

Afriza Umami, Ph.D.
Putri Dwi Eriyanti



Biostatistik dalam Ilmu Kesehatan

TEORI, METODE, DAN APLIKASI

Biostatistik

dalam Ilmu Kesehatan

----- Teori, Metode, dan Aplikasi -----

Afriza Umami, Ph.D.
Putri Dwi Eriyanti



PENERBIT KBM INDONESIA

adalah penerbit dengan misi memudahkan proses penerbitan buku-buku penulis di tanah air Indonesia, serta menjadi media *sharing* proses penerbitan buku.

BIOSTATISTIK DALAM ILMU KESEHATAN

Teori, Metode, dan Aplikasi

Copyright @ 2026 By Afriza Umami, Ph.D. & Putri Dwi Eriyanti

All right reserved

Penulis

Afriza Umami, Ph.D.

Putri Dwi Eriyanti

Desain Sampul

Eca Art

Tata Letak

Husnud Diniyah

Editor

Dr. Muhamad Husein Maruapey, Drs., M.Sc.

Background isi buku di ambil dari <https://www.freepik.com/>

Official

Depok, Sleman-Jogjakarta (Kantor)

Penerbit Karya Bakti Makmur (KBM) Indonesia

Anggota IKAPI/No. IKAPI 279/JTI/2021

081357517526 (Tlpn/WA)

Website

<https://penerbitkbm.com>

www.penerbitbukumurah.com

Email

naskah@penerbitkbm.com

Distributor

<https://penerbitkbm.com/toko-buku/>

Youtube

Penerbit KBM Sastrabook

Instagram

@penerbit.kbmindonesia

@penerbitbukujogja

ISBN: 978-634-278-247-7

Cetakan ke-1, Maret 2026

15,5 x 23 cm, iv + 96 halaman

Isi buku diluar tanggungjawab penerbit
Hak cipta merek KBM Indonesia sudah terdaftar di DJKI-Kemenkumham dan isi
buku dilindungi undang-undang.

Dilarang keras menerjemahkan, memfotokopi, atau
memperbanyak sebagian atau seluruh isi buku ini
tanpa seizin penerbit karena beresiko sengketa hukum

Sanksi Pelanggaran Pasal 113
Undang-Undang No. 28 Tahun 2014 Tentang Hak Cipta

- i. Setiap Orang yang dengan tanpa hak melakukan pelanggaran hak ekonomi sebagaimana dimaksud dalam Pasal 9 ayat (1) huruf i untuk Penggunaan Secara Komersial dipidana dengan pidana penjara paling lama 1 (satu) tahun dan/atau pidana denda paling banyak Rp 100. 000. 000 (seratus juta rupiah).
- ii. Setiap Orang yang dengan tanpa hak dan/atau tanpa izin Pencipta atau pemegang Hak Cipta melakukan pelanggaran hak ekonomi Pencipta sebagaimana dimaksud dalam Pasal 9 ayat (1) huruf c, huruf d, huruf f, dan/atau huruf h untuk Penggunaan Secara Komersial dipidana dengan pidana penjara paling lama 3 (tiga) tahun dan/atau pidana denda paling banyak Rp 500. 000. 000,00 (lima ratus juta rupiah).
- iii. Setiap Orang yang dengan tanpa hak dan/atau tanpa izin Pencipta atau pemegang Hak Cipta melakukan pelanggaran hak ekonomi Pencipta sebagaimana dimaksud dalam Pasal 9 ayat (1) huruf a, huruf b, huruf e, dan/atau huruf g untuk Penggunaan Secara Komersial dipidana dengan pidana penjara paling lama 4 (empat) tahun dan/atau pidana denda paling banyak Rp 1. 000. 000. 000,00 (satu miliar rupiah).
- iv. Setiap Orang yang memenuhi unsur sebagaimana dimaksud pada ayat (3) yang dilakukan dalam bentuk pembajakan, dipidana dengan pidana penjara paling lama 10 (sepuluh) tahun dan/atau pidana denda paling banyak Rp 4. 000. 000. 000,00 (empat miliar rupiah).

KATA PENGANTAR

Puji dan syukur penulis panjatkan ke hadirat Tuhan Yang Maha Esa atas segala rahmat dan karunia-Nya sehingga buku yang berjudul “Biostatistik Dalam Ilmu Kesehatan: Teori, Metode, dan Aplikasi” ini dapat diselesaikan dengan baik. Buku ini disusun sebagai salah satu upaya untuk memberikan pemahaman yang sistematis, praktis, dan aplikatif mengenai konsep serta penggunaan biostatistik dalam bidang ilmu kesehatan.

Biostatistik merupakan salah satu komponen penting dalam penelitian dan praktik ilmu kesehatan. Melalui pendekatan statistik, data kesehatan dapat dianalisis secara ilmiah sehingga menghasilkan informasi yang valid, objektif, dan dapat digunakan sebagai dasar pengambilan keputusan. Oleh karena itu, pemahaman terhadap konsep dasar biostatistik menjadi sangat penting bagi mahasiswa, peneliti, maupun praktisi di bidang kesehatan. Buku ini dirancang secara bertahap mulai dari konsep dasar hingga analisis yang lebih aplikatif.

Penulis menyadari bahwa buku ini masih memiliki keterbatasan dan belum sepenuhnya sempurna. Oleh karena itu, kritik dan saran yang konstruktif dari para pembaca sangat diharapkan untuk perbaikan dan penyempurnaan di masa yang akan datang.

Akhir kata, penulis berharap buku ini dapat memberikan manfaat serta menjadi salah satu referensi yang membantu mahasiswa, peneliti, dan praktisi kesehatan dalam memahami serta menerapkan biostatistik dalam kegiatan akademik maupun penelitian.

Penulis

DAFTAR ISI

KATA PENGANTAR	i
DAFTAR ISI	iii

BAB 1

PENDAHULUAN & KONSEP DASAR.....	1
1.1 Peran Biostatistik dalam Kesehatan.....	1
1.2 Jenis Data.....	2
1.3 Jenis Variabel dan Hubungannya	5
1.4 Populasi dan Sampel	7

BAB 2

ANALISIS DESKRIPTIF & DISTRIBUSI DATA.....	15
2.1 Analisis Deskriptif.....	15
2.2 Visualisasi Dasar	15
2.3 Distribusi Normal	17
2.4 Uji Normalitas (Shapiro Wilk / Kolmogorof Smirnov) ..	18

BAB 3

UJI HIPOTESIS & KONSEP INFERENSIAL.....	35
3.1 Konsep Hipotesis	35
3.2 Error Type I dan Type II.....	38
3.3 P Value dan Confidence Interval.....	39
3.4 Parametrik dan Non Parametrik (Konseptual)	42

BAB 4

UJI DUA KELOMPOK 45

- 4.1 Independent t Test vs Mann Whitney 45
- 4.2 Paired t Test vs Wilcoxon Signed Rank 49

BAB 5

UJI LEBIH DARI DUA KELOMPOK 53

- 5.1 One Way ANOVA 53
- 5.2 Kruskal Wallis..... 56
- 5.3 Post Hoc Test..... 58

BAB 6

ANALISI HUBUNGAN & ASOSIASI 61

- 6.1 Tabulasi Silang..... 61
- 6.2 Chi Square 63
- 6.3 Fisher Exact Test 64
- 6.4 Pearson Correlation vs Spearman Rank..... 66

BAB 7

ANALISIS REGRESI 69

- 7.1 Regresi Linier 69
- 7.2 Regresi Logistik 71
- 7.3 Asumsi dan Diagnostik Model..... 73

BAB 8

UJI INSTRUMEN PENELITIAN 77

- 8.1 Uji Validitas 77
- 8.2 Uji Reabilitas 80
- 8.3 Contoh Kuesioner 85

DAFTAR PUSTAKA 93

PROFIL PENULIS..... 95



BAB 1

PENDAHULUAN & KONSEP DASAR

1.1 Peran Biostatistik dalam Kesehatan

Biostatistik didefinisikan sebagai penerapan prinsip statistik dalam kedokteran, kesehatan masyarakat, atau biologi. Prinsip statistik itu sendiri berakar pada matematika terapan, menyediakan serangkaian alat untuk mengumpulkan data, meringkas, menganalisis, serta menafsirkannya (Fahlevi et al., 2024). Peran utama statistik dalam kesehatan adalah sebagai *tools of reasoning* yang kritis dalam seluruh siklus penelitian kuantitatif kesehatan, mulai dari perencanaan desain studi, pengumpulan dan analisis data, hingga interpretasi hasil yang valid berdasarkan kesehatan manusia dan populasi.

Ruang lingkup biostatistik diantaranya adalah pengembangan metode statistik untuk analisis data klinis dan epidemiologis, perancangan studi penelitian, pengolahan

dan surveilans, evaluasi efektivitas intervensi kesehatan, mengidentifikasi faktor risiko penyakit, hingga menilai keberhasilan program kesehatan Masyarakat berbasis bukti (Putra et al., 2025). Tanpa pemahaman prinsip biostatistik yang memadai, peneliti dan praktisi kesehatan berisiko membuat kesalahan dalam pengambilan Keputusan yang dapat berdampak pada kualitas pelayanan dan keberhasilan program kesehatan.

1.2 Jenis Data

Secara etimologis, kata Data merupakan bentuk jamak dari “Datum” yang berasal dari bahasa latin yang berarti “Sesuatu yang diberikan”. Data juga dapat diartikan sebagai fakta dari suatu objek yang diobservasi atau diamati. Sedangkan dalam segi statistika, data merupakan fakta fakta yang akan digunakan sebagai bahan untuk menarik kesimpulan dari topik berdasarkan data yang diambil (Setyawan, 2022a).

Data harus memenuhi syarat agar dapat dianalisis dan diinterpretasikan dengan baik. Syaratnya adalah sebagai berikut:

- a. Obyektif
Data yang diperoleh dari sumbernya atau hasil dari pengukuran, harus dilaporkan apa adanya.
- b. Relevan
Pengumpulan, pengolahan, dan penyajian data harus sesuai dengan permasalahan yang sedang diteliti atau dihadapi.
- c. *Up to Date* (Sesuai Perkembangan)
Data yang baik adalah data yang mutakhir atau berdasarkan penelitian atau sumber yang terbaru.
- d. Representatif
Data yang diperoleh dari sumber yang tepat dapat memberikan Gambaran secara lengkap terhadap suatu konsisi yang benar benar dapat mewakili suatu kelompok tertentu.

Menurut (Wijaya et al., 2024), data dikualifikasikan menjadi:

1.2.1 Menurut sifatnya

a. Data kualitatif

Data yang tidak berbentuk angka, namun berisi Analisa konsisi seperti data wawancara, data observasi, catatan catatan dari permasalahan, dan lain lain.

b. Data kuantitatif

Data yang berbentuk angka. Misalnya data produksi, data pengidap penyakit tiap tahun. Data jumlah penduduk, dan lain lain. Pada data kauntitatif terbagi dua:

- 1) Data deskrit, yakni data dengan angka yang terbatas pada bilangan bulat dan hanya menampilkan angka yang dapat dihitung dalam jumlah bilangan bulat. Contohnya jumlah kunjungan pasien per hari, jumlah penderita diabetes di suatu kelurahan, jumlah kematian ibu melahirkan per tahun, jumlah tenaga kesehatan di rumah sakit, dan jumlah bayi yang diimunisasi dalam satu bulan.
- 2) Data kontinyu, adalah data yang nilainya dapat berubah secara kontinu dan tidak terbatas pada bilangan bulat. Contohnya kadar gula darah puasa (mg/dL), tekanan darah sistolik (mmHg), tinggi badan balita (cm), berat badan ibu hamil (kg), suhu tubuh pasien (derajat Celsius), kadar hemoglobin (g/dL), dan lama rawat inap pasien (hari).

1.2.2 Data menurut sumbernya

a. Data internal

Data yang didapatkan dari dalam atau pihak internal dalam suatu organisasi, perusahaan, atau tempat dilakukannya penelitian. Contohnya data pasien, data SDM, data produksi, dan lain lain.

b. Data eksternal

Data yang diperoleh dari luar organisasi atau perusahaan. Contohnya data angka harapan hidup dari Badan Pusat Statistik, data prevalensi penyakit dari Kementerian Kesehatan, data inflasi yang mempengaruhi harga obat, dan data pendapatan masyarakat yang berkaitan dengan akses pelayanan kesehatan.

1.2.3 Data menurut cara mendapatkannya

a. Data primer

Data yang dikumpulkan langsung dengan maksud untuk diketahui atau diteliti. Contohnya data yang didapatkan melalui kuesioner.

b. Data sekunder

Data yang dikumpulkan dengan maksud selain untuk menyelesaikan masalah yang dihadapi. Data ini diperoleh dari peneliti dari proses pencarian melalui jaringan internet seperti data inflasi, GDP, nilai tukar, dan lain lain.

1.2.4 Data menurut waktu pengumpulan

a. Data *cross section*

Data yang dikumpulkan pada suatu waktu tertentu yang dapat menggambarkan keadaan pada waktu tertentu. Contohnya sensus penduduk tahun 2024 yang akan menggambarkan keadaan penduduk Indonesia pada tahun 2024, hasil survey kepuasan pasien pada rumah sakit A tahun 2020, dan lain lain.

b. Data *time series*

Data yang dikumpulkan dari kurun waktu ke waktu untuk memberikan Gambaran tentang perkembangan suatu kegiatan. Contohnya adalah data angka kematian selama periode pandemi Covid 19.

1.3 Jenis Variabel dan Hubungannya

Variabel penelitian adalah suatu hal yang ditetapkan oleh seorang peneliti untuk dipelajari hingga memperoleh informasi tentang hal tersebut dengan tujuan untuk menarik kesimpulan (Mustafa et al., 2022). Menurut (Sugiyono, 2013), variabel penelitian merupakan suatu obyek yang memiliki variasi antara satu orang dengan yang lain, yang dapat diartikan sebagai atribut dari bidang keilmuan atau kegiatan tertentu, seperti; berat badan, tinggi badan, sikap, motivasi, bentuk, warna, dan lain lain.

Menurut (Anggreni, 2022), jenis variabel terbagi beberapa jenis.

1.3.1 Variabel berdasarkan sifatnya

a. Variabel kategorik (kualitatif)

Variabel hasil dari pengkategorian atau klasifikasi data. Ciri cirinya berbentuk kata kata, biasanya berisi variabel yang berskala nominal atau ordinal. Contohnya agama, jenis kelamin, pekerjaan, pendidikan, dan lain lain,

b. Variabel numerik

Variabel hasil pengukuran atau perhitungan secara langsung. Ciri cirinya berbentuk angka, biasanya berisi skala interval atau rasio. Contohnya adalah tinggi badan, berat badan, mas otot, GDP/GDS, dan lain lain.

1.3.2 Variabel berdasarkan pengukurannya

a. Skala nominal, yakni skala yang disusun berdasarkan kategori, digunakan sebagai pembeda antara karakteristik yang satu dengan yang lain. Cirinya adalah tidak mengandung tingkatan (ranking) dan tidak memiliki nol mutlak. Contohnya adalah jenis kelamin (1) laki laki; (2) perempuan.

b. Skala ordinal, yakni skala yang berdasarkan urutan atau tingkatan (ranking) dari yang tertinggi ke yang terendah, atau sebaliknya. Contohnya pengetahuan (1) kurang; (2) cukup; (3) baik.

- c. Skala interval, yakni skala yang memiliki jarak atau interval antara satu data dengan data yang lain. Data yang dihasilkan berupa angka yang memiliki besar jarak satu dengan data yang lainnya. Namun skala interval tidak memiliki nilai nol mutlak. Contohnya apabila suhu 0 derajat akan dikonversikan ke nilai fahrenheit, sehingga akan menjadi 32.
- d. Skala Rasio, yakni skala yang memiliki nol mutlak. Data yang dihasilkan dapat berupa angka, sehingga uji statistiknya menggunakan uji parametrik.

Dalam statistic, variabel memiliki hubungan atau peran sebagai sebab akibat. Sehingga variabel dapat dikategorikan jenisnya berdasarkan hubungan antar variabel.

a. Variabel *Independent*

Variabel *independent* atau variabel bebas merupakan variabel yang dapat memengaruhi variabel lain. Apabila variabel *independent* berubah, maka akan menyebabkan perubahan jika terhadap variabel lain. Nama lain dari variabel ini adalah prediktor, risiko, determinan, dan kausa.

b. Variabel *Dependent*

Variabel *Dependent* atau variabel terikat merupakan variabel yang dipengaruhi oleh variabel *independent*, artinya variabel *dependent* berubah karena adanya perubahan pada variabel *independent*.

c. Variabel Perancu atau *Confounding*

Variabel Perancu atau variabel *confounding* merupakan variabel yang berhubungan dengan variabel *Independent* dan *Dependent*, tapi bukan merupakan variabel antara. Variabel perancu dapat memengaruhi validitas penelitian karena dapat menyebabkan bias pada hasil penelitian. Untuk meminimalisirnya, diperlukan identifikasi terhadap variabel ini. Sebagai contoh, dalam penelitian tentang hubungan antara konsumsi kopi (*independent*) dan tekanan darah (*dependent*), usia

dapat menjadi variabel perancu (*confounding*) karena usia memengaruhi kebiasaan minum kopi dan juga tekanan darah. Pengendalian *confounding* dapat dilakukan pada tahap desain penelitian, misalnya dengan *randomizing*, *matching*, atau restriksi. Atau pada tahap analisis data, misalnya pada saat stratifikasi atau analisis multivariat.

1.4 Populasi dan Sampel

1.4.1 Populasi

Menurut (Anggreni, 2022), populasi adalah generalisasi yang terdiri atas objek atau subjek yang memiliki kualitas dan karakteristik tertentu yang ditetapkan oleh peneliti untuk dipelajari dan kemudian akan ditarik kesimpulan. Sedangkan menurut (Sugiyono, 2013), Populasi dapat diartikan sebagai objek atau benda benda, bukan hanya manusia yang sekadar jumlah yang ada pada objek/subjek yang dipelajari, namun juga meliputi karakteristik atau sifat yang dimiliki. Sehingga dapat disimpulkan bahwa populasi merujuk pada seluruh entitas, baik manusia maupun bukan beserta seluruh atribut, sifat, atau cirinya yang relevan dengan pertanyaan penelitian yang dapat ditarik sebuah kesimpulan ilmiah. Menurut (Anggreni, 2022), Populasi dibagi menjadi tiga klasifikasi:

- 1) Populasi berdasarkan jenis
 - a) Populasi terbatas, yakni populasi yang mempunyai sumber data yang jelas batasnya secara kuantitatif sehingga dapat dihitung jumlahnya.
 - b) Populasi tak terbatas, yakni populasi yang sumber datanya tidak dapat ditentukan batas batasnya, sehingga bersifat relative.

- 2) Populasi berdasarkan sifatnya
 - a) Populasi homogen, yakni unsurnya memiliki sifat yang sama dan tidak perlu mempersoalkan jumlahnya secara kuantitatif.
 - b) Populasi heterogeny, yakni sumber data yang unsurnya memiliki sifat atau keadaan yang berbeda (bervariasi) sehingga perlu ditetapkan batas batas secara kualitatif dan kuantitatif.
- 3) Populasi berdasarkan kelompoknya
 - a) Populasi umum, yakni populasi yang sumber datanya merupakan seluruh objek pada lokasi penelitian.
 - b) Populasi target, yakni populasi yang menjadi sasaran dalam mengeneralisasi sebagai kesimpulan sebuah penelitian.

1.4.2 Sampel

Mempelajari seluruh anggota populasi seringkali tidak dimungkinkan karena keterbatasan biaya, waktu, dan tenaga. Di sinilah konsep sampel hadir sebagai solusi pragmatis yang ilmiah. Jika populasi adalah danau yang ingin kita ketahui kedalaman dan komposisi ikannya, maka sampel adalah beberapa kali penebaran jala di titik titik yang terencana. Ketepatan seluruh inferensi statistik yakni penarikan kesimpulan tentang populasi berdasarkan data sampel bergantung sepenuhnya pada prinsip keterwakilan ini. Menurut (Amirudin et al., 2022), sampel adalah bagian dari populasi yang terpilih menjadi sasaran penelitian yang memiliki jumlah dan karakter. Sampel harus mencerminkan keadaan populasi, sehingga hasil penelitian akan menjadikan kesimpulan atas populasi (Hardani et al., 2020). Pada dasarnya ada dua syarat yang harus dipenuhi dalam menentukan sampel, yaitu *representatif* (mewakili populasi yang ada) dan jumlah sampel cukup banyak.

Oleh karena itu, tantangan metodologis utama setelah menetapkan populasi adalah bagaimana cara memilih bagian dari populasi tersebut agar benar benar dapat mewakilinya. Proses

inilah yang disebut penarikan sampel (sampling), dan keilmiahannya terletak pada penggunaan metode atau teknik sampling yang sistematis dan dapat dipertanggungjawabkan. Tanpa metode yang tepat, sekecil apa pun risikonya, sampel dapat mengalami bias seleksi yaitu distorsi sistematis yang membuat sampel tidak lagi mirip dengan populasinya sehingga seluruh kesimpulan penelitian menjadi tidak valid. Maka pemilihan teknik sampling sangat penting untuk meminimalkan bias dan mengoptimalkan keterwakilan dengan kerangka logika. Menurut (Sugiyono, 2013), teknik sampling pada dasarnya dikelompokkan menjadi dua kelompok, yaitu *Prbabilitas Sampling* dan *Non Probabilitas Sampling*.

1) *Probabilitas Sampling*

Pengambilan sampel yang memberikan peluang yang sama bagi setiap unsur populasi untuk menjadi anggota sampel. *Probabilitas sampling* dibagi menjadi beberapa jenis:

a) *Simple random sampling*

Teknik ini memastikan setiap anggota populasi memiliki peluang yang sama dan independen untuk terpilih, umumnya melalui undian atau angka acak. Keunggulannya terletak pada konsepnya yang sederhana dan potensi keterwakilan. Namun, teknik ini kurang efektif jika kerangka sampel tidak lengkap, populasi sangat beragam, atau ukuran sampel kecil, serta bisa tidak praktis untuk populasi besar.

b) *Proportionate stratified sampling*

Proportionate stratified sampling adalah teknik sampling yang digunakan untuk populasi heterogen yang memiliki subkelompok berbeda. Teknik ini membagi populasi ke dalam strata yang homogen terlebih dahulu, lalu mengambil sampel dari setiap strata secara proporsional. Keunggulannya adalah presisi dan keterwakilan strata yang lebih baik. Kelemahannya adalah kebutuhan akan pengetahuan mendalam tentang

populasi untuk menentukan strata, usaha yang lebih besar, dan ketergantungan pada relevansi karakteristik stratifikasi.

c) *Disproportionate stratified random sampling*

Teknik ini digunakan pada populasi berstrata yang tidak proporsional. Sampel dari setiap strata diambil dengan jumlah yang tidak sebanding dengan ukuran strata di populasi, seringkali untuk memastikan strata minoritas terwakili dengan memadai untuk dianalisis. Contohnya, dalam sebuah rumah sakit, semua dokter spesialis jantung dan spesialis saraf yang jumlahnya sangat sedikit dapat diambil seluruhnya sebagai sampel, sementara perawat umum yang jumlahnya besar cukup diambil sebagian. Kelemahannya, analisis data memerlukan pembobotan khusus.

d) *Cluster sampling*

Teknik ini cocok untuk populasi yang sangat besar dan tersebar secara geografis. Sampel diambil dalam dua tahap: pertama memilih sejumlah kelompok alami (klaster) seperti desa atau sekolah secara acak, lalu mengambil sampel individu dari dalam klaster terpilih tersebut. Keuntungan utamanya adalah efisiensi biaya dan waktu. Kelemahannya adalah data kurang presisi dan berisiko bias jika karakteristik anggota dalam satu klaster terlalu serupa.

2) *Non probabilitas Sampling*

Non probabilitas Sampling adalah teknik pengambilan sampel yang memberi peluang pada setiap unsur secara tidak merata untuk dipilih sebagai sampel. Teknik ini dibagi menjadi beberapa jenis:

a) *Sampling sistematis*

Teknik ini memilih sampel berdasarkan urutan nomor dari daftar populasi dengan interval tetap (misalnya, setiap kelipatan ke 5). Pelaksanaannya praktis dan cepat. Namun, jika

pola tertentu tersembunyi dalam daftar, sampel berisiko bias karena elemen di antara interval terabaikan.

b) Sampling insedintal

Sampel diambil dari individu yang kebetulan tersedia dan dijumpai peneliti pada saat penelitian. Teknik ini sangat mudah, murah, dan cepat dilaksanakan. Keterbatasan utamanya adalah sampel cenderung tidak mewakili populasi (*non representative*).

c) Sampling *purposive*

Anggota sampel dipilih secara spesifik berdasarkan kriteria dan tujuan penelitian, seperti memilih pakar untuk topik tertentu. Teknik ini umum dalam penelitian kualitatif untuk menggali kasus secara mendalam, namun hasilnya tidak untuk digeneralisasi.

d) Sampling kuota

Peneliti menetapkan kuota jumlah sampel dengan ciri ciri tertentu (misalnya, usia atau profesi). Pengumpulan data berlanjut hingga kuota setiap kategori terpenuhi. Ini memastikan keragaman sampel tetapi berisiko bias seleksi karena pemilihan dalam kuota sering non acak.

e) Sampling jenuh

Seluruh anggota populasi dijadikan sampel. Teknik ini ideal jika populasi sangat kecil (misalnya, kurang dari 30) untuk meminimalkan kesalahan. Namun, tidak efisien dan jarang layak untuk populasi besar.

f) Sampling snowball

Dimulai dari beberapa subjek awal, yang kemudian merekrut subjek lain dari jaringan mereka. Teknik ini efektif untuk menjangkau populasi tersembunyi atau sulit diakses (misalnya, pengguna narkoba). Kelemahannya adalah sampel sulit dikendalikan dan mungkin tidak mewakili variasi penuh populasi.

Setelah menentukan cara memilih sampel, peneliti harus menentukan jumlah sampel yang diperlukan. Sampel terlalu kecil berisiko tidak memiliki kekuatan statistik (power) untuk mendeteksi efek yang bermakna, sementara sampel terlalu besar boros sumber daya. Oleh karena itu, menghitung besar sampel merupakan langkah krusial untuk memastikan hasil penelitian memiliki tingkat kepercayaan (confidence) dan kekuatan uji (power) yang memadai. Perhitungan ini bergantung pada beberapa parameter statistik kunci. Namun yang perlu diperhatikan adalah semakin besar jumlah sampel maka semakin kecil peluang kesalahan generalisasi. Begitupun sebaliknya, semakin kecil jumlah sampel, maka akan memperbesar peluang kesalahan generalisasi.

Menurut (Anggreni, 2022), ukuran sampel (*Sample size*) meliputi beberapa metode:

a. Rumus Slovin

$$n = \frac{N}{1 + Ne^2}$$

Keterangan:

N = Jumlah populasi

n = Jumlah sampel

e = Tingkat kesalahan dalam penelitian

b. Penelitian Cross Sectional

Untuk penelitian survei, rumus yang dapat digunakan adalah dengan rumus estimasi proporsi. Jika besar populasi (N) diketahui, maka dapat menggunakan rumus berikut:

$$n = \frac{Z^2 P(1 - p)N}{d^2(N - 1) + Z^2 P(1 - P)}$$

Namun apabila besar populasi (N) tidak diketahui, maka besar sampel dapat dihitung menggunakan rumus berikut:

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

n = Jumlah sampel

N = Jumlah populasi

Z = Derajat kepercayaan (biasanya pada tingkat 95% = 1,96)

p = Proporsi suatu kasus tertentu terhadap populasi, bila tidak diketahui proporsinya, ditetapkan 50% (0,50)

d = Derajat penyimpangan terhadap populasi yang diinginkan: 10% (0,10), 5% (0,05).

c. Sampel Berstrata

Teknik stratifikasi dipilih untuk populasi yang bersifat heterogen. Dari populasi tersebut kemudian dibagi ke dalam strata yang karakteristiknya sama. Rumus yang digunakan dalam sampel berstrata adalah sebagai berikut:

$$n_i = \frac{N_i}{N} \times n$$

Keterangan:

N_i = Jumlah populasi pada Tratum

N = Jumlah populasi seluruhnya

n_i = Jumlah sampel pada stratum

n = Jumlah sampel seluruhnya

d. *Isaac and Michael*

Menyediakan pedoman ukuran sampel berdasarkan populasi tertentu dengan mempertimbangkan taraf kesalahan (1%, 5%, atau 10%) dalam bentuk tabel siap pakai.

$$s = \frac{\lambda^2 N \cdot P \cdot Q}{d^2(N - 1) + \lambda^2 \cdot P \cdot Q}$$

Keterangan:

λ² dengan dk = 1, taraf kesalahan bisa 1%, 5%, 10%.

P = Q = 0,5

d = 0,05

s = jumlah sampel

e. Krejcie

Rumus ini membantu peneliti mendapatkan sampel yang valid dan optimal tanpa harus menggunakan perhitungan rumit. Biasanya digunakan dalam penelitian sosial, pendidikan, dan kesehatan.

$$n = \frac{\chi^2 N \cdot P(1 - P)}{d^2(N - 1) + \chi^2 \cdot P(1 - P)}$$

Keterangan:

n = ukuran sampel

N = ukuran populasi

χ^2 = nilai chi kuadrat

P = proporsi populasi

d = galat pendugaan

DAFTAR PUSTAKA

- Amirudin, P. R., et al. (2022). *Metodologi Penelitian Kuantitatif*. Pradina Pustaka.
- Anggreni, D. (2022). *Buku Ajar: Metodologi Penelitian Kesehatan*. STIKes Majapahit Mojokerto.
- Dahlan, M. S. (2014). Statistik untuk kedokteran dan kesehatan seri 1 edisi 6. *Jakarta: Epidemiologi Indonesia*, 6(1), 100–234.
- Fahlevi, R., et al. (2024). *Dasar Biostatistika Untuk Peneliti*.
- Glaser, A. N. (2014). *High Yield: Biostatistics, Epidimiology, and Public Health* (Fourth). Lippincott Williams & Wilkins (LWW).
- Hardani, Andriani, H., et al. (2020). *Metode Penelitian Kuantitatif & Kualitatif*. Penerbit Pustaka Ilmu.
- Hidayat, A. (2013). *Uji Normalitas*. <https://www.statistikian.com/2013/01/uji-normalitas.html>
- Isnaini, M., Afgani, W.A., Al Haqqi, & Azhari, I. (2025). Teknik Analisis Data Uji Normalitas. *J CEKI: Jurnal Cendekia Ilmiah*, 4(2), 1377–1384.
- Mustafa, P. S., et al. (2022). *Metodologi Penelitian Kualitatif PTK Pendidikan Olahraga*. Insight Mediatama.
- Najmah. (2017). *Statistika Kesehatan Aplikasi Stata dan SPSS*. Salemba Medika.
- Nuryadi, Astuti, T. D., Utami, E. S., & Budiantara, M. (2017). *Dasar Dasar Statistik Penelitian*. Gramasurya.
- Nuryani, N. (2019). Validitas dan Reliabilitas Kuesioner Pengetahuan, Sikap dan Perilaku Gizi Seimbang Pada Remaja. *Ghidza: Jurnal Gizi Dan Kesehatan*, 3(2), 37–46.
- Putra, E. S., et al. (2025). *Buku Ajar Biostatistik Dasar Untuk Mahasiswa Kesehatan*.

- Santoso, S. (2014). *Panduan lengkap SPSS versi 20*. Elex Media Komputindo.
- Setyawan, D. A. (2022a). *Buku Ajar Statistika Kesehatan Analisis Bivariat Pada Hipotesis Penelitian*.
- Setyawan, D. A. (2022b). *Statistika Kesehatan: Analisis Bivariat Pada Hipotesis Penelitian*. Tahta Media.
- Sugiyono. (2013). *Metode Penelitian Kuantitatif, Kualitatif, dan R&D*. Alfabeta, CV.
- Trafimow, D. (2021). A Frequentist Alternative to Significance Testing, p Values, and Confidence Intervals. In *Towards a New Paradigm for Statistical Evidence*. MDPI.
- Wijaya, E., et al. (2024). *Pengantar Statistika: Konsep Dasar Untuk Analisis Data*.

PROFIL PENULIS



Afriza Umami., PhD lahir di Lamongan pada 22 April. Ia merupakan akademisi dan peneliti di bidang kesehatan masyarakat dengan kepakaran dalam epidemiologi, biostatistik, serta determinan sosial kesehatan. Pendidikan doktoralnya diselesaikan di University of Szeged, Hungaria (2024). Sebelumnya, ia meraih gelar Magister Kesehatan Masyarakat dengan konsentrasi Epidemiologi dan Biostatistik dari Universitas Sebelas Maret (2018). Karier profesionalnya mencakup pengalaman sebagai peneliti, dosen, dan epidemiolog. Ia menjadi dosen di STIKES Muhammadiyah Bojonegoro serta sebelumnya menjabat sebagai Research Fellow dan Teaching Assistant di University of Szeged (2024-2025). Keahlian risetnya meliputi kesehatan masyarakat, perilaku kesehatan, biopsikososial, akulturasi, serta analisis epidemiologi dan biostatistik. Karya ilmiahnya telah dipublikasikan di berbagai jurnal internasional bereputasi dan dipresentasikan dalam forum ilmiah di tingkat internasional.



Putri Dwi Eriyanti: Penulis merupakan mahasiswi aktif di STIKes Muhammadiyah Bojonegoro, Program Studi S1 Gizi yang saat ini sedang menempuh semester 3. Sebelum memulai perjalanan akademiknya di perguruan tinggi, penulis menyelesaikan pendidikan menengah di SMAN 1 Sugihwaras, Bojonegoro dengan konsentrasi jurusan Matematika dan Ilmu Pengetahuan Alam (MIPA).

Buku Biostatistik ini disusun sebagai bagian dari proses belajar dan pengembangan diri penulis dalam memperdalam pemahaman di bidang kesehatan, khususnya terkait pengolahan data. Melalui penyusunan buku ini, penulis berharap dapat menjadikannya sebagai guidebook yang bermanfaat sekaligus sarana penambah wawasan bagi diri sendiri maupun rekan-rekan mahasiswa lainnya. Penulis berharap karya ini dapat membantu memudahkan pemahaman materi dan menjadi sarana untuk menjalin silaturahmi serta bertukar pikiran melalui email puttrieery@gmail.com

Buku ini merupakan panduan komprehensif bagi mahasiswa dan praktisi kesehatan dalam memahami serta menerapkan biostatistik secara sistematis dan aplikatif. Disusun dari konsep dasar hingga analisis lanjutan, buku ini menjembatani teori statistik dengan kebutuhan penelitian kesehatan.

Materi mencakup konsep data dan variabel, analisis deskriptif, distribusi dan uji normalitas, hingga uji hipotesis dan pendekatan parametrik maupun non-parametrik. Berbagai metode perbandingan kelompok, analisis asosiasi, korelasi, serta regresi linier dan logistik dibahas secara praktis sesuai karakteristik data.

Sebagai pelengkap, buku ini juga mengulas uji validitas dan reliabilitas instrumen penelitian. Dengan penyajian yang ringkas dan aplikatif, buku ini menjadi referensi penting dalam meningkatkan kompetensi analisis data di bidang ilmu kesehatan.

Dengan bahasa yang sistematis, jelas, dan aplikatif, buku ini dirancang untuk membantu pembaca tidak hanya memahami konsep biostatistik, tetapi juga mampu mengaplikasikannya secara tepat dalam penelitian kesehatan berbasis data. Buku ini menjadi referensi penting bagi siapa saja yang ingin meningkatkan kompetensi analisis data dalam bidang ilmu kesehatan.



PENERBIT KBM INDONESIA
Anggota IKAPI
0813 5751 7526 / (0353) 3234874
Kantor : Maguwoharjo, Depok, Sleman-Yogyakarta
@penerbitbukujogja
@penerbit.kbm

