/26523530/, <https://doi.org/10.1289/ehp.1510345>. Accessed 20 Feb. 2022.
- Lee, Sandra Soo-Jin. "Personalized Medicine and Pharmacogenomics: Ethical and Social Challenges." *Personalized Medicine*, vol. 2, no. 1, Mar. 2005, pp. 29–35, <https://doi.org/10.1517/17410541.2.1.29>. Accessed 21 Aug. 2019.

- Lewis, Jamie, and Andrew Bartlett. "Inscribing a Discipline: Tensions in the Field of Bioinformatics." *New Genetics and Society*, vol. 32, no. 3, Sept. 2013, pp. 243–263, <https://doi.org/10.1080/14636778.2013.773172>. Accessed 22 Jan. 2021.
- Lin, R., and J. Lee. "Novel Bayesian Adaptive Designs and Their Applications in Cancer Clinical Trials." *Consensus.app*, 2020, pp. 395–426, consensus.app/papers/novel-bayesian-adaptive-designs-their-applications-lin/bf8c39d807075793ac17160c2079f3a3/?utm_source=chatgpt), https://doi.org/10.1007/978-3-030-42196-0_17. Accessed 27 Jan. 2024.
- Maisonobe, Marion, et al. "The World Network of Scientific Collaborations between Cities: Domestic or International Dynamics?" *Journal of Informetrics*, vol. 10, no. 4, 1 Nov. 2016, pp. 1025–1036, <https://doi.org/10.1016/j.joi.2016.06.002>. Accessed 27 Apr. 2023.
- Mandrekar, Jayawant N., and Sumithra J. Mandrekar. "Biostatistics: A Toolkit for Exploration, Validation, and Interpretation of Clinical Data." *Journal of Thoracic Oncology*, vol. 4, no. 12, Dec. 2009, pp. 1447–1449, <https://doi.org/10.1097/jto.0b013e3181c0a329>. Accessed 5 Mar. 2022.
- Manzoni, Claudia, et al. "Genome, Transcriptome and Proteome: The Rise of Omics Data and Their Integration in Biomedical Sciences." *Briefings in Bioinformatics*, vol. 19, no. 2, 22 Nov. 2016, pp. 286–302, <https://doi.org/10.1093/bib/bbw114>. Accessed 17 Oct. 2019.

- María Margoth Bonilla-Jiménez, et al. "Teaching Strategies against Artificial Intelligence in the Learning of 21st Century Students." *Journal of Law and Sustainable Development*, vol. 11, no. 9, 13 Oct. 2023, pp. e1635–e1635, <https://doi.org/10.55908/sdgs.v11i9.1635>. Accessed 26 Jan. 2024.
- Marioni, Riccardo E., et al. "GWAS on Family History of Alzheimer's Disease." *Translational Psychiatry*, vol. 8, no. 1, 18 May 2018, pp. 1–7, www.nature.com/articles/s41398-018-0150-6, <https://doi.org/10.1038/s41398-018-0150-6>. Accessed 21 Sept. 2020.
- Neal D, Goldstein, et al. "On the Convergence of Epidemiology, Biostatistics, and Data Science." *Harvard Data Science Review*, vol. 2, no. 2, 30 Apr. 2020, <https://doi.org/10.1162/99608f92.9f0215e6>.
- Ong, Joshua, et al. "A Guide to Accessible Artificial Intelligence and Machine Learning for the 21st Century Retina Specialist." *Ophthalmic Surgery, Lasers & Imaging Retina*, vol. 52, no. 7, 1 July 2021, pp. 361–365, <https://doi.org/10.3928/23258160-20210628-01>. Accessed 28 Oct. 2023.
- Ozdemir, V., et al. "Beyond ELSIs: Where to from Here? From "Regulating" to Anticipating and Shaping the Innovation Trajectory in Personalized Medicine." *Consensus.app*, 2013, pp. 405–428, consensus.app/papers/beyond-elsis-where-here-from-regulating-anticipating-ozdemir/4a5f628d9a4a54e790599e8aa654895c/?utm_source=chatgpt, <https://doi.org/10.1016/B978-0-12-391918-2.00011-1>. Accessed 27 Jan. 2024.

- Parker, Robert A. "Estimating the Value of an Internal Biostatistical Consulting Service." *Statistics in Medicine*, vol. 19, no. 16, 2000, pp. 2131–2145, [https://doi.org/10.1002/1097-0258\(20000830\)19:16%3C2131::aid-sim510%3E3.0.co;2-n](https://doi.org/10.1002/1097-0258(20000830)19:16%3C2131::aid-sim510%3E3.0.co;2-n). Accessed 17 Dec. 2020.
- Parpoula, C., et al. "Epidemic Intelligence Statistical Modelling for Biosurveillance." *Consensus.app*, 2017, pp. 349–363, consensus.app/papers/intelligence-statistical-modelling-biosurveillance-parpoula/ec3c430b1589530aa5ec96c84eacb53e/?utm_source=chatgpt, https://doi.org/10.1007/978-3-319-72453-9_29. Accessed 26 Jan. 2024.
- Pastor-Escuredo, David. "Digital Epidemiology: A Review." *ArXiv (Cornell University)*, 8 Apr. 2021, <https://doi.org/10.18103/mra.v10i10.3175>. Accessed 27 Jan. 2024.
- Pearson, Karl. *The History of Statistics in the 17th and 18th Centuries, against the Changing Background of Intellectual, Scientific and Religious Thought: Lecture*. Edited by E. S. Pearson. Amazon, London, Lubrecht & Cramer Ltd, 1 Jan. 1978, www.amazon.com/Statistics-Centuries-Background-Intellectual-Scientific/dp/0852642504. Accessed 25 Jan. 2024.
- Qin, J. "An Investigation of Research Collaboration in the Sciences through the Philosophical Transactions 1901–1991." *Scientometrics*, vol. 29, 1994, pp. 219–238, consensus.app/papers/investigation-research-collaboration-sciences-qin/de30a8878b3b50749c304efa81feda04/?utm_source=chatgpt, <https://doi.org/10.1007/BF02017974>. Accessed 26 Jan. 2024.

- Rajula, Hema Sekhar Reddy, et al. "Comparison of Conventional Statistical Methods with Machine Learning in Medicine: Diagnosis, Drug Development, and Treatment." *Medicina*, vol. 56, no. 9, 8 Sept. 2020, p. 455, <https://doi.org/10.3390/medicina56090455>.
- Rosenfield, Patricia L. "The Potential of Transdisciplinary Research for Sustaining and Extending Linkages between the Health and Social Sciences." *Social Science & Medicine*, vol. 35, no. 11, Dec. 1992, pp. 1343–1357, [https://doi.org/10.1016/0277-9536\(92\)90038-r](https://doi.org/10.1016/0277-9536(92)90038-r). Accessed 21 Nov. 2019.
- Salathé, Marcel, et al. "Digital Epidemiology." *PLoS Computational Biology*, vol. 8, no. 7, 26 July 2012, p. e1002616, journals.plos.org/ploscompbiol/article?id=10.1371/journal.pcbi.1002616, <https://doi.org/10.1371/journal.pcbi.1002616>.
- . "Digital Epidemiology: What Is It, and Where Is It Going?" *Life Sciences, Society and Policy*, vol. 14, no. 1, 4 Jan. 2018, <https://doi.org/10.1186/s40504-017-0065-7>.
- Sanseau, Philippe. "Impact of Human Genome Sequencing for in Silico Target Discovery." *Drug Discovery Today*, vol. 6, no. 6, Mar. 2001, pp. 316–323, [https://doi.org/10.1016/s1359-6446\(01\)01724-x](https://doi.org/10.1016/s1359-6446(01)01724-x). Accessed 8 Feb. 2019.
- Sasaki, Yayoi. "Guest Editor's Note: 1st Annual Biostatistics Meeting: Similarities and Differences in Statistical Drug Evaluation Methodology among Japan, the United States, and Europe, Including Current Issues in Biostatistics." *Drug Information Journal*, vol. 29, no. 4, Oct. 1995, pp. 1329–1330, <https://doi.org/10.1177/009286159502900436>. Accessed 25 Jan. 2022.

- Shortreed, Susan M, et al. "Challenges and Opportunities for Using Big Health Care Data to Advance Medical Science and Public Health." *American Journal of Epidemiology*, vol. 188, no. 5, 16 Mar. 2019, pp. 851–861, <https://doi.org/10.1093/aje/kwy292>. Accessed 7 June 2021.
- Simon, Noah, and Richard Simon. "Using Bayesian Modeling in Frequentist Adaptive Enrichment Designs." *Biostatistics*, vol. 19, no. 1, 17 May 2017, pp. 27–41, <https://doi.org/10.1093/biostatistics/kxw054>. Accessed 25 Apr. 2020.
- Smith, Barry, et al. "The OBO Foundry: Coordinated Evolution of Ontologies to Support Biomedical Data Integration." *Nature Biotechnology*, vol. 25, no. 11, 1 Nov. 2007, pp. 1251–1255, www.nature.com/articles/nbt1346, <https://doi.org/10.1038/nbt1346>. Accessed 11 Mar. 2022.
- Srivastava, Rakesh K, et al. *Genome-Wide Association Studies and Genomic Selection in Pearl Millet: Advances and Prospects*. Vol. 10, 28 Feb. 2020, <https://doi.org/10.3389/fgene.2019.01389>. Accessed 19 May 2023.
- Stigler, Stephen M. *The History of Statistics*. Harvard University Press, 1 Mar. 1990.
- Wallace, Michael P., et al. "Dynamic Treatment Regimen Estimation via Regression-Based Techniques: Introducing R Package **DTRreg**." *Journal of Statistical Software*, vol. 80, no. 2, 2017, <https://doi.org/10.18637/jss.v080.i02>. Accessed 9 May 2023.
- Welty, Leah J., et al. "Strategies for Developing Biostatistics Resources in an Academic Health Center." *Academic Medicine*, vol. 88, no. 4, Apr. 2013, pp. 454–460, <https://doi.org/10.1097/acm.0b013e31828578ed>. Accessed 28 Jan. 2020.

- White, B. G. "Initial Statistical Considerations." *Consensus.app*, 1990, pp. 17–27, consensus.app/papers/statistical-considerations-white/4eefddbc83f55f9199500fc0c3a8835e/?utm_source=chatgpt, https://doi.org/10.1007/978-1-4613-1495-0_2. Accessed 26 Jan. 2024.
- Wittkowski, Knut M, et al. "From Single-SNP to Wide-Locus: Genome-Wide Association Studies Identifying Functionally Related Genes and Intragenic Regions in Small Sample Studies." *Pharmacogenomics*, vol. 14, no. 4, Mar. 2013, pp. 391–401, <https://doi.org/10.2217/pgs.13.28>. Accessed 8 June 2021.
- Wright, Sewall. "Evolution in Mendelian Populations." *Genetics*, vol. 16, no. 2, 1 Mar. 1931, pp. 97–159, www.ncbi.nlm.nih.gov/pmc/articles/PMC1201091/.
- Wu, C. H. "Integrative Bioinformatics for Genomics and Proteomics." *Journal of Biomolecular Techniques*, vol. 22, 2011, consensus.app/papers/bioinformatics-genomics-proteomics-wu/4e51f52d98c95bf38e7b8a99e9ad6730/?utm_source=chatgpt. Accessed 27 Jan. 2024.
- Wu, Rongling, and Zuoheng Wang. "Mathematical Modeling of Systems Pharmacogenomics towards Personalized Drug Delivery." *Advanced Drug Delivery Reviews*, vol. 65, no. 7, June 2013, pp. 903–904, <https://doi.org/10.1016/j.addr.2013.05.004>. Accessed 13 Apr. 2020.
- Xia, Yinglin, and Jun Sun. "Hypothesis Testing and Statistical Analysis of Microbiome." *Genes & Diseases*, vol. 4, no. 3, Sept. 2017, pp. 138–148, <https://doi.org/10.1016/j.gendis.2017.06.001>. Accessed 15 Nov. 2020.

Zhongping, Zhang. "BioDW: A Bioinformatics Data Integration System." *Microcomputer Applications*, 2005, consensus.app/papers/biodw-bioinformatics-data-integration-system-zhongping/b93cf849623d51ca94a09765c9054127/?utm_source=chatgpt. Accessed 27 Jan. 2024.

BAB 3

STATISTIKA DESKRIPTIF

Oleh Wara Alfa Syukrilla

3.1 Pendahuluan

Pada bab ini akan dibahas pengolahan data menggunakan statistika deskriptif yang dapat digunakan oleh pembaca untuk merangkum atau mendeskripsikan karakteristik dari suatu kumpulan data. Statistika deskriptif diantaranya terdiri dari ukuran pemusatan data (*measures of central tendency*), ukuran penyebaran data (*measures of variability*), dan ukuran distribusi frekuensi (*measures of frequency distribution*). Ukuran pemusatan data menggambarkan posisi pusat dari sebuah kumpulan data melalui mean, median, dan modus. Ukuran penyebaran data mengilustrasikan bagaimana sebuah kumpulan data menyebar dan bervariasi yang dapat diukur melalui varians, deviasi standar, *range*, dan *interquartile range*. Sedangkan ukuran distribusi frekuensi mempermudah pemahaman akan pola sebaran data. Penjelasan tentang ketiga jenis ukuran tersebut akan dilengkapi dengan contoh penggunaannya dalam bidang biostatistika.

3.2 Ukuran Pemusatan Data

Ukuran pemusatan data adalah statistik yang memberikan gambaran tentang nilai tengah atau pusat dari suatu set data. Ukuran-ukuran ini membantu merangkum data dan memberikan informasi tentang letak pusat distribusi data. Beberapa ukuran pemusatan data yang umum digunakan meliputi (Khan, 2011):

1. *Mean* (Rata-rata): Rata-rata aritmetika dari semua nilai dalam suatu set data. Rata-rata dihitung dengan menjumlahkan semua nilai dan kemudian dibagi dengan banyaknya data.
2. *Median* (Median): Nilai tengah dari suatu set data yang telah diurutkan. Jika jumlah data ganjil, median adalah nilai tengah. Jika jumlah data genap, median adalah rata-rata dari dua nilai tengah. Median adalah ukuran pemusatan data yang cocok digunakan saat data memiliki nilai ekstrem (data yang nilainya sangat besar atau sangat kecil) atau ketika data tidak terdistribusi normal (White, 2018).
3. *Mode* (Modus): Nilai atau nilai-nilai yang muncul paling sering dalam suatu set data. Sebuah set data bisa memiliki mode tunggal (unimodal) atau lebih dari satu mode (multimodal).

Ukuran-ukuran ini memberikan informasi yang berbeda tentang pusat distribusi data. Rata-rata sensitif terhadap nilai-nilai ekstrem, sedangkan median lebih tahan terhadap nilai-nilai ekstrem. Mode memberikan informasi tentang nilai yang paling sering muncul dalam set data.

Pemilihan ukuran pemusatan data yang tepat tergantung pada sifat distribusi data dan tujuan analisis statistik yang dilakukan. Beberapa distribusi mungkin lebih baik direpresentasikan oleh median daripada rata-rata jika terdapat nilai-nilai ekstrem yang signifikan (White, 2018).

Berikut ini adalah contoh penggunaan ukuran pemusatan data pada data tentang jumlah penderita penyakit tuberkulosis (TBC) setiap kelurahan/desa di sebuah provinsi X dalam satu tahun.

Tabel 3. 1. Data ilustrasi jumlah penderita TBC

ID Kelurahan/ Desa	Jumlah penderita TBC
1.	2
2.	5
3.	7
4.	2
5.	3
6.	4
7.	10
8.	12
9.	2
10.	1

ID Kelurahan/ Desa	Jumlah penderita TBC
11.	3
12.	4
13.	7
14.	9
15.	13
16.	11
17.	20
18.	9
19.	3
20.	5

Sumber: Data Pribadi

Mean pada data yang disajikan Tabel 3.1 adalah

$$\frac{2+5+\dots+2+1+3+4+\dots+3+5}{20} = 6,6.$$

Artinya rata-rata banyaknya penderita TBC di provinsi X adalah sekitar 6 orang tiap kelurahan/desa. Sedangkan median dari data tersebut diperoleh dengan mengurutkan data dari nilai terkecil ke terbesar terlebih dahulu lalu menghitung median menggunakan rumus $Me = \frac{1}{2}(X_{\frac{n}{2}} + X_{\frac{n}{2}+1})$ mengingat banyaknya data adalah genap. Jika banyaknya data ganjil maka median dihitung dengan rumus $Me = \frac{X_{n+1}}{2}$.

Data yang telah diurutkan dari kecil ke besar:

1,2,2,2,3,3,3,3,4,4,5,5,7,7,9,9,10,11,12,13,20

$$\begin{aligned} Me &= \frac{1}{2} \left(X_{\frac{n}{2}} + X_{\frac{n}{2}+1} \right) = \frac{1}{2} \left(X_{\frac{20}{2}} + X_{\frac{20}{2}+1} \right) = \frac{1}{2} (X_{10} + X_{11}) \\ &= \frac{1}{2} (5 + 5) = 5 \end{aligned}$$

Dengan demikian diperoleh median atau nilai tengah dari data Tabel 3.1 adalah 5. Dengan kata lain, dalam satu tahun terdapat 5 orang menderita penyakit TBC di tiap kelurahan/desa pada provinsi X. Tampak jelas disini bahwa walaupun sama-sama menggambarkan posisi pusat data, namun nilai rata-rata (*mean*) berbeda dengan nilai tengah median. Ketika terdapat nilai ekstrem pada data sebaiknya menggunakan median sebagai ukuran pusat data.

Ukuran pemusatan ketiga adalah modus (*mode*). Pada data Tabel 3.1, bilangan 2 muncul tiga kali dan bilangan 3 juga muncul tiga kali, sedangkan bilangan lainnya muncul dua atau satu kali. Artinya, modus dari data ini adalah 2 dan 3 karena dua bilangan tersebut yang paling sering muncul.

3.3 Ukuran Variabilitas Data

Ukuran variabilitas data mengukur seberapa jauh atau seberapa tersebar nilai-nilai dalam suatu set data. Ukuran ini memberikan informasi tentang seberapa heterogen atau homogen distribusi data. Beberapa ukuran variabilitas yang umum digunakan dalam meliputi (Osborn, 2006; Pace, 2012):

1. Rentang (*Range*): Rentang adalah perbedaan antara nilai maksimum dan nilai minimum dalam suatu set data. Rentang memberikan gambaran kasar tentang seberapa besar variasi data.

2. Varians (*Variance*): Varians mengukur sebaran nilai di sekitar rata-rata. Varians dihitung dengan menjumlahkan kuadrat selisih antara setiap nilai dengan rata-rata, kemudian dibagi dengan jumlah total nilai. Varians yang lebih besar menunjukkan variasi yang lebih besar dalam data.
3. Deviasi Standar (*Standard Deviation*): Deviasi standar adalah akar kuadrat dari varians. Deviasi standar memberikan pengukuran variabilitas dalam satuan yang sama dengan data asli. Deviasi standar yang kecil menunjukkan bahwa sebagian besar nilai cenderung mendekati rata-rata.
4. Jangkauan Antar Kuartil (*Interquartile Range/IQR*): IQR adalah perbedaan antara kuartil atas (Q3, posisi 75% data) dan kuartil bawah (Q1, posisi 25% data). Kuartil adalah nilai-nilai yang membagi data menjadi empat bagian sama besar. IQR memberikan informasi tentang sebaran nilai-nilai di sekitar nilai tengah dan lebih tahan terhadap pengaruh nilai-nilai ekstrem dibandingkan dengan rentang.
5. Koefisien Variasi (*Coefficient of Variation*): Koefisien variasi adalah rasio antara deviasi standar dan rata-rata, kemudian dikalikan dengan 100 untuk mendapatkan persentase. Ukuran ini memberikan ukuran relatif variabilitas data dalam satuan persentase dari rata-rata.

Ukuran-ukuran ini akan memberikan informasi tentang seberapa besar variasi atau seberapa seragam data. Dalam memilih ukuran variabilitas mana yang sesuai digunakan untuk mendeskripsikan data yang kita miliki akan bergantung pada sifat distribusi data dan tujuan analisis statistik yang dilakukan. Nilai variabilitas yang tinggi dapat mengindikasikan keragaman atau ketidakpastian dalam data, sementara variabilitas yang rendah menunjukkan bahwa data cenderung lebih seragam atau konsisten.

Contoh penerapan ukuran variabilitas data pada contoh data di Tabel 3.1 adalah sebagai berikut:

- Rentang data adalah 19, diperoleh dari selisih nilai terbesar dengan nilai terkecil yaitu $20-1=19$
- Varians = $\frac{\sum(X_i - \bar{X})^2}{n-1} = \frac{(2-6,6)^2 + \dots + (5-6,6)^2}{20-1} = 23,41$
- Deviasi Standar = $\sqrt{\sigma^2} = \sqrt{\text{varians}} = \sqrt{23,41} = 4,84$
- IQR = $Q_3 - Q_1 = 9,25 - 3 = 6,25$
- Koefisien variasi = $\frac{\text{deviasi standar}}{\text{rata-rata}} \times 100\% = \frac{4,84}{6,6} \times 100\% = 73,31\%$

3.4 Ukuran Distribusi Frekuensi

Ukuran distribusi frekuensi digunakan dalam statistika untuk merangkum dan menampilkan sebaran nilai-nilai dalam suatu set data. Distribusi frekuensi memberikan gambaran tentang seberapa sering nilai-nilai tertentu muncul dalam data. Beberapa ukuran distribusi frekuensi yang umum digunakan melibatkan konsep frekuensi, persentase, dan tabel distribusi frekuensi (Dunn & Davis, 2017; Thomas, 2018).

1. Frekuensi (*Frequency*): Frekuensi adalah jumlah berapa kali suatu nilai muncul dalam suatu set data. Dalam distribusi frekuensi, frekuensi direkam untuk setiap interval atau nilai tertentu, dan total frekuensi memberikan gambaran tentang sebaran data.
2. Persentase (*Percentage*): Persentase adalah bagian dari total frekuensi yang dinyatakan sebagai persentase. Persentase dapat dihitung dengan membagi frekuensi untuk suatu nilai atau interval dengan total frekuensi, lalu dikalikan dengan 100. Persentase memberikan pemahaman yang lebih intuitif tentang kontribusi relatif suatu nilai terhadap keseluruhan distribusi.

3. Tabel Distribusi Frekuensi (*Frequency Distribution Table*): Tabel distribusi frekuensi adalah tabel yang mengorganisir data ke dalam kelompok-kelompok atau interval dengan mencantumkan frekuensi masing-masing. Tabel ini sering melibatkan kolom-kolom seperti nilai, frekuensi, persentase, dan kumulatif (jumlah kumulatif).
4. Histogram: Histogram adalah representasi grafis dari distribusi frekuensi. Histogram menggambarkan sebaran data dalam bentuk batang-batang yang mewakili interval atau kategori tertentu. Histogram membantu visualisasi pola distribusi data, apakah simetris, miring, atau memiliki bentuk khusus.
5. Diagram Batang (*Bar Chart*): Diagram batang sering digunakan untuk menampilkan distribusi frekuensi kategori atau variabel kualitatif. Setiap kategori atau interval direpresentasikan oleh batang vertikal, dan tinggi batang mencerminkan frekuensi atau persentase kategori tersebut.

Ukuran distribusi frekuensi membantu penyajian data secara sistematis dan memudahkan kita lebih memahami pola sebaran data yang dimiliki. Pemilihan ukuran distribusi frekuensi yang tepat dan berguna bagi eksplorasi data yang kita miliki bergantung pada jenis data (kualitatif atau kuantitatif) dan tujuan analisis statistik yang akan kita lakukan.

Apabila data pada Tabel 3.1 disajikan ukuran distribusi frekuensinya maka diperoleh hasil sebagai berikut:

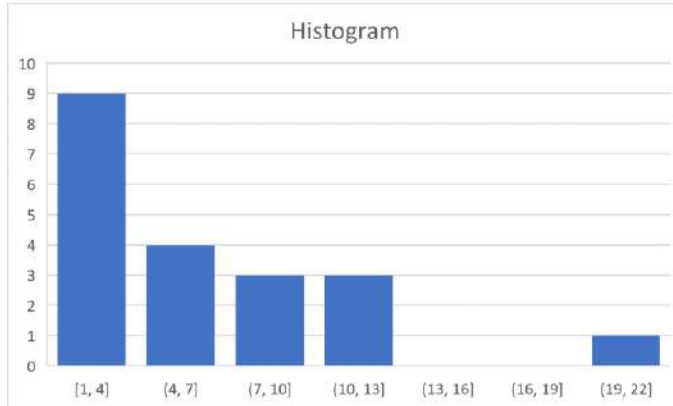
1. Frekuensi munculnya bilangan 2 adalah tiga kali, frekuensi munculnya bilangan 20 adalah satu kali, dst. Dengan kata lain, ditemukan tiga kelurahan/desa yang memiliki 3 orang penderita TBC di daerahnya dan hanya satu kelurahan/desa yang memiliki 20 orang penderita TBC di areanya.

2. Banyaknya kelurahan/desa yang memiliki 2 orang penderita TBC di daerahnya adalah $\frac{3}{20} = 15\%$ dari total kelurahan di provinsi X. Sementara itu, hanya 5% kelurahan di provinsi X yang memiliki 20 orang penderita TBC di daerahnya (diperoleh dari $\frac{1}{20} = 5\%$).
3. Tabel distribusi frekuensi yang tertera pada Tabel 3.2

Tabel 3. 2. Tabel distribusi frekuensi

Nilai	Frekuensi
1	1
2	3
3	3
4	2
5	2
7	2
9	2
10	1
11	1
12	1
13	1
20	1

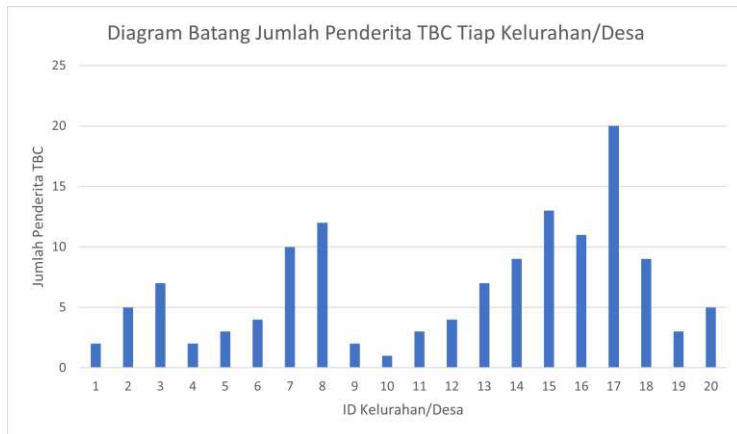
4. Histogram untuk data yang memiliki distribusi frekuensi seperti Tabel 3.2 dapat disajikan seperti Gambar 3.1. Pada Gambar 3.1, batang pertama menunjukkan frekuensi munculnya bilangan 1 sampai 4, yaitu 9. Nilai 9 diperoleh dari perhitungan frekuensi, di mana bilangan 1 muncul satu kali, bilangan 2 muncul 3 kali, bilangan 3 muncul tiga kali, dan bilangan 4 muncul dua kali, sehingga totalnya adalah 9. Sedangkan batang kedua adalah frekuensi munculnya bilangan 5 dan 7, yaitu masing-masing dua kali sehingga frekuensinya adalah 4.



Gambar 3. 1. Histogram data

Sumber: Data Pribadi

5. Apabila Tabel 3.1 disajikan dalam bentuk diagram batang, maka diagram batang jumlah penderita TBC tiap daerah dapat diilustrasikan sebagaimana Gambar 3.2



Gambar 3. 2. Diagram Batang

Sumber: Data Pribadi

3.5 Statistika Deskriptif *Univariate* dan *Bivariate*

Istilah "Univariate" merujuk pada analisis data yang hanya melibatkan satu variabel (Munro, 2005). Dalam konteks ini, pemilik data berfokus pada distribusi dan karakteristik statistik dari satu variabel tunggal tanpa mempertimbangkan korelasi dengan variabel lain. Sebagai contoh, dalam pengumpulan data berat siswa di sebuah kelas, analisis univariate akan terfokus pada distribusi berat badan itu sendiri, menggunakan metode seperti histogram, rata-rata, deviasi standar, khusus untuk variabel tinggi badan.

Sementara itu, istilah "Bivariate" mengacu pada analisis data yang melibatkan dua variabel. Dalam analisis yang melibatkan dua variabel, pemilik data bermaksud untuk mengeksplorasi hubungan antara keduanya (Plichta & Garzon, 2009). Analisis statistika deskriptif bivariate melibatkan teknik seperti korelasi, regresi, diagram pencar (*scatter plot*) antara dua variabel, dan visualisasi dua dimensi lainnya. Sebagai contoh, dalam situasi kita memiliki dua set data, satu untuk kadar kolesterol dan yang lainnya untuk tekanan darah, analisis bivariate dapat membantu kita memahami apakah ada hubungan antara kedua variabel tersebut. Misalnya, dapat dilakukan pembuatan diagram pencar (*scatter plot*) dan perhitungan korelasi antara kolesterol dan tekanan darah.

Contoh kasus lainnya untuk analisis statistika deskriptif bivariate adalah misalkan ada sebuah penelitian yang mencoba memahami hubungan antara umur dan kadar gula darah seseorang. Dalam hal ini, variabel pertama (X) adalah umur dan variabel kedua (Y) adalah kadar gula darah. Analisis statistika deskriptif bivariate yang dapat dilakukan untuk dua variabel ini dapat mencakup:

1. Pengumpulan Data: Mengumpulkan data umur dan kadar gula darah dari sejumlah individu.
2. Visualisasi Data: Membuat scatter plot yang menunjukkan titik-titik data untuk setiap individu, dengan sumbu x mewakili umur dan sumbu y mewakili kadar gula darah.
3. Korelasi: Menghitung koefisien korelasi untuk menentukan apakah ada hubungan linier antara umur dan kadar gula darah. Sebagai contoh, koefisien korelasi positif dapat menunjukkan bahwa semakin tua usia, semakin tinggi kadar gula darah.
4. Regresi: Melakukan analisis regresi untuk memodelkan hubungan antara umur dan kadar gula darah. Analisis Regresi bisa memberikan persamaan garis regresi yang menggambarkan tren umum dalam data.

Analisis bivariate seperti ini membantu pemilik data untuk memahami bagaimana dua variabel berinteraksi dan memberikan informasi tentang hubungan di antara keduanya.

DAFTAR PUSTAKA

- Dunn, P. F., & Davis, M. P. (2017). *Measurement and Data Analysis for Engineering and Science*. CRC Press.
- Khan. (2011). *Research Methodology*. APH Publishing Corporation.
- Munro, B. H. (2005). *Statistical Methods for Health Care Research*. Lippincott Williams & Wilkins.
- Osborn, C. E. (2006). *Statistical Applications for Health Information Management*. Jones and Bartlett Publishers.
- Pace, L. (2012). *Beginning R: An Introduction to Statistical Programming*. Apress.
- Plichta, S. B., & Garzon, L. S. (2009). *Statistics for Nursing and Allied Health*. Wolters Kluwer/Lippincott Williams & Wilkins Health.
- Thomas, D. (2018). *Beginne R: Introductory Statistics Using R*. SuJinSoLa.
- White, L. H. (2018). *Statistics Straight Up :The Essential Guide to Understanding and Using Statistics*. Linus Publications, Incorporated.

BAB 4

STATISTIKA INFERENSIAL

Oleh Putu Erma Pradnyani

4.1 Pendahuluan

Statistik deskriptif dan inferensial adalah dua cabang utama dalam statistika yang memiliki tujuan dan metode yang berbeda. Statistika inferensial adalah cabang statistika yang berkaitan dengan proses pengambilan keputusan, membuat perkiraan, atau melakukan generalisasi tentang populasi berdasarkan data yang diperoleh dari sampel (Susilawati and Dharmawansyah, 2019).

Perbedaan antara statistik deskriptif dan inferensial dilihat dari tujuan, cara analisis data dan penggunaan data. Statistik Deskriptif bertujuan untuk Menyajikan dan merangkum data dengan cara yang informatif. Tujuannya adalah untuk memberikan gambaran yang jelas tentang karakteristik atau pola yang ada dalam data. Analisis data statistika deskriptif adalah Fokus pada pengukuran tendensi sentral (seperti rata-rata, median, dan modus), penyebaran data (seperti rentang dan deviasi standar), serta representasi visual (grafik, tabel, dan diagram). Data hasil statistika deskriptif digunakan untuk memberikan gambaran umum tentang dataset, membuat ringkasan statistik, dan menyajikan data dengan cara yang mudah dimengerti (Martias, 2021). Sedangkan Statistik Inferensial bertujuan Mengambil kesimpulan atau membuat perkiraan tentang populasi berdasarkan sampel data. Tujuannya adalah untuk membuat generalisasi atau inferensi tentang suatu populasi berdasarkan informasi yang diperoleh dari sampel. Analisis data statistika

inferensial Melibatkan teknik-teknik seperti uji hipotesis, interval kepercayaan, dan analisis regresi untuk membuat inferensi tentang parameter populasi. Data hasil statistika inferensial digunakan untuk membuat klaim atau prediksi tentang karakteristik populasi yang lebih besar berdasarkan analisis sampel data.

Contoh penggunaan statistika Deskriptif adalah pengukuran tinggi badan siswa dalam suatu kelas dan ingin tahu rata-rata tinggi badan, menggunakan statistik deskriptif seperti *mean* dan median akan memberikan informasi tentang tinggi badan rata-rata di kelas tersebut. Sedangkan contoh statistika Inferensial adalah membuat klaim tentang tinggi badan rata-rata di seluruh sekolah berdasarkan sampel dari satu kelas, Anda dapat menggunakan inferensi statistik seperti uji hipotesis untuk membuat klaim yang lebih umum.

Dalam praktiknya, statistik deskriptif dan inferensial sering digunakan bersamaan untuk memberikan pemahaman menyeluruh tentang data. Statistik deskriptif digunakan untuk merangkum dan menyajikan data, sementara statistik inferensial digunakan untuk membuat generalisasi lebih lanjut atau membuat prediksi berdasarkan sampel data.

Dengan kata lain, statistika inferensial digunakan untuk membuat inferensi atau kesimpulan tentang populasi berdasarkan informasi yang diperoleh dari sampel. Statistika inferensial digunakan dalam berbagai disiplin ilmu seperti ilmu sosial, ekonomi, kedokteran, dan lain-lain untuk membuat kesimpulan dan generalisasi yang dapat diterapkan pada populasi berdasarkan data yang terbatas yang diperoleh dari sampel (Mardhotillah, Asyhar and Elisa, 2022).

4.2 Kegunaan Statistika Inferensial

Beberapa kegunaan statistika inferensial adalah sebagai berikut (Sholikhah, 2016; Fitriatien, 2017).

1. Generalisasi

Statistika inferensial memungkinkan kita untuk membuat generalisasi tentang populasi berdasarkan data yang diperoleh dari sampel. Ini membantu dalam membuat pernyataan yang lebih luas dan mendasar tentang karakteristik populasi.

2. Pengambilan Keputusan

Dengan menggunakan teknik statistika inferensial, kita dapat membuat keputusan atau memberikan rekomendasi berdasarkan informasi yang terkandung dalam sampel. Misalnya, apakah perbedaan antara dua kelompok adalah hasil kebetulan atau mewakili perbedaan yang sebenarnya.

3. Perkiraan Parameter Populasi

Statistika inferensial memungkinkan kita untuk membuat perkiraan tentang parameter populasi, seperti rata-rata atau proporsi, dengan menggunakan data sampel.

4. Uji Hipotesis

Statistika inferensial membantu dalam menguji hipotesis atau asumsi tentang parameter populasi. Ini berguna untuk memvalidasi atau menolak klaim yang dibuat tentang populasi.

5. Analisis Risiko

Dengan menggunakan interval kepercayaan, kita dapat menganalisis risiko kesalahan yang mungkin terjadi dalam membuat estimasi tentang populasi. Hal ini membantu dalam menyadari tingkat ketidakpastian yang terkait dengan generalisasi dari sampel ke populasi.

4.3 Syarat Statistika Inferensial

Beberapa Syarat Statistika Inferensial adalah sebagai berikut.

1. Sampel Representatif

Sampel yang digunakan untuk inferensi statistika harus representatif terhadap populasi. Ini berarti sampel tersebut harus mencerminkan karakteristik yang sama dengan populasi secara keseluruhan.

2. Random Sampling

Proses pemilihan sampel harus dilakukan secara acak untuk menghindari bias dalam pemilihan dan memastikan bahwa setiap anggota populasi memiliki kesempatan yang setara untuk menjadi bagian dari sampel.

3. Ukuran Sampel yang Cukup Besar

Ukuran sampel yang besar dapat menghasilkan hasil inferensial yang lebih dapat diandalkan. Ukuran sampel yang kecil dapat meningkatkan tingkat ketidakpastian dalam generalisasi ke populasi.

4. Data Berskala yang Sesuai

Teknik statistika inferensial yang digunakan harus sesuai dengan jenis data yang dianalisis. Misalnya, uji t digunakan untuk data berdistribusi normal, sedangkan uji non-parametrik mungkin lebih sesuai untuk data yang tidak memenuhi asumsi normalitas.

5. Ketelitian Pengukuran

Data yang digunakan harus diukur dengan akurat dan konsisten. Ketelitian ini penting untuk memastikan hasil inferensial yang tepat

6. Memahami Asumsi Statistik

Beberapa teknik statistika inferensial memiliki asumsi tertentu yang harus dipenuhi agar hasilnya dapat dianggap valid. Misalnya, uji t membutuhkan asumsi normalitas data.

7. Pengendalian Variabel Lain

Variabel lain yang dapat memengaruhi hasil harus dikendalikan atau diukur secara sistematis untuk meminimalkan pengaruhnya terhadap hasil inferensial.

Penting untuk memahami bahwa penggunaan statistika inferensial memerlukan pemahaman yang baik tentang metode statistik, interpretasi hasil, dan asumsi yang mendasarinya. Selain itu, kesadaran tentang batasan dan ketidakpastian dalam inferensi statistika juga diperlukan.

4.4 Jenis Statistika Inferensial

Beberapa jenis statistika inferensial melibatkan teknik-teknik berikut:

1. Estimasi Parameter

a. Interval Estimasi

Menciptakan interval kepercayaan untuk parameter populasi. Sebagai contoh, kita dapat membuat interval kepercayaan untuk rata-rata populasi berdasarkan data sampel.

b. Estimasi Titik

Menggunakan data sampel untuk menghitung nilai perkiraan titik dari parameter populasi.

2. Uji Hipotesis

a. Uji Parameter

Melibatkan pengujian hipotesis tentang parameter populasi. Misalnya, uji hipotesis tentang rata-rata populasi.

b. Uji Perbedaan

Digunakan untuk menguji apakah terdapat perbedaan yang signifikan antara dua kelompok atau lebih.

c. Uji Asosiasi

Digunakan untuk menguji apakah ada hubungan atau asosiasi antara dua variabel.

3. Analisis Variansi (ANOVA):

a. One-Way ANOVA, Digunakan untuk menguji apakah terdapat perbedaan signifikan antara rata-rata lebih dari dua kelompok.

b. Two-Way ANOVA, Digunakan untuk menganalisis efek dua faktor terhadap suatu variabel.

4. Regresi

a. Regresi Linier, Digunakan untuk memahami hubungan linier antara satu variabel dependen dengan satu atau lebih variabel independen.

b. Regresi Logistik, Digunakan ketika variabel dependen bersifat biner (dua kategori).

5. Analisis Chi-Square

Uji Chi-Square Digunakan untuk menguji apakah terdapat hubungan antara dua variabel kategori.

6. Analisis Korelasi

a. Korelasi Pearson, digunakan untuk mengukur sejauh mana dua variabel berhubungan secara linier.

- b. Korelasi Spearman, Mengukur hubungan monotone antara dua variabel.

4.5 Implementasi Statistika Inferensial yang sering digunakan dalam Penelitian

4.5.1 Uji T-Tes

T-Test adalah metode statistik yang digunakan untuk membandingkan rata-rata dua kelompok. T-Test dapat digunakan dalam berbagai konteks, tetapi umumnya digunakan ketika kita ingin mengetahui apakah ada perbedaan signifikan antara dua kelompok dalam hal rata-rata. Ada beberapa jenis T-Test, tetapi dua yang paling umum adalah T-Test Independen dan T-Test Dependen.

A. T-Test Independen

Digunakan ketika kita memiliki dua kelompok yang tidak saling terkait atau tidak berasal dari pengamatan yang sama. Syarat-syarat untuk T-Test Independen adalah sebagai berikut (Usmadi, 2020).

1. Data Berdistribusi Normal, Data dalam setiap kelompok harus mengikuti distribusi normal. Jika ukuran sampel cukup besar (biasanya $n > 30$), T-Test kurang sensitif terhadap pelanggaran asumsi distribusi normal.
2. Homogenitas Varians, Varians (sebaran data) dalam kelompok-kelompok yang dibandingkan harus homogen atau setidaknya mendekati homogen.
3. Data Skala Interval atau Rasio, T-Test sebaiknya digunakan pada data yang berskala interval atau rasio.

Langkah-langkah T-Test Independen adalah sebagai berikut.

1. Menyusun hipotesis nol (H_0) dan hipotesis alternatif (H_1).
2. Menentukan tingkat signifikansi (α).
3. Mengumpulkan data dari dua kelompok yang akan dibandingkan.
4. Menghitung rata-rata, deviasi standar, dan ukuran sampel dari masing-masing kelompok.
5. Menghitung nilai T-Test menggunakan formula yang sesuai.
6. Membandingkan nilai T-Test dengan nilai kritis dari tabel distribusi T atau menghitung nilai p-value.
7. Menarik kesimpulan berdasarkan hasil uji statistik.

B. T-Test Dependen

Digunakan ketika kita memiliki dua kelompok yang saling terkait, misalnya pengukuran yang diambil pada waktu yang berbeda pada kelompok yang sama. Syarat-syarat untuk T-Test Dependen adalah sebagai berikut.

1. Data Berdistribusi Normal, artinya data dalam setiap kelompok harus mengikuti distribusi normal.
2. Data Berpasangan dikenal dengan T-Test Dependen yang digunakan pada data yang berpasangan atau saling terkait.
3. Skala Interval atau Rasio, T-Test sebaiknya digunakan pada data yang berskala interval atau rasio.

Langkah-langkah T-Test Dependen adalah sebagai berikut.

1. Menyusun hipotesis nol (H_0) dan hipotesis alternatif (H_1).
2. Menentukan tingkat signifikansi (α).
3. Mengumpulkan data dari kelompok yang sama yang diukur pada dua waktu yang berbeda.
4. Menghitung selisih antara pengukuran yang berpasangan.

5. Menghitung rata-rata, deviasi standar, dan ukuran sampel dari selisih tersebut.
6. Menghitung nilai T-Test menggunakan formula yang sesuai.
7. Membandingkan nilai T-Test dengan nilai kritis dari tabel distribusi T atau menghitung nilai p-value.
8. Menarik kesimpulan berdasarkan hasil uji statistik.

Penting untuk memahami bahwa hasil T-Test tidak membuktikan atau menolak hipotesis nol, tetapi memberikan bukti yang memungkinkan kita untuk membuat keputusan statistik. Keputusan akhir harus dibuat berdasarkan konteks riset dan pertimbangan praktis.

4.5.2 Uji Chisquare

Uji Chi-Square (χ^2) adalah metode statistik yang digunakan untuk menentukan apakah ada hubungan antara dua variabel kategorikal dan untuk menguji signifikansi perbedaan antara distribusi pengamatan yang diamati dengan distribusi yang diharapkan. Uji Chi-Square memiliki beberapa variasi tergantung pada jenis data dan desain penelitian. Dua variasi yang umum adalah Uji Chi-Square untuk Uji Independensi dan Uji Chi-Square untuk Uji Kesesuaian.

Penggunaan Uji Independensi untuk menentukan apakah ada hubungan antara dua variabel kategorikal. Sedangkan Uji Kesesuaian Digunakan untuk menentukan sejauh mana distribusi pengamatan sesuai dengan distribusi yang diharapkan.

Uji Chi-Square dapat diterapkan dalam berbagai bidang seperti ilmu sosial, kedokteran, ekonomi, dan lainnya untuk menjawab pertanyaan terkait hubungan atau kesesuaian antara variabel kategorikal. Penting untuk memahami bahwa hasil uji ini memberikan informasi tentang signifikansi statistik, tetapi bukan

mengenai sebab-akibat atau kekuatan hubungan antar variabel tersebut.

A. Uji Chi-Square untuk Uji Independensi

Syarat-syarat uji chisquare adalah sebagai berikut (Kawulur, 2022).

1. Data Kategorikal, Variabel yang diukur harus bersifat kategorikal, artinya data dikelompokkan ke dalam kategori atau kelompok.
2. Tidak Ada Pengelompokan yang Terlalu Kecil, Harapannya, tidak ada lebih dari 20% sel yang memiliki nilai harapan kurang dari 5.

Langkah-langkah:

1. Menyusun hipotesis nol (H_0) dan hipotesis alternatif (H_1).
2. Menentukan tingkat signifikansi (α).
3. Membuat tabel kontingensi yang memperlihatkan frekuensi pengamatan untuk setiap kombinasi kategori.
4. Menghitung nilai Chi-Square dengan rumus yang sesuai.
5. Menghitung derajat kebebasan (df) untuk uji Chi-Square.
6. Menentukan nilai kritis Chi-Square pada tingkat signifikansi tertentu atau menghitung p-value.
7. Menarik kesimpulan berdasarkan hasil uji statistik.

B. Uji Chi-Square untuk Uji Kesesuaian

Syarat-syarat uji Chisquare adalah sebagai berikut.

1. Data Kategorikal, Variabel yang diukur harus bersifat kategorikal.
2. Data Diobservasi dan Dihipotesiskan, Ada perbandingan antara distribusi pengamatan yang diamati dengan distribusi yang diharapkan.

Langkah-langkah uji Chisquare adalah sebagai berikut.

1. Menyusun hipotesis nol (H_0) dan hipotesis alternatif (H_1).
2. Menentukan tingkat signifikansi (α).
3. Membuat tabel distribusi yang diharapkan berdasarkan hipotesis nol.
4. Menghitung nilai Chi-Square dengan rumus yang sesuai.
5. Menghitung derajat kebebasan (df) untuk uji Chi-Square.
6. Menentukan nilai kritis Chi-Square pada tingkat signifikansi tertentu atau menghitung p-value.
7. Menarik kesimpulan berdasarkan hasil uji statistik.

4.5.3 Uji Anova

Uji ANOVA (*Analysis of Variance*) digunakan untuk membandingkan rata-rata dari tiga atau lebih kelompok yang berbeda (Marpaung, Sutrisno and Lumintang, 2017). ANOVA memeriksa apakah ada perbedaan yang signifikan antara setidaknya dua kelompok rata-rata. Terdapat beberapa jenis ANOVA, dan syarat serta langkah-langkahnya dapat bervariasi tergantung pada jenisnya. ANOVA umumnya digunakan ketika kita ingin mengetahui apakah ada perbedaan yang signifikan antara rata-rata kelompok dalam eksperimen atau penelitian. Contoh penggunaan ANOVA dapat ditemui dalam bidang-bidang seperti ilmu sosial, kedokteran, bisnis, dan lainnya. Misalnya, dalam ilmu sosial, ANOVA dapat digunakan untuk membandingkan rata-rata kepuasan pelanggan antara tiga merek yang berbeda.

A. ANOVA Satu Arah (*One-Way ANOVA*)

Syarat-syarat uji Anova adalah sebagai berikut.

1. Data Normal, Data dalam setiap kelompok harus mengikuti distribusi normal.
2. Homogenitas Varians, Varians dari kelompok-kelompok yang dibandingkan harus sebanding.

3. Independensi Observasi, Observasi dalam satu kelompok tidak boleh memengaruhi observasi di kelompok lain.

Langkah-langkah uji Anova adalah sebagai berikut

1. Menyusun hipotesis nol (H_0) dan hipotesis alternatif (H_1).
2. Menentukan tingkat signifikansi (α).
3. Mengumpulkan data dari setiap kelompok.
4. Menghitung rata-rata, deviasi standar, dan ukuran sampel untuk masing-masing kelompok.
5. Menghitung variansi antar kelompok (Variance Between Groups) dan variansi dalam kelompok (Variance Within Groups).
6. Menghitung nilai F-Test (F-ratio) dengan membagi variansi antar kelompok dengan variansi dalam kelompok.
7. Menentukan nilai kritis F atau menghitung p-value.
8. Menarik kesimpulan berdasarkan hasil uji statistik.

B. ANOVA Dua Arah (*Two-Way ANOVA*)

ANOVA Dua Arah digunakan ketika ada dua faktor yang mempengaruhi variabel dependen. Syarat-syarat dan langkah-langkah Sama dengan ANOVA Satu Arah, dengan tambahan Independensi Faktor yaitu Faktor-faktor yang mempengaruhi kelompok-kelompok harus independen serta Interaksi Antara Faktor yang Melakukan analisis interaksi untuk melihat apakah ada efek interaksi antara dua faktor.

C. MANOVA (*Multivariate Analysis of Variance*)

MANOVA memperluas ANOVA untuk variabel dependen yang lebih dari satu. Syarat-syarat dan langkah-langkah sama dengan ANOVA Satu Arah, dengan tambahan Multivariate Normality yaitu Variabel dependen multivariat harus mengikuti distribusi

multivariat normal serta Homogenitas Kovarians yaitu Matriks kovarians antar kelompok harus homogen.

4.5.4 Uji Korelasi

Uji korelasi adalah metode statistik yang digunakan untuk mengukur sejauh mana hubungan antara dua variabel. Tujuannya adalah untuk menentukan apakah ada korelasi atau hubungan antara variabel-variabel tersebut. Uji korelasi sering menggunakan koefisien korelasi Pearson, tetapi ada juga metode lain seperti Spearman dan Kendall, terutama digunakan ketika data tidak memenuhi asumsi distribusi normal. Uji korelasi digunakan ketika kita ingin mengetahui apakah ada hubungan antara dua variabel dan seberapa kuat hubungan tersebut. Contoh penggunaan dapat ditemui dalam berbagai bidang seperti ilmu sosial, ekonomi, ilmu kedokteran, dan sebagainya. Misalnya, dalam ilmu sosial, uji korelasi dapat digunakan untuk menentukan apakah ada hubungan antara tingkat pendidikan dan pendapatan individu.

A. Korelasi Pearson

Syarat-syarat uji korelasi pearson adalah

1. Data Berskala Interval atau Rasio
Korelasi Pearson sebaiknya digunakan pada data yang berskala interval atau rasio.
2. Distribusi Normal
Data dari setiap variabel harus mengikuti distribusi normal. Korelasi Pearson cukup sensitif terhadap pelanggaran asumsi normalitas, tetapi ketika ukuran sampel cukup besar, pelanggaran ini mungkin tidak signifikan.
3. Linearitas

Korelasi Pearson mengukur hubungan linear antara dua variabel. Jadi, jika hubungan antara variabel tidak bersifat linier, metode lain seperti Spearman dapat lebih sesuai.

Langkah-langkah melakukan uji korelasi pearson adalah sebagai berikut.

1. Menyusun hipotesis nol (H_0) dan hipotesis alternatif (H_1).
2. Menentukan tingkat signifikansi (α).
3. Mengumpulkan data dari dua variabel.
4. Menghitung koefisien korelasi Pearson menggunakan rumus yang sesuai.
5. Menghitung nilai statistik uji (t atau z , tergantung pada ukuran sampel) dan nilai p -value.
6. Membandingkan nilai p -value dengan tingkat signifikansi yang ditentukan.
7. Menarik kesimpulan berdasarkan hasil uji statistik.

B. Korelasi Spearman dan Korelasi Kendall

Syarat-syarat uji korelasi spearman dan kendal adalah sebagai berikut.

1. Data Berskala Ordinal atau Terskala
Baik Korelasi Spearman maupun Korelasi Kendall dapat digunakan pada data yang berskala ordinal atau terskala.
2. Tidak Bergantung pada Distribusi Normal
Tidak memiliki asumsi tentang distribusi normal, sehingga lebih tahan terhadap data yang tidak terdistribusi normal.

Langkah-langkah melakukan uji korelasi spearman dan korelasi kendall.

1. Menyusun hipotesis nol (H_0) dan hipotesis alternatif (H_1).
2. Menentukan tingkat signifikansi (α).

3. Mengumpulkan data dari dua variabel.
4. Menghitung koefisien korelasi Spearman atau Kendall menggunakan rumus yang sesuai.
5. Menghitung nilai statistik uji (z untuk Korelasi Spearman, τ -b atau τ -c untuk Korelasi Kendall) dan nilai p -value.
6. Membandingkan nilai p -value dengan tingkat signifikansi yang ditentukan.
7. Menarik kesimpulan berdasarkan hasil uji statistik.

4.5.5 Uji Regresi

Uji regresi adalah metode statistik yang digunakan untuk memeriksa hubungan antara satu atau lebih variabel independen (biasanya disebut sebagai variabel prediktor) dengan variabel dependen. Tujuannya adalah untuk menilai sejauh mana variabel independen dapat digunakan untuk memprediksi atau menjelaskan variasi dalam variabel dependen. Uji regresi sering digunakan dalam analisis data untuk mengevaluasi hubungan antara variabel-variabel tersebut. Uji regresi dapat diterapkan dalam berbagai bidang, termasuk ekonomi, ilmu sosial, ilmu kesehatan, dan lainnya, untuk menjawab pertanyaan tentang hubungan antara variabel-variabel yang relevan.

Penggunaan Uji Regresi diantaranya adalah sebagai berikut.

1. **Prediksi**
Uji regresi digunakan untuk membangun model prediktif yang memungkinkan kita memprediksi nilai variabel dependen berdasarkan nilai variabel independen.
2. **Eksplanasi Hubungan**
Uji regresi membantu menjelaskan hubungan antara variabel-variabel tersebut. Misalnya, dapat memberikan informasi tentang sejauh mana perubahan dalam variabel independen terkait dengan perubahan dalam variabel dependen.

3. Evaluasi Pengaruh Variabel

Digunakan untuk menilai seberapa besar pengaruh masing-masing variabel independen terhadap variabel dependen. Ini dapat membantu dalam memahami kontribusi relatif dari setiap variabel.

4. Identifikasi Outlier

Uji regresi dapat membantu mengidentifikasi pencilan (outlier) atau observasi yang berada jauh dari pola umum, yang dapat mempengaruhi hasil analisis.

5. Validasi Model

Berguna untuk menguji validitas model regresi, termasuk kecocokan model terhadap data dan asumsi-asumsi yang terkait dengan analisis regresi.

Langkah-langkah Uji Regresi adalah sebagai berikut.

1. Menyusun Hipotesis

Menentukan hipotesis nol (H_0) dan hipotesis alternatif (H_1) yang berkaitan dengan hubungan antara variabel independen dan variabel dependen.

2. Pemilihan Model

Memilih model regresi yang paling sesuai untuk data. Ini dapat mencakup model linier atau non-linier, bergantung pada karakteristik data.

3. Mengumpulkan Data

Mengumpulkan data dari variabel independen dan variabel dependen yang ingin dianalisis.

4. Estimasi Parameter

Menghitung koefisien regresi untuk menilai sejauh mana variabel independen memprediksi variabel dependen.

5. Uji Signifikansi

Melakukan uji signifikansi seperti uji F dan uji t untuk menentukan apakah hubungan antara variabel independen dan variabel dependen signifikan secara statistik.

6. Evaluasi Asumsi

Mengevaluasi asumsi-asumsi yang terkait dengan analisis regresi, seperti homoskedastisitas, normalitas residual, dan independensi residual.

7. Interpretasi Hasil

Menafsirkan hasil analisis, termasuk koefisien regresi, tingkat signifikansi, dan kecocokan model terhadap data.

Ada beberapa jenis uji regresi yang dapat digunakan tergantung pada karakteristik data dan hubungan antar variabel yang diamati. Setiap jenis uji regresi memiliki kegunaan dan aplikasi khususnya sendiri. Pemilihan jenis regresi harus didasarkan pada karakteristik data, tujuan analisis, dan asumsi-asumsi yang dapat terpenuhi.

Berikut adalah beberapa jenis uji regresi dan kegunaannya adalah sebagai berikut.

1. Regresi Linear Sederhana

Kegunaan Menilai hubungan linier antara satu variabel independen dan satu variabel dependen serta Memprediksi nilai variabel dependen berdasarkan nilai variabel independen.

2. Regresi Linear Berganda

Menilai hubungan linier antara dua atau lebih variabel independen dan satu variabel dependen serta Mengidentifikasi kontribusi relatif masing-masing variabel independen terhadap variabel dependen.

3. Regresi Logistik

Digunakan ketika variabel dependen bersifat biner (dua kategori). Mengukur hubungan antara variabel independen dan probabilitas kejadian suatu kejadian.

4. Regresi Polinomial

Menilai hubungan antara variabel independen dan variabel dependen yang tidak linier. Dapat digunakan untuk memodelkan hubungan yang lebih kompleks daripada regresi linear.

5. Regresi Ridge dan Lasso

Digunakan dalam regresi berganda untuk mengatasi masalah multikolinearitas. Membantu memilih variabel yang paling berpengaruh dan mengontrol overfitting.

6. Regresi Nonparametrik

Digunakan ketika asumsi-asumsi regresi parametrik tidak terpenuhi. Tidak mengasumsikan distribusi khusus untuk variabel dependen dan independen.

7. Regresi Bayesian

Memodelkan distribusi probabilitas dari parameter regresi menggunakan pendekatan Bayesian. Memperhitungkan ketidakpastian dalam estimasi parameter.

8. Regresi Time Series

Digunakan untuk menganalisis hubungan antara variabel dependen dan waktu. Memprediksi nilai-nilai masa depan berdasarkan tren dan pola waktu.

9. Regresi Kuantil

Digunakan untuk memahami perbedaan dalam efek variabel independen pada berbagai kuantil distribusi variabel dependen. Mempelajari variasi efek dalam kondisi yang berbeda.

10. Regresi Robust

Mengatasi masalah outlier dan gangguan dalam data. Memberikan estimasi parameter yang lebih stabil dalam kasus-kasus di mana asumsi-asumsi regresi klasik tidak terpenuhi.

DAFTAR PUSTAKA

- Fitriatien, S. R. (2017) 'Pengantar Statistika Untuk Penelitian', *Jurnal Buana Pendidikan*, (23), pp. 49–55.
- Kawulur, G. E. (2022) 'Hubungan Program Online Onsite Dengan Kepuasan Masyarakat Pada Badan Penyelenggara Jaminan Sosial Ketenagakerjaan Cabang Tondano', *Jurnal Administrasi Publik*, VIII(1), pp. 33–40.
- Mardhotillah, B., Asyhar, R. and Elisa, E. (2022) 'Philosophy of Applied Statistical Science in the Era of Smart Society 5.0', *Multi Proximity: Jurnal Statistika Universitas Jambi*, 1(1), pp. 57–70. Available at: <https://mail.online-journal.unja.ac.id/multiproximity/article/view/22441>.
- Marpaung, J. L., Sutrisno, A. and Lumintang, R. (2017) 'Penerapan Metode Anova Untuk Analisis Sifat Mekanik Komposit Serabut Kelapa', *Jurnal Online Poros Teknik Mesin*, 6(2), pp. 151–162.
- Martias, L. D. (2021) 'Statistika Deskriptif Sebagai Kumpulan Informasi', *Fihris: Jurnal Ilmu Perpustakaan dan Informasi*, 16(1), p. 40. doi: 10.14421/fhrs.2021.161.40-59.
- Sholikhah, A. (2016) 'Statistik Deskriptif dalam Penelitian Kualitatif', *Komunika*, 10(2), pp. 342–362.
- Susilawati, T. and Dharmawansyah, D. (2019) 'Metode Penerapan Keselamatan Dan Kesehatan Kerja Terhadap Kinerja Proyek Konstruksi (Studi Kasus Proyek Pembangkit Listrik Tenaga Mesin Gas Sumbawa)', *Jurnal TAMBORA*, 3(3), pp. 107–114. doi: 10.36761/jt.v3i3.403.
- Usmadi, U. (2020) 'Pengujian Persyaratan Analisis (Uji Homogenitas Dan Uji Normalitas)', *Inovasi Pendidikan*, 7(1), pp. 50–62. doi: 10.31869/ip.v7i1.2281.

BAB 5

MACAM-MACAM DATA

Oleh Nur Al-faida

5.1 Pendahuluan

Bab ini akan membahas beberapa macam data, antara lain berdasarkan sumber, cara perolehan, macam-macam kategorisasi data, kategori-kategori data berdasarkan sifatnya, dan jenis-jenis data tergantung pada waktu pengumpulannya.

Data penelitian yang diperoleh harus memenuhi persyaratan tertentu, antara lain validitas, ketergantungan, dan objektivitas. Apabila terdapat tingkat kebenaran atau kesesuaian antara data yang diperoleh peneliti dengan data sebenarnya tentang hal yang diteliti, maka data tersebut dianggap sah. Karena mungkin sulit untuk hanya mendapatkan data yang valid untuk penelitian, uji reliabilitas dan objektivitas dilakukan pada data yang telah diperoleh sebelum validitas ditetapkan. Data yang dapat diandalkan dan tidak memihak sering kali valid. Namun, data yang valid harus dapat dipercaya dan tidak memihak.

Data digambarkan sebagai representasi global aktual, diwakili oleh angka, huruf, simbol, teks, gambar, suara, atau campuran semuanya, yang dicatat dalam kaitannya dengan suatu item, seperti orang, hewan, peristiwa, konsep, keadaan, dan seterusnya. Dengan kata lain, data adalah aktualitas yang mencirikan kejadian tunggal dan spesifik. Data adalah materi atau hal standar yang harus diolah untuk menghasilkan sesuatu yang

lebih bermakna karena saat ini kurang memiliki makna atau dampak langsung bagi pengguna.

5.2 Jenis Data Menurut Cara Memperolehnya

Data adalah informasi yang diketahui tentang apa pun; itu mungkin anggapan atau fakta yang diperlukan tentang suatu item. Data merupakan hal yang penting dan sangat vital dalam pembelajaran. Sebab hasil dan analisis data adalah yang pada akhirnya menentukan objektivitas penelitian. Dalam buku Penelitian Analisis Data Menggunakan Statistika tahun 2013, Misbahuddin dan Iqbal Hasan menyatakan bahwa data dapat dipisahkan menjadi dua kategori: data primer dan data sekunder, tergantung cara memperolehnya.

5.2.1 Data Primer

Data primer adalah data yang hanya dikumpulkan atau diperoleh peneliti pada saat berada di lapangan. Data asli adalah istilah umum lainnya untuk informasi semacam ini.

Data primer harus diproses ulang karena seringkali tersedia dalam bentuk yang paling mentah. Namun karena mereka akan meneliti sumber primer untuk data yang mereka inginkan, peneliti bisa lebih tepat dalam pengumpulan datanya. Hasilnya, peneliti dapat memodifikasi pilihan sumbernya untuk mendapatkan informasi yang mereka butuhkan.

Sifat-sifat data primer adalah sebagai berikut, diambil dari buku Statistika Pendidikan karya Sri Rizqi Wahyuningrum (2020):

- a. Para peneliti mengumpulkan dan menangani datanya sendiri.
- b. Langsung dari sumbernya (*first-hand source*).
- c. Secara eksklusif berasal dari item tersebut.
- d. Dapat diperoleh dari laboratorium atau lapangan.

- e. Data diperoleh tanpa menggunakan perantara apa pun.

Dalam metode penelitian kualitatif seperti studi lapangan, fenomenologi, dan observasi, jenis data ini lebih sering digunakan.

Fungsi Data Primer

Primary data is needed to provide explanations about certain problems. The following are the functions of primary data:

- a. Memberikan tanggapan langsung dari objek penelitian terhadap rumusan masalah penelitian.
- b. Sebagai bahan kajian untuk mengatasi suatu permasalahan.
- c. Sebagai landasan bagi penelitian yang bertanggung jawab.
- d. Mendapatkan secara langsung data yang akurat dan terkini dari pokok bahasan.
- e. Membantu menghindari kesalahan yang dilakukan peneliti.

Contoh-contoh dari data primer juga harus menjadi bahan pertimbangan peneliti guna meningkatkan pemahamannya terhadapnya. Berikut ilustrasinya:

- a. **Penelitian Pasar**

Diambil dari buku Metodologi Penelitian Bisnis, Konsep dan Aplikasi Juliandi, A., & Manurung, S. (2014). Agar dapat berfungsi dengan baik, perusahaan harus melakukan penelitiannya sendiri. Namun, perusahaan dapat mengumpulkan informasi langsung tentang dinamika pasar. Misalnya, ketika sebuah bisnis memperkenalkan produk ponsel pintar, bisnis tersebut memiliki kemampuan langsung untuk mengumpulkan data dari banyak pelanggan. Informasi tertentu yang dapat diperoleh dari penelitian ini mencakup fitur-fitur ponsel cerdas yang disukai konsumen dan daya belinya.

b. Penelitian Akademisi

Peneliti juga dapat mencari dan mengumpulkan data dari sumber asli ketika melakukan penelitian akademis. Jenis yang diminta dapat dimodifikasi agar sesuai dengan subjek penelitian. Misalnya, ketika seorang peneliti ingin mengetahui seberapa stres mahasiswa semester lima, maka peneliti dapat menanyakannya sendiri kepada mahasiswa tersebut. Pengumpulan data untuk hal ini dapat dilakukan dengan memberikan kuesioner. Sebagai contoh tambahan, survei dapat digunakan untuk mengetahui seberapa besar pengetahuan generasi muda tentang fenomena perubahan iklim. Tentu saja generasi muda yang menjadi subjek penelitian harus menjadi fokus survei.

c. Cara Mendapatkan Data Primer

Memperoleh data primer dapat dilakukan dengan beberapa cara. Kriyantono (2010) menyatakan ada beberapa pendekatan yang bisa diterapkan:

d. Wawancara

Mewawancarai orang adalah metode pengumpulan informasi. Wawancara adalah suatu teknik pengumpulan data yang melibatkan narasumber dan pewawancara. Ada beberapa cara untuk melakukan wawancara, antara lain secara langsung, melalui telepon, melalui pesan WhatsApp, dan cara lainnya. Salah satu metode terbaik untuk mendapatkan data yang dibutuhkan adalah dengan melakukan wawancara. Namun, kemampuan peneliti dalam mengajukan pertanyaan menentukan berhasil atau tidaknya suatu wawancara. Oleh karena itu, untuk mendapatkan wawasan yang komprehensif dari sumber, peneliti perlu menyusun pertanyaan mereka dengan cermat. Wawancara mempunyai manfaat karena memungkinkan pengumpulan informasi yang mendalam. Sebaliknya, kelemahannya adalah

wawancara membutuhkan waktu lebih lama dan perlu fleksibilitas penjadwalan antara narasumber dan peneliti.

e. Survei

Dengan menggunakan survei, seseorang dapat mengumpulkan informasi dari audiens target dan mempelajari preferensi, sudut pandang, keputusan, dan komentar mereka mengenai suatu subjek. Survei online adalah pendekatan populer yang digunakan oleh akademisi saat ini karena kemudahan penggunaannya. Survei online memungkinkan peneliti mengumpulkan sejumlah besar data sekaligus. Peneliti mungkin menanyakan serangkaian pertanyaan kepada responden dalam survei. Penelitian dapat dilakukan dengan menggunakan data-data yang telah dikumpulkan sebagai landasan.

f. *Focus Group Discussion*

Dengan metode pengumpulan data ini, ukuran sampel yang kecil biasanya tidak lebih dari 6 atau 10 digunakan. Orang-orang yang tertarik dengan topik penelitian dikumpulkan dalam kelompok ini. Moderator memfasilitasi percakapan antar peserta dalam kelompok fokus sehingga diperoleh pemahaman yang lebih komprehensif. Sambil mendengarkan percakapan, peneliti akan membuat catatan mengenai ide-ide penting yang akan dijadikan data penelitian.

5.2.2 Data Sekunder

Menurut buku Kajian Matematika Sekolah Dasar (2019) karya Erna Yayuk dan Suko Prasetyo, data sekunder merupakan informasi yang dikumpulkan dari sumber lain. Dengan kata lain, informasi semacam ini dikumpulkan melalui penelitian berdasarkan materi yang diterbitkan sebelumnya. Laporan penelitian, misalnya. Khususnya mengenai atribut data sekunde:

- a. Data yang diperoleh disiapkan, dikumpulkan, dan diolah oleh pihak lain;
- b. Biasanya, bidang penelitian yang relevan akan mengintegrasikan data.
- c. Informasi ini berasal dari sumber sekunder.
- d. Sumber data tidak terbatas pada lapangan.

Karena sifat deskriptif data sekunder, maka data ini lebih sering digunakan dalam penelitian kuantitatif. Oleh karena itu, para ilmuwan hanya perlu menangani dan menampilkan data yang sesuai dengan kebutuhan penelitian mereka.

Fungsi Data Sekunder

Data sekunder memiliki berbagai fungsi dan manfaat bagi peneliti. Inilah beberapa di antaranya:

- a. Meningkatkan Pemahaman Konteks.
Melalui penggunaan data yang ada, peneliti dapat mengumpulkan rincian yang diperlukan untuk memperoleh pemahaman yang lebih mendalam tentang konteks seputar permasalahan mereka, dengan menggunakan data kualitatif dan kuantitatif.
- b. Menguji Hipotesis
Peneliti mungkin menguji korelasi antar variabel tertentu dengan menggunakan data sekunder untuk mendukung teori mereka.
- c. Melakukan Perbandingan
Para sarjana mungkin melihat variasi dan persamaan dalam fenomena yang mereka kaji dengan membandingkan informasi dari berbagai sumber.
- d. Mengidentifikasi Tren dan Pola
Para sarjana dapat menemukan pola, tren, atau modifikasi data dengan menggunakan data sekunder.

e. Mendukung Pengambilan Keputusan

Peneliti dapat mengambil kesimpulan berdasarkan informasi yang akurat dan dapat dipercaya dengan cara mengevaluasi data yang telah tersedia.

f. Memicu Riset Lanjutan

Inovasi dan penelitian lebih lanjut mungkin mendapatkan inspirasi dari data sekunder. Karena para sarjana mampu mengidentifikasi bidang pengetahuan yang belum lengkap atau kemungkinan penelitian yang belum tereksplorasi.

Para sarjana sering kali memanfaatkan berbagai sumber data sekunder, termasuk publikasi ilmiah, catatan perusahaan, dan data survei. Berikut adalah beberapa contoh data sekunder:

a. Data Sensus Penduduk

informasi demografis yang dikumpulkan dari survei populasi sebelumnya yang dilakukan oleh organisasi pemerintah, termasuk populasi, usia, jenis kelamin, pekerjaan, dan pencapaian pendidikan.

b. Data Keuangan dan Laporan Perusahaan

Laporan keuangan atau laporan tahunan yang diterbitkan oleh Perseroan.

c. Data Publikasi Ilmiah

Artikel jurnal, tesis, disertasi, dan publikasi lain yang memuat informasi ilmiah.

d. Data Administrasi Pemerintah

Informasi dari lembaga pemerintah, termasuk statistik demografi, kesehatan, dan ekonomi serta informasi lain yang dikumpulkan untuk keperluan administratif.

e. Data Riset Pasar

Informasi dari riset pasar terkait analisis pasar, seperti preferensi pelanggan, tren pasar, atau fakta lainnya.

f. Data Arsip dan Koleksi

Informasi disimpan dalam arsip oleh perusahaan seperti perpustakaan, museum, dan pusat kebudayaan lainnya. Koleksi sejarah, informasi arkeologi, catatan sejarah, gambar, dan sumber referensi lainnya termasuk dalam kategori ini.

g. Data Lingkungan

Informasi tentang lingkungan yang dikumpulkan oleh kelompok nirlaba atau lembaga penelitian lingkungan, seperti kualitas udara, dampak lingkungan, dan sebagainya.

h. Data Pendidikan

Informasi yang dikumpulkan oleh lembaga pendidikan atau organisasi statistik, seperti tingkat kelulusan, tingkat keterlibatan siswa, dan statistik pendidikan.

i. Data Ekonomi

Informasi ekonomi yang dipublikasikan oleh organisasi atau lembaga pemerintah yang melakukan penelitian ekonomi, seperti statistik ekonomi nasional, ketenagakerjaan, inflasi, dan data perdagangan.

j. Data Teknologi Informasi

Statistik penggunaan teknologi informasi, seperti yang berasal dari organisasi riset mengenai media sosial, eCommerce, dan internet.

5.3 Klasifikasi Data Berdasarkan Sifatnya

Penting untuk memahami jenis data yang diproses serta skala pengukuran data yang digunakan. Data dapat dikategorikan menjadi kuantitatif, kualitatif, dan/atau kategori berdasarkan sifatnya. Kontinuitas juga dapat digunakan untuk memisahkan lebih lanjut data kuantitatif ke dalam kategori diskrit dan kontinu.

Data dipisahkan menjadi 4 (empat) tingkatan pengukuran berdasarkan skala pengukurannya: data nominal, ordinal, interval, dan rasio. Meskipun data interval dan rasio berasal dari hasil pengukuran sehingga dianggap sebagai data kuantitatif, sedangkan data nominal dan ordinal diklasifikasikan sebagai data kualitatif.

5.3.1 Data Kuantitatif

Sugiyono (2018) mengartikan data kuantitatif sebagai metodologi penelitian positivis (disebut juga data nyata). Merupakan data penelitian yang berupa angka-angka yang relevan dengan permasalahan yang diteliti yang akan diukur dengan menggunakan statistik sebagai instrumen tes kualifikasi untuk menarik suatu kesimpulan. Dalam penelitian kuantitatif, ada dua kategori data;

1. Data Diskrit

Jika data hanya sebatas menghitung angka, maka data tersebut tidak dapat dikuantifikasi dan disebut sebagai data diskrit dalam penelitian kuantitatif. Data ini mencakup ukuran-ukuran yang tidak termasuk dalam kategori data diskrit, misalnya berat badan, panjang badan, dan tinggi badan.

2. Data Kontinu

Tipe data yang dikenal sebagai data kontinu menggunakan nilai numerik yang dapat dibagi menjadi beberapa bagian yang lebih kecil. Data kontinyu dapat dimasukkan ke dalam skala pengukuran, berbeda dengan data diskrit yang tidak dapat dikuantifikasi. Misalnya panjang, durasi, berat, dan lain-lain.

5.3.2 Data Kualitatif

Sugiyono (2015) mendefinisikan data kualitatif sebagai informasi yang berbentuk kata, diagram, atau gambar. Dalam penelitian kualitatif, ada dua kategori data;

1. Data nominal

Data nominal, terkadang disebut sebagai skala nominal dalam statistik, adalah klasifikasi variabel yang tidak memiliki nilai kuantitatif. Kadang-kadang dikenal sebagai data yang diidentifikasi atau diberi label.

Kata nominal digunakan untuk mengidentifikasi atau merangkul variabel tanpa memberikan nilai kuantitatif karena berasal dari tata nama Latin “Nomen” yang berarti nama. Dalam kasus tertentu, ketika data nominal mengasumsikan nilai kuantitatif, hal ini tidak benar. Namun, tidak ada sifat numerik yang terkait dengan nilai kuantitatif ini.

2. Data Ordinal

Contoh data kualitatif dengan kategori variabel yang teratur dan alami serta jarak antar kategori yang tidak dapat ditentukan disebut data ordinal. Berbagai macam variabel ordinal inilah yang disebut dengan data ordinal. Misalnya informasi yang diperoleh dengan menggunakan skala likert untuk mengajukan pertanyaan bersifat ordinal.

Data ordinal terkadang dikategorikan sebagai bentuk data kuantitatif atau berada dalam ruang kualitatif-kuantitatif perantara. Alasannya adalah karena data ordinal menunjukkan ciri-ciri kuantitatif dan kualitatif.

5.4 Jenis Data Menurut Waktu Pengumpulannya

Kategori informasi Data dapat dikategorikan ke dalam dua kelompok berdasarkan waktu pengumpulannya: data cross-section dan data time series. Karena data cross-sectional dikumpulkan

sekaligus, isi variabel awal hanya terdiri dari satu tanggal. Secara umum, terdapat banyak variasi data cross section. Data deret waktu terdiri dari informasi yang dikumpulkan selama periode waktu yang telah ditentukan, misalnya satu tahun. Seringkali hanya ada dua variabel dalam data deret waktu: waktu dan data. karena faktanya kami akan menekankan arah waktu yang lebih banyak.

5.4.1 Data Cross Section

Data cross-sectional merupakan suatu penelitian yang hanya dilakukan satu kali pengukuran atau observasi untuk memperoleh data (Suprajitno, 2013).

Studi cross-sectional pada hakikatnya adalah suatu metodologi penelitian observasional yang mengkaji sekumpulan data variabel penelitian yang dikumpulkan pada waktu tertentu dari semua jenis populasi dan sampel. Oleh karena itu, metodologi ini umumnya dikenal menggunakan subset yang dipengaruhi sebelumnya.

Namun, hal yang tidak dapat diubah adalah studi cross-section mendokumentasikan data dari suatu populasi tanpa menyesuaikan variabel apa pun. Hasilnya, studi-studi ini dapat digunakan untuk mengkarakterisasi ciri-ciri komunitas tanpa menetapkan hubungan sebab-akibat antara variabel-variabel yang tidak setara.

Survei adalah strategi yang sering digunakan untuk memperoleh data lintas sektor. Ada tiga cara untuk melakukan survei: melalui telepon, online, atau melalui wawancara langsung. Beberapa responden yang mewakili demografi yang ingin Anda pantau dapat berpartisipasi dalam jajak pendapat ini. Analisis statistik dapat digunakan untuk menemukan tren dan hubungan antara variabel yang dapat diamati setelah data survei dikumpulkan.

Selain survei, sumber sekunder mencakup jurnal ilmiah, makalah penelitian, atau data yang dikumpulkan oleh badan pemerintah atau organisasi lain dan juga dapat menyediakan data lintas bagian. Data ini berguna untuk analisis dan pengambilan keputusan karena sering kali dapat diakses oleh masyarakat umum.

Data survei kepuasan pelanggan restoran merupakan salah satu jenis data cross-section. Pada saat yang sama, data ini mencatat balasan pelanggan terhadap layanan restoran. Tingkat kepuasan, kualitas layanan, biaya, dan faktor lainnya mungkin diamati.

Data dari sensus penduduk suatu negara merupakan contoh lain. Data ini menangkap rincian demografi pada suatu waktu tertentu, termasuk usia, jenis kelamin, pekerjaan, dan pendidikan.

5.4.2. Data *Time Series*/ Berkala

Deret waktu adalah kumpulan titik data untuk suatu variabel yang sering dicatat pada periode waktu tertentu (Robinson & Sciences, 2020). Deret waktu menurut pendapat para ahli lainnya adalah seluruh nilai dalam rentang skala yang sama yang disusun secara kronologis sebagai suatu indeks (Brillinger, 2015).

Nampaknya secara praktis setiap kegiatan didasarkan pada konsep deret waktu. Untuk komunikasi, deskripsi, dan visualisasi, deret waktu sering kali digunakan dengan cara yang sama. Karena sifat fisik waktu, model matematika deret waktu seperti parameter karakteristik dapat diterapkan secara praktis (Brillinger, 2015).

Analisis deret waktu, sering juga dikenal sebagai analisis pola data, adalah tindakan menafsirkan data deret waktu dan memproyeksikan nilai berdasarkan pola data tersebut (Robinson & Sciences, 2020). Dengan menerapkan alat analisis deret waktu pada deret waktu, seseorang dapat lebih memahami data dan

polanya. Menurut Brillinger (2015), tujuan terdiri dari delapan ringkasan, penilaian, deskripsi, dan prediksi. Di antara banyak tujuan analisis deret waktu adalah sebagai berikut (Robinson & Sciences, 2020):

1. Menelaah deskripsi data primer, yang sering dilakukan setelah pembuatan diagram data.
2. Dengan menggunakan hubungan antar variabel dengan variabel lain sebagai pedoman, jelaskan isi datanya.
3. Peramalan, atau prediksi, seperti yang biasa diketahui.
4. Pemantauan data biasanya digunakan untuk pengendalian kualitas data.

Data deret waktu dapat diperoleh dengan berbagai metode. Sumber utamanya adalah organisasi pemerintah yang mengumpulkan dan menyebarkan statistik ekonomi, termasuk angka produksi industri, inflasi, dan tingkat kemiskinan. Sumber reguler data runtun waktu mencakup Bank Sentral, Kementerian Keuangan, dan Badan Pusat Statistik.

Selain itu, banyak bisnis menyimpan data internal mereka sendiri, termasuk angka penjualan harian atau informasi penggunaan produk. Dalam situasi ini, bisnis mungkin menggunakan data tersebut untuk memeriksa tren dan pola operasional.

Data harga saham harian adalah contoh umum dari data deret waktu. Harga penutupan saham setiap hari perdagangan didokumentasikan dalam data ini. Harga saham merupakan variabel yang diamati dalam data ini, sedangkan waktu merupakan variabel bebas.

Data cuaca harian adalah contoh tambahan. Setiap hari, curah hujan, kelembaban, dan suhu dicatat dalam data ini. Waktu merupakan variabel bebas dan variabel yang diamati adalah suhu, kelembaban, dan curah hujan.

DAFTAR PUSTAKA

- Brillinger, D. R. (2015). Time Series: General. In International Encyclopedia of the Social & Behavioral Sciences: Second Edition (Second Edi, Vol. 23). Elsevier. <https://doi.org/10.1016/B978-0-08-097086-8.42084-2>
- (2015). Time Series: General. In International Encyclopedia of the Social & Behavioral Sciences: Second Edition (Second Edi, Vol. 23). Elsevier. <https://doi.org/10.1016/B978-0-08-097086-8.42084-2>
- Juliandi, A., & Manurung, S. (2014). Metodologi Penelitian Bisnis, Konsep dan Aplikasi: Sukses Menulis Skripsi & Tesis Mandiri. Umsu Press.
- Robinson, NJ et al. (2020): Temperature, salinity and acoustic backscatter observations and tidal model output in McMurdo Sound, Antarctica in 2016 and 2017. *PANGAEA*.
- (2020) :Temperature, salinity and acoustic backscatter observations and tidal model output in McMurdo Sound, Antarctica in 2016 and 2017. *PANGAEA*
- Kriyantono, Rachmat. 2010. Teknik Praktis Riset Komunikasi: Disertai Contoh Praktis Riset Media, Public Relations, Advertising, Komunikasi Organisasi, Komunikasi Pemasaran. Jakarta: Kencana.
- Misbahudin, Iqbal Hasan, (2013), Analisis Data Penelitian Dengan Statistik, Jakarta, Bumi Aksara.
- Sugiyono. 2015. Metode Penelitian Kuantitatif, Kualitatif, dan R&D. Bandung : ALFABETA.
- Sugiyono. 2018. Metode Penelitian Kuantitatif, Kualitatif, dan R&D, penerbit Alfabeta, Bandung
- Suprajitno. 2013. Asuhan Keperawatan Keluarga Aplikasi dalam Praktik. Jakarta : EGC.

- Wahyuningrum, Sri Rizqi. 2020. Statistika Pendidikan (Konsep Data dan Peluang). Surabaya: CV. Jakad Media Publishing.
- Yayuk, Erna. 2019. Pembelajaran Matematika SD. Malang: Universitas Muhammadiyah Malang.

BAB 6

SKALA UKURAN

Oleh Ardiana Fatma Dewi

6.1 Skala Pengukuran

Skala pengukuran data merupakan cara untuk menentukan ukuran dari suatu variabel berdasarkan jenis data yang terkait dengan penelitian. Penggunaan alat ukur atau satuan tertentu untuk merepresentasikan kualitas atau kuantitas dari suatu variabel merupakan bagian penting dari proses pengukuran. Skala pengukuran ini membantu peneliti atau analis statistik untuk memahami sifat data yang mereka hadapi dan menentukan jenis analisis statistik yang paling sesuai untuk digunakan (Nazir, 2020). Penting untuk diingat bahwa setiap jenis skala memiliki kelebihan dan kelemahan tersendiri. Oleh karena itu, sebagai peneliti harus memilih skala pengukuran yang paling sesuai dengan tujuan penelitiannya serta karakteristik dari objek-objek yang akan diukur. Macam-macam skala data dapat dibagi menjadi skala kategorik dan numerik, dimana kategorik terdiri dari skala nominal dan ordinal, serta numerik terdiri dari skala interval, dan rasio (Juliandi & Manurung, 2014).

6.2 Skala Nominal

Skala nominal dapat dikatakan sebagai jenis skala pengukuran yang paling rendah di antara jenis-jenis skala pengukuran lainnya. Skala nominal merupakan skala data yang diperoleh melalui klasifikasi (pengelompokan) yakni

dikelompokkan dengan ciri-ciri yang sama, lalu masing masing kelompok diberi symbol sebagai pembeda agar dapat diolah dalam statistic (Purnomo. 2020). Skala ini hanya dapat diklasifikasikan ke dalam kategori atau kelompok tanpa adanya urutan atau nilai-nilai yang dapat dibandingkan. Biasanya digunakan untuk mengklasifikasikan objek, individu, atau kelompok dalam bentuk kategori. Angka-angka atau symbol tersebut tidak memiliki nilai intrinsik sehingga operasi matematika standart seperti perkalian, pembagian, penjumlahan dan pengurangan pada skala ini tidak bisa dilakukan. Peralatan statistik yang dapat digunakan untuk mengolah data dengan skala nominal adalah peralatan statistik yang berbasis jumlah dan proporsi seperti modus, distribusi frekuensi chi square dan beberapa alat statistik non-parametrik lainnya.

Pemberian angka atau simbol pada skala nominal hanya berfungsi sebagai label kategori atau grup dan tidak memiliki nilai intrinsik. Contoh: Jenis kelamin, misalnya kita memberikan kode laki-laki=0 dan perempuan=1, angka ini tidak memiliki nilai intrinsik dan artinya tidak bisa diartikan bahwa perempuan lebih banyak daripada laki-laki. Contoh lain dari skala nominal adalah agama misal kita memberikan kode Islam=1, Kristen=2, Katolik=3, Hindu=4, Budha=5, angka ini juga tidak memiliki nilai intrinsik dan tidak bisa diartikan bahwa islam adalah agama nomor 1, karena angka disini hanyalah symbol saja. Angka tersebut tidak memiliki nilai yang berarti sehingga tidak dapat dilakukan operasi matematika.

6.3 Skala Ordinal

Skala ordinal adalah jenis pengukuran yang digunakan untuk mengurutkan atau menyusun objek berdasarkan karakteristik tertentu tanpa memberikan informasi tentang jarak atau

perbedaan antara nilai-nilai. Dengan kata lain, dalam skala ordinal, kita mengetahui bahwa satu objek memiliki nilai yang lebih tinggi atau lebih rendah dari objek lainnya, tetapi kita tidak tahu seberapa jauh perbedaan antara nilai-nilai tersebut. Dalam skala ordinal, hanya urutan atau peringkat yang relevan, sedangkan jarak antara nilai-nilai tidak memiliki interpretasi yang konsisten atau bermakna. Skala ordinal memiliki tingkatan yang lebih tinggi daripada skala nominal. Karena selain menunjukkan perbedaan antara objek, juga menunjukkan posisi relatif mereka dalam urutan yang telah ditetapkan. Sehingga, sama halnya seperti skala nominal, skala ini juga tidak dapat dilakukan operasi matematika standart yaitu perkalian, pembagian, penjumlahan dan pengurangan.

Contoh umum dari skala ordinal adalah rating film, seperti memberi peringkat film dari yang paling tidak disukai hingga yang paling disukai dengan menggunakan angka atau simbol. Misalnya, sebuah film bisa mendapatkan rating "sangat buruk", "buruk", "sedang", "baik", atau "sangat baik". Meskipun kita tahu bahwa "sangat baik" lebih tinggi dari "baik", dan sebagainya, tetapi kita tidak memiliki informasi tentang seberapa besar perbedaan antara setiap peringkat tersebut.

6.4 Skala Interval

Skala interval dalam statistik adalah jenis skala data yang memiliki sifat sama seperti skala nominal dan ordinal, tetapi dengan tambahan bahwa interval pada nilai pengamatan memiliki makna yang konsisten. Pada skala interval terdapat informasi tentang perbedaan antara nilai, tetapi tidak memberikan informasi rasio antara nilai. Karena di dalam skala interval tidak memiliki nilai nol mutlak sehingga tidak dapat dilakukan rasio antar nilai pengamatan (Jusmiana & Herianto, 2020). Dalam data dengan

skala interval, kita dapat dilakukan operasi matematika yang standart seperti penjumlahan dan pengurangan, tetapi tidak ada nilai nol mutlak yang memiliki arti yang signifikan.

Contoh yang paling umum dari skala interval adalah pengukuran suhu dalam derajat Celsius ($^{\circ}\text{C}$) dan Fahrenheit ($^{\circ}\text{F}$). Dalam kedua skala ini, setiap satuan yang digunakan untuk mengukur suhu memiliki interval yang sama, artinya perbedaan antara 20°C dan 40°C adalah sama dengan perbedaan antara 40°C dan 60°C , hal yang sama berlaku juga untuk Fahrenheit. Meskipun selisihnya sama besar, tetapi bukanlah kelipatan satu sama lain dan tidak bisa dibandingkan karena antara derajat celcius dan fahrenheit memiliki titik beku yang berbeda. Ketika suhu 0° dalam celcius tidak sama artinya dengan 0° dalam Fahrenheit, dimana nol dalam celcius adalah titik beku air sedangkan nol dalam fahreheit merupakan titik beku es, hal ini yang dikatakan bahwa skala interval tidak memiliki nilai nol yang mutlak.

6.4 Skala Rasio

Skala rasio merupakan skala yang paling tinggi tingkatnya diantara skala nominal, ordinal, dan interval (Irianto, 2020). Skala data ini memiliki semua karakteristik yang ada pada ketiga skala lainnya. Data Rasio merupakan data pengukuran dari hasil pengukuran yang dapat diurutkan, dibedakan dan dapat dibandingkan. Selain itu pada skala ini memiliki nilai nol yang mutlak, artinya pada skala ini nilai dasar untuk acuan tidak dapat diubah meskipun menggunakan skala lainnya. Karena skala rasio memiliki nilai nol yang mutlak, maka antar nilai di dalam pengamatan dapat dilakukan perbandingan. Sehingga pada skala data rasio dapat dilakukan semua operasi matematika standart yaitu perkalian, pembagian, penjumlahan dan pengurangan.

Contoh pengukuran dalam skala rasio adalah pengukuran tinggi dan berat. Misalnya berat benda A adalah 10 kg sedangkan benda B adalah 20 kg. Dari hal tersebut dapat dihitung selisih berat benda A dan B adalah 10 kg, serta dapat dikatakan bahwa benda B dua kali lebih berat dibandingkan dengan benda A. Karena kedua benda tersebut memiliki nilai nol yang mutlak artinya ketika 0 kg pada benda A sama dengan 0 kg pada benda B, sehingga kedua benda tersebut dapat dibandingkan. Selain dapat dihitung selisihnya melalui operasi matematika penjumlahan dan pengurangan, dapat dihitung juga rasionya melalui operasi matematika perkalian dan pembagian. Sehingga dapat dikatakan bahwa skala rasio memiliki sifat yang lengkap dibandingkan dengan skala data yang lainnya.

DAFTAR PUSTAKA

- Asari, A., dkk. 2023. *Pengantar Statistika*. PT Mafy Media Literasi Indonesia, Solok Sumatera Barat.
- Irianto, Agus. 2020. *Statistik*. Kencana, Jakarta.
- Nazir, M. P. 2020. *Metode Penelitian*. Ghalia Indonesia, Bogor.
- Nuryadi., dkk. 2017. *Dasar-dasar Statistik Penelitian*. Sibuku Media, Bantul Yogyakarta.
- Juliandi, A., & Manurung, S. 2014. *Metodologi Penelitian Bisnis, Konsep dan Aplikasi: Sukses Menulis Skripsi & Tesis Mandiri*. UMSU Press.
- Jusmiana, A., & Herianto, H. 2020. *Suplemen Materi Statistik Terapan Dalam Ilmu Kesehatan*. Universitas Pejuang Republik Indonesia (UPRI), Makassar.
- Purnomo. 2020. *Bahan Ajar Variabel dan Hipotesis*. Poltekkes, Semarang.
- Sugiyono. 2017. *Metode Penelitian Bisnis*. Alfabeta, Bandung.
- Walpole, R. E. 1995. *Pengantar Statistika* edisi ke-3. Penerbit PT. Gramedia Pustaka Utama, Jakarta.

BAB 7

POPULASI

Oleh Ayatullah Harun

Implementasi penelitian selalu melibatkan objek yang diteliti atau diselidiki. Objek tersebut bisa berupa manusia, hewan, tumbuhan, benda mati lainnya, serta peristiwa dan fenomena yang terjadi di masyarakat atau alam. Dalam pelaksanaan penelitian, terkadang peneliti menginvestigasi keseluruhan objek, tetapi seringkali peneliti hanya memilih sebagian dari objek tersebut. Meskipun hanya mengambil sebagian dari objek yang diteliti, hasilnya dapat mencerminkan atau mewakili keseluruhan objek tersebut. (Notoatmodjo, 2010)

Bab ini membahas tentang pengertian populasi, karakteristik Populasi dan jenis-jenis Populasi. Pentingnya mempelajari populasi sebagai bentuk fokus dalam menentukan penelitian yang akan dilakukan, Populasi juga merupakan batasan responden yang menjadi batasan-batasan peneliti dalam mengembangkan penelitiannya, Populasi dalam sebuah penelitian sangatlah penting mengingat tanpa adanya populasi yang membatasi suatu area penelitian mengakibatkan suatu kesalahan peneliti dalam menentukan objek yang akan diteliti.

Dalam kehidupan sehari-hari, kita sering mengalami situasi yang terkait dengan populasi dan sampel. Misalnya, ketika kita memasak sup, kita akan menambahkan garam yang kemudian diaduk untuk memastikan garam larut dalam sup. Untuk memastikan bahwa sup yang dimasak tidak terlalu asin, kita akan mengambil sesendok kuah sup untuk dicicipi. Jika kita merasa

rasanya sudah cukup asin, kita dapat menyimpulkan bahwa seluruh panci sup sudah cukup asin. Dengan demikian, sesendok sup yang kita ambil untuk dicicipi merupakan sampel, sedangkan seluruh panci sup merupakan populasi (Rizki and Nawangwulan, 2018).

Contoh yang sesuai adalah ketika kita membuat teh manis dengan mencampurkan teh dan gula, kita akan mencicipi sedikit terlebih dahulu dengan sendok. Rasa teh yang kita cicipi dalam sendok tersebut mewakili keseluruhan minuman dalam gelas. Sendok teh yang kita cicipi merupakan sampel, sedangkan gelas teh tersebut merupakan populasi. Jika kita ingin membuat teh lainnya, kita mungkin akan mencicipinya lagi karena kita tidak yakin seberapa manis atau kuat rasanya. Namun, jika kita membuat minuman yang berbeda seperti susu, tentu kita tidak akan mencicipi susu dengan menggunakan teh yang telah ada. Ini menunjukkan bahwa populasi yang berbeda akan menghasilkan sampel yang berbeda pula. (Machfoedz, 2010)

Tidak hanya untuk satu gelas teh, jika kita membuat teh dalam jumlah besar untuk menyajikan kepada tamu, kita juga akan mengambil sampel dengan sendok untuk memastikan rasanya sudah sesuai. Sendok teh yang diambil merupakan sampel, sedangkan teh dalam jumlah besar tersebut adalah populasi. (Machfoedz, 2010)

Contoh-contoh populasi dalam bidang kesehatan merujuk pada keseluruhan objek yang menjadi fokus penelitian. Misalnya, data jumlah kelahiran dan kematian bayi di Indonesia selama periode tertentu, jumlah gedung polindes di suatu kabupaten, jumlah pasangan usia subur, jumlah rumah sakit dalam satu provinsi, jumlah balita yang menderita gizi buruk, jumlah perawat atau bidan, jumlah dokter spesialis anak, jumlah penggunaan alat

kontrasepsi dari tahun ke tahun, serta jumlah ibu hamil. (Sunyoto and Setiawan, 2013).

Beberapa sumber menjelaskan bahwa populasi merupakan area generalisasi yang terdiri dari objek atau subjek yang memiliki kuantitas dan karakteristik khusus yang ditetapkan oleh peneliti untuk dianalisis dan kemudian ditarik kesimpulannya. Misalnya, jika akan melakukan penelitian di Puskesmas X, maka Puskesmas X tersebut memiliki populasi yang bisa terdiri dari jumlah subjek atau individu serta karakteristiknya. Ini mengindikasikan bahwa populasi adalah kumpulan individu atau subjek serta objek yang diamati. Populasi tidak hanya terbatas pada manusia, tetapi juga mencakup objek dan fenomena alam lainnya. Lebih dari sekadar jumlah, populasi melibatkan seluruh karakteristik atau sifat yang dimiliki oleh subjek atau objek yang sedang diteliti. Konsep ini menyoroti bahwa individu di Puskesmas X memiliki beragam karakteristik seperti motivasi kerja, disiplin, kepemimpinan, iklim organisasi, dan sebagainya (Sunyoto and Setiawan, 2013).

Seorang individu dapat dianggap sebagai populasi karena memiliki berbagai karakteristik yang berbeda, seperti gaya bicara, disiplin, hobi, cara bergaul, kepemimpinan, dan lain sebagainya. Dalam konteks kedokteran, sering kali satu individu dijadikan populasi, seperti contohnya darah yang ada pada setiap individu dianggap sebagai populasi. Untuk pemeriksaan, cukup diambil sebagian kecil darah sebagai sampel. Data yang diperoleh dari sampel darah tersebut kemudian dapat diterapkan untuk menyimpulkan tentang seluruh darah yang dimiliki oleh individu tersebut (Sugiyono, 2015).

7.1 Pengertian Populasi

Pentingnya memahami populasi berdasarkan beberapa referensi yang menjelaskan definisi Populasi diantaranya :

1. Menurut Ircham Machfoedz (2010), populasi merujuk pada keseluruhan subjek yang menjadi fokus penelitian. Subjek tersebut dapat berupa benda, dan semua benda yang memiliki sifat atau ciri tertentu menjadi subjek yang diteliti..
2. Menurut Soenarto (1987) Populasi adalah kumpulan manusia, bangunan, hewan, dan entitas lainnya yang minimal memiliki satu ciri atau karakteristik tertentu (Susila and Suryanto, 2014)
3. Populasi merujuk pada keseluruhan objek atau himpunan objek yang memiliki ciri yang serupa. Ini bisa terdiri dari individu, kelompok individu, organisasi, perusahaan, benda (baik hidup maupun mati), kejadian, kasus, waktu, atau tempat dengan sifat atau ciri yang serupa (Rizki and Nawangwulan, 2018).
4. Menurut Cooper (2003), populasi merujuk pada keseluruhan kelompok orang, peristiwa, atau benda yang menjadi fokus perhatian peneliti untuk dijadikan objek penelitian (Sudaryono, 2017).
5. Menurut Kurniawan Populasi adalah suatu kawasan generalisasi yang mencakup sejumlah objek atau subjek dengan kualitas dan karakteristik tertentu yang telah ditetapkan oleh penelitian untuk menjadi fokus kajian. Melalui pendekatan ini, peneliti dapat menggali informasi dan data yang relevan dari populasi tersebut, sehingga memungkinkan mereka untuk menyusun kesimpulan yang berdasarkan pada hasil-hasil temuan yang telah ditemukan selama proses penelitian. Dengan demikian, pemahaman yang lebih mendalam tentang karakteristik populasi dapat

diperoleh melalui pengumpulan dan analisis data yang sistematis (Sudaryono, 2017).

6. Populasi tidak hanya merujuk pada manusia, melainkan juga mencakup objek dan elemen alam lainnya, yang melibatkan seluruh karakteristik atau sifat yang dimiliki oleh subjek atau objek tersebut. Oleh karena itu, pada tahap awal perlu melakukan identifikasi populasi dengan tepat dan akurat. Jika definisi populasi tidak jelas, hasil kesimpulan dari penelitian dapat cenderung tidak akurat (Sudaryono, 2017).
7. Populasi dalam konteks penelitian adalah semua subjek (contohnya manusia, pasien) yang memenuhi kriteria tertentu yang telah ditetapkan, seperti semua pasien yang telah menjalani operasi jantung di rumah sakit tertentu (Nursalam, 2003).
8. Populasi adalah keseluruhan subjek (baik manusia, binatang, percobaan, data laboratorium, dan lain-lain) yang akan diteliti dan memenuhi karakteristik yang telah ditetapkan (Riyanto, 2013).
9. Cooper dan Emory (1996), yang mengemukakan bahwa populasi merujuk pada keseluruhan kumpulan elemen yang dapat digunakan untuk menghasilkan kesimpulan (Susila and Suryanto, 2014).
10. Menurut Sudjana (1996) Populasi merupakan totalitas semua nilai yang mungkin, baik hasil perhitungan maupun pengukuran, baik secara kualitatif maupun kuantitatif, terkait dengan karakteristik suatu kelompok objek yang lengkap dan jelas (Susila and Suryanto, 2014).
11. Populasi adalah domain generalisasi yang mencakup objek atau subjek yang memiliki kuantitas dan karakteristik spesifik yang telah ditentukan oleh peneliti untuk dianalisis, dan dari situ, kesimpulan dapat ditarik (Sugiyono, 2015)

12. Menurut Nasir (1988) Populasi dapat diartikan sebagai sekelompok individu yang memiliki kualitas dan ciri-ciri yang telah ditetapkan. Kualitas dan ciri-ciri ini ditentukan oleh variabel yang terlibat. Batas populasi bukanlah terkait dengan lokasi dan waktu penelitian, melainkan terkait dengan karakteristik elemen atau individu dalam populasi. Oleh karena itu, tidak semua subjek dalam lokasi dan waktu penelitian dianalisis, melainkan hanya subjek-subjek yang memiliki karakteristik yang sama (Susila and Suryanto, 2014).
13. Populasi adalah domain generalisasi yang terdiri dari objek/subjek yang memiliki jumlah dan karakteristik tertentu yang ditetapkan oleh peneliti untuk dipelajari dan kemudian ditarik kesimpulannya. Oleh karena itu, populasi tidak hanya terbatas pada manusia, tetapi juga mencakup benda-benda alam lainnya. Selain itu, populasi juga tidak hanya berhubungan dengan jumlah yang ada pada objek/subjek yang diteliti, tetapi mencakup semua karakteristik/sifat yang dimiliki oleh subjek atau objek tersebut (Setiawan and Saryono, 2011)

7.2 Jenis Populasi

Suatu populasi merujuk kepada sekelompok subjek yang menjadi fokus atau sasaran dari suatu penelitian. Sasaran penelitian ini dapat berupa manusia maupun non-manusia, seperti wilayah geografis, jenis penyakit, faktor penyebab penyakit, program kesehatan, gejala penyakit, dan lain sebagainya. Tanpa adanya pembatasan yang jelas terhadap populasi, kesimpulan yang diambil dari hasil penelitian mungkin tidak mencerminkan atau mewakili seluruh populasi. Ketiadaan pembatasan yang jelas terhadap anggota populasi dapat mengakibatkan sampel yang diambil tidak mewakili populasi secara keseluruhan. Oleh karena

itu, dalam setiap penelitian, penting untuk membatasi populasi, misalnya berdasarkan wilayah administratif seperti kelurahan, kecamatan, atau kabupaten, kelompok usia tertentu, jenis penyakit tertentu, dan sebagainya.

Perlu diingat bahwa nilai dari hasil penelitian tidak ditentukan oleh besarnya atau kecilnya populasi, tetapi ditentukan oleh bagaimana peneliti menggunakan dasar pengambilan kesimpulan atau teknik sampling. Jika kesimpulan ditarik dari sampel yang diambil dengan teknik yang tidak tepat, maka hasil penelitian tersebut tidak dapat diterapkan pada seluruh populasi. Sebaliknya, jika penelitian dilakukan dengan sampel yang representatif, maka kesimpulan tersebut dapat diterapkan pada populasi yang lebih luas. Oleh karena itu, pembatasan populasi sangatlah penting untuk memperoleh sampel yang representatif. (Notoatmodjo, 2010) Beberapa referensi menjelaskan bahwa terdapat beberapa jenis populasi, antara lain :

7.2.1 Populasi Target (Target populasi)

Populasi target atau populasi sasaran adalah himpunan unit atau entitas yang menjadi subjek inferensi atau generalisasi dalam suatu penelitian, yang juga sering disebut sebagai fokus penelitian (Surahman, Rachmat and Supardi, 2016). Populasi target ini adalah populasi yang menjadi tujuan akhir dari penerapan hasil penelitian. Populasi target cenderung bersifat umum dan luas, seperti dalam penelitian tentang perilaku pemeriksaan kehamilan oleh ibu hamil di Makassar, di mana populasi targetnya adalah semua ibu hamil di wilayah kota Makassar. Ini berarti bahwa semua ibu hamil di wilayah Kota Makassar menjadi populasi target (Riyanto, 2013).

Menurut (Nursalam, 2003), populasi target adalah kelompok populasi yang memenuhi kriteria sampling dan menjadi tujuan akhir dari penelitian. Polit and Hungler (1999), seperti yang

dikutip oleh Nursalam (2003), menjelaskan bahwa populasi target umumnya memiliki batasan demografis, seperti jenis kelamin dan usia, terutama dalam penelitian klinis. Sebagai contoh, kita dapat memiliki populasi target yang merupakan pasien diabetes mellitus di Makassar. Contoh lain dari populasi target adalah pasien karsinoma paru-paru yang berusia di bawah 40 tahun atau remaja pengguna narkoba (Susila and Suryanto, 2014).

Menurut (Dharma, 2011), populasi target adalah unit di mana hasil penelitian akan diterapkan atau digeneralisir. Idealnya, penelitian dilakukan pada seluruh populasi karena dapat memberikan gambaran menyeluruh tentang populasi sebagai unit di mana hasil penelitian akan diterapkan. Namun, peneliti sering dibatasi oleh karakteristik demografis seperti lokasi geografis, keterbatasan waktu untuk menjangkau seluruh anggota populasi, keterbatasan dana untuk melaksanakan penelitian pada seluruh anggota populasi, dan keterbatasan sumber daya manusia sebagai pelaksana penelitian.

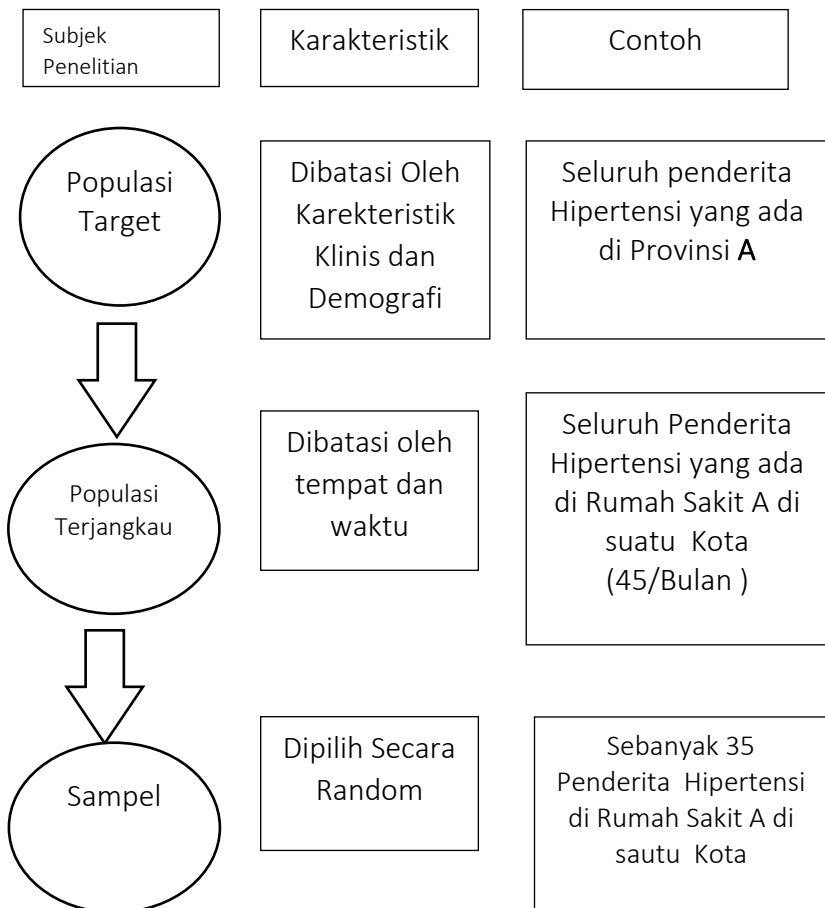
7.2.2 Populasi Terjangkau (*Accessible Population, Source Population*)

Populasi terjangkau atau yang sering disebut sebagai populasi sumber merupakan bagian dari populasi target yang dapat diakses oleh peneliti. Dengan demikian, populasi terjangkau ini merupakan segmen dari populasi target yang dibatasi oleh tempat dan waktu yang lebih spesifik, dan dari populasi terjangkau ini peneliti akan mengambil sampel untuk penelitian. Sebagai contoh, dalam penelitian mengenai perilaku pemeriksaan kehamilan oleh ibu hamil di Kecamatan Biringkanaya Kota Makassar, populasi terjangkau adalah semua ibu hamil di Kecamatan Biringkanaya kota Makassar. Ini berarti bahwa semua ibu hamil di wilayah Kecamatan Biringkanaya Kota Makassar menjadi populasi terjangkau (Riyanto, 2013).

Populasi terjangkau harus mampu mencerminkan kondisi yang lebih luas dari populasi target, termasuk karakteristik dan kondisi lainnya. Peneliti harus memastikan bahwa populasi terjangkau yang akan digunakan dalam penelitian dapat secara tepat mencerminkan situasi yang sebenarnya dari populasi target. Penetapan populasi terjangkau tidak semata berdasarkan pertimbangan metodologi atau perhitungan statistik, melainkan lebih pada kebutuhan praktis dalam penelitian. Sebagai contoh, dalam penelitian tentang orang tua yang memiliki anak tuna grahita di Indonesia sebagai populasi target, jika peneliti menghadapi keterbatasan demografis dalam mengakses populasi ini, maka populasi terjangkau dapat diidentifikasi dalam orang tua yang memiliki anak tuna grahita di Kalimantan Barat, dengan asumsi bahwa karakteristik populasi terjangkau tidak jauh berbeda dengan populasi target (Dharma, 2011).

Menurut Susila dan Suyanto (2014), populasi terjangkau merupakan bagian dari populasi target yang terbatas oleh lokasi dan periode waktu tertentu, seperti contohnya pasien hipertensi yang berobat di Rumah Sakit A selama periode tahun 2019-2023. Dalam konteks penelitian klinis, populasi terjangkau dibatasi oleh karakteristik klinis, demografis, serta lokasi dan waktu

Menurut Nursalam (2003), populasi terjangkau merujuk pada populasi yang memenuhi kriteria penelitian dan biasanya dapat diakses oleh peneliti dari kelompoknya. Sebagai contoh, semua pasien hipertensi yang terdaftar sebagai anggota BPJS di sebuah rumah sakit di suatu Kota. Peneliti biasanya mengambil sampel dari populasi target ini dan berharap sampel tersebut dapat mewakili keseluruhan populasi pasien hipertensi di Kota tersebut. Sastroasmoro dan Ismael (1995) menjelaskan bahwa skema hubungan antara populasi target, populasi terjangkau, subjek terpilih, dan subjek yang sebenarnya diteliti dapat diilustrasikan sebagai berikut :



Gambar 7. 1. Skema hubungan antara populasi Target dan Populasi Terjangkau

Pemilihan populasi terjangkau biasanya tidak mengikuti suatu metode sistematis, tetapi lebih didasarkan pada pertimbangan praktis. Subjek yang dipilih adalah mereka yang memenuhi kriteria penelitian dan dipilih dengan cara tertentu sehingga dianggap mewakili populasi terjangkau. Sebagian subjek terpilih mungkin tidak dapat menyelesaikan penelitian karena

berbagai alasan, sehingga akhirnya data hanya diperoleh dari subjek yang benar-benar diteliti. Hasil penelitian pada subjek yang diteliti ini digeneralisasikan ke populasi terjangkau secara statistika, sedangkan generalisasi dari populasi terjangkau ke populasi target tidak dapat dilakukan secara statistika, melainkan dilakukan dengan menggunakan logika dan penilaian yang masuk akal. (Nursalam, 2003)

Beberapa referensi juga menjelaskan bahwa jenis populasi juga dibedakan menjadi dua yaitu populasi Finit dan populasi Infinit :

7.2.3 Populasi Infinit (Tidak terbatas)

Populasi Infinit merujuk pada populasi yang jumlahnya tidak pernah diketahui dengan pasti dan cakupannya sangat luas (Susila and Suryanto, 2014).

Contohnya :

- a. Serangkaian pelemparan mata dadu yang berlangsung terus-menerus dianggap sebagai populasi infinit (Rinaldi and Mujiyanto, 2017).
- b. Orang-orang yang menjenguk pasien di suatu rumah sakit, yang tidak jelas jumlahnya dan karakteristiknya karena tidak ada catatan atau identifikasi yang resmi, dapat dikategorikan sebagai populasi yang tidak pasti atau tidak terdefinisi dengan jelas (Machfoedz, 2010).

7.2.4 Populasi finit (Terbatas)

Populasi Finit adalah suatu populasi yang secara konkrit diberi batasan ruang dan waktu, seperti tempat, periode, dan karakteristik lainnya, sehingga jumlahnya, lokasi, periode, dan karakteristiknya dapat diketahui dengan pasti (Susila and Suryanto, 2014).

Contoh :

- a. Jumlah petani di sebuah desa merupakan contoh dari populasi finit (Rinaldi and Mujiyanto, 2017).
- b. Ircham Machfoedz (2014) menjelaskan bahwa populasi finit adalah populasi yang memiliki jumlah pasti, di mana identitas subjeknya sudah jelas, yakni pasien. Bahkan, identitas atau ciri-ciri mereka dapat ditelusuri, seperti usia, jenis kelamin, dan jenis penyakit yang tercatat pada kartu status mereka.

Menurut Margono (2004) Populasi juga dibedakan atas dua yaitu (Rinaldi and Mujiyanto, 2017) :

- a. Populasi teoretis (*Theoritical Population*)

Populasi teoritis merupakan sekelompok populasi yang batasannya ditetapkan secara kualitatif. Tujuannya adalah agar hasil penelitian dapat diberlakukan juga pada populasi yang lebih luas. Sebagai contoh, populasi teoritis dapat terdiri dari guru yang berusia antara 30 hingga 45 tahun, mahasiswa program S1, dan sebagainya.

- b. Populasi yang tersedia (*Accessible Population*)

Populasi yang tersedia adalah sekelompok populasi yang dapat diukur secara kuantitatif dengan jelas. Sebagai contoh, jumlah mahasiswa sebanyak 150 orang di sebuah kampus, yang memiliki karakteristik yang telah ditetapkan dalam populasi teoritis.

Selain itu Margono menyatakan bahwa persoalan populasi penelitian harus dibedakan kedalam sifat berikut ini (Rinaldi and Mujiyanto, 2017) :

a. Populasi yang bersifat homogen

Populasi yang homogen adalah populasi di mana unsur-unsurnya memiliki sifat yang serupa, sehingga jumlahnya tidak perlu diperhitungkan secara kuantitatif. Sebagai contoh, jika seorang dokter akan memeriksa golongan darah seseorang, dia hanya perlu mengambil satu tetes darah saja. Dokter tersebut tidak perlu mengambil satu botol darah, karena hasilnya akan sama baik dari satu tetes maupun satu botol darah.

b. Populasi yang bersifat heterogen

Populasi yang heterogen adalah populasi di mana unsur-unsurnya memiliki variasi dalam sifat atau keadaan, sehingga perlu ditetapkan batas-batasnya, baik secara kualitatif maupun kuantitatif. Penelitian dalam bidang sosial yang berfokus pada manusia atau fenomena kehidupan manusia seringkali menghadapi populasi yang heterogen. Sebagai contoh, jika ingin mengetahui distribusi berat badan penduduk di suatu kampung atau distribusi sosial ekonomi penduduk di suatu desa, terlebih lagi di suatu kota, populasi tersebut tentu sangat heterogen. Pengambilan sampel dari berat badan atau situasi sosial ekonomi orang-orang tersebut akan sulit untuk memenuhi syarat representatif (Machfoedz, 2010).

7.3 Kriteria Populasi

Menurut (Nursalam, 2003), dalam menentukan populasi, peneliti perlu memusatkan perhatian pada kriteria yang telah ditetapkan. Dasar pertimbangan untuk menetapkan kriteria populasi meliputi:

1. Biaya

Jika ingin melakukan penelitian pada populasi suku tertentu, peneliti harus mengeluarkan biaya tambahan karena perlu waktu yang cukup lama untuk mempelajari budaya dan bahasa tertentu agar dapat berinteraksi dengan baik, yang tentunya memerlukan biaya tambahan.

2. Praktik / Keterbatasan Akses

Kesulitan muncul dalam melibatkan populasi sebagai subjek penelitian ketika mereka berasal dari daerah terpencil atau sulit dijangkau, seperti masyarakat di daerah pegunungan.

3. Kemampuan orang berpartisipasi dalam penelitian/ Kelayakan Partisipasi Individu

Kondisi kesehatan subjek penelitian, seperti gangguan mental, tidak sadar, atau ketidakstabilan mental, harus dipertimbangkan dalam menetapkan kriteria populasi yang sesuai.

4. Desain penelitian

Dalam penelitian dengan menggunakan rancangan eksperimen, diperlukan populasi yang homogen untuk mengontrol variabel acak, perancu, dan faktor lain yang mungkin mempengaruhi hasil penelitian.

Penggunaan kriteria tersebut di atas dapat digunakan untuk mendefinisikan suatu populasi dalam penelitian dan memiliki dampak dalam menginterpretasikan dan melakukan generalisasi hasil. Menurut (Rizki and Nawangwulan, 2018), dalam suatu populasi terdapat persyaratan yang harus dipenuhi dimana populasi dalam penelitian harus dibatasi dengan jelas. Oleh karena itu, sebelum sampel diambil, kriteria atau batasan populasi harus ditetapkan dengan jelas. Hal ini akan memastikan pengambilan

sampel dilakukan dengan tepat, dengan menentukan kriteria inklusif dan eksklusifnya.

Contoh Kriteria Inklusi

- a. Ibu hamil yang tinggal di wilayah kerja Puskesmas X sekurang kurangnya 1 tahun
- b. Ibu yang mempunyai bayi yang berumur diatas 6 bulan
- c. Memiliki HP Android
- d. Bersedia ikut serta dalam penelitian.

Contoh Kriteria Eksklusi

- a. Ibu hamil yang tidak tinggal di wilayah kerja Puskesmas X kurang 1 tahun
- b. Ibu yang mempunyai bayi 0-5 bulan
- c. Tidak memiliki Hp Android
- d. Tidak bersedia ikut serta dalam penelitian.

Sasaran penelitian dapat berupa manusia atau objek lainnya, seperti wilayah geografis, penyakit, penyebab penyakit, dan program-program kesehatan. Jika tidak ada pembatasan terhadap populasi, maka kesimpulan dari hasil penelitian tidak akan mencerminkan seluruh populasi. Dengan kata lain, tanpa pembatasan yang jelas terhadap anggota populasi, sampel yang diambil tidak akan mewakili populasi secara adekuat. Oleh karena itu, dalam setiap penelitian, populasi harus dibatasi, misalnya menjadi suatu wilayah kelurahan, kecamatan, kabupaten, kelompok usia tertentu, ibu hamil, balita, lansia, atau penyakit tertentu. Penting untuk diingat bahwa kualitas hasil penelitian tidak ditentukan oleh ukuran populasi, melainkan oleh cara peneliti menggunakan dasar pengambilan keputusan dan teknik sampling yang tepat. Jika penelitian dilakukan dengan sampel yang mewakili populasi dan menggunakan teknik yang sesuai, maka

kesimpulan atau generalisasi yang dihasilkan dapat dianggap mewakili populasi secara umum (Rizki and Nawangwulan, 2018).

Menurut Rizki and Nawangwulan, (2018) setidaknya ada tiga hal yang harus diperhatikan dalam menentukan populasi penelitian yaitu :

1. Mendaftarkan seluruh unit yang menjadi anggota populasi
Dengan melakukan pencatatan yang tepat, kita dapat mengidentifikasi unit-unit yang termasuk dalam populasi serta yang tidak. Sebagai contoh, dalam penelitian tentang pengunjung pasien hipertensi di Rumah Sakit "A", pencatatan semua pengunjung pasien hipertensi di RS A pada tahun 2023 harus dilakukan sebelum pengambilan sampel.
2. Menentukan Sampel yang akan dipilih
Memilih sampel dari daftar anggota populasi seperti yang telah disebutkan sebelumnya. Besarnya atau kecilnya ukuran sampel bukanlah penentu apakah sampel tersebut representatif atau tidak. Hal ini tergantung pada karakteristik populasi, seperti homogenitas dan heterogenitas populasi.
3. Menentukan teknik sampling
Pemilihan teknik pengambilan sampel sangat krusial, karena kesalahan dalam menggunakan teknik sampling dapat menghasilkan data yang tidak akurat.

7.4 Populasi dalam Penelitian Kualitatif

Menurut Zuchri Abdussamad (2021), terdapat perbedaan mendasar dalam konsep populasi dan sampel antara penelitian kualitatif dan kuantitatif. Dalam penelitian kualitatif, istilah "populasi" tidak digunakan karena penelitian berfokus pada kasus-kasus tertentu yang muncul dalam konteks sosial tertentu. Hasil penelitian tersebut tidak dimaksudkan untuk diberlakukan secara

umum pada populasi, tetapi untuk dipindahkan atau ditransfer ke konteks sosial lain yang memiliki kesamaan dengan situasi sosial pada kasus yang diteliti. Dalam pendekatan kualitatif, istilah "populasi" lebih tepat disebut sebagai sumber data dalam suatu situasi sosial tertentu yang menjadi fokus penelitian. Ini bisa berupa objek, peristiwa, atau individu di mana data tentang objek penelitian terdapat pada mereka. Karena itu, subjek penelitian mendapat posisi sentral dalam penelitian karena data mengenai fenomena atau masalah yang diselidiki terdapat pada subjek penelitian tersebut. Unit khusus yang dianggap sebagai subjek penelitian yang dipelajari disebut sebagai unit analisis, unit elemen, atau elemen penelitian (Abdussamad, 2021).

Dalam penelitian kualitatif, istilah "sampel" tidak mengacu pada responden, melainkan pada narasumber, partisipan, informan, teman, guru, atau konsultan yang terlibat dalam penelitian. Mereka tidak hanya memberikan jawaban secara pasif terhadap pertanyaan, tetapi secara aktif berinteraksi dengan peneliti sesuai dengan dinamika yang diciptakan oleh peneliti. Dalam konteks penelitian kualitatif, sampel juga tidak disebut sebagai sampel statistik, melainkan sebagai sampel teoritis karena fokus penelitian kualitatif adalah untuk mengembangkan teori (Abdussamad, 2021).

Menurut I Made Laut Merta Jaya (2020), istilah "populasi" berasal dari bahasa Inggris "Population", yang merujuk pada jumlah keseluruhan penduduk atau data yang terdiri dari subjek dan objek penelitian. Namun, dalam konteks penelitian kualitatif, konsep populasi tidak diterapkan seperti dalam penelitian kuantitatif. Fokus penelitian kualitatif berada pada domain tunggal atau beberapa domain yang terkait dengan situasi sosial. Situasi sosial tersebutlah yang dianggap sebagai populasi dalam penelitian kualitatif. Penetapan hasil penelitian kualitatif tidak hanya didasarkan pada variabelnya, melainkan harus memperhatikan

secara menyeluruh situasi sosial yang terkait. Situasi sosial yang diteliti meliputi aspek tempat, sumber informasi, dan aktivitas yang saling berinteraksi secara sinergis. Penelitian kualitatif tidak mengadopsi konsep populasi karena berangkat dari fenomena yang ada dalam situasi sosial. Oleh karena itu, hasil penelitiannya tidak berdampak pada suatu populasi, tetapi lebih bermanfaat dalam konteks objek penelitian dan situasi sosial yang diteliti (Jaya, 2020).

Berbeda dengan penelitian kuantitatif yang memiliki metode dan rumus yang jelas terkait jumlah dan representativitas sampel, penelitian kualitatif menggunakan metode *purposive sampling*. *Purposive sampling* merupakan jenis sampling non probabilitas yang umumnya digunakan dalam penelitian kualitatif, di mana tujuannya bukan untuk menggeneralisasikan temuan sampel kepada populasi. Sehingga, sifat penelitian kualitatif bersifat ideografis atau kasuistik. Dalam penelitian kasuistik, pendekatan yang digunakan bersifat naturalistik atau alamiah, sehingga perlu dilakukan dengan hati-hati dalam mentransfer temuan dari satu kasus individu atau kelompok ke kasus individu atau kelompok lainnya. Tantangannya adalah bagaimana peneliti dapat mengidentifikasi responden yang potensial dan bersedia untuk diwawancarai, bahkan mungkin beberapa kali selama periode beberapa bulan. Salah satu cara untuk melakukannya adalah dengan menemukan satu atau beberapa responden terlebih dahulu, baik melalui kebetulan, melalui jaringan kenalan, melalui iklan, atau cara lainnya, dan kemudian merekrut responden berikutnya melalui referensi dari responden sebelumnya. Melalui proses ini, peneliti dapat menemukan lebih banyak responden lagi, dan seterusnya. Oleh karena itu, setelah menggunakan *purposive sampling*, penelitian sering dilanjutkan dengan menggunakan metode *Snowball sampling* (Machfoedz, 2010).

7.5 Hubungan Populasi dan Sampel

Populasi merujuk pada keseluruhan objek yang menjadi fokus penelitian. Masalah terkait populasi muncul ketika objek penelitian telah secara jelas diformulasikan, seperti contohnya populasi penderita flu burung di perkotaan atau pedesaan, populasi ayam yang mati mendadak di perkotaan, populasi ibu hamil berusia 35 tahun, populasi peserta KB jenis suntik, atau populasi remaja putri. Formulasi yang jelas terhadap populasi sangat penting karena sampel yang akan dipilih harus sesuai dengan informasi yang ingin ditemukan. Selain itu, penegasan terhadap populasi ini juga diperlukan jika ingin membandingkannya dengan populasi lain, sehingga kesimpulan dari penelitian sesuai dengan tujuan yang telah ditetapkan (Sunyoto and Setiawan, 2013).

Walaupun pembahasan mengenai sampel akan dijelaskan lebih lanjut pada bab berikutnya, namun penting untuk memahami bahwa populasi dan sampel merupakan dua konsep yang saling terkait dan tidak dapat dipisahkan. Dengan demikian, memahami hubungan antara populasi dan sampel menjadi hal yang penting, di mana sampel merupakan bagian dari populasi yang karakteristiknya hendak diselidiki dan dianggap dapat mewakili keseluruhan populasi (dengan jumlah yang lebih sedikit daripada jumlah populasi) (Sunyoto, 2012). Sampel penelitian merupakan sekelompok individu yang lebih kecil lagi dibanding populasi yang terjangkau, di mana peneliti secara langsung mengumpulkan data atau melakukan pengamatan/pengukuran pada unit ini. Pada dasarnya, penelitian dilakukan pada sampel yang dipilih dari populasi yang terjangkau. Menurut Sastroasmoro (2008), penggunaan sampel dalam penelitian memiliki keuntungan karena

lebih ekonomis, lebih mudah, lebih cepat, lebih akurat, mampu mewakili populasi, dan lebih spesifik (Dharma, 2011).

Menurut Margono (1997), penggunaan populasi relevan ketika penelitian bertujuan untuk memperoleh pemahaman yang pasti tentang keadaan sesungguhnya dari populasi yang membutuhkan tingkat ketelitian dan kecermatan yang tinggi, terutama ketika sumber informasi bersifat heterogen di mana karakteristik masing-masing sumber sulit dibedakan (Setiawan and Saryono, 2011). Penggunaan sampel dalam penelitian, seperti yang disebutkan oleh Aminuddin Rosyid (1987), menjadi penting ketika jumlah populasi yang akan diteliti terlalu besar, wilayah populasi sangat luas dan tersebar, waktu penelitian terbatas, dana terbatas, tenaga peneliti terbatas, fasilitas yang tersedia terbatas, serta situasi untuk melakukan penelitian tidak memadai. Selain itu, menurut Agus Suradika (2000), ada beberapa alasan untuk menggunakan sampel dalam penelitian, yaitu ketika jumlah individu yang akan diselidiki tidak terbatas, ketika penelitian bersifat destruktif, ketika objek yang diteliti homogen, dan ketika ketelitian yang mutlak tidak diperlukan atau hasil penelitian diperlukan dengan segera (Setiawan and Saryono, 2011).

DAFTAR PUSTAKA

- Abdussamad, Z. (2021) Metode Penelitian Kualitatif. Syakir Media Press.
- Dharma, K. K. (2011) Metode Penelitian Keperawatan Panduan Melaksanakan dan Menerapkan Hasil Penelitian. Trans Info Media, Jakarta.
- Jaya, I. M. L. M. (2020) Metode Penelitian Kuantitatif dan Kualitatif Teori , Penerapan dan Riset Nyata. Quadrant, Yogyakarta.
- Machfoedz, I. (2010) Metode Penelitian (Kuantitatif dan Kualitatif) Bidang Kesehatan , Keperawatan, Kebidanan, Kedokteran. Fitramaya, Yogyakarta.
- Notoatmodjo, S. (2010) Metode Penelitian Kesehatan. Rineka Cipta Jakarta.
- Nursalam (2003) Konsep dan Penerapan Metodologi Penelitian Ilmu Keperawatan Pedoman Skripsi Tesis dan Instrumen Penelitian Keperawatan. Salemba Medika, Jakarta.
- Rinaldi, S. F. and Mujianto, B. (2017) Metode Penelitian dan Statistik. Pusat Pendidikan Sumber Daya Manusia Kesehatan Badan Pengembangan dan Pemberdayaan Sumber Daya Manusia Kesehatan Kementerian Kesehatan Republik Indonesia.
- Rizki, R. and Nawangwulan, S. (2018) Metode Penelitian Kesehatan. Sidoarjo: Indomedia Pustaka, Sidoarjo.
- Riyanto, A. (2013) Statistik Deskriptif untuk Kesehatan. Nuha Medika, Yogyakarta.
- Setiawan, A. and Saryono (2011) Metode Penelitian Kebidanan DIII, DIV, S1 dan S2. Nuha Medika, Yogyakarta.
- Sudaryono (2017) Metode Penelitian Kuantitatif, Kualitatif dan Mix Method. Rajawali Pers, Divisi Buku Perguruan Tinggi, PT Raja Grafindo Persada Depok.

- Sugiyono (2015) Statistik Nonparametris untuk Penelitian. Alfabeta Bandung.
- Sunyoto, D. (2012) Biostatistik Untuk Kebidanan. Nuha Medika Yogyakarta.
- Sunyoto, D. and Setiawan, A. (2013) Statistik Kesehatan Parametrik, Non Paramatik, Validitas, Reliabilitas. Nuha Medika, Yogyakarta.
- Surahman, Rachmat, M. and Supardi, S. (2016) Metodologi Penelitian. Pusat Pendidikan Sumber Daya Manusia Kesehatan Badan Pengembangan dan Pemberdayaan Sumber Daya Manusia Kesehatan Kementerian Kesehatan Republik Indonesia.
- Susila and Suryanto (2014) Metode Penelitian Epidemiologi Bidang Kedokteran dan Kesehatan. Bursa Ilmu Yogyakarta.

BAB 8

SAMPEL

Oleh Arif Rachman

8.1 Pendahuluan

8.1.1 Definisi Biostatistika

Biostatistika adalah cabang statistika yang berkaitan dengan analisis data dalam konteks ilmu biomedis dan kesehatan. Dalam bidang ini, teknik-teknik statistika digunakan untuk merancang studi penelitian, menganalisis data percobaan, dan menyimpulkan informasi dari hasil pengamatan atau eksperimen di bidang biomedis (Pagano, M., & Gauvreau, K, 2000).

8.1.2 Penggunaan statistika dalam ilmu kedokteran dan ilmu kesehatan

Penggunaan statistika dalam ilmu kedokteran dan ilmu kesehatan sangat luas dan penting untuk analisis data, interpretasi hasil penelitian, dan pengambilan keputusan klinis. Berikut adalah beberapa contoh penggunaan statistika dalam kedokteran dan kesehatan (Rothman, K. J., *et.all*, 2008) (Levin, K. A., 2006) :

- a. Analisis Epidemiologi
- b. Uji Klinis dan Penelitian Kedokteran
- c. Statistik dalam Diagnostik Medis
- d. Analisis Regresi dan Model Biostatistik
- e. Analisis Survival (Analisis Kelangsungan Hidup)
- f. Statistik dalam Penelitian Genomik
- g. Analisis Data Kesehatan Masyarakat.

8.1.3 Peran biostatistika dalam analisis data penelitian

Peran biostatistik dalam analisis data penelitian sangat penting untuk menghasilkan temuan yang akurat dan dapat diandalkan. Berikut adalah beberapa peran biostatistik dalam analisis data penelitian (Hulley, S. B., *et.all.*, 2013) (Armitage, P., *et.all.*, 2008):

- a. Desain Penelitian dan Pengumpulan Data
- b. Deskripsi dan Karakterisasi Data
- c. Uji Hipotesis dan Pengujian Signifikansi
- d. Analisis Survival (Analisis Kelangsungan Hidup)
- e. Meta-analisis
- f. Pengelolaan dan Analisis Data Genomik
- h. Penggunaan Teknologi Informasi dalam Biostatistika.

8.2 Sampel dalam Konteks Penelitian

8.2.1 Definisi Sampel

Sampel adalah sejumlah elemen atau unit yang diambil dari suatu populasi untuk dianalisis atau diuji agar dapat membuat inferensi atau generalisasi tentang populasi tersebut. (Creswell, J. W., 2014) (Levin, B., & Fox, J. A., 2009).

8.2.1.1 Subset dari populasi yang diambil untuk penelitian

Subset dari populasi yang diambil untuk penelitian disebut "sampel." Sampel merupakan representasi sebagian dari populasi yang digunakan untuk mengumpulkan data, melakukan analisis, dan membuat inferensi atau generalisasi terhadap keseluruhan populasi. Penggunaan sampel dalam penelitian memungkinkan peneliti untuk menghemat waktu, biaya, dan sumber daya, sementara masih dapat menyediakan informasi yang cukup untuk

membuat estimasi atau kesimpulan yang dapat diterapkan pada populasi secara umum (Creswell, J. W.,2014) (Trochim, W. M., & Donnelly, J. P., 2008).

- a. Contoh penggunaan sampel dalam penelitian:
 - Penelitian Tinggi Badan Siswa:
 - Populasi: Seluruh siswa di suatu sekolah.
 - Sampel: Sejumlah siswa yang dipilih secara acak dari seluruh sekolah tersebut.
- b. Contoh Survei Opini Publik:
 - Populasi: Semua warga negara di suatu negara.
 - Sampel: Sejumlah warga yang dipilih secara acak untuk diwawancarai.
- c. Contoh Pengukuran Kualitas Produk:
 - Populasi: Semua produk yang diproduksi oleh suatu perusahaan.
 - Sampel: Sejumlah produk yang diambil secara acak dari jalur produksi.

8.2.1.2 Tujuan pengambilan sampel

Pengambilan sampel dalam penelitian memiliki beberapa tujuan utama yang dirancang untuk meningkatkan efisiensi dan kepraktisan penelitian. Beberapa tujuan umum pengambilan sampel melibatkan (Creswell, J. W.,2014) (Trochim, W. M., & Donnelly, J. P., 2008):

- a. Efisiensi Riset.
- b. Praktisitas.
- c. Generalisasi.
- d. Akurasi Statistik .
- e. Keterwakilan.
- f. Fleksibilitas.
- g. Pengendalian Variabel.

h. Keamanan dan Etika.

Tujuan pengambilan sampel dalam penelitian melibatkan beberapa aspek, dan para peneliti merinci tujuan-tujuan ini untuk mencapai hasil yang valid dan dapat diandalkan. Pemahaman mendalam tentang tujuan pengambilan sampel sangat penting untuk merancang penelitian yang efektif dan valid.

8.2.2 Populasi dan Sampel

Populasi merujuk pada kelompok keseluruhan yang menjadi fokus dari suatu penelitian atau observasi. Ini adalah kumpulan individu, objek, atau peristiwa yang memiliki karakteristik umum atau kesamaan tertentu.

Sampel adalah sebagian kecil dari populasi yang diambil untuk diuji atau diamati. Sampel dipilih dengan hati-hati untuk mencerminkan karakteristik umum dari keseluruhan populasi (Creswell, J. W., 2014) (Trochim, W. M., & Donnelly, J. P., 2008).

Populasi dan sampel adalah konsep-konsep dasar dalam penelitian, dan pemahaman yang baik tentang keduanya penting untuk merancang penelitian yang valid dan hasil yang dapat diandalkan.

8.2.2.1 Hubungan antara populasi dan sampel

Populasi dan Sampel memiliki hubungan yang erat dalam konteks penelitian.

- a. Definisi dan Ruang Lingkup.
- b. Tujuan Penelitian.
- c. Efisiensi dan Praktisitas.
- d. Analisis Statistik.
- e. Representativitas.
- f. Inferensi Statistik.

g. Kontrol Variabel.

Keakuratan hasil dari sampel tergantung pada seberapa baik sampel tersebut merepresentasikan karakteristik populasi secara keseluruhan (Creswell, J. W., 2014) (Trochim, W. M., & Donnelly, J. P., 2008).

8.2.2.2 Representasi sampel terhadap populasi

Representasi sampel terhadap populasi adalah sejauh mana sampel yang diambil mencerminkan karakteristik dan variasi yang ada di dalam populasi secara keseluruhan. Berikut adalah beberapa konsep terkait dengan representasi sampel terhadap populasi:

- a. Ketidakpastian dan Error Sampling.
- b. Karakteristik Utama Populasi.
- c. Metode Pengambilan Sampel.
- d. Konteks Penelitian.
- e. Keterwakilan Secara Proporsional.
- f. Validitas Internal dan Eksternal.

Representasi sampel yang baik adalah kunci untuk mendapatkan hasil penelitian yang dapat diandalkan dan relevan dengan populasi yang lebih besar (Creswell, J. W., 2014) (Trochim, W. M., & Donnelly, J. P., 2008).

8.3 Proses Pemilihan Sampel

Proses pemilihan sampel merupakan langkah kritis dalam desain penelitian yang dapat mempengaruhi validitas dan generalisabilitas hasil penelitian. Ada beberapa metode yang dapat digunakan untuk memilih sampel, dan pilihan tergantung pada

tujuan penelitian, populasi yang diinginkan, serta ketersediaan sumber daya. Berikut adalah beberapa metode umum dalam proses pemilihan sampel, yaitu: Pemilihan Sampel Acak, Pemilihan Sistematis, Pemilihan Kluster, Pemilihan Sampel Non-Acak, Pemilihan Sampel Gemuk (*Convenience Sampling*), Pemilihan Sampel Kuota (*Quota Sampling*), Pemilihan Sampel Stratifikasi, Pemilihan Sampel *Multistage*, Pemilihan Sampel *Matched* (Pengaitan), Pemilihan Sampel *Purposive* (Penghakiman). Proses pemilihan sampel memerlukan pertimbangan hati-hati untuk memastikan representativitas dan validitas hasil penelitian (Creswell, J. W., 2014) (Trochim, W. M., & Donnelly, J. P., 2008).

8.3.1 Random Sampling

Random Sampling (Pemilihan Acak) adalah metode pengambilan sampel di mana setiap anggota dari populasi yang akan diambil memiliki peluang yang sama untuk dipilih. Ini berarti bahwa setiap elemen dalam populasi memiliki probabilitas yang dikenal dan setara untuk menjadi bagian dari sampel. Random sampling memberikan dasar untuk membuat inferensi statistik yang lebih kuat tentang populasi karena mengurangi bias dan memastikan representasi yang lebih baik. Ada beberapa teknik random sampling yang umum digunakan, yaitu: pemilihan acak sederhana (*simple random sampling*), pemilihan acak sistematis (*systematic random sampling*), pemilihan acak berstrata (*stratified random sampling*), pemilihan acak kluster (*cluster random sampling*). Keuntungan Random Sampling adalah memastikan representasi yang lebih baik dari populasi dan meminimalkan bias dan memungkinkan generalisasi yang lebih baik. Pemilihan acak memberikan dasar metodologis yang kuat dalam merancang penelitian dan membantu memastikan hasil penelitian lebih umum berlaku untuk populasi yang lebih besar (Creswell, J. W., 2014) (Trochim, W. M., & Donnelly, J. P., 2008).

8.3.1.1 Pengambilan sampel acak

Pengambilan sampel acak adalah proses memilih elemen-elemen dari populasi yang akan dijadikan sampel, di mana setiap elemen memiliki peluang yang sama untuk dipilih. Berikut adalah langkah-langkah umum dalam pengambilan sampel acak: identifikasi populasi, penetapan ukuran sampel, pemilihan metode pengambilan acak (pemilihan acak sederhana dan pemilihan acak sistematis), pelaksanaan pengambilan sampel, pengumpulan data dan analisis data.

Pengambilan sampel acak adalah pendekatan yang kuat untuk memastikan representativitas dan keadilan dalam pengumpulan sampel, yang esensial untuk generalisasi hasil penelitian ke populasi yang lebih besar (Creswell, J. W., 2014) (Trochim, W. M., & Donnelly, J. P., 2008).

8.3.1.2 Kelebihan dan kelemahan Random Sampling

Kelebihan Random Sampling:

- a. Representativitas yang Baik.
- b. Kemungkinan Generalisasi yang Baik.
- c. Kemungkinan Kesalahan Statistik yang Rendah.
- d. Sederhana dan Transparan.
- e. Penghilangan Bias Subyektif.

Kelemahan Random Sampling:

- a. Biaya dan Waktu.
- b. Keterbatasan Akses ke Data Populasi.
- c. Keterbatasan Informasi pada Subpopulasi Kecil.
- d. Potensi untuk Kesalahan Pengambilan Sampel.
- e. Tidak Praktis untuk Populasi Tersebar Luas.
- f. Tidak Memperhatikan Struktur Populasi.

Meskipun memiliki kelebihan signifikan, pengambilan sampel acak juga memiliki keterbatasan dan harus dipertimbangkan dengan cermat dalam desain penelitian (Creswell, J. W., 2014) (Trochim, W. M., & Donnelly, J. P., 2008).

8.3.2 Stratified Sampling

Stratified Sampling (Pemilihan Sampel Berstrata) adalah metode pengambilan sampel di mana populasi dibagi menjadi subkelompok atau strata yang homogen berdasarkan karakteristik tertentu, dan kemudian sampel diambil dari setiap strata. Ini membantu memastikan bahwa setiap subkelompok dalam populasi diwakili dalam sampel, sehingga meningkatkan keakuratan dan representativitas hasil penelitian. Langkah-langkah dalam *stratified sampling*, yaitu; identifikasi populasi, penetapan strata, penentuan ukuran sampel strata, pemilihan metode pengambilan sampel, pemilihan sampel dari setiap strata, penggabungan sampel, pengumpulan data, analisis data.

Berikut adalah kelebihan dan kelemahan dari *stratified sampling*:

Kelebihan *Stratified Sampling*:

- a. Representasi yang Baik.
- b. Analisis Lebih Mendalam.
- c. Ketepatan Statistik yang Tinggi.

Kelemahan *Stratified Sampling*:

- a. Kompleksitas Perencanaan.
- b. Kesulitan Identifikasi Strata yang Relevan.
- c. Peningkatan Kompleksitas Analisis Data.

Stratified sampling merupakan teknik yang bermanfaat untuk memastikan representasi yang baik dari setiap kelompok dalam populasi dan meningkatkan validitas hasil penelitian. (Creswell, J. W., 2014) (Trochim, W. M., & Donnelly, J. P., 2008).

8.3.3 Systematic Sampling

Systematic Sampling adalah metode pengambilan sampel di mana elemen-elemen populasi dipilih dengan interval tetap, dimulai dari elemen yang dipilih secara acak. Metode ini bergantung pada urutan atau pola tertentu dalam populasi.

Langkah-langkah dalam *systematic sampling*, yaitu: identifikasi populasi, penetapan ukuran sampel, perhitungan *interval sampling*, pemilihan elemen pertama, pengambilan sampel sistematis, pengumpulan data. Rumusnya *Interval Sampling* adalah:

$$k = \frac{\text{Jumlah elemen dalam populasi}}{\text{Ukuran sampel}}$$

Kelebihan *systematic sampling* adalah sederhana dan efisien; dan cocok untuk populasi teratur.

Kelemahan *systematic sampling* adalah ketidakcocokan untuk populasi non-teratur dan risiko pola tersembunyi.

Keuntungan dari *systematic sampling* meliputi efisiensi waktu dan biaya karena tidak perlu mengidentifikasi dan mengakses seluruh populasi. Meskipun demikian, terdapat beberapa pertimbangan dan risiko potensial yang perlu diperhatikan adalah ketidakcocokan dengan pola populasi, pengaruh outlier, keterbatasan dalam situasi tertentu (Creswell, J. W., 2014) (Trochim, W. M., & Donnelly, J. P., 2008).

8.4 Ukuran Sampel

Ukuran sampel (*sample size*) adalah jumlah elemen atau individu yang dipilih dari populasi untuk diikuti sertakan dalam penelitian atau percobaan. Ukuran sampel merujuk pada jumlah elemen atau individu yang dipilih dari populasi untuk dimasukkan dalam sampel penelitian. Menentukan ukuran sampel yang tepat merupakan langkah penting dalam perancangan penelitian, dan faktor-faktor tertentu perlu dipertimbangkan. Beberapa pertimbangan dalam menentukan ukuran sampel melibatkan, yaitu: tingkat kepercayaan (*confidence level*), *margin of error* (kesalahan margin), variabilitas dalam populasi, tingkat signifikansi (*significance level*), kebutuhan untuk subgrup analisis, kompleksitas analisis statistik, biaya dan waktu, efek besar yang diinginkan, sumber daya yang tersedia.

Menentukan ukuran sampel yang tepat melibatkan keseimbangan antara kebutuhan untuk akurasi statistik dan keterbatasan sumber daya. Penggunaan kalkulator ukuran sampel atau konsultasi dengan ahli statistik dapat membantu dalam proses ini (Creswell, J. W., 2014) (Trochim, W. M., & Donnelly, J. P., 2008).

8.4.1 Faktor-faktor yang mempengaruhi ukuran sampel

Ukuran sampel yang tepat sangat penting dalam merancang penelitian karena dapat mempengaruhi validitas dan generalisabilitas hasil penelitian. Beberapa faktor yang mempengaruhi ukuran sampel adalah ukuran populasi (n), tingkat kepercayaan (*confidence level*), *margin of error* (kesalahan margin), variabilitas dalam populasi, efek besar yang diinginkan (*effect size*), deviasi standar (*standard deviation*), analisis statistik yang digunakan, kebutuhan untuk subgrup analisis, biaya dan waktu, karakteristik penelitian (desain dan tujuan) dan kompleksitas populasi.

Faktor-faktor yang mempengaruhi ukuran sampel dapat bervariasi tergantung pada konteks penelitian dan jenis analisis statistik yang akan dilakukan (Creswell, J. W., 2014).

8.4.2 Rumus dan metode perhitungan

Berikut adalah beberapa rumus dan metode umum untuk menghitung ukuran sampel berdasarkan beberapa situasi khusus:

a. Ukuran Sampel untuk Proporsi:

$$n = \frac{Z^2 \times p \times (1-p)}{E^2}$$

- n = Ukuran sampel yang dibutuhkan.
- Z = Nilai z-skor yang sesuai dengan tingkat kepercayaan.
- p = Perkiraan proporsi dalam populasi.
- E = Margin of error yang diinginkan.

b. Ukuran Sampel untuk Rata-rata:

$$n = \frac{Z^2 \times \sigma^2}{E^2}$$

- n = Ukuran sampel yang dibutuhkan.
- Z = Nilai z-skor yang sesuai dengan tingkat kepercayaan.
- σ = Deviasi standar populasi.
- E = Margin of error yang diinginkan.

c. Ukuran Sampel untuk Perbedaan Dua Proporsi:

$$n = \frac{(Z^2 \times p_1 \times (1-p_1)) + (Z^2 \times p_2 \times (1-p_2))}{E^2}$$

- n = Ukuran sampel yang dibutuhkan.
- Z = Nilai z-skor yang sesuai dengan tingkat kepercayaan.
- p_1 dan p_2 = Proporsi dalam kedua populasi yang dibandingkan.
- E = Margin of error yang diinginkan.

d. Ukuran Sampel untuk Perbedaan Dua Rata-rata:

$$n = \frac{(Z^2 \times \sigma_1^2) + (Z^2 \times \sigma_2^2)}{E^2}$$

- n = Ukuran sampel yang dibutuhkan.
- Z = Nilai z-skor yang sesuai dengan tingkat kepercayaan.
- σ_1 dan σ_2 = Deviasi standar kedua populasi yang dibandingkan.
- E = Margin of error yang diinginkan.

e. Ukuran Sampel untuk Regresi Linear Sederhana:

$$n = \frac{N}{1 + \frac{N}{k}}$$

- n = Ukuran sampel yang dibutuhkan.
- N = Jumlah elemen dalam populasi.
- k = Jumlah variabel independen dalam analisis regresi.

f. Ukuran Sampel untuk Studi Kualitatif (Metode Jenuh Wawancara): Ukuran sampel dalam penelitian kualitatif sering ditentukan oleh konsep "jenuh" atau "kenyangan informasi" yang dikenal sebagai titik di mana penelitian tidak lagi menghasilkan informasi baru. Namun, ukuran sampel sering berkisar antara 20 hingga 30 partisipan.

Catatan Penting:

Nilai Z untuk tingkat kepercayaan tertentu dapat ditemukan dalam tabel distribusi normal standar. Ukuran sampel yang lebih besar umumnya diperlukan untuk tingkat kepercayaan yang lebih tinggi dan margin of error yang lebih kecil (Creswell, J. W., 2014).

8.4.3 Ketentuan Pengambilan Sampel yang Representatif

Pengambilan sampel yang representatif adalah langkah kritis dalam penelitian, karena sampel yang baik akan mencerminkan karakteristik populasi secara keseluruhan. Berikut adalah beberapa ketentuan dan prinsip yang dapat membantu dalam pengambilan sampel yang representatif adalah random sampling (pengambilan sampel acak), *stratified sampling* (pengambilan sampel berlapis), *cluster sampling* (pengambilan sampel kluster), proporsional sampling teknik, pengambilan sampel sistematis, pertimbangkan faktor-faktor demografis penting, perhatikan karakteristik khusus populasi, jumlah sampel yang cukup, konsistensi dalam pengumpulan data, validasi hasil dengan data populasi, transparansi dan laporan metode pengambilan sampel (Lohr, S., 2019) (Thompson, S. K., 2012).

8.4.4 Keberagaman dan homogenitas

Keberagaman (*Diversity*) adalah merujuk pada variasi atau perbedaan di antara elemen atau individu dalam suatu kelompok atau populasi. Ini dapat mencakup berbagai karakteristik seperti usia, jenis kelamin, suku bangsa, pendidikan, keyakinan agama, dan banyak lagi. Keberagaman bisa menjadi indikator kemajemukan dalam suatu kelompok dan bisa memiliki dampak positif, seperti meningkatkan inovasi dan perspektif.

Keberagaman sering kali dianggap sebagai sumber kekayaan dan ketahanan suatu sistem atau kelompok. Dalam banyak konteks, keberagaman dianggap positif karena dapat memperkaya pengalaman, mempromosikan inovasi, dan memperkuat keberlanjutan. Penting untuk memahami dan menghargai keberagaman dalam berbagai aspek kehidupan untuk menciptakan lingkungan yang inklusif dan berkelanjutan (Creswell, J. W., 2014).

Homogenitas adalah keadaan di mana elemen atau individu dalam suatu kelompok memiliki kesamaan atau keseragaman tertentu. Sebaliknya dengan keberagaman, homogenitas menunjukkan kurangnya variasi atau perbedaan yang signifikan antara elemen-elemen dalam kelompok tersebut.

Hubungan antara Keberagaman dan Homogenitas adalah sebuah kelompok atau populasi dapat menjadi homogen dalam beberapa aspek dan heterogen dalam aspek lainnya. Keberagaman dapat memberikan kekayaan dalam ide, pandangan, dan keterampilan, sementara homogenitas dapat membantu menciptakan keseragaman dan koherensi dalam suatu kelompok.

8.4.5 Pentingnya ukuran sampel yang memadai

Ukuran sampel yang memadai sangat penting dalam penelitian karena memengaruhi validitas dan reliabilitas hasil penelitian. Berikut adalah beberapa alasan mengapa ukuran sampel yang memadai sangat penting, yaitu; representatifitas populasi, keakuratan dan ketepatan hasil, keandalan statistik inferensial, kemampuan untuk menemukan perbedaan yang signifikan, penggunaan metode statistik yang lebih kompleks, kemampuan untuk membuat generalisasi, pertimbangan etika, efisiensi penggunaan sumber daya, meningkatkan kekuatan statistik, pertimbangan untuk analisis subkelompok. Ukuran sampel yang memadai sangat penting dalam penelitian karena mempengaruhi validitas dan reliabilitas hasil penelitian.

8.5 Teknik Pengumpulan Data dari Sampel

Teknik pengumpulan data dari sampel dapat bervariasi tergantung pada jenis penelitian, tujuan, dan karakteristik sampel yang terlibat. Berikut adalah beberapa teknik pengumpulan data yang umum digunakan antara lain wawancara (Kvale, S., 2007),

kuesioner atau survei (Dillman, D. *et.all.*, 2014), pengamatan (Fetterman, D. M., 2010), studi kasus (Yin, R. K., 2014), *focus group* (Krueger, R. A., & Casey, M. A., 2014), pengukuran fisik atau tes (DeVon, H. A., *et.all.*, 2007), dokumentasi atau analisis konten (Neuendorf, K. A., 2002), pengambilan sampel biologis (Burtis, C. A., *et.all.*, 2012), sensor atau peralatan otomatis (Trochim, W., & Donnelly, J. P., 2008), survei online atau pengumpulan data elektronik (Couper, M. P., *et.all* (Eds.), 2017). Pemilihan teknik pengumpulan data harus didasarkan pada pertimbangan metodologis dan kebutuhan penelitian. Kombinasi beberapa teknik seringkali memberikan hasil yang lebih holistik dan dapat mendukung validitas data.

8.5.1 Metode Observasional

Metode observasional adalah suatu pendekatan penelitian di mana peneliti mengamati dan mencatat perilaku, kejadian, atau fenomena secara langsung tanpa intervensi atau manipulasi. Metode ini biasanya digunakan untuk memahami karakteristik alami suatu lingkungan atau perilaku manusia dan hewan.

Metode observasional dapat digunakan dalam berbagai konteks, termasuk penelitian psikologi, sosiologi, antropologi, pendidikan, dan ilmu lingkungan. Keunggulan metode ini termasuk keakuratan dalam pengamatan langsung dan potensi untuk menangkap konteks yang kompleks. Namun, kelemahannya termasuk keterbatasan dalam generalisasi hasil dan potensi pengaruh subjektivitas pengamat. Oleh karena itu, perencanaan yang cermat dan pelatihan yang baik diperlukan untuk memastikan kualitas pengamatan (Jerry W. W., 2007).

8.6 Analisis Statistik pada Sampel

Analisis statistik pada sampel adalah langkah penting dalam penelitian untuk membuat inferensi atau generalisasi terhadap populasi yang lebih besar. Berikut adalah beberapa langkah umum yang terlibat dalam analisis statistik pada sampel adalah deskripsi data, visualisasi data, uji normalitas (uji kolmogorov-smirnov atau shapiro-wilk), uji hipotesis, regresi, analisis varians (ANOVA), analisis *chi-square*, analisis multivariat, analisis kesesuaian model, *confidence interval*, ketepatan pengukuran, interpretasi hasil (Field, A., 2013) (Tabachnick, B. G., & Fidell, L. S., 2013).

8.6.1 Inferensi Statistik

Inferensi Statistik adalah cabang statistika yang berkaitan dengan penggunaan sampel data untuk membuat kesimpulan atau inferensi tentang karakteristik atau parameter populasi secara lebih luas. Dalam inferensi statistik, penelitian dilakukan pada sampel data dengan tujuan untuk menyimpulkan atau mengambil kesimpulan tentang populasi yang lebih besar dari mana sampel tersebut diambil. Terdapat dua aspek utama dalam inferensi statistik: estimasi dan pengujian hipotesis.

8.6.2 Penggunaan sampel untuk membuat inferensi terhadap populasi

Penggunaan sampel untuk membuat inferensi terhadap populasi adalah prinsip dasar dalam statistika inferensial. Berikut adalah beberapa langkah umum yang terlibat dalam proses ini, yaitu; pemilihan sampel yang representatif, eksplorasi dan analisis sampel, pembuatan inferensi, estimasi parameter populasi, penggunaan interval kepercayaan, pengukuran ketidakpastian, kontrol kesalahan tipe I dan tipe II, validasi dan generalisasi.

Penting untuk diingat bahwa inferensi statistik membutuhkan pertimbangan yang hati-hati terhadap asumsi, metode pengambilan sampel, dan teknik analisis statistik. Kesalahan yang mungkin terjadi selalu harus diidentifikasi dan dinilai (Triola, M. F., 2017).

8.7 Pentingnya Konsep Sampel dalam Penelitian dan Biostatistika

Konsep sampel memiliki peran penting dalam penelitian dan biostatistika karena berkaitan erat dengan bagaimana kita menggeneralisasi hasil dari suatu studi atau percobaan ke populasi yang lebih besar. Sampel merupakan subset dari populasi yang diambil untuk memperoleh informasi tentang populasi secara keseluruhan, yaitu; representativitas populasi (Cochran, W. G., 1977), efisiensi dan hemat waktu (Lohr, S., 2010), biaya penelitian (Thompson, S. K., 2012), keterbatasan sumber daya (Israel, G. D., 2003), ketidakpastian statistik (Hogg, R. V., *et.all.*, 2018), kemudahan pengumpulan data (Fowler, F. J., 2013), ketepatan dan validitas penelitian (Trochim, W. M., & Donnelly, J. P., 2008), pengendalian variabilitas (Sheskin, D. J., 2007), generalisabilitas temuan (Gravetter, F. J., & Forzano, L.-A. B., 2018), analisis statistik yang akurat (Agresti, A., & Finlay, B., 2009), pengambilan keputusan yang informasional (Aron, A., *et.all.*, 2016), etimologi statistik (Levin, R. I., *et.all.*, 2018).

Pentingnya konsep sampel dalam penelitian dan biostatistika mencerminkan upaya untuk mengoptimalkan penggunaan sumber daya, memastikan validitas hasil, dan memberikan dasar yang kokoh untuk pengambilan keputusan berdasarkan data. Dengan memahami dan menerapkan konsep sampel dengan baik, peneliti dapat meningkatkan kualitas penelitian mereka dan memberikan kontribusi yang lebih berarti dalam pemahaman ilmiah.

DAFTAR PUSTAKA

- Agresti, A., & Finlay, B. (2009). *Statistical Methods for the Social Sciences* (4th ed.). Pearson.
- Armitage, P., Berry, G., & Matthews, J. N. (2008). *Statistical Methods in Medical Research*. Blackwell Science.
- Aron, A., Aron, E. N., & Coups, E. J. (2016). *Statistics for Psychology* (7th ed.). Pearson.
- Burtis, C. A., Ashwood, E. R., & Bruns, D. E. (2012). *Tietz Textbook of Clinical Chemistry and Molecular Diagnostics*. Saunders.
- Couper, M. P., Tourangeau, R., & Conrad, F. G. (Eds.). (2017). *The Science of Web Surveys*. Oxford University Press.
- Creswell, J. W. (2014). *Research Design: Qualitative, Quantitative, and Mixed Methods Approaches*. SAGE Publications.
- DeVon, H. A., Block, M. E., Moyle-Wright, P., Ernst, D. M., Hayden, S. J., Lazzara, D. J., & Kostas-Polston, E. (2007). A psychometric toolbox for testing validity and reliability. *Journal of Nursing Scholarship*, 39(2), 155-164.
- Dillman, D. A., Smyth, J. D., & Christian, L. M. (2014). *Internet, Phone, Mail, and Mixed-Mode Surveys: The Tailored Design Method*. Wiley.
- Fetterman, D. M. (2010). *Ethnography: Step-by-Step*. SAGE Publications.
- Field, A. (2013). *Discovering Statistics Using IBM SPSS Statistics*. Sage Publications.
- Fowler Jr, F. J. (2013). *Survey research methods*. Sage publications.
- Gravetter, F. J., & Forzano, L. B. (2018). *Research Methods for the Behavioral Sciences*. Cengage Learning.
- Hogg, R. V., McKean, J., & Craig, A. (2018). *Introduction to Mathematical Statistics* (8th ed.). Pearson.

- Hulley, S. B., Cummings, S. R., Browner, W. S., Grady, D. G., & Newman, T. B. (2013). *Designing Clinical Research*. Lippincott Williams & Wilkins.
- Israel, G. D. (2003). *Sampling the Evidence of Extension Program Impact*. University of Florida IFAS Extension.
- Jerry W. W, (2007). *Observational Research Methods*. SAGE Publications.
- Krueger, R. A., & Casey, M. A. (2014). *Focus Groups: A Practical Guide for Applied Research*. SAGE Publications.
- Kvale, S. (2007). *Doing Interviews*. SAGE Publications.
- Levin, B., & Fox, J. A. (2009). *Elementary Statistics in Social Research*. Allyn & Bacon.
- Levin, K. A. (2006). Study design III: Cross-sectional studies. *Evidence-Based Dentistry*, 7(1), 24-25.
- Levin, R. I., Fox, J. A., & Forzano, L. (2018). *Elementary Statistics in Social Research* (12th ed.). Pearson.
- Lohr, S. (2010). *Sampling: Design and Analysis*. Cengage Learning.
- Lohr, S. (2019). *Sampling: Design and Analysis*. Cengage Learning.
- Neuendorf, K. A. (2002). *The Content Analysis Guidebook*. SAGE Publications.
- Pagano, M., & Gauvreau, K. (2000). *Principles of Biostatistics*. Duxbury Press.
- Rothman, K. J., Greenland, S., & Lash, T. L. (2008). *Modern Epidemiology*. Lippincott Williams & Wilkins.
- Sheskin, D. J. (2007). *Handbook of Parametric and Nonparametric Statistical Procedures* (4th ed.). Chapman & Hall/CRC.
- Tabachnick, B. G., & Fidell, L. S. (2013). *Using Multivariate Statistics* (6th ed.). Pearson.
- Thompson, S. K. (2012). *Sampling* (3rd ed.). John Wiley & Sons.

Triola, M. F. (2017). Elementary Statistics. Pearson.

Trochim, W. M., & Donnelly, J. P. (2008). Research Methods: The Essential Knowledge Base. Cengage Learning.

Yin, R. K. (2014). Case Study Research: Design and Methods. SAGE Publications.

BAB 9

METODE PENGUMPULAN DATA

Oleh Asruria Sani Fajriah

9.1 Pendahuluan

Bab ini membahas lebih rinci tentang bagaimana kita sebagai peneliti membangun sebuah kuesioner yang digunakan untuk mendapatkan data-data dari para subjek penelitian atau responden. Teknik pengumpulan data ini sangat penting untuk menghimpun data dari sekian banyak responden yang jumlahnya besar dan secara geografik lokasinya tersebar (Zaki *et al.*, 2018).

Salah satu tahapan utama dalam melakukan penelitian adalah pengumpulan data yang memungkinkan peneliti untuk menemukan jawaban atas pertanyaan penelitian. Pengumpulan data adalah proses pengumpulan data yang bertujuan untuk mendapatkan wawasan mengenai topik penelitian. Ada berbagai jenis data dan metode pengumpulan data yang berbeda dan mungkin beberapa peneliti merasa sulit untuk memilih jenis pengumpulan data yang paling tepat (Taherdoost, 2021).

Metode pengumpulan data digunakan untuk mengumpulkan informasi mengenai variabel tertentu dari penelitian yang bertujuan untuk menggunakan data tersebut pada saat analisis data untuk mencapai hasil penelitian, mendapatkan jawaban dari pertanyaan penelitian atau menguji hipotesis. Pengumpulan data sebagai tahap utama dalam penelitian dapat mempengaruhi kualitas hasil penelitian dengan meminimalisir kemungkinan kesalahan yang mungkin terjadi selama proyek penelitian. Oleh

karena itu, selain desain penelitian yang baik, berkualitas penelitian juga dapat dipengaruhi oleh pengumpulan data untuk mendapatkan hasil yang tepat karena data yang tidak memadai dan tidak akurat dapat menurunkan tingkat keakuratan temuan pada hasil penelitian. Namun di sisi lain, meskipun metode pengumpulan data yang sesuai dapat membantu merencanakan penelitian yang baik, tetapi tidak bisa menjamin keberhasilan keseluruhan proyek penelitian (Taherdoost, 2021). Karena perubahan norma memerlukan penentuan dari persepsi dan cara berpikir peneliti tentang isu-isu tertentu, metode pengumpulan data biasanya akan mencakup survei, kelompok fokus atau pertanyaan wawancara (Reisman, Gienapp and Stachowiak, 2023).

Saat ini, sudah sangat jelas bahwa penelitian harus melibatkan pengumpulan data dan ada berbagai cara untuk melakukannya. Misalnya, jika kita ingin mengukur perilaku agresif pada anak-anak, kita dapat mengumpulkan data tersebut dengan mengamati anak-anak dengan mata kita, dengan menggunakan peralatan untuk mengukur kekuatan yang mereka gunakan untuk memukul suatu objek, dengan memeriksa catatan kejahatan remaja, dengan mensurvei orang tua dan guru, atau dengan memberikan skala agresi kepada anak-anak. Ini hanyalah contoh dari metode yang memungkinkan; Kami yakin Anda bisa membayangkan banyak contoh yang lain. Namun, contoh-contoh di atas menggambarkan beberapa metode yang berbeda yang dapat digunakan untuk mengumpulkan data. Seperti kebanyakan desain penelitian, teknik setiap metode memiliki kelebihan dan keterbatasan (Lammers and Badia, 2020).

9.2 Jenis Data

Data adalah bentuk jamak dari datum yang berarti "sepotong informasi". Informasi dapat dikumpulkan selama penelitian, atau sebagai hasil percobaan, atau selama pengamatan, atau melalui sensus atau survei (Mazhar *et al.*, 2021).

Sebelum memilih metode pengumpulan data, jenis data yang diperlukan untuk penelitian harus ditentukan terlebih dahulu. Bagian ini bertujuan untuk memberikan ringkasan tipe data untuk melalui metode pengumpulan data. Namun, kita perlu memahami apa sebenarnya data itu? Informasi yang terkandung dalam bentuk angka atau fakta yang digunakan untuk menganalisis untuk perhitungan yang berbeda dan akhirnya mendapatkan hasil untuk menjawab pertanyaan penelitian atau pengujian hipotesis dikenal sebagai data (Taherdoost, 2021). Data dapat dikategorikan menggunakan berbagai cara termasuk kuantitatif dan kualitatif.

9.2.1 Data Kualitatif

Baik data non-numerik nominal maupun deskriptif yang tidak dapat ditampilkan sebagai angka dan disediakan dalam format kata atau kalimat disebut juga data kualitatif. Jenis data ini menjawab pertanyaan "bagaimana dan mengapa" dalam studi penelitian dan sebagian besar mencakup data mengenai perasaan, persepsi, dan emosi menggunakan pendekatan tidak terstruktur seperti wawancara untuk pengumpulan data. Para peneliti menggunakan metode yang berbeda seperti menggunakan kaset audio, sketsa, catatan, dan foto untuk mengumpulkan data ini (Taherdoost, 2021).

Data kualitatif dapat disesuaikan agar peneliti dapat mengumpulkan informasi lebih lanjut untuk mengeksplorasi dan menentukan data yang baru. Tidak hanya itu, data kualitatif dapat digunakan untuk mengumpulkan informasi mengenai efek dan

konsekuensi dari program pada penelitian, dan akhirnya meningkatkan kualitas hasil kuantitatif. Implementasinya tergantung pada pengeluaran sejumlah besar biaya dan waktu dan hasilnya mungkin tidak dapat digeneralisasikan. Ini berarti temuan studi kasus dapat digunakan hanya untuk masalah yang sama dengan pola umum untuk studi yang berbeda (Taherdoost, 2021)

9.2.2 Data Kuantitatif

Data Kuantitatif adalah data numerik yang dihasilkan dan dihitung secara matematis. Ada skala yang berbeda untuk mengukur data kuantitatif termasuk nominal, ordinal, interval, dan rasio. Terdapat skala yang berbeda untuk mengukur data kuantitatif termasuk skala nominal, ordinal, interval, dan rasio. Skala dapat dikategorikan menjadi dua jenis umum yaitu "Skala Peringkat dan Skala Sikap". Skala peringkat menetapkan nilai numerik ke poin atau kategori untuk mengevaluasinya. Di sisi lain, metode yang lebih kompleks adalah skala sikap yang menentukan kecenderungan orang terhadap individu, fenomena, atau objek apa pun. Metode kuantitatif membahas jenis pertanyaan "apa" dalam sebuah penelitian. Pendekatan ini menggunakan metode pengumpulan data terstruktur dan didasarkan pada pengambilan sampel acak. Dibandingkan dengan metode kualitatif, metode ini dianggap lebih murah, dan temuan dapat distandarisasi untuk mencapai hasil lain berdasarkan beberapa kriteria seperti ukuran. Temuan ini dapat dengan mudah digeneralisasi dan diringkas juga. Perbandingan sederhana antara hasil juga dimungkinkan. Namun demikian, metode ini juga dapat menghadapi perbedaan yang tidak terduga dan beberapa kesulitan karena kapasitas implementasi dan investigasi terbatas dalam metode ini. Pendekatan ini menggunakan metode yang berbeda seperti eksperimen dan wawancara terstruktur untuk pengumpulan data yang dibahas di bagian metode pengumpulan data (Taherdoost, 2021).

9.3 Metode Pengumpulan Data

9.3.1 Kuesioner

Jenis prosedur pengumpulan data ini sangat populer dalam kasus survei besar. Metode ini diadopsi oleh peneliti, tokoh swasta, organisasi swasta dan bahkan oleh pemerintah. Dalam metode ini, kuesioner ditujukan kepada individu yang bersangkutan kemudian diminta untuk menanggapi pertanyaan dan mengumpulkan kembali kuesioner yang telah diisi. Kuesioner terdiri dari sejumlah pertanyaan yang dicetak atau diketik dalam urutan tertentu pada formulir atau serangkaian formulir. Kuesioner dikirimkan kepada responden yang diharapkan mampu menjawab pertanyaan, memahami isi pertanyaan, dan memberikan tanggapan mereka terhadap pertanyaan-pertanyaan di ruangan dan waktu yang dialokasikan untuk mengisi kuesioner. Para responden diwajibkan untuk memberikan jawaban mereka sendiri atas pertanyaan yang ada (Mazhar et al., 2021).

Kuesioner adalah serangkaian pertanyaan yang telah ditentukan sebelumnya yang diberikan kepada sejumlah responden. Instrumen ini bagus untuk mendapatkan informasi dari banyak orang, tersebar di area yang luas, dan yang tidak mudah untuk kontak tatap muka. Kuesioner harus memiliki penjelasan singkat tentang seperti apa penelitian yang sedang dilakukan. Kuesioner harus selalu mematuhi kode etik dan moral perilaku.

Fitur Kuesioner yang Baik antara lain:

1. Pertanyaan harus lebih kecil dan to the point
2. Pertanyaan harus berurutan
3. Pertanyaan harus dilanjutkan dalam urutan logis bergerak dari pertanyaan yang mudah hingga yang lebih sulit.
4. Istilah teknis tidak boleh digunakan dalam kuesioner
5. Pertanyaan mungkin dikotomis, pilihan ganda atau terbuka

6. Pertanyaan yang dapat mempengaruhi jawaban responden harus dihindari
7. Pertanyaan harus mencakup semua spesifikasi
(Mazhar *et al.*, 2021)

9.3.2 Wawancara

Berdasarkan pernyataan Muhammad Sajjad Kabir (2016), wawancara adalah kegiatan pengumpulan data dengan cara mengajukan pertanyaan dan mendapatkan jawaban dari peserta dalam sebuah penelitian. Wawancara memiliki berbagai bentuk termasuk wawancara individu, tatap muka, dan wawancara kelompok tatap muka. Mengajukan dan menjawab pertanyaan dapat dimediasi oleh telepon atau perangkat elektronik lainnya (misalnya komputer). Wawancara dapat dibedakan sebagai berikut;

1. Wawancara Terstruktur

Karakteristik wawancara terstruktur adalah;

- a. Pewawancara menanyakan serangkaian pertanyaan yang sama kepada setiap responden.
- b. Pertanyaan dibuat sebelum wawancara dan seringkali memiliki serangkaian kategori respons terbatas.
- c. Biasanya ada sedikit ruang untuk variasi dalam tanggapan dan ada beberapa pertanyaan terbuka yang termasuk dalam panduan wawancara.
- d. Pertanyaan distandarisasi dan urutan maupun ungkapan pertanyaan tetap konsisten dari wawancara ke wawancara.
- e. Pewawancara memainkan peran netral dan bertindak santai dan ramah, tetapi tidak memasukkan pendapatnya dalam wawancara.
- f. Kuesioner yang dikelola sendiri adalah jenis wawancara terstruktur

2. Wawancara Semi-terstruktur

Karakteristik wawancara semi-terstruktur adalah;

- a. Pewawancara dan responden terlibat dalam wawancara formal.
- b. Pewawancara mengembangkan dan menggunakan 'panduan wawancara'. Ini adalah daftar pertanyaan dan topik yang perlu dibahas selama percakapan, biasanya dalam urutan tertentu.
- c. Pewawancara mengikuti panduan, tetapi dapat mengikuti lintasan topikal dalam percakapan yang mungkin menyimpang dari panduan ketika dia merasa ini tepat

3. Wawancara Tidak Terstruktur

Karakteristik wawancara tidak terstruktur adalah;

- a. Pewawancara dan responden terlibat dalam wawancara formal karena mereka memiliki waktu yang dijadwalkan untuk duduk dan berbicara satu sama lain dan kedua belah pihak mengakui ini sebagai wawancara.
- b. Pewawancara memiliki rencana yang jelas dalam pikiran mengenai fokus dan tujuan wawancara. Ini memandu diskusi.
- c. Tidak ada panduan wawancara terstruktur. Sebaliknya, pewawancara membangun hubungan dengan responden, membuat responden terbuka dan mengekspresikan diri dengan cara mereka.
- d. Pertanyaan cenderung terbuka dan mengungkapkan sedikit kendali atas tanggapan informan.
- e. Wawancara etnografi dan mendalam tidak terstruktur. Tiga jenis wawancara mendalam etnografi tidak terstruktur - sejarah lisan, wawancara kreatif, dan wawancara postmodern.

9.3.3 Focus Group Discussion (FGD)

Berdasarkan pernyataan dari Kumar Sahoo (2022), bahwa focus Group Discussion merupakan kegiatan wawancara kelompok di mana peneliti mencoba mengumpulkan pemahaman bersama dari beberapa individu pada saat yang bersamaan. Dalam jenis wawancara ini, pewawancara mengundang orang-orang yang tertarik pada topik umum yang sama untuk berkumpul dan membahasnya. Di dalam metode ini, responden yakin bahwa mereka akan bebas mengekspresikan diri dengan kata-kata mereka sendiri dan untuk menanggapi tidak hanya kepada pewawancara tetapi juga kepada peserta lain. Ketika melakukan FGD, penting juga untuk memastikan bahwa semua peserta memiliki suara mereka dan bukan sesuatu yang didominasi oleh satu atau dua peserta saja. Dengan menggunakan jadwal wawancara terstruktur atau semi-terstruktur, pewawancara dapat mengajukan pertanyaan kepada kelompok dan mendorong semua peserta untuk merespons. Beberapa keuntungan dari FGD adalah bahwa mereka merespon tidak hanya kepada peneliti tetapi juga untuk peserta lain. Interaksi antara peserta biasanya mengungkapkan lebih banyak tentang sudut pandang subjek daripada yang akan terjadi dengan wawancara yang didominasi pewawancara. Ini memungkinkan peneliti untuk melihat bagaimana subjek menggabungkan sudut pandang orang lain dalam menyusun pemahaman mereka. Selain keuntungan yang disebutkan di atas, wawancara kelompok mungkin bermanfaat dalam situasi berikut:

1. Untuk mengembangkan pertanyaan penelitian dengan memanfaatkan kebijaksanaan kolektif peserta.
2. Ketika peneliti tidak memiliki keahlian atau keterampilan yang diperlukan untuk melakukan wawancara individu yang mendalam dan bermaksud untuk mendorong peserta untuk berkomunikasi.

3. Di mana mungkin ada berbagai sudut pandang, dan jumlah peserta setuju atau tidak setuju pada apa pun yang menarik bagi peneliti.
4. Untuk mengevaluasi proposisi teoritis peneliti, berdasarkan pengumpulan data dan analisis masa lalu.

9.3.4 Observasi

Observasi digunakan dalam penelitian sebagai metode untuk mengumpulkan data tentang orang, proses, dan budaya. Observasi juga merupakan alat yang digunakan secara teratur untuk mengumpulkan data oleh peneliti, pekerja sosial di lingkungan masyarakat, dan oleh psikolog yang merekam perilaku manusia. Observasi adalah deskripsi sistematis dari peristiwa, perilaku, dan artefak dari penelitian. Ada dua jenis observasi partisipan dan observasi langsung. Observasi partisipan yang melibatkan pengamat dan peserta. Sedangkan observasi langsung merupakan kegiatan untuk mengamati tanpa berinteraksi dengan objek atau orang yang diteliti (Kawulich, 2012).

Observasi digunakan dalam studi kuantitatif dan kualitatif. Misalnya, dalam percobaan titrasi kimia, peneliti dapat mengamati tingkat produk yang dikeluarkan ke dalam tabung reaksi atau gelas kimia yang dikalibrasi (kuantitatif). Dalam pendidikan, guru/peneliti dapat mengamati siswa untuk menentukan berapa kali pelanggaran perilaku terjadi atau kegiatan yang memicu pelanggaran tertentu (kuantitatif atau kualitatif). Dalam psikologi, peneliti dapat mengamati perilaku untuk menentukan bagaimana pasien/klien bereaksi terhadap rangsangan tertentu (kualitatif). Dalam linguistik, peneliti dapat mengamati bagaimana responden mengekspresikan diri dalam situasi tertentu (kuantitatif atau kualitatif). Apapun disiplin ilmunya, observasi banyak digunakan sebagai metode pengumpulan data (Kawulich, 2012).

9.3.5 Survei

Survei adalah metode yang tepat untuk menentukan perasaan, pendapat, dan pikiran. Tujuan dari Survei dapat bersifat global dan khusus. Mereka dapat memberikan sejumlah besar data menggunakan panggilan telepon, email, atau wawancara tatap muka. Di sisi lain, data dapat dikumpulkan dalam survei penyelesaian sendiri atau oleh pewawancara. Survei dapat digunakan untuk mengeksplorasi perilaku sosial dan populasi target yang besar serta ketika data rinci dan mendalam tidak diperlukan di dalam penelitian. Dalam survei, serangkaian pertanyaan disediakan untuk memberikan sampel yang dipilih dari populasi target tertentu. Sampel ini menyajikan karakteristik dan perilaku populasi. Pemilihan sampel merupakan tahap penting dalam proses ini yang dapat sangat mempengaruhi temuan. Ukuran sampel harus dipilih berdasarkan kemungkinan memilih setiap peserta dengan peluang bukan nol (Taherdoost, 2021).

9.3.6 Studi Kasus

Studi kasus adalah metode pengumpulan data di mana informasi deskriptif mendalam tentang entitas tertentu, atau kasus, dikumpulkan, diatur, ditafsirkan, dan disajikan dalam format naratif. Laporan studi kasus pada dasarnya adalah sebuah cerita. Subjek kasus dapat berupa individu, keluarga, lingkungan, kelompok kerja, ruang kelas, sekolah, organisasi, program, atau entitas lainnya. Studi kasus juga dapat berfokus pada peristiwa sosial atau alam seperti enam bulan pertama supervisor baru di tempat kerja, reaksi karyawan terhadap akuisisi organisasi mereka oleh perusahaan lain, atau respons masyarakat terhadap bencana alam. Sebagai pendekatan pengumpulan data, ini diterapkan secara luas dalam sosiologi, antropologi, psikologi, pendidikan, dan kedokteran dan menawarkan banyak nilai potensial untuk teknologi kinerja. Studi kasus menawarkan perspektif dan

wawasan yang kaya dan dapat mengarah pada pemahaman mendalam tentang variasi dan masalah (Marelli, 2007).

9.3.7 Eksperimen

Eksperimen juga merupakan salah satu metode utama pengumpulan data. Dalam metode ini, peneliti dapat membuat situasi penelitian, memilih peserta, dan memanipulasi variabel independen (seringkali terbatas dan jumlah kecil) yang dapat mempengaruhi variabel dependen, dan mengeksplorasi bagaimana mereka menaungi variabel dependen. Secara singkat, ia menemukan hubungan sebab akibat di antara mereka yang dikenal sebagai hubungan efek dan penyebab. Metode ini objektif, dan pemikiran para peneliti tidak mempengaruhi hasil. Oleh karena itu, validitasnya tinggi dan efek bias sangat minimal. Namun, bias yang mungkin masih dapat berasal dari berbagai aspek diantaranya; Seleksi; Efek plasebo, ketika keyakinan tentang subjek penelitian dapat mempengaruhi temuan; Eksperimen, ketika Anda sebagai peneliti memiliki keyakinan khusus tentang subjek (Taherdoost, 2021).

9.3.8 Diaries

Penelitian Diary atau *Experience Sampling Method* merupakan metode pengambilan data secara berulang dalam jangka waktu tertentu yang dilakukan dalam kegiatan sehari-hari subjek penelitian. Metode pengumpulan data menggunakan ESM memungkinkan peneliti untuk mengambil data real-time, di luar batasan laboratorium, dan dalam konteks kehidupan sehari-hari. ESM bisa digunakan untuk merekam variabel-variabel seperti memori, pikiran, perasaan, *mood*, dan perilaku yang bersifat episodik (terdiri atas beberapa peristiwa/ kejadian yang saling berhubungan) di kehidupan sehari-hari. ESM berbeda dengan tiga metode pengumpulan momentary data lain, yaitu: *thought-*

sampling (i.e., secara khusus mengukur pemikiran batin dengan mengabaikan peristiwa eksternal); *descriptive-experience sampling* (i.e., hanya mengumpulkan data kualitatif) dan *ecological momentary assessment* (i.e., memiliki karakter sama dengan ESM hanya saja, EMA juga mempertimbangkan elemen lingkungan yang berkaitan dengan pengalaman). Meskipun hampir empat dekade telah berlalu semenjak ESM pertama kali dilakukan, riset-riset menggunakan metode ini di Indonesia masih terbatas. Pencarian menggunakan Google scholar dengan kata kunci “experience sampling methods” menghasilkan artikel yang hampir semuanya merupakan artikel berbahasa inggris dengan riset yang dilakukan di negara-negara barat. Sementara itu, melalui kata kunci “diary studies”, “diary study”, experience sampling method”, dan “ESM” peneliti tidak menemukan satu artikel pun di <https://sinta.ristekbrin.go.id/>. Secara khusus, ESM masih terhitung sangat baru di Indonesia. Padahal berdasarkan riset-riset yang telah dilakukan, ESM memiliki potensi yang besar dalam mengembangkan penelitian-penelitian di bidang perilaku, emosi, mood bahkan kecanduan zat adiktif semacam alkohol, narkoba dan rokok (Apriliawati, 2020).

9.3.9 Aktivitas Sampling Teknik

Metode ini menggunakan pengamatan berturut-turut pada individu, kelompok, dan bahkan mesin dalam suatu periode tertentu kerangka waktu untuk merekam peristiwa yang terjadi dan mencapai peringkat jika diperlukan. Di sini, tujuan utamanya adalah untuk mencatat penundaan dan persentase pekerjaan yang dilakukan selama periode tertentu. Ini menggunakan teknik pengambilan sampel untuk menentukan aktivitas subjek pengganti dan kemudian menganalisisnya menggunakan metode statistik. Jumlah pengamatan merupakan faktor kunci karena dapat mempengaruhi keakuratan penelitian. Namun, karena ada batasan

karena waktu dan biaya, jumlah optimal harus digunakan berdasarkan perkiraan jumlah kemunculan faktor dan keakuratan penelitian yang dapat diterima (Taherdoost, 2021).

9.3.10 Motion Memo Study

Metode ini juga dikenal sebagai *spaced-shot photography*, dan menggunakan kamera untuk menganalisis proses yang panjang. Sebagai alat waktu, ia menggunakan bahan film untuk tujuan belajar. Saat ini, teknik ini digunakan secara umum, karena banyak pengaturan saat ini dapat mengatasi kondisi yang diperlukan (Taherdoost, 2021).

9.3.11 Data Sekunder

Data sekunder didefinisikan sebagai data yang dikumpulkan dari sumber yang dipublikasikan. Data yang dikumpulkan ini dapat berasal dari sumber kualitatif seperti laporan wawancara dan data sumber kuantitatif seperti sensus. Metode pengumpulan data sekunder, secara umum, dapat dikategorikan menurut tabel Tabel 9.1 berikut ini.

Data Sekunder	Deskripsi dan Kredibilitas
Sumber cetak yang dipublikasikan	Dalam jenis ini, penulis, publikasi, dan waktu penerbitan adalah penting
Buku	Yang paling otentik di antara sumber-sumber sekunder yang lain
Jurnal	Metode yang paling penting karena up-to-date dan dapat menyediakan data atau informasi yang sangat spesifik.

Majalah/ Koran	Sangat tidak dapat diandalkan tetapi dapat efektif juga dalam penelitian
E-Jurnal	Sangat cepat dan mudah diakses
Website	Metode yang umum digunakan dan menyediakan buku harian yang andal
Rekam medis	Berguna dalam beberapa kasus
Diaries	Metode yang umum digunakan dan menyediakan buku harian yang andal
Surat	Harus diperiksa dalam hal keandalan sebelum menggunakan
Catatan pemerintah	Sangat berguna dalam berbagai cabang ilmu seperti penelitian manusia, ilmu sosial, dll.
Data sensus/ statistik	Sertakan catatan pendidikan dan kesehatan
Catatan Sektor Publik	Disediakan oleh perusahaan swasta dan survei LSM

Sumber: (Taherdoost, 2021)

DAFTAR PUSTAKA

- Apriliawati, D. (2020) 'Diary Study sebagai Metode Pengumpulan Data pada Riset Kuantitatif: Sebuah Literature Review', *Journal of Psychological Perspective*, 2(2), pp. 79–89. doi: 10.47679/jopp.022.12200007.
- Kawulich, B. (2012) 'Collecting Data Through Observation', in C. Wagner, B. Kawulich, and M. Garner (eds) *Doing Social Research: A global context*. Georgia: McGraw Hill, pp. 150–160. Available at: https://www.researchgate.net/publication/257944783_Collecting_data_through_observation (Accessed: 15 January 2024).
- Kumar Sahoo, R. (2022) 'Interview as a Tool for Data Collection in Educational Research', in. Delhi: Lucky International. Available at: https://www.researchgate.net/publication/360313105_Interview_as_a_Tool_for_Data_Collection_in_Educational_Research (Accessed: 15 January 2024).
- Lammers, J. W. and Badia, P. (2020) 'Methods of Data Collection', in *Fundamentals of Behavioral Research*, pp. 1–26. Available at: <https://uca.edu/psychology/files/2013/08/Ch6-Methods-of-Data-Collection.pdf>.
- Marelli, A. F. (2007) 'Collecting Data Through Case Study', *Performance Improvement*, 46(9), pp. 9–16. doi: 10.1002/pfi.
- Mazhar, S. A. *et al.* (2021) 'Methods of Data Collection: A Fundamental Tool of Research', *Journal of Integrated Community Health*, 10(01), pp. 6–10. doi: 10.24321/2319.9113.202101.
- Muhammad Sajjad Kabir, S. (2016) 'Methods of Data Collection', in. Bangladesh: Book Zone Publication, pp. 202–276. Available at: https://www.researchgate.net/publication/325846997_METHODS_OF_DATA_COLLECTION (Accessed: 15 January 2024).

- Reisman, J., Gienapp, A. and Stachowiak, S. (2023) 'A Handbook of Data Collection Tools: Companion To "A Guide To Measuring Advocacy and Policy"', p. 45. Available at: https://www.orsimpact.com/DirectoryAttachments/132018_31904_823_a_handbook_of_data_collection_tools.pdf.
- Taherdoost, H. (2021) 'Data Collection Methods and Tools for Research; A Step-by-Step Guide to Choose Data Collection Technique for Academic and Business Research Projects', *International Journal of Academic Research in Management (IJARM)*, 2021(1), pp. 10–38. Available at: <https://hal.science/hal-03741847>.
- Zaki, A. et al. (2018) *Metoda Pengumpulan dan Teknik Analisis Data*. Yogyakarta: ANDI.

BIODATA PENULIS



Reza Fahlevi, S.K.M., M.M., M.Psi., Psikolog

Dosen Prodi Psikologi

Universitas Tarumanagara Jakarta

Penulis lahir di Palembang pada tanggal 23 Desember 1991. Penulis merupakan Dosen Tetap Program Studi Psikologi Universitas Tarumanagara dan Praktisi Psikolog Klinis dengan bidang minat terhadap Psikologi Kesehatan dan Psikologi Positif. Penulis telah menyelesaikan pendidikan S2 di Magister Profesi Psikologi Universitas Tarumanagara pada bidang Psikologi Klinis. Penulis juga telah menyelesaikan pendidikan S2 di Magister Manajemen Universitas Tridinanti bidang Sumber Daya Manusia serta S1 Kesehatan Masyarakat di Universitas Sriwijaya dan S1 Psikologi di Universitas Bina Darma. Saat ini kesibukan dari Penulis selain sebagai Dosen dan Psikolog, Penulis juga aktif sebagai Trainer dan Pengurus Asosiasi Psikologi Kesehatan Indonesia (APKI HIMPSI) sebagai Wakil Ketua APKI. Penulis juga sudah menulis beberapa book chapter antara lain: Psikologi klinis, Psikologi Positif, Perilaku Manusia, Psikologi Abnormal, Psikologi Kesehatan, Pengantar Ilmu Komunikasi, Psikologi Sosial, Psikologi Kepemimpinan, Psikologi Perkembangan, Pendidikan Anti Korupsi di Perguruan Tinggi, Psikologi Sekolah, Gerontolgi, Psikologi

Perkembangan Peserta Didik, Manajemen Kinerja dan Psikologi
Kepribadian Anak

Email: reifahlevipsy@gmail.com

Instagram : reifahlevipsy

Linkedin : <https://www.linkedin.com/in/reifahlevi/>

BIODATA PENULIS



Dr.Ir. Patrich Phill Edrich Papilaya. M.Sc.Forest.Trop.

Dosen Program Studi Pascasarjana Manajemen Hutan
Universitas Pattimura

Penulis lahir di Ambon Tanggal 14 Pebruari 1967. Penulis adalah dosen tetap pada Program Studi Manajemen Hutan Pascasarjana. Universitas Pattimura. Ambon. Menyelesaikan pendidikan S1 pada Jurusan Kehutanan Fakultas Pertanian Universitas Pattimura. menyelesaikan Pendidikan S2 pada Georg-August Universitas. Faculty of Forestry. Göttingen. dalam bidang Remote sensing dan GIS. Pendidikan S3 diselesaikan pada IPB University. Bogor

BIODATA PENULIS



Wara Alfa Syukrilla, M.Sc.
Dosen Statistika
UIN Syarif Hidayatullah Jakarta

Penulis adalah dosen tetap statistika pada Fakultas Psikologi UIN Syarif Hidayatullah Jakarta. Penulis menyelesaikan pendidikan S1 pada Jurusan Statistika di Universitas Brawijaya dan melanjutkan S2 Master of Statistics spesialisasi *Biostatistics* di Universiteit Hasselt, Belgia dengan beasiswa penuh dari pemerintah Belgia yaitu beasiswa VLIR UOS. Penulis menekuni bidang statisika, spatial statistics, psychometrics, dan biostatistics. Saran yang membangun terkait tulisan ini dapat dikomunikasikan dengan penulis melalui email waraalfa@gmail.com.

BIODATA PENULIS



Putu Erma Pradnyani, S.KM., M.Kes.

Dosen Program Studi Manajemen Informasi Kesehatan
Politeknik Kesehatan Kartini Bali

Penulis lahir di Sleman tanggal 2 Maret 1995. Penulis adalah dosen tetap pada Program Studi Manajemen Informasi Kesehatan, Politeknik Kesehatan Kartini Bali. Menyelesaikan pendidikan S1 pada Jurusan ilmu Kesehatan masyarakat tahun 2016 dan melanjutkan S2 pada Jurusan Kesehatan masyarakat Universitas Airlangga, Surabaya. Penulis menekuni bidang Kesehatan masyarakat terutama biostatistika dan epidemiologi, manajemen pelayanan Kesehatan, serta mengarah pada Kesehatan ibu dan anak.

Penulis aktif dalam mengikuti kegiatan organisasi baik diperguruan tinggi maupun tingkat Provinsi. Penulis bergabung menjadi volunteer di Kita Sayang Remaja PKBI Bali, bergabung menjadi pengurus IAKMI Bali serta menjadi Ketua Inkubator Bisnis di Politeknik Kesehatan Kartini Bali. Penulis juga aktif menekuni bidang Menulis dan memiliki publikasi Buku, SCOPUS maupun Jurnal Nasional Terakreditasi. Beberapa publikasi yang penulis terbitkan adalah The effects of socio-demographic characteristics on Indonesian women's knowledge of HIV/AIDS: A cross-sectional study, Vulnerability to domestic physical violence among married women in Indonesia. Determinan Keberhasilan Pengobatan Pasien Tuberkulosis di Kota Denpasar Tahun 2021, dan 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



Ardiana Fatma Dewi, S.tat., M.Stat.

Dosen Program Studi Tadris Matematika

Fakultas Tarbiyah

Institut Agama Islam Negeri (IAIN) Kediri

Penulis lahir di Kediri pada tanggal 23 Desember 1996. Anak kedua dari tiga bersaudara. Penulis adalah dosen tetap pada Program Studi Tadris Matematika Fakultas Tarbiyah IAIN Kediri. Menyelesaikan pendidikan S1 Jurusan Statistika di Universitas Brawijaya (UB) Malang pada tahun 2019. Penulis melanjutkan S2 pada Jurusan Statistika di Institut Teknologi Sepuluh Nopember (ITS) Surabaya menyelesaikan studi pada tahun 2021.

Penulis aktif menjadi dosen sejak tahun 2022, dan awal mula berkarir di IAIN Kediri dengan beberapa mata kuliah statistika dan matematika yang diampu. Termasuk dipercaya untuk mengampu mata kuliah trigonometri. Berawal dari bangku kuliah yang dipercaya dosen menjadi asisten mata kuliah *Time Series* sehingga membawa peneliti memilih berkarir menjadi seperti yang sekarang ini menjadi seorang pengajar. Penulis menekuni bidang meneliti dan menulis dalam berbagai bidang statistika dan matematika. Penulis dapat dihubungi melalui *e-mail* berikut ardianafatmadewi@iainkediri.ac.id

BIODATA PENULIS



Dr. Ayatullah Harun, S.Si.,SKM.,M.Kes.

Dosen Program Studi DIII Kebidanan
Institut Ilmu Kesehatan Pelamonia Makassar

Penulis lahir di Pinrang tanggal 13 April 1981. Penulis adalah dosen tetap pada Program Studi D III Kebidanan Institut Ilmu Kesehatan Pelamonia Makassar. Menyelesaikan pendidikan S1 pada Jurusan Matematika Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Negeri Makassar Selesai tahun 2004 dan melanjutkan S2 pada program studi Kesehatan Masyarakat Universitas Hasanuddin selesai tahun 2011, S3 pada Program studi Ilmu Kesehatan Masyarakat Fakultas Kesehatan Masyarakat Universitas Hasanuddin selesai tahun 2023.

BIODATA PENULIS



Dr.Arif Rachman,
drg.,SH.,MH.,MM.,MTr.Hanla.,Sp.Pros.,CIQnR.,CIQaR.
Dosen Program Studi Sarjana Kedokteran, Departemen
Biomolekuler, sel dan genetika
Fakultas Kedokteran Militer, Universitas Pertahanan Republik
Indonesia

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

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). Karya Tulis dalam bentuk buku dan jurnal penelitian adalah:

Buku Referensi dan Bahan Ajar :

- 1) Metodologi Penelitian Kuantitatif, Kualitatif, dan R&D, CV. Saba Jaya Publisher ISSN: 978-623-09-7582-0, Januari 2024
- 2) Fakultas Kedokteran Militer Universitas Pertahanan Cita dan Jejak Sejarah, Unhan Press, Januari Tahun 2021.
- 3) Bioetika Multidisplin, Multidisiciplinary Prespectives on Bioethics, Unhan Press, Januari Tahun 2022.
- 4) Aesthethic 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
- 5) Biocompatibility of Yttria-Tetragonal Zirconia Polycrystal Seeded with Human Adipose Derived Mesenchymal Stem Cell, ACTA INFORM MED. 2018 DEC; 26(4): 249-253 (Q3)
- 6) 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)
- 7) 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)
- 8) CBCT: Tracing Dominant Region of Periodontitis on Psychological Stress, Sys Rev Pharm. 2020;11(11):1567-1572 (Q2)
- 9) Ideal Bone Defects Distance on Orthodontic Tooth Movement for Preparation of hADMSC-Scaffold Chitosan Intervention, Sys Rev Pharm. 2020;11(8):571-575 (Q2)

- 10) 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).
- 11) 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).
- 12) 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).

BIODATA PENULIS



Asruria Sani Fajriah, SST., M.KM.

Dosen Program Studi S1 Kebidanan
Fakultas Keperawatan dan Kebidanan
Institut Ilmu Kesehatan STRADA Indonesia

Penulis lahir di Kediri tanggal 29 Juni 1992. Penulis adalah dosen tetap pada Program Studi Studi S1 Kebidanan Fakultas Keperawatan dan Kebidanan, Institut Ilmu Kesehatan STRADA Indonesia. Menyelesaikan pendidikan D3 pada Jurusan Kebidanan, kemudian melanjutkan S1/D4 pada Jurusan Bidan Pendidik, dan melanjutkan pendidikan S2 pada Jurusan Ilmu Kesehatan Masyarakat. Penulis menekuni bidang kesehatan masyarakat dan riset. Penulis juga sebagai pengelola Journal for Quality in Public Health. Pernah mengisi sebagai pembicara dan moderator di beberapa seminar online kolaborasi berbagai institusi negeri dan swasta di Indonesia.