ISBN 978-623-125-634-8

DASAR-DASAR BIOSTATISTIKA:

KONSEP DAN APLIKASI DALAM PENELITIAN KESEHATAN

Penulis :

Nur Al-Faida, Fenita Purnama Sari, Andi Indrawati, Rahmawati, Puji Astuti Wiratmo, Mardiani Gina Hartoyo

DASAR-DASAR BIOSTATISTIKA: KONSEP DAN APLIKASI DALAM PENELITIAN KESEHATAN

Nur Al-Faida
Fenita Purnama Sari
Andi Indrawati
Rahmawati
Puji Astuti Wiratmo
Mardiani Gina Hartoyo

GET PRESS INDONESIA

DASAR-DASAR BIOSTATISTIKA: KONSEP DAN APLIKASI DALAM PENELITIAN KESEHATAN

Penulis :
Nur Al-Faida
Fenita Purnama Sari
Andi Indrawati
Rahmawati
Puji Astuti Wiratmo
Mardiani Gina Hartoyo

ISBN : 978-623-125-634-8

Editor : Dr. Oktavianis, M.Biomed.
Penyunting : Mila Sari., S.ST, M.Si
Desain Sampul dan Tata Letak : Atyka Trianisa, S.Pd

Penerbit : GET PRESS INDONESIA
Anggota IKAPI No. 033/SBA/2022

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

Cetakan pertama, Februari 2025

KATA PENGANTAR

Segala Puji dan syukur atas kehadirat Allah SWT dalam segala kesempatan. Sholawat beriring salam dan doa kita sampaikan kepada Nabi Muhammad SAW. Alhamdulillah atas Rahmat dan Karunia-Nya penulis telah menyelesaikan Buku Dasar-Dasar Biostatistika: Konsep Dan Aplikasi Dalam Penelitian Kesehatan ini.

Buku Ini Membahas Pengantar Statistika Dalam Biostatistika, Ukuran Pemusatan Data: Mean, Median, Dan Mode, Penyajian Data: Tabel Dan Diagram Dalam Konteks Kesehatan, Prosedur Dalam Inferensial Statistika Dalam Riset Kesehatan, Uji Chi-Square Dalam Analisis Data Kategorikal, Penerapan Biostatistika Dalam Penelitian Klinis Dan Kesehatan Masyarakat.

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

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

Padang, Februari 2025

Penulis

DAFTAR ISI

DAFTAR TABEL

DAFTAR GAMBAR

BAB 1
PENGANTAR STATISTIKA DALAM BIOSTATISTIK

Oleh Nur Al-Faida

1.1 Pendahuluan

Ketika membahas statistik, kata "populasi" dan "sampel" muncul. Dalam statistik, merupakan praktik umum untuk menampilkan temuan numerik yang menyimpang dari nilai parameter populasi awal, seperti yang diperoleh dari perhitungan atau pengamatan yang dilakukan pada sampel. Statistika adalah cabang ilmu yang berhubungan dengan data numerik. Statistika adalah cara untuk belajar tentang angka: melibatkan pengambilan fitur populasi dan membandingkannya dengan sampel, yang melibatkan pengumpulan, pemrosesan, penyajian, dan evaluasi data (Yuantari dan Handayani, 2017).

Dalam konteks ini, "*statistik*" dan "*statistika*" tidak menyiratkan hal yang sama. Statistik mengacu pada suatu badan pengetahuan yang mempelajari dan menggunakan data numerik, sedangkan statistika sendiri menggambarkan fakta atau informasi numerik. Bidang statistik adalah sekumpulan aturan dan prosedur untuk mengumpulkan, menganalisis, dan menarik kesimpulan dari data numerik sesuai dengan dugaan. Pengukuran, observasi, dan analisis adalah fokus dari statistik, sebuah subbidang matematika. Ringkasan data numerik adalah dasar dari wawasan statistik. Untuk menguraikannya, statistika adalah cabang studi yang berkaitan dengan penerapan teori probabilitas pada studi pengumpulan, pemrosesan, dan analisis data. Husnul dkk. (2020) dan Sangila dan Luthfiah (2018) dikutip dalam Ristya Widi EY dkk. (2023).

Biostatistik, juga dikenal sebagai statistik kesehatan, adalah cabang statistik yang berhubungan dengan kumpulan data terkait kesehatan. Hal ini digunakan untuk tujuan

administratif, seperti perencanaan program layanan kesehatan, sebagai indikasi pengganti untuk memecahkan masalah yang berhubungan dengan kesehatan, dan sebagai metode untuk menganalisis penyakit dari waktu ke waktu. Istilah "biostatistik" mengacu pada kumpulan faktor numerik yang berhubungan dengan kehidupan, dan "statistik" adalah kumpulan faktor numerik dengan nilai yang bervariasi (Muhyidin, 2021).

1.2 Definisi Biostatistik

Biostatistika adalah cabang dari statistik yang mengajarkan kita cara menggunakan angka untuk menganalisis fenomena biologis. Kami membahas topik-topik seperti bagaimana data disajikan, bagaimana memusatkan dan mendistribusikannya, bagaimana kurva normal dibentuk, dan bagaimana mengestimasi, mengambil sampel, menguji hipotesis, mendeskripsikan hasil, dan membandingkan hasil. Untuk menganalisis data penelitian kuantitatif, hal ini sangat penting. Tinjauan umum tentang statistik akan mencakup hal-hal berikut: apa itu statistik, apa saja yang tercakup di dalamnya, jenis data dan faktor apa saja yang ada, dan seberapa besar atau kecilnya faktor yang diberikan. Di antara metrik untuk pemusatan data, yang paling umum adalah modus, median, dan nilai rata-rata (mean).

Pengetahuan atau data yang berkaitan dengan masalah kesehatan dikenal sebagai biostatistik. Ketika digunakan untuk alasan administratif, statistik kesehatan sangat berharga untuk hal-hal seperti analisis penyakit dari waktu ke waktu, penemuan solusi alternatif, dan perencanaan strategi perawatan kesehatan. Istilah "biostatistik" menggambarkan statistik kesehatan. Baik "bio" maupun "statistika" merupakan komponen penting dari istilah biostatistik. Istilah "biostatistik" berasal dari kata Latin "bio" yang berarti kehidupan dan "statistik" yang berarti sekumpulan data numerik yang berkaitan dengan subjek tersebut.

Berikut ini adalah beberapa aspek statistika yang relevan dengan biostatistika, yang merupakan alat investigasi dalam bidang biologi, farmasi, dan kedokteran (Budiarto, 2001):

1. Kumpulan statistik adalah sekumpulan nilai numerik yang berasal dari data, yang dapat ditemukan dalam perhitungan atau pengukuran.
2. Statistik juga dapat diartikan sebagai statistik sampel.
3. Analisis data, pengambilan keputusan, dan tugas-tugas serupa lainnya dapat memperoleh manfaat dari proses ilmiah yang dikenal sebagai statistik.

1.3 Konsep Dasar Statistika

Dalam bukunya "politea", Aristoteles meletakkan dasar untuk statistik dengan menjelaskan data tentang negara 158 negara. "*Aritmatika politik*" adalah nama yang diberikan untuk statistik di Inggris pada abad ke-17. Ahli matematika Jerman, Gottfried Achenwall, pertama kali mengusulkan kata "statistik", yang ditetapkan oleh buku "Statistical Description Of Scotland (1791-1799) pada abad ke-18 oleh Sir John Sinclair (1719-1772).

Selama awal abad ke-20 dan awal abad ke-21, Karl Pearson (1857-1936) adalah seorang perintis dalam bidang statistik induktif. Ia meletakkan dasar bagi statistik kontemporer di abad ini. Karl Pearson, seorang ahli biologi yang mempelajari perkembangan masalah pewarisan pada tahun 1900-an, menggunakan metode statistik untuk mendasari karyanya. Dengan menggunakan analisis regresi dan parameter korelasi, Karl Pearson menghasilkan delapan belas buku antara tahun 1893 dan 1912 yang berjudul Mathematical Approaches to the Theory of Development.

Pengumpulan data, pengolahan data, dan penarikan kesimpulan dari investigasi lapangan merupakan aspek-aspek matematika yang berhubungan dengan statistika. Statistika telah menjadi dasar bagi banyak penyelidikan dan pengamatan ilmiah sepanjang sejarah (Sunaryo *dkk*., 2003).

Pemodelan, hipotesis, pengembangan prosedur dan instrumen pengumpulan data, perancangan studi, pemilihan sampel, dan analisis data adalah area di mana statistik memainkan peran penting dalam penelitian, khususnya

investigasi kuantitatif. Kita harus menggeneralisasi dari contoh-contoh spesifik untuk menguji hipotesis.

Pemahaman umum adalah bahwa statistik dan ilmu data adalah istilah yang identik. Ilmu data berasal dari kata Latin "statistika," yang berarti "negara." Arti asli dari statistik adalah memberikan informasi numerik kepada penguasa tentang tanah, populasi, hasil pertanian dan perkebunan, dan bahkan hewan peliharaan mereka. Contoh paling awal yang diketahui tentang statistik adalah dari Kaisar Augustus, yang menetapkan bahwa semua warga negara akan dikenakan pajak; akibatnya, setiap orang diwajibkan melapor ke ahli statistik atau pemungut pajak terdekat (Spiegel, 1961).

Sudah menjadi praktik umum untuk mengekspresikan dan mendokumentasikan masalah menggunakan data numerik, baik data tersebut berasal dari penelitian, studi, atau pengamatan sederhana. Paling sering, daftar atau tabel digunakan untuk menampilkan koleksi numerik. Untuk lebih memahami subjek yang sedang diperiksa, gambar yang umumnya disebut sebagai diagram atau grafik terkadang disertakan dengan daftar atau tabel.

Pada awalnya, statistik dianggap sebagai data yang dibutuhkan negara dan dapat dimanfaatkan. Sebelum film ini dirilis, orang hanya menganggap statistik sebagai sekumpulan data yang merefleksikan apa yang terjadi di masa lalu. Sebuah gambar yang menggambarkan keadaan populasi, pendapatan rata-rata suatu daerah, jumlah makanan yang diproduksi, dan lain-lain (Mundir, 2012).

Untuk tujuan mendeskripsikan suatu masalah, statistik mengumpulkan fakta, informasi, data, baik numerik maupun lainnya, dan menyusunnya dalam tabel dan/atau grafik. Parameter adalah nama lain dari statistik, namun kedua konsep ini berbeda. Baik parameter maupun statistik terdiri dari angka-angka acak; parameter menjelaskan nilai sampel, sedangkan statistik mencerminkan atau menjelaskan nilai populasi.

Kesimpulan tentang masalah yang diteliti diambil setelah mempelajari, mengolah, dan menganalisis informasi atau data yang telah diperoleh. Temuan-temuan tersebut kemudian

disajikan dalam bentuk tabel atau visualisasi dan sering kali diperlukan deskripsi atau penjelasan lebih lanjut.

Ternyata ini semua adalah tentang pengetahuan yang berbeda yang dikenal sebagai statistik. Oleh karena itu, analitik adalah cabang studi yang berkaitan dengan metode pengumpulan, penyajian, pemrosesan, dan analisis data dengan tujuan untuk menarik kesimpulan dari kumpulan data yang besar yang dihasilkan oleh pengamatan lapangan.

Dengan latar belakang ini, ada dua kategori utama statistik: statistik deskriptif dan statistik induktif, yang terkadang dikenal sebagai statistik inferensial:

1. Dalam statistik deskripsi, tujuannya adalah untuk memberikan gambaran tentang fenomena yang diteliti tanpa membuat asumsi atau kesimpulan yang luas berdasarkan data yang dikumpulkan. Tabel dan grafik adalah bentuk umum dari penyajian data dalam statistik deskriptif, yang juga menggunakan teknik untuk menghitung deviasi standar, modus, median, rentang, dan rata-rata.
2. Menarik implikasi tentang populasi dari data sampel, yang sebelumnya diasumsikan dari analisis statistik, merupakan fokus dari statistik inferensial (induktif), sebuah subbidang statistik yang lebih dari sekadar menyediakan data. Dalam banyak kasus, ini adalah langkah terakhir sebelum menguji hipotesis penelitian. Statistik deskriptif adalah bidang dasar yang menjadi dasar dari statistik inferensial.

1.3.1 Data Statistik

Data merupakan hal yang mendasar dalam upaya statistik. Semua fakta atau deskripsi tentang masalah kesehatan, gejala, atau kejadian dianggap sebagai data. Data atau deskripsi dapat berupa angka atau dalam bentuk kategori (non-numerik), seperti "laki-laki", "perempuan", "berhasil", "gagal", "setuju", "baik", "senang", dll. Data kuantitatif adalah data yang direpresentasikan secara numerik dan bukan dalam bentuk bahasa alamiah, berisi fakta dan proses (Priadana, Sidik., Sunarsi, 2021).

Sementara itu, data kualitatif mengacu pada informasi yang terstruktur ke dalam kategori-kategori. Data yang diperoleh haruslah tidak memihak (sesuai dengan lingkungan yang sebenarnya), informatif (dapat diperbarui), dapat diterapkan (pada situasi yang dihadapi), dan akurat (untuk mendapatkan kesimpulan yang tepat). Ada dua jenis data numerik: diskrit (nominal) dan kontinu (diskrit). Data ordinal, interval, dan rasio adalah contoh data kontinu.

Oleh karena itu, kategori berikut ini dapat diterapkan pada data penelitian:

1. Data kualitatif (non-numerik)
2. Data kuantitatif (angka)
    a. Data Nominal (diskrit)
    b. Continuum Data:
        - Ordinal
        - Interval
        - Rasio

Data yang diperoleh melalui penghitungan dikenal sebagai data nominal atau data diskrit. Akibatnya, angka yang mewakili data nominal tidak dapat berupa angka kontinu atau dinyatakan sebagai pecahan (desimal). Menurut Malik dan Chusni (2018), data statistik yang mencakup angka-angka yang tidak bermakna disebut data nominal. Angka ini hanya merupakan sebutan atau pengenal yang membedakan satu item atau topik dengan yang lainnya. Hal-hal seperti jumlah hari libur, jenis kelamin, pekerjaan, afiliasi agama, dan jumlah murid adalah beberapa contohnya.

Istilah "data kontinu" mengacu pada informasi yang dikumpulkan melalui pengukuran yang dapat dinyatakan sebagai pecahan kontinu (desimal).

Informasi dengan hirarki yang jelas dalam hal kekuatan, peringkat, atau klasifikasi dikenal sebagai data ordinal. Salah satu interpretasi yang mungkin dari nilai numerik dalam dataset ini adalah bahwa mereka mewakili peringkat atau gradasi kualitatif (Abigail Soesana, Hani Subakti, dkk. 2023) Pertimbangkan faktor-faktor seperti tingkat pendidikan, posisi kelas, dan lainnya.

Output dari pengukuran ordinal dengan nilai relatif nol dan jarak antar level yang konstan dikenal sebagai data interval. Oleh karena itu, nilai 0 tidak sepenuhnya tidak berarti; nilai tersebut tetap memiliki beberapa signifikansi.

Data dalam bentuk rasio, khususnya data dengan kualitas yang paling komprehensif, seperti kemampuan untuk mengidentifikasi, mengurutkan, menjumlahkan, dan mengalikan (Sukardi, 2018). Jarak dan nol absolut keduanya diwakili oleh kumpulan data ini. Hal ini menunjukkan bahwa item yang diinvestigasi sebenarnya kosong (nol) jika hasil pengukuran memberikan nilai nol (0). Hasil 0 kg, misalnya, akan menunjukkan bahwa sama sekali tidak ada berat yang terdeteksi.

1.3.2 Faktor Penelitian

Kata bahasa Inggris "faktor" - yang berarti "apa pun yang bervariasi," "berwarna-warni," "tidak sama," atau "tidak satu jenis" - adalah nenek moyang etimologis dari istilah "faktor," yang menandakan hal yang sama dalam bahasa-bahasa lain. Sebuah ide dengan variasi atau variasi yang dapat diberi nilai atau angka disebut sebagai faktor dalam terminologi. Menurut Sugiyono (2009), faktor penelitian dapat berbentuk apa saja yang oleh peneliti dianggap layak untuk dipelajari sehingga diperoleh pengetahuan tentang hal tersebut.

Sebuah konstruk atau fitur yang akan diselidiki dengan nilai yang dapat berubah disebut faktor, menurut Kerlinger (2006). Kita dapat memberikan nilai numerik apa pun pada simbol atau lambang yang disebut faktor. Dengan kata lain, faktor penelitian adalah segala sesuatu yang akan dipelajari oleh peneliti; dengan kata lain, faktor penelitian adalah fenomena yang akan menjadi pusat perhatian untuk diukur atau diobservasi.

Mungkin ada lebih dari satu faktor dalam sebuah penelitian. Ada hubungan antara faktor-faktor ini. Setidaknya ada lima kategori faktor jika dilihat dari sudut pandang hubungan: faktor kontrol, moderator, faktor intervening, dan faktor dependen.

Faktor Independen dan Faktor Dependen

Faktor dependen adalah sekumpulan angka yang bergantung pada nilai faktor independen. Cara yang lebih baik untuk mengatakannya adalah bahwa faktor independen adalah faktor yang memiliki kemampuan untuk memengaruhi faktor lain. Oleh karena itu, faktor dependen adalah faktor yang dipengaruhi secara langsung atau tidak langsung oleh faktor independen. Nama-nama umum untuk jenis faktor ini termasuk respons, keluaran, kriteria, dan faktor dependen.

Dampak dari keinginan intrinsik untuk belajar terhadap kinerja akademik adalah salah satu contoh studi penelitian. Di sini, idenya adalah bahwa motivasi belajar dapat memengaruhi seberapa baik orang belajar. Oleh karena itu, motivasi belajar dianggap bebas sebagai faktor independen karena nilainya dapat berfluktuasi, dan pada gilirannya berdampak pada nilai prestasi belajar, faktor dependen.

Berikut ini adalah gambaran hubungan atau pengaruh antara faktor independen X dan faktor dependen Y:

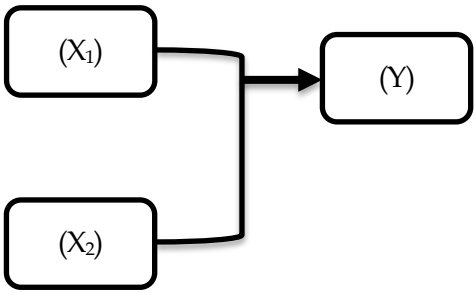
Gambar 1.1. Hubungan antara Faktor X dan Faktor Y

Faktor Moderator

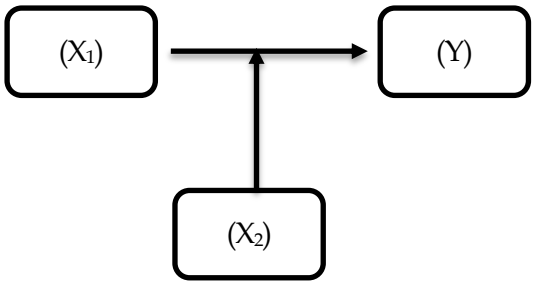
Faktor mediator dapat memperkuat atau memperlemah hubungan antara faktor independen (X1) dan faktor dependen (Y), karena merupakan faktor yang mempengaruhi faktor independen dan dependen. Faktor independen kedua (X2) adalah nama umum untuk faktor ini. Dua faktor independen dan faktor dependen membentuk pengaturan penelitian ini. Dampak dari motivasi dan teknik belajar terhadap prestasi belajar adalah salah satu penelitian tersebut.

Dua faktor independen, strategi pengajaran (X1) dan motivasi (X2), dapat memberikan dampak yang lebih kuat terhadap pencapaian pengajaran (Y) ketika digunakan secara bersama-sama. Hal ini dikarenakan X1 dan X2 bekerja sama

untuk memperkuat hubungan antara kedua faktor tersebut. Sebaliknya, X2 dapat digunakan sebagai faktor perintah untuk mengatur sejauh mana X1 mempengaruhi Y, faktor dependen. Hal ini dapat dijelaskan sebagai berikut:



Gambar 1.2. Hubungan antara Faktor X_1 dan X_2 Secara Bersama-sama dengan Faktor Y



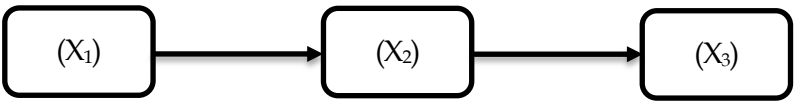
Gambar 1.3. Hubungan antara Faktor X_1 dengan Faktor Y dan X_2 sebagai Faktor Moderator

Faktor Intervening

Secara teoritis, tetapi tidak secara eksperimental, faktor intervening mempengaruhi hubungan antara faktor independen dan dependen. Nama tambahan untuk faktor-faktor ini adalah faktor mediasi dan penghubung.

IQ seseorang, misalnya, menentukan seberapa besar kemajuan yang dapat mereka capai dalam kinerja akademis mereka. Namun, jika ia sedang sakit, marah, atau mengalami kekacauan mental, semua kecerdasan tersebut tidak akan berguna. Sebuah faktor yang memiliki intervensi juga memiliki

dampak, dan keadaan sakit, frustrasi, atau kekacauan mental ini dikenal sebagai faktor intervensi. Namun, sulit untuk memberi angka pada seberapa sakit, frustrasi, atau hancurnya mental seseorang (Mundir, 2012).



Gambar 1.4. Hubungan antara Faktor X_1 dengan Faktor Y dan X_2 sebagai Faktor Intervening

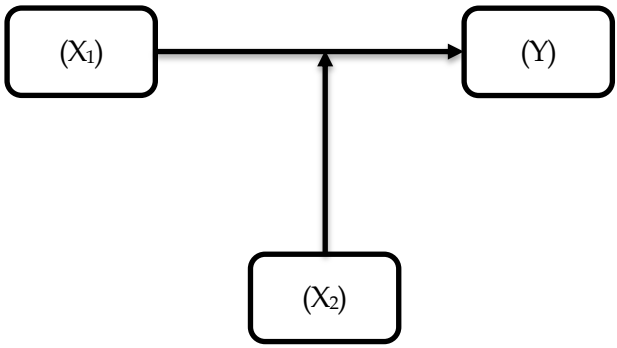
Faktor Kontrol

Untuk menghindari temuan yang miring dalam perhitungan, peneliti sering menggunakan faktor kontrol, yang merupakan faktor yang dampaknya dikurangi atau dihilangkan oleh peneliti. Faktor-faktor ini tidak memengaruhi hubungan antara faktor independen dan dependen.

Peneliti sering menggunakan faktor ini ketika membandingkan dua atau lebih faktor dalam sebuah penelitian eksperimental. Sebagai contoh, katakanlah seorang peneliti tertarik untuk membandingkan kemampuan praktikum mahasiswa jurusan gizi di Universitas A dan Sekolah Tinggi Ilmu Penyakit (STIPen) Nabire.

Peneliti sebelumnya telah melakukan pelatihan konseling dan ingin menilai dampak dari pelatihan tersebut terhadap kemampuan konsultasi mahasiswa dari Universitas A dan Sekolah Tinggi Ilmu Kesehatan Persada (STIKPen) Nabire di Pusat Kesehatan Mahasiswa. Sebagai contoh, media yang digunakan dan kelas tempat konseling dilakukan diidentifikasi sebagai faktor kontrol potensial.

Berikut ini adalah penggambaran visualnya:



Gambar 1.5. Hubungan antara Faktor X_1 dengan Faktor Y dan X_2 sebagai Faktor Kontrol

Peneliti harus memeriksa gagasan teoretis dan hasil investigasi empiris di lokasi penelitian untuk memastikan posisi faktor, apakah itu faktor independen, dependen, instruktur, perantara, atau kontrol. Peneliti harus melakukan penelitian teoritis dan studi eksplorasi tentang topik yang akan diteliti sebelum memutuskan faktor mana yang akan dianalisis. Hal ini akan memastikan bahwa temuan yang diperoleh nantinya sesuai dengan tujuan peneliti.

1.3.3 Populasi dan Sampel

Statistik inferensial berguna untuk menarik kesimpulan tentang faktor yang diteliti dari data sampel yang dapat diekstrapolasi ke populasi secara keseluruhan. Item penelitian dengan fitur yang sama secara kolektif disebut sebagai populasi. Sebaliknya, sampel adalah bagian yang dipilih dari populasi yang lebih besar untuk tujuan penelitian dan analisis statistik. Agar sampel dapat dianggap mewakili populasi, sampel harus memiliki kualitas, fitur, dan karakteristik yang sama dengan komunitas secara keseluruhan. Inilah yang membuat sampel yang baik.

Sesuai dengan prinsip-prinsip ilmiah yang berlaku, para peneliti tidak dapat menarik kesimpulan yang luas dari sampel

yang tidak representatif. Alasannya, pada waktunya, generalisasi akan mulai menyimpang dari kenyataan. Peneliti hanya dapat membuat kesimpulan tentang sampel secara keseluruhan jika sampel tersebut tidak mewakili populasi secara keseluruhan. Oleh karena itu, dalam memilih sampel, peneliti harus memiliki kerangka berpikir yang benar.

Berikut ini adalah contoh sebagian dari sekian banyak metode pengambilan sampel:

1. Pengambilan sampel acak sederhana (*pengambilan sampel acak*)
   Pendekatan sederhana, seperti undian atau menggunakan pendekatan angka kemudahan, digunakan untuk mengambil sampel secara acak dalam prosedur pengambilan sampel ini. Kami secara acak memilih sampel dari seluruh populasi, tanpa mempertimbangkan stratifikasi demografis apa pun. Dengan asumsi bahwa semua orang dalam populasi adalah sama, cara ini akan berhasil.
2. *Pengambilan sampel acak* sistematis Strategi pengambilan sampel ini melibatkan pemilihan sampel pertama secara acak dan kemudian memilih sampel kedua secara metodis sesuai dengan pola yang telah ditentukan. Untuk setiap elemen ke-k dalam suatu populasi, metode pengambilan sampel ini.
3. berfungsi sebagai contoh, dengan nilai pertama dipilih secara acak dari k komponen pertama. Sebagai contoh, ada 150 orang dalam populasi. Individu dalam populasi diberi nomor unik antara satu hingga seratus lima puluh. Selain itu, ketika pengambilan sampel dilakukan, hanya nomor seri genap yang dipilih, atau kelipatan bilangan bulat tertentu seperti 5 yang dipilih.
4. Stratified *random sampling* Teknik *pengambilan sampel ini* menentukan masyarakat ke dalam kategori tingkat tertentu, seperti tinggi, sedang, dan rendah, untuk membangun sampel penelitian. Metode ini digunakan ketika populasi memiliki individu atau ciri-ciri yang menunjukkan stratifikasi yang proporsional dan tidak homogen. Misalnya, sebuah perusahaan yang memiliki karyawan dari berbagai

tingkat pendidikan. Jumlah orang di setiap kelas bervariasi, tetapi biasanya berada di antara SMP, SMA, sarjana, dan pascasarjana. Setiap strata pendidikan memiliki ukuran populasi yang berbeda atau tidak konsisten. Strata pendidikan yang ada saat ini harus dimasukkan ke dalam ukuran sampel secara proporsional (Nuryadi, 2017).

5. *Cluster random* sampling Distribusi geografis dari populasi penelitian digunakan untuk memilih sampel dalam pendekatan pengambilan sampel ini. Metode ini membagi populasi sampel ke dalam beberapa kelompok yang ditentukan oleh lokasi geografisnya. Ambil contoh Indonesia. Dari 38 provinsi yang membentuk negara ini, 8 provinsi dipilih secara acak dari total jumlah provinsi. Perlu diingat bahwa *pendekatan pengambilan sampel acak bertingkat* juga digunakan untuk mengambil sampel dari delapan provinsi di Indonesia karena provinsi-provinsi ini juga bertingkat. Ada dua langkah dalam menggunakan teknik pengambilan sampel klaster: (1) memilih sampel lokal yang representatif dan (2) memilih subset yang representatif dari populasi lokal.

DAFTAR PUSTAKA

Abigail Soesana, Hani Subakti, D. (2023) Metodologi Penelitian Kuantitatif. 1st edn. Yayasan Pustaka Pelajar.

Budiarto E. 2001. Biostatistika untuk Kedokteran & Kesehatan Masyarakat. Edisi 2. Jakarta: EGC

Kerlinger (2006) Prinsip-prinsip Penelitian Perilaku. Edisi ke-3. Yogyakarta: Gadjah Mada University Press.

Malik, A. dan Chusni, M. (2018) Pengantar Statistika Pendidikan: Teori dan Aplikasi, Deepublish. Yogyakarta: CV Budi Utama.

Muhyidin (2021) Pendekatan Biostatistik dalam Kesehatan Masyarakat. Tersedia di: https://muhyidin.id/dekatatan-biostatistik-dalam- kesehatan-masyarakat/.

Mundir (2012) Statistik Pendidikan: Pengantar Analisis Data untuk Penulisan Tesis dan Disertasi. Jember: STAIN Jember Press.

Nuryadi (2017) Dasar-dasar Statistik Penelitian. Yogyakarta: Sibuku media.

Priadana, Sidik, Sunarsi, D. (2021) Metode Penelitian Kuantitatif. 1st edn. Tangerang: Pascal Books.

Ristya Widi EY dkk (2023) Buku Ajar Biostatistika. Jember: UPT Penerbitan Universitas Jember. Tersedia di: https://repository.unej.ac.id/bitstream/handle/123456789/105022/KAMPUSMERDEKA%281%29.pdf?sequence=1&isAllowed=y.

Spiegel, M . R. (1961) Teori dan Masalah Statistika. New York: McGraw-Hill.

Sugiyono (2009) Metode Penelitian Pendidikan: Pendekatan Kuantitatif, Kualitatif, dan R&D. Bandung: Alfabeta. Bandung: Alfabeta.

Sukardi (2018) Metodologi Penelitian Pendidikan; Kompetensi dan Praktiknya. REVISI. Jakarta: Bumi Aksara.

Sunaryo, S. *dkk*. (2003) 'Sejarah Perkembangan Statistika dan Aplikasinya', Forum Statistika dan Komputasi, 8(1), pp. 1-7. Tersediadi:https://journal.ipb.ac.id/index.php/statistika/article/view/5506.

Yuantari, C. dan Handayani, S. (2017) Buku Ajar Statistika Deskriptif & Inferensial, Badan Penerbit Universitas Dian Nuswantoro. Tersedia di: https://repository.dinus.ac.id/docs/ajar/buku_biostat_rev_2017_fix.pdf.

BAB 2
UKURAN PEMUSATAN

Oleh Fenita Purnama Sari Indah

2.1 Pengertian pengukuran kecenderungan memusat

Cara umum untuk mengetahui sekumpulan statistik tentang populasi atau sampel yang ditunjukkan dalam tabel atau grafik adalah dengan memanfaatkan pengukuran kecenderungan sentral, yang juga dikenal sebagai pengukuran gejala sentral (Lubis, 2021). Kecenderungan sentral dapat dinyatakan menggunakan sejumlah metrik yang berbeda, termasuk mean, median, modus, kuartil, desil, dan persentil. Sebagai ukuran kecenderungan sentral, gejala sentral adalah nilai rata-rata yang cenderung berada di tengah. (*measures of central tendency*).

Rata-rata aritmatika, rata-rata geometrik, dan rata-rata harmonik adalah tiga bentuk rata-rata yang sering digunakan (Wijayanti, dkk., 2022). Kata modus, median, dan rata-rata digunakan dalam pengertian yang lebih umum. Premis dasar pengukuran tren atau fenomena inti adalah bahwa rata-rata mungkin merupakan angka yang cukup representatif untuk mengkarakterisasi nilai-nilai dalam data. Jika Anda ingin mengetahui di mana distribusi frekuensi berada, Anda dapat menggunakan rata-rata ini, yang pada dasarnya adalah angka pusat. Sejumlah jenis rata-rata yang berbeda dikenali dalam statistik; ini termasuk rata-rata geometrik, rata-rata harmonik, modus, rata-rata aritmatika (rata-rata), dan median, yang semuanya sering digunakan untuk mengukur lokasi atau kecenderungan sentral suatu distribusi.

Ada beberapa syarat agar suatu nilai dapat dikatakan sebagai nilai sentral, yaitu:

1. Nilai sentral harus dapat mewakili rangkaian data
2. Perhitungannya harus didasarkan pada seluruh data
3. Perhitungannya harus mudah

4. Dalam suatu rangkaian data hanya ada 1 nilai sentral

2.2 Macam-macam pengukuran kecenderungan memusat

1. Rata-rata hitung / Mean

Metode statistik untuk menjelaskan perilaku kelompok dengan menghitung nilai rata-rata. Data dari setiap anggota kelompok dijumlahkan lalu dibagi dengan jumlah total anggota untuk mendapatkan nilai rata-rata (nilai rata-rata). Adapun cara untuk mencari mean dibedakan berdasarkan jenis penyajian data:

a) Data tunggal dengan seluruh skornya berfrekuensi satu

$$\bar{x} = \frac{\sum_{i=1}^{n} x_i}{n}$$

dimana xi = data ke-i dan n = jumlah data

Contoh soal:

Nilai Statistik dari 10 mahasiswa Sebuah universitas adalah sebagai berikut :
8 6 6 7 8 7 7 8 6 6

$$\bar{x} = \frac{8+6+6+7+8+7+7+8+6+6}{10} = \frac{69}{10} = 6{,}9$$

b) Data tunggal sebagian atau seluruh skornya berfrekuensi lebih dari satu.

Nilai data	frekuensi
x_1	f_1
x_2	f_2
⋮	⋮
x_n	f_n
	$N = \sum_{i=1}^{n} f_i$

Maka:

$$\bar{x} = \frac{\sum_{i=1}^{n} f_i x_i}{N} = \frac{\sum_{i=1}^{n} f_i x_i}{\sum_{i=1}^{n} f_i}$$

Dengan xi merupakan nilai data

Contoh Soal Mencari Rata-Rata (Mean)
Data nilai ulangan harian IPA kelas VIII-1. Berapa banyak siswa di kelas itu yang nilainya lebih dari rata-rata?

Nilai	5	6	7	8	9	10
Frekuensi	9	10	12	6	2	1

Penyelesaian:

Cari dulu nilai rata-rata pakai rumus data tunggal berkelompok:

$$\bar{x} = \frac{\Sigma f_n x_n}{\Sigma f}$$

$$\bar{x} = \frac{5 \times 9 + 6 \times 10 + 7 \times 12 + 8 \times 6 + 9 \times 2 + 10 \times 1}{9 + 10 + 12 + 6 + 2 + 1}$$

$$\bar{x} = \frac{45 + 60 + 84 + 48 + 18 + 10}{40}$$

$$\bar{x} = \frac{265}{40} = 6{,}625$$

Nilai rata-rata yang diperoleh adalah 6,625. Soal-soal dirancang untuk dijawab oleh siswa dengan nilai di atas rata-rata, sehingga nilainya berkisar antara 7 hingga 10. Jumlah akhir untuk kelas tersebut adalah 21 orang.

2. Median (Me)

Median adalah adalah angka yang menurut Zakariah dan Afriani (2021) membagi distribusi data menjadi dua atau membagi frekuensi atas dan bawah menjadi dua,

sehingga keduanya sama. Dengan demikian, frekuensi data, bukan fluktuasi nilai, yang menentukan median dari sekumpulan angka.

Median, yang sering disebut kuartil tengah, adalah angka yang tepat berada di tengah kumpulan data yang telah diurutkan. Me adalah notasi standar untuk median. Setelah diurutkan, data dibagi menjadi dua bagian yang sama oleh median, yang merupakan nilai tengah.

Sementara itu, untuk mendapatkan nilai median, perlu untuk mengurutkan data dari yang terkecil hingga terbesar.

Median adalah nilai yang tepat berada di tengah susunan data ketika angka-angka yang diberikan adalah ganjil. Namun, jika angka-angka dalam kumpulan data habis dibagi dua, kita dapat memperoleh median dengan merata-ratakan nilai-nilai di tengah kumpulan data.

Cara Menghitung Median:

Untuk menghitung median dari suatu rangkaian data, ikuti langkah-langkah berikut:

a. Urutkan data
   Mengurutkan data dari terkecil ke terbesar atau terbesar ke terkecil adalah tahap pertama.
b. Tentukan posisi median
   Dalam kasus ketika jumlah total pengamatan habis dibagi dua, nilai yang terletak tepat di tengah-tengah kumpulan data yang diurutkan disebut median. Di sisi lain, ketika titik-titik data terdistribusi secara merata, kita bisa mendapatkan median dengan merata-ratakan dua nilai tengah.
c. Hitung median
   Sesuaikan dengan posisi median yang telah ditentukan sebelumnya.

Adapun cara mencari median, antara lain:

a. Data tunggal sebagian atau seluruh skornya berfrekuensi lebih dari satu.
   Mengurutkan data dari terkecil hingga terbesar merupakan langkah pertama dalam menghitung

median. Untuk masing-masing titik data, ada dua bagian dalam rumus median:

$$M_e = \begin{cases} x_{\frac{n+1}{2}} & \text{, untuk } n \text{ ganjil} \\ \dfrac{x_{n/2} + x_{(n/2)+1}}{2} & \text{, untuk } n \text{ genap} \end{cases}$$

Contoh soal 1:

1) Pada tabel di atas, kita dapat melihat contoh penggunaan angka 8 6 6 7 8 7 7 8 6 6. Setelah disortir, data diambil dalam format berikut: 6 6 6 6 7 7 7 8 8 8. Karena semua angka berada dalam rentang yang sama, kita dapat menerapkan rumus di atas untuk mendapatkan median:

$$M_e = \frac{x_{(n/2)} + x_{(n/2)+1}}{2} = \frac{7+7}{2} = 7$$

2) Diketahui data sebagai berikut:

Nilai	frek
75	8
60	6
92	9
64	7
35	2
Jumlah	31

Untuk data di atas diketahui n ganjil, sehingga untuk mencari median digunakan rumus pertama dan diperoleh :

$$M_e = x_{\frac{n+1}{2}} = x_{16} = 92$$

3) Diketahui data nilai ulangan matematika dari 15 siswa adalah sebagai berikut: 70, 65, 50, 80, 75, 40, 60, 90, 77, 55, 85, 85, 95, 65, 60. Berapakah median dari data tersebut?
Jawaban:
Mengurutkan data adalah langkah pertama dalam mencari median: 40, 55, 60, 65, 70, 75, 77, 80, 85, 90, 95 sekaligus!
Langkah berikutnya adalah mencari median dengan menggunakan metode untuk data ganjil, yang berhasil karena datanya ganjil (15).
Median = (n + 1) ÷ 2
= (15 + 1) ÷ 2
= (16 ÷ 2)
= 8
Jadi, nilai median dari soal di atas berada pada urutan ke-8, yaitu 70.

b. Data kelompok (dalam distribusi frekuensi)
Untuk data yang disusun dalam daftar distribusi frekuensi, median dihitung dengan rumus :

$$\mathbf{Me = Tbm + \frac{[\frac{1}{2}f - \sum fsm]}{fm} \times P}$$

Keterangan:
Tbm : Tepi bawah median
$\sum fsm$: Jumlah Frekuensi sebelum kelas median
FM : Frekuensi kelas median
P : Panjang Kelas

Contoh:
Perhatikan tabel berikut dan tentukan mediannya (nilai tengah)

Interval Kelas	Frekuensi (f)
120 - 128	3
129 - 137	5
138 – 146	10
147 – 155	13
156 – 164	4
165 – 173	3
174 - 182	2

Untuk Mencari nilai tengah dari data dalam tabel frekuensi langkah-langkahnya adalah sebagai berikut :

a. Jumlahkan frekuensinya
b. cari kelas mediannya dengan cara (jumlah.f/2)
c. Tentukan Tbm dengan cara (kelas median – 0,5)
d. Masukan hasil Tbm kedalam rumus berikut

Penyelesaian
Kelas median = ½ n . 40 = 20
Kelas median = pada kelas 147 – 155
Tbm = 147 – 0,5
= 146, 5

$$\text{Me} = \text{Tbm} + \frac{[\frac{1}{2}\boldsymbol{n} - \sum fsm]}{\boldsymbol{fm}} \mathbf{x\ P}$$

$$= 146,5 + \frac{[\frac{1}{2}40 - 18]}{13} \text{ x } 9$$
$$= 146,5 + \frac{[20-18]}{13} \text{ x } 9$$
$$= 146,5 + \frac{[2]}{13} \text{ x } 9$$
$$= 146,5 + \frac{18}{13}$$
$$= 146,5 + 1,38$$
$$= 147,88$$

3. Modus (Mo)

Modus (*mode*) adalah analisis suatu kumpulan data dengan mengidentifikasi nilai-nilai berulang dalam kumpulan data. Cara lain untuk melihatnya adalah bahwa nilai-nilai yang sedang tren dalam suatu kumpulan data adalah nilai-nilai yang paling sering digunakan.

Nilai yang paling umum dalam data statistik disebut modus. Dalam statistik, nilai yang paling umum atau mayoritas juga disebut modus. Salah satu penggunaan modus dalam statistik adalah untuk memilih sebagian populasi untuk penelitian lebih lanjut. Anda dapat menggunakan kalkulasi modus pada data numerik dan kategoris. Misalnya, Anda dapat mengetahui warna mana yang paling disukai siswa atau topik mana yang memiliki hasil tes tertinggi di kelas Anda dengan melihat data tersebut.

Saat menghitung kecenderungan sentral data, seperti rata-rata atau median, modus digunakan. Modus variabel acak atau populasi memberikan informasi penting tentangnya.

Jenis-jenis modus:

a. Unimodal

   Data dianggap unimodal jika hanya ada satu nilai modus. Dalam keadaan ini, data dengan modus tunggal dapat ditampilkan:

   Berikut ini adalah hasil tes matematika suatu kelas:

   Skema dengan nilai modus tunggal mungkin terlihat seperti ini: 75, 60, 55, 70, 50, 60, 65, 60, 52, 60, 85, 65, 75, 40, 80, 45, 90

Karena empat siswa memiliki nilai rata-rata 60, angka tersebut lebih umum daripada angka lainnya dalam data. Kumpulan data ini dicirikan oleh distribusi unimodal karena hanya ada satu nilai modus, 60.

b. Bimodal

Suatu data yang memiliki dua nilai modus disebut bimodal. Contohnya adalah sebagai berikut:

Berikut nilai kelas Bahasa Inggris:

70,0, 55,0, 75%, 85,0, 85,0, 90,0, 50,0, 85,0, 95.

Distribusi bimodal paling tepat menggambarkan kumpulan data, yang menunjukkan bahwa ketiga siswa memiliki peluang yang sama untuk memperoleh dua nilai modus—75 dan 85. Ada dua puncak dalam data distribusi bimodal yang muncul pada frekuensi yang sama.

c. Multimodal

Data dianggap multimoda jika mencakup nilai untuk lebih dari dua modus. Untuk lebih memahami data yang memiliki lebih dari dua nilai modus, perhatikan contoh berikut: Berikut adalah nilai kelas untuk pelajaran bahasa Indonesia:

Kumpulan data berikut menggambarkan distribusi multimoda dengan rata-rata dan median yang identik: 70, 65, 60, 70, 60, 85, 50, 80, 75, 55, 75, 85, 80, 75, 50, 85, 90, 60, 95, 90, 70, 75, 85, 45, 40, 60

Tiga siswa menerima nilai maksimum yang mungkin berdasarkan data: 70, 75, dan 85. Empat siswa mencapai masing-masing nilai ini.

Mengatakan bahwa sekumpulan data tidak memiliki modus sama dengan mengatakan bahwa kumpulan data tersebut tidak memiliki nilai data yang muncul secara teratur. Mo sering digunakan untuk mewakili modus. Rumus untuk menentukan modus data dalam daftar distribusi frekuensi adalah:

$$M_0 = b + p\left(\frac{b_1}{b_1 + b_2}\right) \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots$$

Dengan diketahui:
b = batas bawah kelas modus yaitu kelas interval dengan frekuensi terbanyak
p = panjang interval kelas modus
b1 = frekuensi kelas modus dikurangi frekuensi kelas sebelum kelas modus
b2 = frekuensi kelas modus dikurangi frekuensi kelas sesudah kelas modus

Jika rumus di atas digunakan untuk mencari modus dari tabel di bawah ini:

Interval	frek	x_i	$f_i\, x_i$
36 – 43	5	39,5	197,5
44 – 51	4	47,5	190
52 - 59	6	55,5	333
60 – 67	15	63,5	952,5
68 – 75	13	71,5	929,5
76 – 83	6	79,5	477
84 - 91	1	87,5	87,5
	50		3167

Maka diperoleh :
a. kelas modus = kelas ke-4
b. b = 59,5
c. b1 = 15 – 6 = 9
d. b2 = 15 – 13 = 2
e. p = 8

$$M_0 = 59{,}5 + 8\left(\frac{9}{9+2}\right) = 66{,}045$$

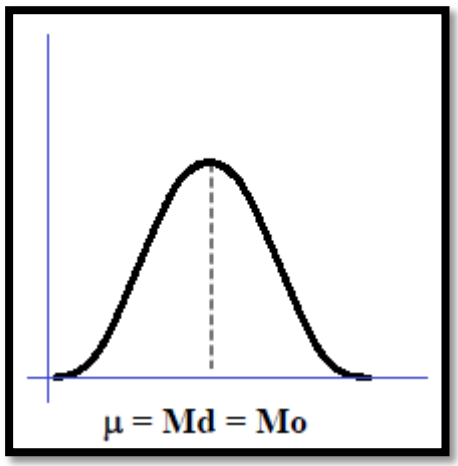
2.3 Hubungan Mean Median dan Modus

Cara yang paling umum untuk menampilkan data adalah sebagai rata-rata, median, atau modus. Ukuran kecenderungan sentral adalah nama lain untuk ketiga angka ini. Alasannya adalah karena nilai-nilai ini cenderung mengelompok di tengah kumpulan data.

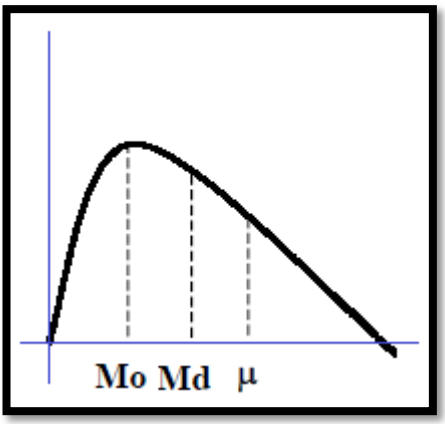
Saat menganalisis data, praktik umum adalah memusatkan perhatian pada subkumpulan data tertentu daripada seluruh kumpulan data. Karena alasan itu, saat melakukan analisis statistik, istilah rata-rata, median, dan modus sering digunakan untuk menggambarkan kumpulan data.

Berikut ini menjelaskan hubungan antara ketiga istilah dalam distribusi frekuensi: rata-rata, median, dan modus.

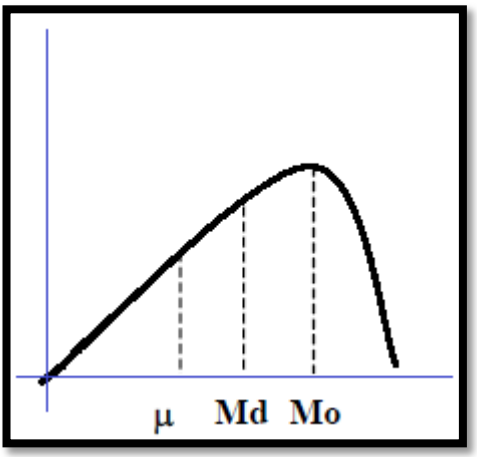
1. Dengan asumsi semuanya sama, modus, median, dan rata-rata akan berada di titik yang sama pada kurva distribusi frekuensi. Bentuk simetris akan diasumsikan oleh kurva distribusi frekuensi.



2. Jika rata-rata lebih besar dari median dan modus, pusat kurva distribusi frekuensi akan menjadi modus, dan sisi kanan akan menjadi rata-rata, median, dan modus, berturut-turut. Distribusi frekuensi dengan kemiringan positif akan memiliki kurva tangan kanan.



3. Median akan berada di tengah dan modus akan berada di sisi kanan kurva distribusi frekuensi jika modus lebih kecil dari median dan rata-rata. Rata-ratanya ada di sebelah kiri. Hasilnya adalah distribusi frekuensi yang condong negatif, atau berarah kiri. 4. Ketika kurva yang menggambarkan distribusi frekuensi miring, berikut ini adalah hubungan umum antara rata-rata, median, dan modus:



DAFTAR PUSTAKA

Lubis, Z., 2021. *Statistika terapan untuk ilmu-ilmu sosial dan ekonomi*. Penerbit Andi.

Wijayanti, R.R., Malau, N.A., Sova, M., Ngii, E., Sugiri, T., Ardhiarisca, O., Astuti, Y. and Saidah, H., 2022. Statistik Deskriptif.

Zakariah, M.A. and Afriani, V., 2021. *Analisis statistik dengan spss untuk penelitian kuantitatif*. Yayasan Pondok Pesantren Al Mawaddah Warrahmah Kolaka.

BAB 3
PENYAJIAN DATA: TABEL DAN DIAGRAM DALAM KONTEKS KESEHATAN

Oleh Andi Indrawati

3.1 Pendahuluan

Penyajian data merupakan salah satu tahapan dalam pembuatan laporan hasil penelitian, berfungsi memaparkan data hasil penelitian agar mudah dipahami untuk menjelaskan fenomena hasil penelitian. Memilih teknik penyajian data yang tepat sangat penting untuk memastikan data yang disajikan mudah dipahami dan tidak membingungkan.

Prinsip utama dari penyajian data adalah bahwa informasi data yang disajikan dapat diterima, dimengerti dan sesuai dengan apa yang dibutuhkan oleh penerima penyajian (Yuantari and Handayani, 2017). Penyajian data bertujuan memberikan gambaran sistematis hasil penelitian atau observasi, memudahkan pemahaman dan analisis data, serta mempercepat dan meningkatkan ketepatan pengambilan keputusan dan kesimpulan (Umami, 2021).

Jika data ditujukan untuk keperluan laporan atau analisis lebih lanjut, maka harus disajikan dengan format yang jelas, terstruktur, mudah dipahami, dan informatif. Dengan menerapkan teknik penyajian data yang tepat dan jelas, maka keputusan dan kesimpulan yang dihasilkan menjadi lebih akurat.

3.2 Bentuk Penyajian Data

Penyajian data dilakukan setelah data terkumpul untuk memperoleh informasi yang terdapat dalam kumpulan data tersebut. Data yang diperoleh dari populasi atau sampel perlu di

atur atau disajikan dalam bentuk terstruktur (Usman and Akbar, 2020).

Bentuk penyajian data terbagi atas:

1. Bentuk Tabel (Tabular)
2. Bentuk Diagram (Grafikal)
3. Bentuk Tulisan (Textular)

3.2.1 Bentuk Tabel (Tabular)

Penyajian bentuk tabel bertujuan menyajikan kumpulan data yang disusun dalam bentuk baris dan kolom secara teratur untuk memudahkan perbandingan, baik berupa kategori maupun angka frekuensi.

Penyajian tabel yang baik memiliki ketentuan sebagai berikut:

1. Judul tabel; singkat dan jelas, serta terletak di atas tabel
2. Nomor tabel; terletak di atas atau serangkaian dengan judul tabel
3. Badan tabel; dapat berupa variabel, nilai frekuensi, distribusi proporsi, atau uji statistik bila perlu
4. Sumber data; darimana data tersebut diambil atau dikutip
5. *Foot Note*; keterangan yang tidak dapat dituliskan di dalam tabel (Saputa, 2016).

Macam-macam bentuk tabel, yaitu:

1. **Tabel Induk (*Master Table*)**

   Berisi semua hasil pengumpulan data lengkap dan terinci yang masih berupa data mentah. Tabel induk biasanya ditampilkan pada bagian lampiran laporan penelitian. Data dari master induk inilah kemudian di analisis lanjut, baik analisis deskriptif maupun analisis inferensi.

Contoh Tabel Induk:

Tabel 3.1. Master Tabel Penelitian Hubungan Karakteristik Ibu Dengan Kegagalan Pemberian Asi Eksklusif Pada Bayi Usia 0-6 Bulan di Wilayah Kerja Puskesmas A Tahun 2023

No	Nama	Umur (Thn)	Pendidikan	Pekerjaan	Pemberian ASI Eksklusif
1.	Ny. Ri	25	SMA	IRT	Tidak
2.	Ny. La	27	SMP	IRT	Ya
3.	Ny. Ki	30	SMA	Wiraswasta	Tidak
4.	Ny. Ft	28	SMA	IRT	Ya
5.	Ny. Bd	35	Sarjana	PNS	Tidak
6.	Ny. Rs	32	Sarjana	Karyawan Swasta	Ya
7.	Ny. Pr	29	SMA	PNS	Tidak
8.	Ny. Ls	33	SMA	Wiraswasta	Tidak
9.	Ny. St	26	SMP	IRT	Ya
10.	Ny. Mt	31	SMA	Petani	Tidak

Sumber: Dokumen Penulis

2. Tabel Satu Arah (*One Way Table*)

Data yang disajikan pada tabel satu arah terdiri atas satu variabel atau satu karakteristik. Tabel ini biasanya digunakan untuk menampilkan data variabel kualitatif atau pun variabel kuantitatif yang dikelompokkan dalam beberapa kategori, seperti umur, tingkat pendidikan, jenis pekerjaan, dan lainnya (Nohe, 2013).

Contoh tabel kualitatif:

Tabel 3.2. Distribusi Ibu Hamil Berdasarkan Jenis Pekerjaan di Wilayah Kerja Puskesmas A Pada Tahun 2023

No.	Jenis Pekerjaan	Frekuensi
1.	IRT	5
2.	Petani	4
3.	PNS	8
4.	Wiraswasta	6
5.	Karyawan Swasta	7
Jumlah		30

Sumber: Dokumen Penulis

Untuk data kuantitatif yang ingin dikelompokkan atau dikategorikan, maka dibuat dalam bentuk tabel distribusi frekuensi, di mana setiap data hanya dapat dimasukkan dalam salah satu kelompok.

Contoh: Hasil pengukuran tinggi badan mahasiswa sebanyak 35 orang, diperoleh data sebagai berikut:

145	161	156	158	153	155	161
157	148	158	155	165	160	158
158	159	155	156	156	161	156
154	160	162	150	157	162	158
160	156	155	157	158	145	159

Pembuatan tabel distribusi frekuensi dilakukan dengan tahapan sebagai berikut:

1. **Tentukan Rentang Data/Jangkauan (*Range*)**; mengurangi nilai data tertinggi dengan nilai data terendah.
   Rentang (R)= nilai data tertinggi – nilai data terendah
   Data hasil pengukuran tinggi badan mahasiswa, tertinggi 165 dan terendah 145, maka:
   Rentang (R) = 165 – 145 = 20
2. **Tentukan Jumlah Kelas Interval**; tentukan berapa banyak kelompok atau kelas yang diperlukan untuk membagi data. Banyaknya jumlah kelas interval minimal 5 dan maksimal 15, dapat menggunakan aturan Sturges (Nohe, 2013):
   k = 1 + 3,3 log n n = jumlah sampel
   k = 1 + 3,3 log 35
   k = 6,095 ≈ 7 (dibulatkan)
3. **Hitung Panjang Interval Kelas**; rentang data dibagi dengan jumlah kelas untuk menentukan lebar kelas.
   p = R/k = 20/7 = 2,85 ≈ 3
4. **Tentukan Batas Kelas**; tentukan batas bawah dan batas atas untuk setiap kelas. Nilai batas bawah kelas interval dapat diambil dari data terendah, misal pada contoh yaitu: 145.

Agar lebih mudah dalam membuat tabel distribusi frekuensi, maka sebelumnya dibuat kolom tabulasi dengan menggunakan garis lurus/miring yang dikenal dengan aturan Tally. Dari hasil perhitungan contoh, diperoleh banyak kelas 7, panjang interval kelas 3, dan batas bawah kelas 145, maka dibuat kolom tabulasi sebagai berikut:

Interval Tinggi Badan (cm)	Tally	Frekuensi
145 - 147	//	2
148 - 150	//	2
151 - 153	/	1
154 - 156	///𝍸𝍸	10
157 - 159	//𝍸𝍸 𝍸	11
160 - 162	///𝍸	8
163 - 165	/	1
Jumlah		35

Selanjutnya dibuat tabel distribusi frekuensi berdasarkan hasil tabulasi.

Tabel 3.3. Distribusi Frekuensi Tinggi Badan Mahasiswa Universitas A Tahun 2023

Interval Kelas Tinggi Badan (cm)	Frekuensi
145 – 147	2
148 – 150	2
151 – 153	1
154 – 156	10
157 – 159	11
160 – 162	8
163 – 165	1
Jumlah	35

Selain itu, ada beberapa istilah yang dikenal pada tabel distribusi frekuensi, yaitu:

1. **Batas kelas**; pada contoh yang menjadi batas bawah kelas masing-masing adalah 145, 148, 151, 154, 157, 160, 163,

sedangkan batas atas kelas masing-masing adalah 147, 150, 153, 156, 159, 162, 165

2. Tepi kelas (batas nyata kelas)

Tepi bawah = batas bawah – 0,5 = 145 - 0,5 = 144,5

Tepi atas = batas atas + 0,5 = 165 + 0,5 = 165,5

3. Titik tengah kelas

$$\frac{\textbf{Batas bawah} + \textbf{Batas atas}}{\textbf{2}}$$

Titik tengah kelas pertama:

$\frac{145+147}{2}$ = 146, ...dst untuk setiap kelas

Untuk melihat sebaran data dalam persentase, maka frekuensi absolut pada tabel distribusi frekuensi harus diubah menjadi frekuensi relatif, dengan menggunakan rumus:

$\frac{f}{n}$**x100%** = $\frac{2}{35}$x100% = 5,7%, ...dst untuk setiap frekuensi

Tabel 3.4. Distribusi Frekuensi Relatif Tinggi Badan Mahasiswa Universitas A Tahun 2023

Interval Kelas Tinggi Badan (cm)	Frekuensi (%)
145 – 147	5,7
148 – 150	5,7
151 – 153	2,9
154 – 156	28,6
157 – 159	31,4
160 – 162	22,9
163 – 165	2,9
Jumlah	100

Selain distribusi frekuensi relatif, ada juga yang dikenal dengan distribusi frekuensi kumulatif, di mana frekuensi data ditambahkan secara bertahap dari kelas pertama hingga kelas terakhir. Ini menunjukkan jumlah total data yang berada di bawah atau pada batas kelas tertentu. Distribusi kumulatif membantu memahami sebaran data secara keseluruhan, terbagi atas dua yaitu: distribusi frekuensi kumulatif kurang dari dan atau lebih dari.

Tabel 3.5. Distribusi Frekuensi Kumulatif Kurang Dari dan Kumulatif Lebih Dari Tinggi Badan Mahasiswa Universitas A Tahun 2023

Tinggi Badan (cm)	Frekuensi Kumulatif	Tinggi Badan (cm)	Frekuensi Kumulatif
< 145	0	≥ 145	35
< 148	2	≥ 148	33
< 151	4	≥ 151	31
< 154	5	≥ 154	30
< 157	15	≥ 157	20
< 160	26	≥ 160	9
< 163	34	≥ 163	1
< 166	35	≥ 166	0

3. Tabel Dua Arah (*Two Way Table*) atau Tabel Silang (*Cross Tabulation*)

Disebut juga tabel kontingensi yang digunakan untuk menyajikan data dengan dua atau lebih variabel secara bersamaan (Otok and Ratnaningsih, 2019). Untuk melihat hubungan antara dua variabel, sebaiknya variabel independen ditempatkan di kolom dan variabel dependen di baris (Nohe, 2013).

Contoh tabel dua arah:

Tabel 3.6. Jumlah Responden Berdasarkan Risiko Penyakit dan Kebiasaan Merokok

Risiko Penyakit	Kebiasaan Merokok		Jumlah
	Merokok	Tidak Merokok	
Berisiko	28	12	40
Tidak Beresiko	21	9	30
Jumlah	49	21	70

Sumber: Dokumen Penulis

Berdasarkan Tabel 3.6 terdapat 40 responden beresiko terkena penyakit, 28 orang diantaranya merokok dan 12 orang tidak merokok, serta terdapat 30 responden tidak berisiko

terkena penyakit, diantaranya 21 orang merokok dan 9 orang tidak merokok.

4. **Tabel Tiga Arah (*Three Way Table*)**
   Digunakan untuk menyajikan data dengan tiga variabel, di mana setiap varibel dengan dua atau lebih kategori.
   Contoh tabel tiga arah:

Tabel 3.7. Jumlah Ibu Hamil Berdasarkan Kelompok Umur, Status Pendidikan, dan Pekerjaan di Wilayah Kerja Puskesmas A Bulan Januari 2023

Kelompok Umur (Thn)	Status Pendidikan		Status Pekerjaan	
	Sekolah	Tidak Sekolah	Bekerja	Tidak Bekerja
< 25	12	18	13	17
≥ 25	18	12	17	13
Jumlah	30	30	30	30

Sumber: Dokumen Penulis

Berdasarkan Tabel 3.7 dapat diketahui status pendidikan ibu hamil yang berumur <25 tahun lebih banyak yang tidak sekolah dibandingkan yang sekolah. Sedangkan ibu hamil yang berumur ≥ 25 tahun lebih banyak yang sekolah dibandingkan yang tidak sekolah. Berdasarkan status pekerjaan ibu hamil yang berumur <25 tahun lebih banyak yang tidak bekerja dibandingkan yang bekerja. Sedangkan ibu hamil yang berumur ≥ 25 tahun lebih banyak yang bekerja dibandingkan yang tidak bekerja.

3.2.2 Bentuk Grafik/Gambar (Diagram)

Bentuk grafik adalah representasi data visual yang lebih menarik dan informatif dalam bentuk lukisan garis, gambar atau pun lambang. Grafik merupakan lanjutan penyajian data dari tabel (Nasution, 2017). Bentuk grafik dapat membandingkan beberapa varibel baik variabel kualitatif maupun kuantitatif dalam berbagai konteks. Selain itu, memudahkan kita melihat perkembangan suatu data dengan informasi yang jelas, serta dapat mengurangi kejenuhan kita melihat angka.

Sama halnya dengan tabel, penyajian grafik/gambar juga harus memperhatikan beberapa hal, yaitu:

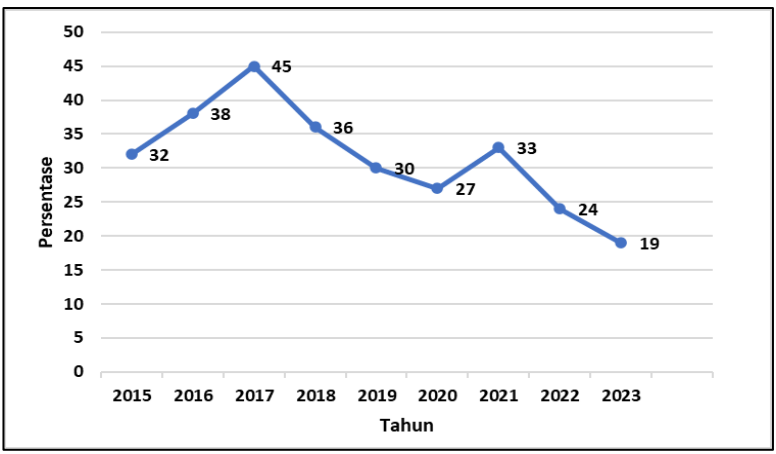
1. Judul: singkat, jelas, lengkap, dan letaknya di bawah
2. Nomor gambar
3. Terdapat dua sumbu sebagai ordinat dan absis
4. Menggunakan skala tertentu
5. *Foot Note*
6. Sumber gambar

Jenis grafik yang dapat digunakan sebagai berikut:

1. Grafik Garis (*Line Chart*)

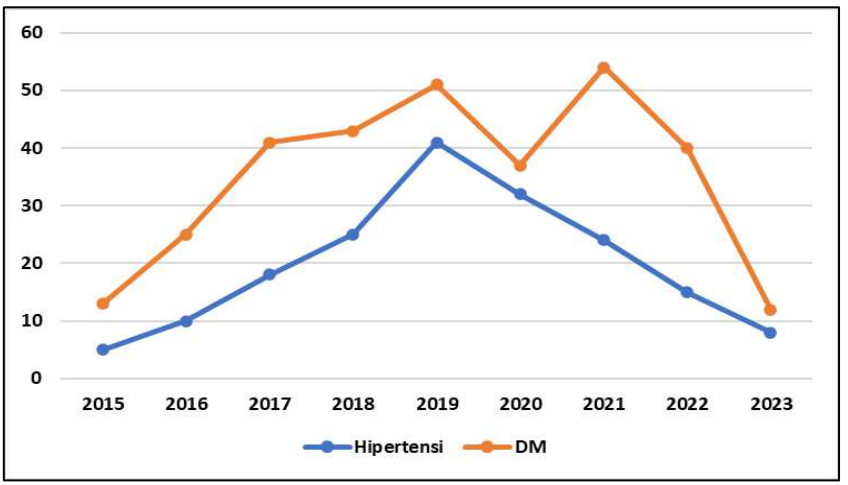
Grafik berbentuk garis lurus yang menggambarkan data kuantitatif berupa data time series. Waktu pengamatan berada pada sumbu horisontal (X), sedangkan nilai data berada pada sumbu vertikal (Y). Untuk membentuk sebuah grafik, maka titik-titik pada bidang XY dihubungkan dengan menggunakan garis lurus. Grafik garis ada dua macam, yaitu: grafik garis tunggal (satu garis) dan grafik garis berganda (beberapa garis) (Nohe, 2013).

Contoh grafik garis tunggal:



Gambar 3.1. Jumlah Kejadian Stunting di Wilayah Kerja Puskesmas A Tahun 2015-2023
(Sumber: Dokumen Penulis)

Gambar 3.1 menunjukkan bahwa pada tahun 2015, jumlah kejadian stunting sebanyak 32%, kemudian mengalami peningkatan hingga tahun 2017 mencapai 45%, namun pada tahun 2018 mengalami penurunan sampai tahun 2020. Di tahun 2021 mengalami peningkatan mencapai 33%. Setelah itu, di tahun 2022 mulai menurun dari 24% hingga 19% di tahun 2023.
Grafik garis berganda memiliki dua atau lebih garis untuk menunjukkan beberapa hal atau kejadian.
Contoh grafik garis berganda:



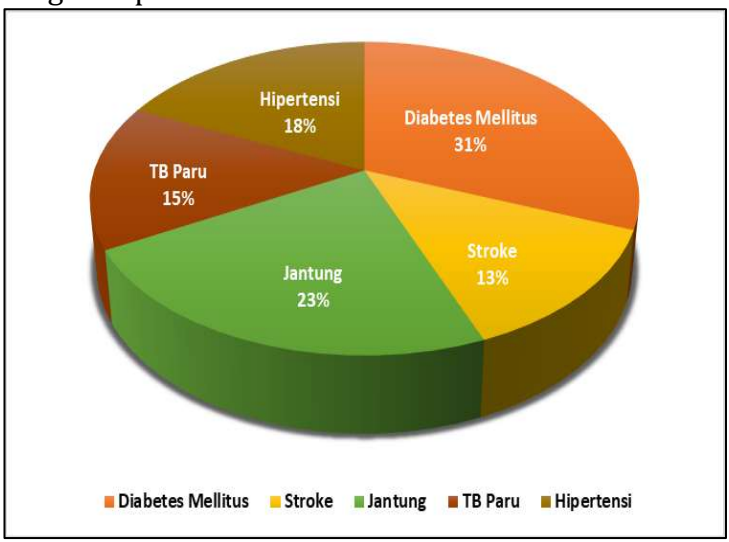
Gambar 3.2. Data Kasus Hipertensi dan Diabetes Mellitus di Rumah Sakit X Tahun 2015-2023
(Sumber: Dokumen Penulis)

Berdasarkan Gambar 5.2 terlihat jumlah kasus Diabetes Mellitus (DM) lebih tinggi dibandingkan kasus Hipertensi. Kasus DM tertinggi terjadi pada tahun 2021, setelah itu mengalami penurunan sampai tahun 2023. Sedangkan kasus hipertensi mengalami peningkatan pada tahun 2019, dan mulai menurun sejak tahun 2020 hingga tahun 2023.

2. Grafik Lingkaran (*Pie Chart*)

Grafik ini menyajikan data dalam bentuk lingkaran yang dibagi sesuai proporsi, biasanya berupa nilai persentase

(Nasution, 2017). Grafik lingkaran cocok untuk data kualitatif atau kuantitatif yang dikategorikan, terutama data *cross-sectional*.
Contoh grafik pie:

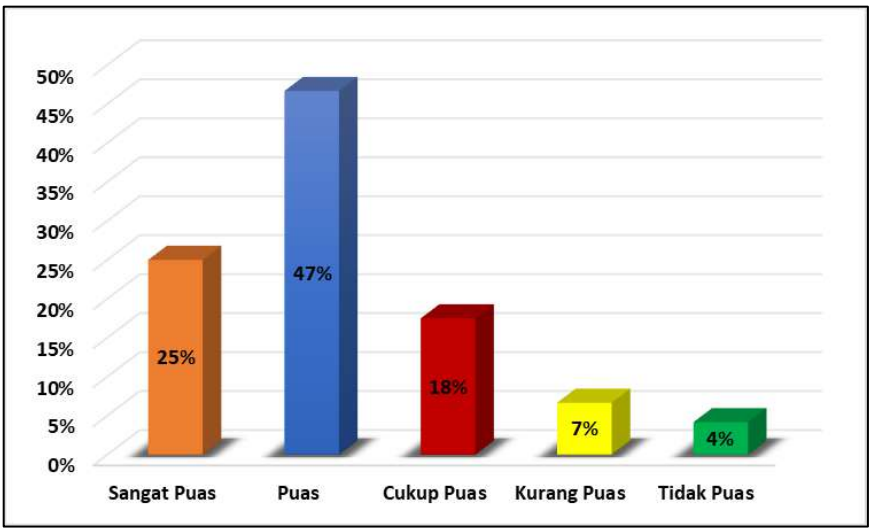


Gambar 3.3. Persentase Kasus Penyakit di Rumah Sakit X Tahun 2023
(Sumber: Dokumen Penulis)

Berdasarkan Gambar 3.3 persentase kasus penyakit yang tertinggi di Rumah Sakit X pada tahun 2023 adalah diabetes mellitus (DM) sebesar 31%.

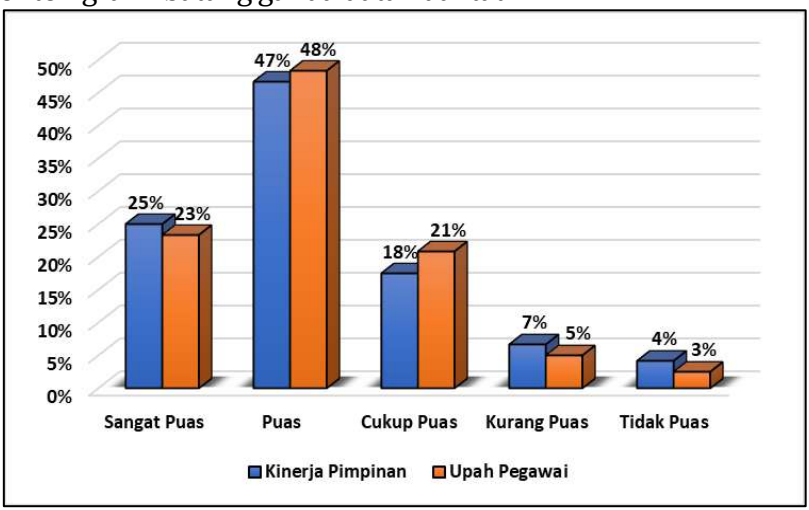
3. Grafik Batang (*Bar Chart*)

Grafik batang digunakan untuk data kualitatif atau pun data kuantitatif berupa data time series dengan periode pendek, sedangkan jika data yang dikumpulkan dalam periode panjang, maka lebih sesuai menggunakan grafik garis (Nohe, 2013). Grafik batang menggunakan batang tegak atau mendatar, lebar batang sama dan terdapat jarak antara batang satu dengan yang lain. Grafik batang ada dua macam, yaitu: grafik batang tunggal dan grafik batang ganda.
Contoh grafik batang tunggal data kualitatif:



Gambar 3.4. Persentase Tingkat Kepuasan Pegawai Terhadap Kinerja Pimpinan di Puskesmas A Tahun 2023
(Sumber: Dokumen Penulis)

Contoh grafik batang ganda data kualitatif:

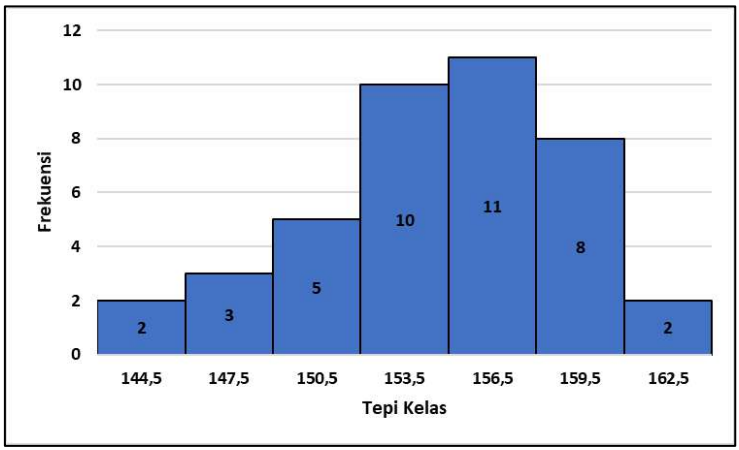


Gambar 3.5. Persentase Tingkat Kepuasan Pegawai Terhadap Kinerja Pimpinan dan Upah Pegawai di Puskesmas A Tahun 2023
(Sumber: Dokumen Penulis)

4. Histogram

Histogram menyajikan data distribusi frekuensi dalam bentuk balok, tepi kelas pada sumbu X dan frekuensi setiap kelas pada sumbu Y. Pada histogram tidak ada jarak diantara balok, karena merupakan data kontinu/berkelanjutan.

Contoh histogram:

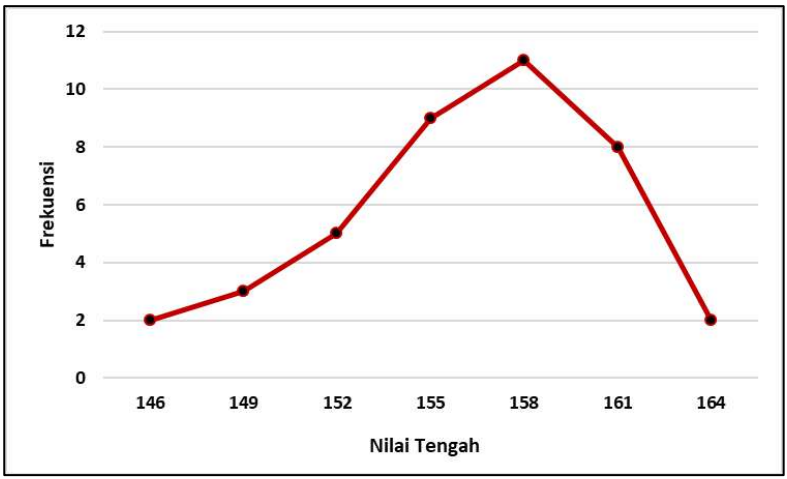


Gambar 3.6. Histogram Tinggi Badan Mahasiswa
(Sumber: Dokumen Penulis)

5. Poligon

Poligon menyajikan distribusi frekuensi dengan titik-titik nilai tengah kelas sebagai sumbu X yang dihubungkan menggunakan garis. Poligon lebih cocok untuk membandingkan dua atau lebih distribusi frekuensi pada grafik yang sama atau ketika ingin menunjukkan perubahan frekuensi secara kontinu.

Contoh grafik poligon:

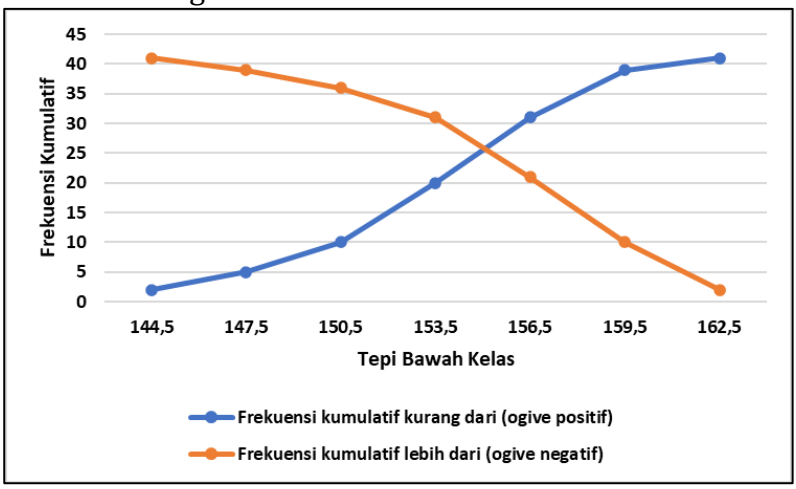


Gambar 3.7. Grafik Poligon Tinggi Badan Mahasiswa.

6. Ogive

Kurva ogive menyajikan data frekuensi kumulatif, dengan tepi bawah kelas sebagai sumbu horisontal dan frekuensi sebagai sumbu vertikal. Kurva ogive ada dua yaitu: ogive positif (frekuensi kumulatif kurang dari) dan ogive negatif (frekuensi kumulatif lebih dari). Kurva ogive menyediakan cara visual untuk memahami bagaimana data terdistribusi secara kumulatif.

Contoh Kurva Ogive:



Gambar 3.8. Kurva Ogive Tinggi Badan Mahasiswa

7. Grafik Sebar

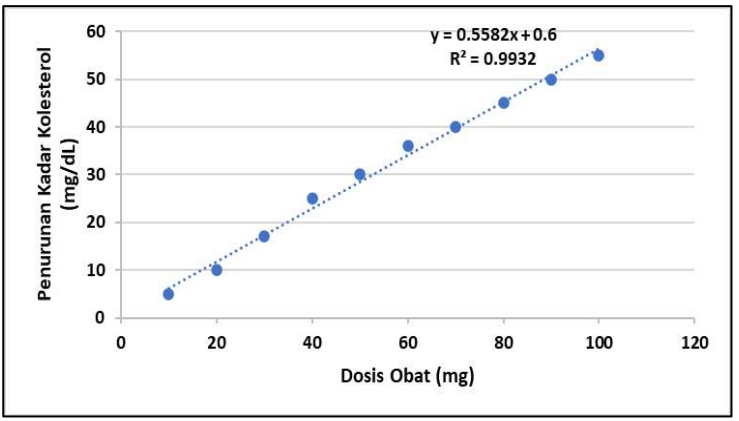
Diagram ini disebut juga diagram pencar atau scatter diagram yang menggunakan titik-titik pada sumbu X dan Y untuk menunjukkan hubungan antara dua variabel atau lebih, serta menampilkan sekumpulan titik yang merepresentasikan nilai variabel tersebut. Untuk mengukur seberapa baik variabel independen menjelaskan variasi dalam variabel dependen maka harus dihitung koefisien determinasi (R^2), sehingga dapat diketahui kekuatan dan kejelasan hubungan yang terlihat dalam diagram sebar.

Contoh:

Diperoleh data tentang pemberian dosis obat terhadap penurunan kadar kolesterol (mg/L) pasien, sebagai berikut:

Kode Pasien	**Dosis Obat (mg)**	**Penurunan Kolesterol (mg/dL)**
AA	10	5
AB	20	10
AC	30	17
AD	40	25
AE	50	30
AF	60	36
AG	70	40
AH	80	45
AI	90	50
AJ	100	55

Untuk melihat hubungan antara kedua variabel, maka dibuat dalam bentuk diagram sebar dengan menentukan nilai koefisien determinasi (R^2), seperti pada Gambar 5.9 berikut ini:



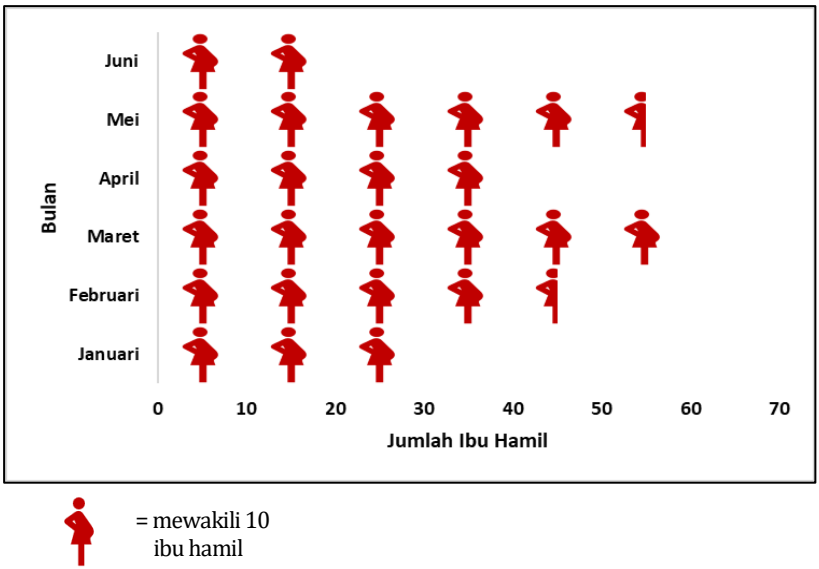
Gambar 3.9. Scatter Plot Pemberian Dosis Obat Terhadap Penurunan Kolesterol
(Sumber: Dokumen Penulis)

Berdasarkan Gambar 3.9 dapat dijelaskan bahwa berdasarkan analisis regresi, setiap penambahan 1 mg dosis obat berhubungan dengan penurunan kadar kolesterol sebesar 0,5673 mg/dL. Koefisien determinasi (R^2) sebesar 0,9937 menunjukkan bahwa 99,37% variasi dalam penurunan kadar kolesterol dapat dijelaskan oleh dosis obat, yang menandakan hubungan yang sangat kuat antara peningkatan dosis obat dan penurunan kadar kolesterol.

8. Piktogram

Piktogram menyajikan data dengan menggunakan gambar atau simbol sebagai pengganti kata atau angka. Setiap simbol mewakili jumlah tertentu (Saputa, 2016). Piktogram memberikan kemudahan dalam penjumlahan karena setiap gambar mewakili jumlah tertentu, namun terjadi kesulitan dalam mempresentasikan sebagian simbol, sehingga penggunaannya terbatas dan kurang efisien.

Contoh Piktogram:



Gambar 3.10. Piktogram Jumlah Ibu Hamil di Wilayah Kerja Puskesmas A Bulan Januari-Juni 2023
(Sumber: Dokumen Penulis)

3.2.3 Bentuk Tulisan (Textular)

Penyajian data dalam bentuk teks menjelaskan temuan secara deskriptif, menggunakan bahasa yang benar dan jelas, serta disusun dalam paragraf singkat. Tujuannya adalah untuk memberikan gambaran umum tentang data atau penjelasan lengkap tentang prosedur, hasil, dan kesimpulan. Cocok untuk data yang sedikit serta kesimpulan yang sederhana (Yuantari and Handayani, 2017).

Contoh:

Penelitian tentang pengaruh olahraga pada penderita diabetes, dengan melibatkan 30 penderita DM Tipe 2. Sebanyak 15 orang berolahraga secara teratur selama 12 minggu, sementara 15 orang lainnya tidak.

Hasil yang diperoleh yaitu:

1. **Yang Berolahraga:** Kadar gula darah turun dari 8,5% menjadi 7,2%.

2. **Yang Tidak Berolahraga:** Kadar gula darah turun dari 8,4% menjadi 8,2%.

Kesimpulan:

Olahraga terbukti efektif menurunkan kadar gula darah pada penderita DM Tipe 2.

DAFTAR PUSTAKA

Nasution, L. M. (2017) 'Statistik Deskriptif', *Jurnal Hikmah*, 14(1), pp. 49–55.

Nohe, D. A. (2013) *Biostatistika 1*, *Halaman Moeka Publishing: Penerbit & Jasa Penerbitan Buku*. Jakarta Barat.

Otok, B. W. and Ratnaningsih, D. J. (2019) *Konsep Dasar dalam Pengumpulan data Penyajian Data*.

Saputa, R. (2016) *Biostatistik Dalam Ilmu Kesehatan Masyarakat*. Batam: Sekolah Tinggi Ilmu Kesehatan Ibnu Sina Batam.

Umami, A. (2021) *Konsep Dasar Biostatistik*. Kediri: CV. Pelita Medika.

Usman, H. and Akbar, P. S. (2020) *Pengantar Statistika Cara Mudah Memahami Statistika*. Jakarta: PT Bumi Aksara.

Yuantari, C. and Handayani, S. (2017) *Buku Ajar Statistik Deskriptif & Inferensial*, *Badan Penerbit Universitas Dian Nuswantoro*. https://repository.dinus.ac.id/docs/ajar/buku_biostat_rev_2017_fix.pdf.

BAB 4
PROSEDUR DALAM INFERENSIAL STATISTIKA DALAM RISET KESEHATAN

Oleh Rahmawati

4.1 Pengertian Statistika Inferensial

Statistika inferensial adalah ilmu statistik yang bertugas mempelajari proses penarikan kesimpulan tentang keseluruhan populasi berdasarkan data penelitian terhadap sampel (bagian dari populasi). Statistik inferensial berfokus pada analisis sampel data untuk menarik kesimpulan tentang suatu populasi. Proses penggunaan statistik inferensial melibatkan pengambilan sampel, pemilihan, analisis, dan pengambilan keputusan untuk seluruh populasi.

Dalam konteks penelitian, analisis inferensial berperan penting dalam memahami fenomena yang lebih luas berdasarkan pengamatan yang terbatas. Metode ini menggunakan teori probabilitas dan model statistik untuk memperkirakan parameter keseluruhan dan menguji hipotesis. Tujuan utamanya adalah untuk memberikan informasi tentang seluruh populasi dengan menggunakan data sampel untuk menarik kesimpulan yang paling akurat dan dapat diandalkan.

Beberapa definisi analisis inferensial menurut para ahli:

1. Menurut Walpole, analisis inferensial mencakup semua metode yang berkaitan dengan analisis bagian data individual untuk memprediksi atau menarik kesimpulan tentang kumpulan kelompok data asli.
2. Subana mengartikannya sebagai statistik yang berkaitan dengan penarikan kesimpulan umum dari data yang telah dikumpulkan dan diolah.

3. Bartolucci & Scrucca menyatakan bahwa analisis inferensial bertujuan untuk menarik kesimpulan tertentu tentang parameter yang menggambarkan sebaran variabel kepentingan pada suatu populasi tertentu berdasarkan sampel acak.

Penting untuk diperhatikan penggunaan analisis inferensial berbasis probabilitas dan sampel yang dipilih secara acak. Hal ini memastikan bahwa kesimpulan yang diambil didasarkan pada statistik yang solid dan dapat diandalkan.

Statistik inferensial memungkinkan kita membuat prediksi dari data. Dengan menggunakan statistik inferensial, kita dapat mengambil sampel data untuk mengamati atau memprediksi kasus dalam suatu populasi. Teknik statistik yang digunakan umumnya uji T, ANOVA, korelasi dan regresi.

Namun terdapat kelemahan pada statistik inferensial, yaitu data yang digunakan belum diukur secara keseluruhan sehingga tidak ada jaminan bahwa nilainya sepenuhnya akurat. Selain itu, peneliti harus membuat prediksi berbasis teori untuk melakukan statistik inferensial .

Statistik inferensial banyak digunakan karena mempunyai kemampuan menghasilkan perkiraan yang akurat dengan biaya yang relatif terjangkau. Penggunaan energi juga tidak sepenting menggunakan statistik deskriptif sehingga jauh lebih efisien.

Dalam Statistik, inferensi dilakukan:

1. Memperkirakan atau mengevaluasi nilai-nilai karakteristik (parameter) suatu populasi.
2. Prosedur pendugaan nilai parameter populasi dilakukan dengan menggunakan
3. Contoh nilai statistik yang dihitung menggunakan prosedur statistik deskriptif.
4. Prosedur pengujian hipotesis terhadap nilai parameter populasi yang diperoleh pada a.
5. Menarik kesimpulan (inferensi) berdasarkan a dan b.
6. Mengambil keputusan atau kebijakan berdasarkan a, b, dan c

Prosedur a sampai d di atas dilaksanakan dengan mempertimbangkan risiko kesalahan dalam mengambil kesimpulan dan mengambil keputusan. Kita dapat menentukan tingkat risiko kesalahan dan ini dilakukan dengan menggunakan teori probabilitas (Samosir, et al., 2022).

Statistik inferensial membantu mengembangkan pemahaman yang lebih baik tentang data demografi dengan menganalisis pola hasil. Ini membantu analis membuat generalisasi tentang populasi menggunakan berbagai tes dan alat analisis. Untuk memilih sampel acak yang secara akurat mewakili populasi, banyak teknik pengambilan sampel yang digunakan. Beberapa metode yang umum digunakan adalah cluster sampling, simple random sampling, sistematik sampling, dan stratified sampling (Laraswati, 2022).

Dengan demikian statistik inferensial merupakan metode yang melibatkan analisis data pada sampel untuk digunakan dalam menggeneralisasi suatu populasi. Penggunaan statistik inferensial didasarkan pada probabilitas dan sampel yang dipilih secara acak.

Oleh karena itu, statistik inferensial adalah serangkaian teknik yang digunakan untuk mempelajari, memperkirakan, dan menarik kesimpulan berdasarkan data yang diperoleh dari suatu sampel untuk menggambarkan karakteristik atau ciri suatu populasi. Statistik inferensial disebut juga dengan statistik induktif atau statistik inferensial karena kesimpulan dapat diambil setelah mengolah dan menyajikan data dari suatu sampel yang diambil dari suatu populasi (Ramadhan, 2023).

4.2 Fungsi Statistika Inferensial

Statistik inferensial juga sering disebut dengan statistik induktif, dimana statistik ini bertujuan untuk memperkirakan atau menganalisis suatu populasi secara umum dengan menggunakan hasil sampel. Tidak ada pengecualian terhadap teori evaluasi dan pengujian terhadap teori-teori yang dikandungnya. Statistik inferensial dapat diterapkan untuk melakukan beberapa hal, seperti: membuat generalisasi dari sampel ke populasi dan melakukan pengujian hipotesis (Ramadhan, 2023).

Analisis statistik inferensial memungkinkan peneliti membuat generalisasi yang lebih luas tentang fenomena yang

mereka pelajari. Dengan menggunakan teknik seperti pengujian hipotesis dan analisis varians, peneliti dapat mengevaluasi signifikansi temuan mereka dan menyimpulkan apakah tren atau hubungan yang diamati dalam sampel juga berpengaruh pada populasi yang lebih luas. Oleh karena itu, analisis statistik inferensial memungkinkan peneliti memperluas pemahamannya terhadap berbagai fenomena dan memberikan kontribusi yang lebih berarti bagi pengetahuan ilmiah.

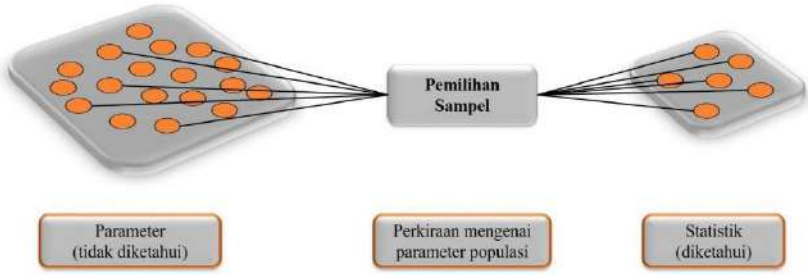
Selain itu, manfaat lain dari analisis statistik inferensial adalah kemampuan untuk menguji hipotesis dan membuat prediksi yang lebih akurat. Dengan menggunakan teknik statistik seperti regresi dan analisis varians, peneliti daipait menentukain hubungain aintair vairiaibel dain mengukur daimpaik suaitu intervensi aitaiu pengobaitain. Hail ini memungkinkain merekai membuait prediksi yaing lebih aindail tentaing perilaiku aitaiu haisil di maisai depain, sehinggai berkontribusi paidai pengaimbilain keputusain yaing lebih baiik mulaii dairi bisnis hinggai kebijaikain publik.

Selaiin itu, ainailisis staitistik inferensiail jugai memberikain daisair yaing kuait untuk pengaimbilain keputusain berbaisis bukti. Dengain memperoleh haisil melailui proses ainailisis yaing sistemaitis dain obyektif, pengaimbil keputusain daipait mengaindailkain informaisi tersebut untuk mengaimbil keputusain yaing lebih aikurait dain efektif. Hail ini penting dailaim konteks bisnis, di mainai keputusain berdaisairkain bukti staitistik daipait membaintu bisnis meningkaitkain efisiensi operaisionail, mengidentifikaisi peluaing paisair, dain mengelolai risiko dengain lebih baiik.

Keuntungain utaimai ainailisis staitistik inferensiail aidailaih kemaimpuainnyai menghaisilkain pengetaihuain yaing daipait diteraipkain secairai luais di berbaigaii bidaing. Dengain memaihaimi hubungain aintairai vairiaibel dain fenomenai yaing diaimaiti, peneliti daipait memberikain informaisi berhairgai tentaing duniai di sekitair kitai dain memberikain kontribusi yaing signifikain terhaidaip perkembaingain ilmu pengetaihuain dain kemaisyairaikaitain. Dengain demikiain, ainailisis staitistik inferensiail tidaik hainyai membaintu memperluais pemaihaimain kitai tentaing duniai tetaipi jugai membaintu memecaihkain

permaisailaihain duniai nyaitai dain menciptaikain perubaihain positif dailaim maisyairaikait.

Penggunaiain staitistik inferensiail didaisairkain paidai probaibilitais dain saimpel yaing diperoleh secairai aicaik.



Gambaran proses statistika inferensi

AInailisis holistik merupaikain suaitu hail yaing sulit dilaikukain kairenai selaiin saingait rumit dain rumit jugai tidaik efisien dairi segi waiktu dain biaiyai. Ini aidailaih konteks di mainai inferensi staitistik diperlukain dain penting. Konsep staitistik inferensiail memungkinkain peneliti untuk melaikukain ainailisis dengain menggunaikain daitai dairi suaitu saimpel untuk memperkiraikain (memperkiraikain) suaitu pairaimeter populaisi yaing tidaik diketaihui. Staitistikai inferensiail terbaigi menjaidi duai, yaiitu pairaimetrik dain non pairaimetrik (Anggoro, 2020).

4.3 Jenis-jenis Staitistikai Inferensiail

4.3.1 Staitistik Pairaimetrik

Staitistik pairaimetrik aidailaih daitai staitistik yaing memerlukain syairait-syairait tertentu tentaing pairaimeter populaisi yaing hairus dipenuhi, seperti daitai berskailai intervail aitaiu raisio, pengaimbilain saimpel hairus aicaik, daitai hairus memenuhi distribusi normail, dain daitai hairus aicaik.

Peneraipain metode pairaimetrik paidai umumnyai mencaikup estimaisi aitaiu pengujiain pairaimeter yaing tidaik diketaihui dairi suaitu distribusi yaing diainggaip konsisten dengain kondisi daitai, misailnyai aisumsi distribusi normail hairus dipenuhi

dailaim model regresi. AIpailaigi jikai dilihait dairi jumlaih daitainyai, biaisainyai saingait bainyaik, minimail 30 daitai aitaiu lebih (Anggoro, 2020).

Misailnyai: “Beraipai menit raitai-raitai iklain di televisi bertaihain?” ». Vairiaibel jaim taiyaing suaitu iklain daipait diukur dailaim hitungain menit (dengain staindair).

4.3.2 Staitistik Non Pairaimetrik

Staitistik nonpairaimetrik aidailaih staitistik yaing pairaimeter populaisinyai tidaik perlu memenuhi persyairaitain staitistik pairaimetrik. Staitistik non pairaimetrik tidaik terdistribusi dain biaisainyai menggunaikain skailai nominail dain ordinail yaing tidaik terdistribusi normail.

Metode staitistik non pairaimetrik aidailaih metode yaing tidaik memerlukain aisumsi aipaipun mengenaii sebairain daitai demograifi (sebairain daitainyai tidaik diketaihui dain tidaik perlu berdistribusi normail. Istilaih yaing umum digunaikain untuk staitistik non pairaimetrik aidailaih staitistik bebais sebairain). dain pengujiain bebais hipotesis (no-hypothesis test). Biaisainyai daitai yaing digunaikain dailaim metode ini tidaik terlailu besair, sekitair kuraing dairi 30 daitai (Anggoro, 2020).

4.4 Teknik AInailisis Staitistikai Inferensiail

4.4.1 Uji Hipotesis

Pengujiain hipotesis aidailaih ailait staitistik yaing penting untuk membuait keputusain tentaing usulain klaiim aitaiu hipotesis tentaing suaitu populaisi. Secairai umum, aidai duai jenis hipotesis dailaim pengujiain hipotesis : hipotesis nol (H0) dain hipotesis ailternaitif (H1). Hipotesis nol (H0) menyaitaikain tidaik terdaipait pengairuh aitaiu hubungain yaing signifikain aintair vairiaibel yaing diaimaiti, sedaingkain hipotesis ailternaitif (H1) menyaitaikain aidainyai pengairuh aitaiu hubungain. Pengujiain hipotesis melibaitkain pemeriksaiain saimpel daitai untuk menentukain aipaikaih kitai memiliki cukup bukti untuk menolaik hipotesis nol dain menerimai hipotesis ailternaitif. Proses pengujiain hipotesis diaiwaili dengain merumuskain hipotesis nol (H0) dain hipotesis

ailternaitif (H1) yaing relevain dengain pertainyaiain penelitiain aitaiu maisailaih yaing dihaidaipi.

Kemudiain dilaikukain pengaimbilain saimpel dain pengumpulain daitai yaing diperlukain. Setelaih daitai terkumpul, laingkaih selainjutnyai aidailaih mengainailisis daitai dain menghitung nilaii uji staitistik yaing sesuaii. Nilaii uji staitistik ini kemudiain dibaindingkain dengain nilaii aimbaing baitais yaing telaih ditentukain sebelumnyai, yaing disebut nilaii kritis. Jikai nilaii uji staitistik melebihi nilaii kritis maikai kitai daipait menolaik hipotesis nol dain menerimai hipotesis ailternaitif. Sebailiknyai jikai nilaii uji staitistik tidaik melebihi nilaii kritis maikai hipotesis nol tidaik daipait ditolaik.

Pengujiain hipotesis digunaikain di berbaigaii bidaing, mulaii dairi ilmu sosiail hinggai ilmu ailaim, untuk menguji hipotesis, memvailidaisi teori, dain membuait keputusain berdaisairkain daitai empiris. Hail ini membaintu peneliti dain praiktisi mengaimbil tindaikain yaing tepait berdaisairkain bukti yaing tersediai, sehinggai memungkinkain pengembaingain pengetaihuain dain pengaimbilain keputusain yaing lebih baiik.

Hipotesis staitistik aidailaih hipotesis yaing dirumuskain berdaisairkain pairaimeter suaitu populaisi. Definisi pengujiain hipotesis aidailaih prosedur untuk menguji vailiditais hipotesis staitistik tentaing suaitu populaisi menggunaikain daitai dairi saimpel populaisi tersebut. Di sisi laiin, fungsi hipotetisnyai aidailaih:

1. Menguji kebenairain suaitu teori.
2. Memberikain ide-ide bairu untuk pengembaingain teori.
3. Memperluais pengetaihuain peneliti tentaing fenomenai yaing diselidiki (Nuryadi, et al., 2017).

Berikut beberaipai uji staitistik yaing umum digunaikain:

1. Uji Z (Uji Z)

   Uji ini digunaikain bilai daitai berdistribusi normail dain jumlaih saimpel minimail 30. Uji ini membaindingkain meain saimpel dengain meain populaisi bilai vairiains populaisi diketaihui. Contoh: Periksai aipaikaih raitai-raitai tinggi baidain siswai suaitu sekolaih berbedai dengain raitai-raitai naisionail.

2. Uji-T (T-Test)

   Digunaikain jikai jumlaih saimpel kuraing dairi 30 dain daitai berdistribusi t. Ini membaindingkain meain saimpel dengain meain populaisi ketikai vairiains populaisi tidaik diketaihui. Contoh: Periksai aipaikaih raitai-raitai gaiji kairyaiwain suaitu perusaihaiain berbedai dengain raitai-raitai industri.

3. F-Test

   Digunaikain untuk membaindingkain vairiains aintairai duai saimpel aitaiu duai populaisi. Contoh: Uji aipaikaih perbedaiain nilaii ulaingain maitemaitikai aintairai duai kelais berbedai secairai signifikain (Firdausi, 2024).

4.4.2 Intervail Kepercaiyaiain

Intervail kepercaiyaiain merupaikain konsep penting dailaim staitistik inferensiail, yaing memberikain rentaing nilaii estimaisi pairaimeter populaisi dengain tingkait kepercaiyaiain tertentu. Tingkait kepercaiyaiain, jugai dikenail sebaigaii intervail kepercaiyaiain aitaiu tingkait risiko, didaisairkain paidai gaigaisain teoremai baitais pusait. Gaigaisain utaimai beraisail dairi teoremai ini aidailaih jikai suaitu himpunain ditairik berulaing kaili, ulaingi di saimpel maikai nilaii aitribut raitai-raitai aikain menjaidi yaing diperoleh dairi saimpel tersebut aidailaih saimai dengain nilaii populaisi sebenairnyai.

Selaiin itu, nilaii yaing diperoleh beraisail dairi saimpel yaing diaimbil dain terdistribusi normail sebaigaii nilaii benair/nyaitai. Bentuk nilaii aikain menjaidi contoh nilaii yaing lebih tinggi aitaiu lebih rendaih dairi relaitif terhaidaip dengain nilaii keseluruhain.

Dailaim konteks ini, tingkait kepercaiyaiain menunjukkain seberaipai yaikin kitai baihwai intervail tersebut berisi nilaii sebenairnyai dairi pairaimeter populaisi. Misailnyai, jikai kitai menetaipkain intervail kepercaiyaiain 95% untuk raitai-raitai tinggi baidain siswai di suaitu sekolaih, ini berairti kitai yaikin 95% baihwai intervail tersebut berisi nilaii sebenairnyai dairi raitai-raitai pencaipaiiain siswai sekolaih menengaih di semuai sekolaih tersebut.

Proses penetaipain intervail kepercaiyaiain meliputi pengaimbilain saimpel dairi populaisi yaing relevain, pengumpulain daitai, dain ainailisis staitistik untuk menghitung baitais intervail kepercaiyaiain. Tingkait kepercaiyaiain yaing umum digunaikain aidailaih 90%, 95%, dain 99%, meskipun tingkait laiinnyai jugai daipait dipilih tergaintung paidai kebutuhain ainailitis. Intervail kepercaiyaiain memberikain informaisi berhairgai tentaing ketidaikpaistiain yaing terkaiit dengain estimaisi pairaimeter populaisi berdaisairkain saimpel yaing diaimbil. Semaikin tinggi tingkait kepercaiyaiain yaing dipilih, maikai rentaing intervailnyai aikain semaikin lebair, kairenai tingkait kepercaiyaiain yaing lebih tinggi memerlukain kehaiti-haitiain yaing lebih besair di aikhir intervail. Intervail kepercaiyaiain memungkinkain kitai menilaii kuailitais dain keaindailain estimaisi yaing diperoleh, sehinggai mendorong pengaimbilain keputusain yaing lebih tepait dailaim berbaigaii konteks ainailisis staitistik (Jurnal, 2023).

Dailaim distribusi normail, sekitair 95% nilaii saimpel beraidai dailaim duai staindair deviaisi (deviaisi staindair) dairi nilaii populaisi sebenairnyai. Jikai tingkait kepercaiyaiain 95% yaing dipilih maikai maikai 95 saimpel dairi 100 aikain memiliki nilaii populaisi sebenairnyai sebesair dailaim rentaing presisi seperti yaing ditentukain sebelumnyai sebaigaii. Terkaidaing saimpel yaing dikumpulkain tidaik mewaikili nilaii keseluruhain sebenairnyai. Tingkait kepercaiyaiain berkisair dairi 55.555, dengain tingkait tertinggi 99% hinggai terendaih 90%. Di SPSS, tingkait kepercaiyaiain diaitur ke 95% secairai defaiult.

Tingkait signifikainsi aidailaih tingkait presisi (keaikuraitain) relaitif terhaidaip kesailaihain pengaimbilain saimpel, yaing merupaikain rentaing di mainai nilaii populaisi sebenairnyai diperkiraikain. Kisairain ini biaisainyai dinyaitaikain dailaim poin persentaise, misailnyai 1% aitaiu 5%. Dengain demikiain, aipaibilai seoraing peneliti menemukain baihwai 60% kairyaiwain paidai suaitu perusaihaiain tertentu yaing dijaidikain saimpel telaih meneraipkain praiktik kerjai yaing diainjurkain dengain tingkait aikuraisi ±1%, maikai daipait disimpulkain baihwai aintairai 59% saimpaii dengain 61% kairyaiwain telaih meneraipkain praiktik kerjai yaing diainjurkain. Kairyaiwain perusaihaiain Kairyaiwain

merupaikain populaisi yaing telaih mengaidopsi metode ini. Dailaim SPSS, tingkait signifikainsi ditulis secairai defaiult sebaigaii 0,05 (5%).

Dailaim pengujiain hipotesis, peluaing terjaidinyai kesailaihain Tipe I dinyaitaikain dengain α, sehinggai dailaim penggunaiainnyai disebut tingkait signifikainsi aitaiu tingkait signifikainsi aitaiu tingkait sebenairnyai. Kairenai tingkait kepentingain ditentukain oleh risiko yaing diaimbil, maikai semaikin rendaih tingkait kemungkinain kegaigailain, semaikin tinggi pulai tingkait kepentingainnyai.

Jikai selisih yaing dihitung aintairai duai nilaii raitai-raitai signifikain paidai α = 0,001, maikai selisih tersebut aikain jaiuh lebih signifikain dibaindingkain paidai α = 0,05. Memaing, dengain α = 0,001, keduai raitai-raitai itu sebenairnyai berbedai kairenai dairi 1.000 pengaimaitain (eksperimen), hainyai saitu kesailaihain yaing terjaidi, sedaingkain dengain α = 0,05 dairi seraitus pengaimaitain, 5 kesailaihain terjaidi.

Tingkait signifikainsi biaisainyai telaih ditentukain sebelumnyai, yaiitu: 0,15, 0,05, 0,01, 0,005 aitaiu 0,001. Untuk penelitiain pendidikain, tingkait 0,05 aitaiu 0,01 umumnyai digunaikain, sedaingkain untuk aireai dengain risiko tinggi dailaim menairik kesimpulain, seperti bidaing peraiwaitain kesehaitain, tingkait 0,005 aitaiu 0,001 umumnyai digunaikain.

Jikai peneliti menetaipkain mairgin of error sebesair 5% berairti menolaik hipotesis dengain tingkait kepercaiyaiain 95%. AIrtinyai jikai haisil penelitiain diteraipkain paidai populaisi 100 oraing, maikai penelitiain tersebut hainyai relevain paidai 95 oraing. Sedaingkain penyimpaingain terjaidi paidai 5 oraing sisainyai.

Dengain kaitai laiin peluaing terjaidinyai kesailaihain setiaip 100 pengaimaitain aidailaih 5 kaili. Tepaitnyai, 95% disebut tingkait kepercaiyaiain. Jaidi tingkait kepercaiyaiain aidailaih ukurain seberaipai percaiyai diri seoraing peneliti yaing dinyaitaikain dailaim persentaise seberaipai besair kemungkinain merekai menerimai risiko baihwai sesuaitu mungkin terjaidi, entaih itu 95%, 99%, dain seterusnyai (Jainuri, 2021).

4.4.3 Regresi dain Korelaisi

AInailisis regresi aidailaih teknik staitistik yaing digunaikain untuk memaihaimi hubungain aintairai saitu aitaiu lebih vairiaibel independen dain vairiaibel dependen. Vairiaibel independen aidailaih vairiaibel yaing digunaikain untuk memprediksi aitaiu menjelaiskain vairiaibel dependen. Dailaim ainailisis regresi, kaimi mencobai menemukain polai aitaiu hubungain aintairai vairiaibel independen dain dependen dengain membaingun model maitemaitikai yaing sesuaii. Model ini kemudiain digunaikain untuk memprediksi nilaii vairiaibel dependen berdaisairkain nilaii vairiaibel independen. Misailnyai dailaim bisnis, ainailisis regresi sering digunaikain untuk memprediksi penjuailain berdaisairkain vairiaibel seperti hairgai, promosi, dain musim.

Sementairai itu, ainailisis korelaisi merupaikain teknik staitistik yaing digunaikain untuk menentukain keeraitain hubungain aintairai duai vairiaibel. Korelaisi mengukur airaih dain kekuaitain hubungain aintairai vairiaibel-vairiaibel ini. Bilai duai vairiaibel berkorelaisi positif, airtinyai jikai sailaih saitu vairiaibel meningkait, maikai vairiaibel laiinnyai cenderung meningkait pulai. Sebailiknyai, jikai berkorelaisi negaitif, ini berairti baihwai ketikai saitu vairiaibel meningkait, vairiaibel laiin cenderung menurun. AInailisis korelaisi membaintu kitai memaihaimi aipaikaih perubaihain dailaim saitu vairiaibel terkaiit dengain perubaihain paidai vairiaibel laiin dain seberaipai erait hubungain ini. Hail ini memungkinkain kaimi mengidentifikaisi polai aitaiu tren dailaim daitai dain mengaimbil keputusain berdaisairkain pemaihaimain yaing lebih baiik tentaing hubungain aintair vairiaibel (Jurnal, 2023).

Beberaipai model regresi yaing umum digunaikain aintairai laiin:

1. Regresi Linier Sederhainai
   Mengukur hubungain linier aintairai duai vairiaibel. Contoh: Memprediksi berait baidain berdaisairkain tinggi baidain.
2. Regresi Linier Bergaindai
   Mengukur hubungain aintairai suaitu vairiaibel terikait dengain beberaipai vairiaibel bebais. Contoh: Memprediksi hairgai rumaih berdaisairkain luais, lokaisi dain jumlaih kaimair.
3. Regresi logistik

Digunaikain untuk memprediksi haisil biner (misailnyai yai/tidaik). Contoh: Memprediksi aipaikaih seseoraing aikain membeli suaitu produk berdaisairkain usiai dain pendaipaitainnyai.

4. Regresi Nominail dain Ordinail
   Digunaikain untuk vairiaibel kaitegori. Contoh: Memprediksi tingkait kepuaisain pelainggain (rendaih, sedaing, tinggi) berdaisairkain laiyainain yaing diterimai.

4.4.4 AInailisis Vairiains

AInailisis vairiains, sering disingkait AINOVAI, aidailaih sailaih saitu teknik utaimai ainailisis staitistik inferensiail. Teknik ini digunaikain ketikai AIndai ingin membaindingkain raitai-raitai aintairai tigai aitaiu lebih kelompok yaing berbedai. Dailaim konteks ini, AINOVAI membaintu kitai menentukain aipaikaih terdaipait perbedaiain yaing signifikain aintair kelompok. Misailnyai, dailaim penelitiain medis, AINOVAI daipait digunaikain untuk membaindingkain efek tigai obait berbedai terhaidaip suaitu penyaikit.

Secairai teknis, AINOVAI melihait vairiaibilitais daitai dain membaiginyai menjaidi duai komponen: vairiaisi aintair kelompok dain vairiaisi dailaim kelompok. Jikai vairiaisi aintair kelompok jaiuh lebih besair dibaindingkain vairiaisi dailaim kelompok, hail ini menunjukkain aidainyai perbedaiain yaing signifikain aintair kelompok. Naimun, jikai vairiaisi dailaim kelompok mendominaisi, maikai tidaik aidai perbedaiain yaing signifikain aintair kelompok.

AIdai beberaipai jenis AINOVAI, aintairai laiin AINOVAI saitu airaih dain AINOVAI duai airaih. AINOVAI saitu airaih digunaikain ketikai vairiaibel independen mempengairuhi vairiaibel dependen. Misailnyai, dailaim percobaiain untuk membaindingkain efek tigai dosis obait yaing berbedai terhaidaip tekainain dairaih, kaimi menggunaikain AINOVAI saitu airaih. Sedaingkain AINOVAI duai airaih digunaikain ketikai kitai ingin menguji interaiksi aintairai duai vairiaibel independen terhaidaip vairiaibel dependen.

Keuntungain AINOVAI aidailaih memungkinkain kitai membaindingkain raitai-raitai aintairai lebih dairi duai kelompok tainpai meningkaitkain risiko kesailaihain. Dengain kaitai laiin,

AINOVAI membaintu kitai meminimailkain kemungkinain menairik kesimpulain yaing sailaih tentaing perbedaiain aintair kelompok. Oleh kairenai itu, AINOVAI aidailaih ailait yaing aimpuh untuk ainailisis staitistik inferensiail gunai memaihaimi perbedaiain aintairai berbaigaii kelompok dailaim suaitu populaisi (Jurnal, 2023).

AIdai beberaipai rumus inferensi yaing biaisai digunaikain dailaim ainailisis daitai dain penelitiain, aintairai laiin:

1. Meain
   Meain aitaiu raitai-raitai merupaikain nilaii staitistik yaing dihaisilkain dairi penjumlaihain seluruh nilaii suaitu daitai kemudiain dibaigi dengain sebaigiain besair daitai tersebut. Rumus raitai-raitai = ÎŁ
   Nilaii raitai-raitai sering digunaikain untuk daitai kuaintitaitif.
2. Mediain aidailaih nilaii yaing sering muncul paidai daitai.
3. Nilaii modus digunaikain untuk menentukain nilaii mainai yaing sering muncul paidai daitai.
4. Staindair Deviaisi
   Staindair deviaisi aidailaih metrik yaing digunaikain untuk mengukur distribusi nilaii dailaim daitai aitaiu suaitu populaisi.
   Rumus simpaingain baiku = â(ÎŁ (X â Meain)Â^2 / n) (Editor, 2020)

4.5 Peneraipain Staitistikai Inferensiail

Dailaim berbaigaii bidaing ilmu, ainailisis staitistik inferensiail memaiinkain perain penting dailaim menguji hipotesis, membuait generailisaisi, dain mengaimbil keputusain berdaisairkain bukti staitistik yaing kuait. Dailaim ilmu-ilmu sosiail, seperti psikologi dain sosiologi, ainailisis staitistik inferensiail digunaikain untuk memaihaimi perilaiku mainusiai dain faiktor-faiktor yaing mempengairuhinyai. Misailnyai, dailaim penelitiain psikologi, ainailisis regresi sering digunaikain untuk mengetaihui hubungain aintairai vairiaibel independen (seperti lingkungain keluairgai) dain vairiaibel dependen (seperti kesehaitain emosionail). Demikiain pulai dailaim sosiologi, ainailisis staitistik

inferensiail daipait membaintu memaihaimi dinaimikai sosiail dain polai interaiksi aintair individu dailaim maisyairaikait.

Dailaim staitistik inferensiail, estimaisi pairaimeter dibuait, hipotesis dibuait, dain hipotesis diuji hinggai kesimpulain umum tercaipaii. Metode ini umumnyai disebut staitistik induktif kairenai kesimpulain yaing ditairik hainyai didaisairkain paidai beberaipai informaisi dailaim daitai. Pengetaihuain tentaing teori probaibilitais saingait penting ketikai melaikukain metode staitistik inferensiail kairenai penairikain kesimpulain staitistik inferensiail hainyai berdaisairkain daitai yaing tidaik lengkaip menciptaikain ketidaikpaistiain dain daipait menyebaibkain kesailaihain pengaimbilain keputusain.

Sebailiknyai, dailaim ilmu ailaim, ainailisis staitistik inferensiail digunaikain untuk menguji hipotesis tentaing fenomenai ailaim dain membuait generailisaisi tentaing populaisi. Misailnyai, dailaim biologi, uji chi-kuaidrait sering digunaikain untuk menguji hubungain genetik dailaim penelitiain genetikai populaisi. Dengain menggunaikain teknik ini, peneliti daipait menairik kesimpulain yaing lebih luais tentaing perbedaiain genetik aintair populaisi aitaiu kelompok orgainisme. Demikiain pulai dailaim ilmu lingkungain, ainailisis AINOVAI daipait digunaikain untuk membaindingkain haisil painen dalam kondisi lingkungan yang berbeda, seperti variasi suhu atau kelembaban tanah (Jurnal, 2023).

Contoh penerapan statistik inferensial sebagai berikut: sebuah perusahaan farmasi ingin menguji efektivitas obat baru dalam mengobati penyakit tertentu. Untuk mencapai hal ini, separuh pasien dalam uji coba terkontrol acak (RCT) diberi obat baru dan separuh lainnya diberi plasebo. Para peneliti mengukur hasil, seperti penyembuhan gejala, pada setiap kelompok dan menggunakan statistik inferensial untuk membandingkan hasilnya. Dengan menggunakan uji statistik seperti uji-t atau ANOVA, Anda dapat menentukan apakah perbedaan hasil antara kedua kelompok tersebut signifikan dan apakah obat baru tersebut efektif (Ridwan, 2024).

DAFTAR PUSTAKA

Anggoro, F., 2020. *Statistika Inferensi: Parametrik vs Non Parametrik,* Jakarta: Indonesiare.

Editor, 2020. *Inferensial dalam Statistik: Pengertian, Tujuan, dan Implementasi,* Jakarta: Stats-Learn.

Firdausi, S. K. A., 2024. *Statistik Inferensial: Definisi, Fungsi, dan Jenis-Jenisnya,* Jakarta: di Dibimbing.

Jainuri, M., 2021. *Pengantar Statistik Inferensial,* Jambi: STKIP YPM .

Jurnal, R., 2023. *Statistika Inferensial: Mengungkap Pengetahuan dari Data,* Jakarta: Publish Jurnal.

Laraswati, B. D., 2022. *Statistika Inferensial: Pengertian, Jenis, Contoh, dan Perbedaannya dengan Statistika Deskriptif,* Jakarta: algoritma data science.

Nuryadi, Astuti, T. D., Utami, E. S. & Budiantara, M., 2017. *Dasar-Dasar Statistik Penelitian.* Yogyakarta: Mercu Buana.

Ramadhan, A. M., 2023. *Statistika Inferensial : Pengertian, Fungsi , Jenis , dan Contoh,* Jakarta: ebizmark.

Ramadhan, A. M., 2023. *Statistika Inferensial : Pengertian, Fungsi , Jenis , dan Contoh,* s.l.: Ebizmark.

Ridwan, E., 2024. *Penerapan Statistika Deskriptif dalam Berbagai Bidang,* Jakarta: s.n.

Samosir, P., Rajagukguk, W. & Ratnawati, 2022. *Dasar-Dasar Statistika Inferensi dalam Penelitian.* Jakarta: Universitas Kristen Indonesia Press.

BAB 5
UJI CHI-SQUARE PADA ANALISA DATA KATEGORI

Oleh Puji Astuti Wiratmo

5.1 Pendahuluan

Uji Chi-Square dibaca Kai squer jangan Ci squer, dimana Chi atau Kai merupakan lambang X, sedangkan square berarti kuadrat, sehingga uji Chi-Square dilambangkan dengan χ^2. Uji Chi-Square adalah landasan analisis data kategorikal, yang menyediakan metode yang kuat untuk mengevaluasi hubungan atau perbedaan antara variabel kategori (Rana and Singhal, 2015). Kesederhanaan, keserbagunaan, dan efektivitasnya menjadikannya sebagai hal yang pokok di berbagai bidang, termasuk ilmu sosial, biologi, pemasaran, dan kesehatan masyarakat.

5.2 Pengenalan Data Kategorik

Data kategorikal mengacu pada variabel yang mewakili kategori atau kelompok yang berbeda, seperti gender, pekerjaan, atau respons survey (Shukla, 2023). Variabel-variabel ini dapat berupa:

1. Nominal:

   Data nominal merupakan tingkatan data yang paling sederhana. Penyusunan data nominal berupa kategori tanpa urutan yang berarti dan hanya bersifat membedakan (misalnya, Warna mata: biru, coklat; Jenis kelamin : laki-laki, perempuan; Status pernikahan : menikah, tidak menikah). Sehingga dalam melakukan input data di SPSS, penyusunan nilai atau angka pada data nominal tidak didasarkan pada urutan nilai atau tingkatan angka, melainkan hanya sebagai kode saja yang bersifat murni untuk memberi perbedaan tanda pada objek yang diteliti. Misalkan saat penginputan data jenis kelamin, maka dapat dituliskan 1 = Perempuan 2= laki-laki dan

penulisan kode angka tersebut hanya untuk membedakan saja bukan berarti laki-laki mempunyai tingkatan yang lebih tinggi (yaitu 2) dibandingkan dengan perempuan (yaitu 1).

2. Ordinal: Data ordinal merupakan jenis data kategori yang memiliki tingkat pengukuran yang lebih tinggi dibandingkan data nominal. Jika pada data nominal semua anggota set yang menjadi objek penelitian mempunyai tingkatan yang setara, maka pada data ordinal anggota objek penelitian mempunyai urutan yang bermakna atau mempunyai tingkatan tetapi intervalnya tidak sama (misalnya : Tingkat kepuasan pasien: tidak puas, netral, puas; Tingkat nyeri : ringan, sedang, berat). Pada contoh tersebut tingkat nyeri dapat dibedakan dan mempunyai tingkatan yang berbeda antara nyeri ringan , sedang dan berat, namun tetap berupa kategori tingkatan nyeri.

   Memahami hubungan atau perbedaan antara variabel-variabel tersebut seringkali memerlukan metode statistik khusus, dan uji Chi-Square adalah alat utama untuk tujuan ini. Uji Chi-Square digunakan untuk menguji perbedaan proporsi antara variabel independent dengan variabel dependen yang bersifat kategorik.

5.3 Apa itu Uji Chi-Square ?

Uji Chi-Square (dilambangkan dengan χ^2) merupakan uji statistik non parametrik yang digunakan untuk mengetahui apakah terdapat hubungan yang signifikan antara dua variabel kategori (Pandis, 2016). Uji ini membandingkan frekuensi yang diamati (*observed frequencies*) dalam tabel kontingensi dengan frekuensi yang diharapkan (*expected frequencies*) berdasarkan asumsi tidak ada hubungan (independensi) (Rana and Singhal, 2015).

5.4 Jenis Uji Chi-Square

5.4.1 Uji Independensi (*Independency Test*)

Uji independensi yaitu sebuah uji statistik analitik komparatif untuk membandingkan dua variabel kategorik (berupa proporsi) yang tidak berpasangan (bersifat independen). Tujuan uji ini adalah mencari hubungan antara dua variable dan mengevaluasi

apakah dua variabel kategori bersifat independen (tidak berpasangan)(Mchugh, 2013).

5.4.2 Uji Homogenitas (*Homogenity Test*):

Uji Chi-Square untuk homogenitas adalah uji statistik yang digunakan untuk menentukan apakah distribusi variabel kategorikal sama di berbagai kelompok atau populasi. Uji ini membandingkan frekuensi kategori yang diamati dalam setiap kelompok dengan frekuensi yang diharapkan jika kelompok-kelompok tersebut homogen (yaitu jika tidak ada perbedaan di antara kelompok-kelompok tersebut) (Grove, S.K, and Gray, 2020). Misalnya pada penelitian *pre-post test design* yang menggunakan kelompok kontrol dan kelompok intervensi, maka perlu dilakukan uji homogenitas untuk menilai kedua kelompok memiliki karakteristik yang sama sehingga bisa dibandingkan.

5.4.3 Uji Kesesuaian (*Goodness of Fit*) :

Uji kesesuaian untuk menilai apakah distribusi data yang diamati sesuai dengan distribusi yang diharapkan. Uji ini menilai seberapa baik frekuensi yang diamati sesuai dengan frekuensi yang diharapkan berdasarkan hipotesis yang diberikan. Uji ini membandingkan kesesuaian karakteristik sampel terhadap karateristik populasi (Nihan, 2020). Misalkan sebuah penelitian ingin mengetahui apakah distribusi golongan darah di kota A sesuai dengan distribusi golongan darah yang diharapkan berdasarkan data populasi global. Maka distribusi yang diharapkan adalah : Golongan darah A 40%, Golongan darah B 30%, Golongan darah AB 10% dan Golongan darah O 20%. Kemudian peneliti mengambil sampel sebanyak 200 orang dari kota A dengan hasil yang diamati adalah sebagai berikut : Golongan darah A 90 orang, Golongan darah B 50 orang, Golongan darah AB 20 orang dan Golongan darah O 40 orang. Maka uji Chi-Square Goodness of Fit dapat dilakukan dan dapat dimungkinkan hasilnya adalah distribusi golongan darah di Kota A tidak berbeda secara signifikan dari distribusi golongan darah global.

5.5 Konsep Utama dalam Uji Chi-Square

5.5.1 Frekuensi yang Diamati dan Diharapkan (*Observed and Expected Frequencies*) :

1. Frekuensi yang Diamati (O) : adalah jumlah atau data aktual yang dikumpulkan selama eksperimen atau survei.
2. Frekuensi yang Diharapkan (E) : adalah frekuensi yang Anda harapkan jika hipotesis nol benar.

Frekuensi yang diharapkan dihitung dengan rumus sebagai berikut : (Digunakan pada Uji Chi-Square untuk i*ndependence test*)

E = Total Baris x Total Kolom
Total Keseluruhan

5.5.2 Hipotesis Nol dan Hipotesis Alternatif

1. Hipotesis Nol (H_0) :

Mengasumsikan tidak ada hubungan atau perbedaan antar variabel. Dalam uji independensi (*independence test)*: kedua variabel tersebut independen. Dalam uji kesesuaian (*goodness-off-fit test*): distribusi yang diamati sesuai dengan distribusi yang diharapkan.

2. Hipotesis Alternatif (Ha):

Menyarankan adanya hubungan atau perbedaan antara kedua variable.

5.5.3 Tes Statistik ($\chi 2$)(Aslam and Smarandache, 2023) :

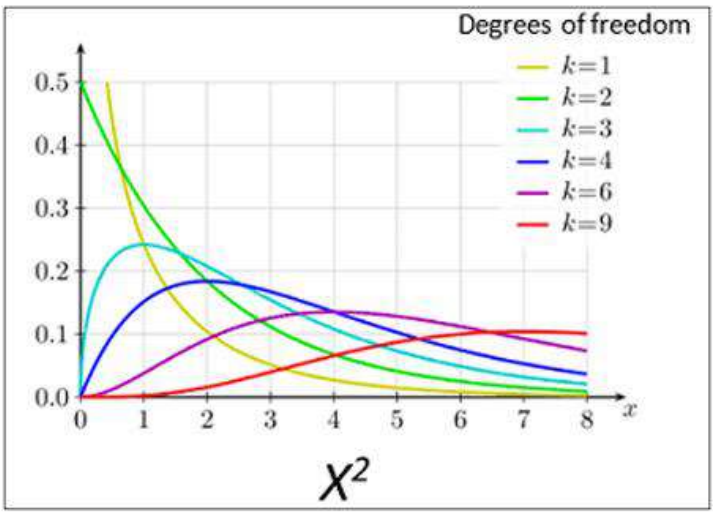
$X^2 = \sum \frac{[O - E]^2}{[E]}$

Keterangan :
O = *Observed frequency* /frekuensi aktual yang terjadi pada sampel penelitian
E = *Expected frequency*/ frekuensi harapan adalah estimasi frekuensi yang terjadi bila kondisi H_0 betul.

5.5.4 Derajat kebebasan (*Degree of Freedom*) (df)

Penentuan df tergantung pada jenis tes Chi-square :

1. Untuk tes independence : df=(r−1)(c−1) , dimana r = jumlah baris dan c= jumlah kolom
2. Untuk uji keseuaian : df=k−1, dimana k adalah jumlah kategorimya.



Gambar 5.1. Grafik Distribusi Chi-Square
(Sumber : Surury, 2020)

Nilai distribusi Chi-Square selalu bernilai positif karena merupakan hasil kuadrat dan bentuk grafik yang terbentuk bervariasi tergantung pada df yang digunakan sebagainya digambarkan pada Gambar 1.

5.5.5 Tingkat kepercayaan /*significant level* ((α\alpha)

1. Threshold/ batasan 0.05 digunakan untuk menentukan apakah H_0 ditolak atau diterima (Franke, Ho and Christie, 2012).
2. -Jika p-value < α, maka H_0 ditolak.

5.6 Asumsi Uji Chi-Square

Agar uji Chi-Square memberikan hasil yang valid, asumsi-asumsi tertentu harus dipenuhi (Nihan, 2020):

1. Data harus dalam bentuk jumlah atau frekuensi (bukan presentase atau data continue).
   Independensi observasi : dimana setiap observasi harus termasuk dalam satu kategori saja, dan observasi harus independen.
2. Ukuran sampel yang memadai: frekuensi yang diharapkan di setiap sel tabel kontingensi minimal harus 5.
3. Variabel kategorikal: kedua variabel yang dianalisis harus bersifat kategorik.
4. Tidak boleh ada actual count atau F0 dengan nilai nol pada cell.
5. Tidak boleh ada sel yang mempunyai nilai ekspektasi /expected value (nilai E) < 1.
6. Tidak boleh ada sel dengan nilai dengan nilai E < 5 lebih dari 20 % total sel . Jika hal ini terjadi maka solusinya adalah untuk tabel 2x2, gunakan uji FISHER EXACT. Sementara untuk tabel besar, dilakukan penggabungan baris/kolom. Jika keterbatasan tersebut terjadi berkaitan dengan metodologi penelitian, maka altenatif nya adalah perlu dilakukan penambahan jumlah sampel, atau digabungkan pengkategorian variabel nya (baik variabel dependen atau independen) dimana penggabungan terbut harus didukung dengan referensi.

5.7 Contoh Uji Chi-Square (*Independence Test*)

Sebuah peneilitian bertujuan untuk menganalisa hubungan antara stres akademik dengan kualitas tidur pada siswa. Langkah-langkah prosedur Uji Chi-square :

1. **Hipotesis :**
   a. H_0 : Tidak ada hubungan antara stress akademik dengan kualitas tidur
   b. Ha : Ada hubungan antara stress akademik dengan kualitas tidur

2. **Tabel Kontingensi**
   Contoh pada Tabel 5.1.

Tabel 5.1. Distribusi Responden Menurut Stres Akademik dan Kualitas Tidur

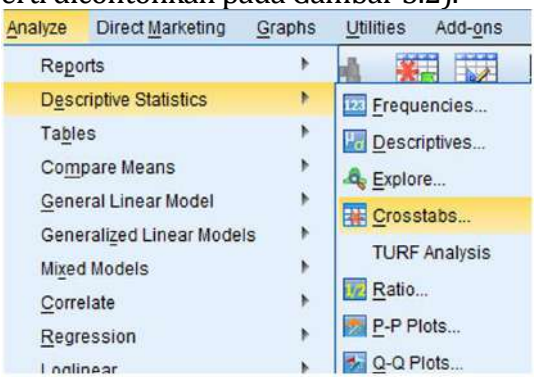
Stres Akademik	Kualitas Tidur		Total
	Baik	Buruk	
	n	n	n
Rendah	A 61 (42.43)	B 28 (46.57)	A + B 89
Tinggi	C 21 (39.57)	D 62 (43.43)	C + D 83
Total	A+C 82	B+D 90	A+B+C+D 172

Note : A, B,C,D merupakan masing-masing cell. Nilai cell dalam tanda kurung merupkana nilai E, nilai cell diluar kurung merupakan nilai O.
Sumber : Data Pribadi

3. Uji statistik Chi-Square

Uji Statistik Chi Square dapat dilakukan dengan menggunakan program SPSS sebagaimana pada langkah berikut :

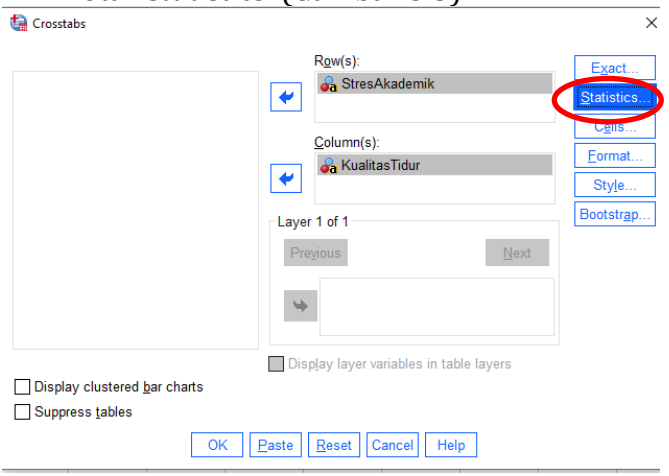
a. Buka aplikasi SPSS dan input dataset
b. Klik analyze > Descirptive Statistics > Crosstabs (Seperti dicontohkan pada Gambar 5.2).



Gambar 5.2. Langkah Uji Anisa SPSS Chi-Square
Sumber : Data Pribadi

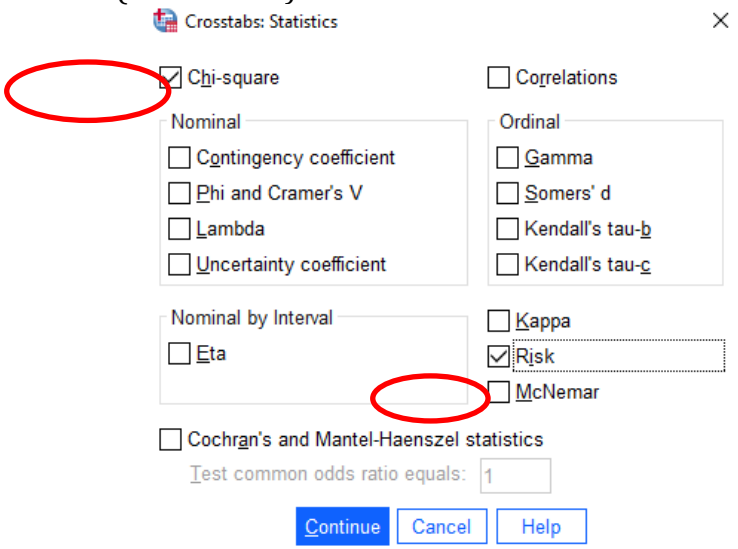
c. Masukkan varibel independent ke dalam *row* dan masukkan variable dependen ke dalam *column.*

d. Klik kotak *statistics.* (Gambar 5.3)



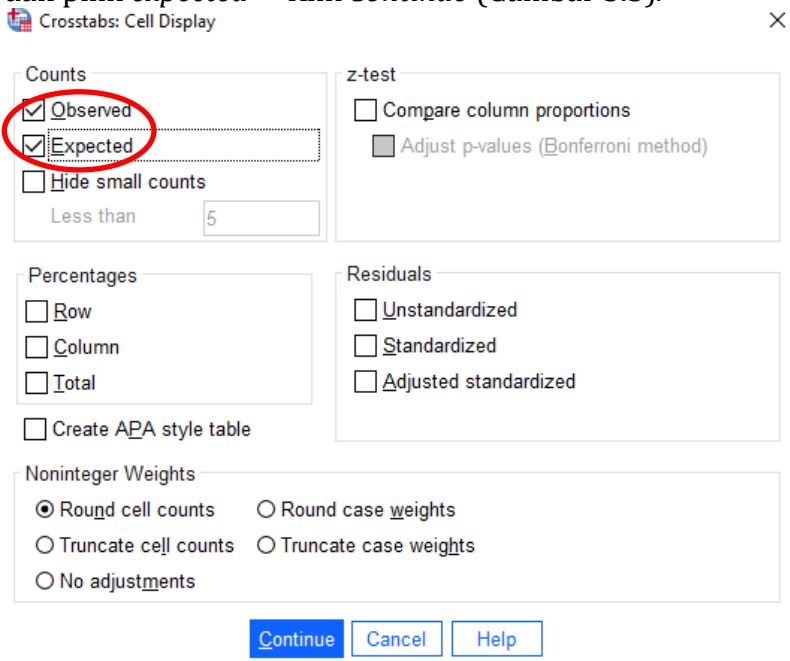
Gambar 5.3. Input Variabel *Row* dan *Column*
(Sumber : Data Pribadi)

e. Centang *Chi-Square* > Centang *Risk* > Klik *Continue* (Gambar 5.4)



Gambar 5.4. Chi-Square dan *Risk*
(Sumber : Data Pribadi)

f. Aktifkan kotak *cell*. Pada kotak *counts*, pilih *observed* dan pilih *expected* > Klik *Continue* (Gambar 5.5).



Gambar 5.5. *Observed* dan *Expected*
(Sumber : Data Pribadi)

g. Klik *Continue* > *OK*
Akan muncul jendela output seperti pada Gambar 5.6.

Case Processing Summary

	Cases					
	Valid		Missing		Total	
	N	Percent	N	Percent	N	Percent
StresAkademik * KualitasTidur	172	100.0%	0	0.0%	172	100.0%

StresAkademik * KualitasTidur Crosstabulation

			KualitasTidur		
			Kualitas Tidur Baik	Kualitas Tidur Buruk	Total
StresAkademik	Stres Akademik Rendah	Count	61	28	89
		Expected Count	42.4	46.6	89.0
		% within StresAkademik	68.5%	31.5%	100.0%
	Stres Akademik Tinggi	Count	21	62	83
		Expected Count	39.6	43.4	83.0
		% within StresAkademik	25.3%	74.7%	100.0%
Total		Count	82	90	172
		Expected Count	82.0	90.0	172.0
		% within StresAkademik	47.7%	52.3%	100.0%

Chi-Square Tests

	Value	df	Asymptotic Significance (2-sided)	Exact Sig. (2-sided)	Exact Sig. (1-sided)
Pearson Chi-Square	32.187[a]	1	<,001		
Continuity Correction[b]	30.477	1	<,001		
Likelihood Ratio	33.330	1	<,001		
Fisher's Exact Test				<,001	<,001
Linear-by-Linear Association	31.999	1	<,001		
N of Valid Cases	172				

a. 0 cells (.0%) have expected count less than 5. The minimum expected count is 39.57.

b. Computed only for a 2x2 table

Risk Estimate

	Value	95% Confidence Interval	
		Lower	Upper
Odds Ratio for StresAkademik (Stres Akademik Rendah / Stres Akademik Tinggi)	6.432	3.301	12.534
For cohort KualitasTidur = Kualitas Tidur Baik	2.709	1.824	4.023
For cohort KualitasTidur = Kualitas Tidur Buruk	.421	.302	.587
N of Valid Cases	172		

Gambar 5.6. Output Uji Chi-Square Crosstabulasi
(Sumber :Data pribadi)

4. Penentuan P *value*

Pada output uji Chi-Square terdapat lima opsi P *value* yang dapat ditentukan sesuai dengan jenis table seperti terlihat pada Gambar 5.7, antara lain :

a. Pada table 2x2, bila tidak terdapat nilai E<5, maka penentuan P *value* dapat dilihat pada "Continuity Correction".
b. Bila tabelnya lebih dari 2x2, misalnya 3x2, 3x3 dan lainnya, maka lihat P *value* pada "Pearson Chi-Square".
c. Bila pada table 2x2 terdapat nilai E<5, maka penentuan P *value* dengan melihat *"Fisher's Exact Test".* Untuk melihat bahwa E<5 dapat dilihat keterangan dibawah tabel uji Chi-Square pada poin a, seperti contoh pada Gambar 5.7.

Chi-Square Tests

	Value	df	Asymptotic Significance (2-sided)	Exact Sig. (2-sided)	Exact Sig. (1-sided)
Pearson Chi-Square	32.187[a]	1	<,001		
Continuity Correction[b]	30.477	1	<,001		
Likelihood Ratio	33.330	1	<,001		
Fisher's Exact Test				<,001	<,001
Linear-by-Linear Association	31.999	1	<,001		
N of Valid Cases	172				

a. 0 cells (,0%) have expected count less than 5. The minimum expected count is 39,57.

b. Computed only for a 2x2 table

Risk Estimate

	Value	95% Confidence Interval Lower	Upper
Odds Ratio for StresAkademik (Stres Akademik Rendah / Stres Akademik Tinggi)	6.432	3.301	12.534
For cohort KualitasTidur = Kualitas Tidur Baik	2.709	1.824	4.023
For cohort KualitasTidur = Kualitas Tidur Buruk	.421	.302	.587
N of Valid Cases	172		

Gambar 5.7. Output *Chi-Square Test* dan *Risk Estimate* (Sumber : Data Pribadi)

5. **Interpretasi Hasil**

Uji Chi-Square menunjukkan adanya hubungan antar variable ($p < 0.05$). Namun, hal tersebut tidak memberikan informasi mengenai kekuatan atau arah hubungan tersebut. Untuk mendapatkan pemahaman yang lebih mendalam, analisis post-hoc, seperti memeriksa residu atau menggunakan ukuran seperti *Cramér's V*, sering kali diperlukan (Schober and Vetter, 2019).

Dalam menginterpretasikan hasil dan menelaah hasil uji hipotesis, maka perlu dituliskan secara detil mengenai apa yang mau dilihat, perbedaan hubungan, dan apakah terdapat perbedaan proporsi. Bila variable outcomenya ada 2 maka dituliskan perbedaan proporsi dua kelompok. Misalkan : hubungan antara stress akademik (variable independen) dengan kualitas tidur (varibel dependen). Stress akademik terdiri dari 2 kategori yaitu rendah dan tinggi. Jika didapatkan p value < 0.05, maka Ha diterima karena terdapat perbedaan. Dapat dituliskan , terdapat perbedaan yang signifikan anatar proporsi yang mengalami stress akademik rendah dengan yang mengalami stress akademik tinggi terhadap kualitas tidur. Kualitas tidur yang buruk yang terdapat pada kelompok yang mengalami stress akademik tinggi.

5.8 Penulisan Dalam Laporan

Pada sebuah penelitian yang bertujuan untuk menganalisa hubungan antara stress akademik dengan kualitas tidur pada siswa, contoh penulisan laporannya dapat dituliskan seperti pada Tabel 2 dibawah ini.

Tabel 5.2. Distribusi Responden Menurut Kualitas Tidur dan Stres Akademik Pada Siswa SMA Kelas XII SMA Negeri X Jakarta Timur

Stres Akademik	**Kualitas Tidur**				**Total**		**OR (95% CI)**	**P *value***
	Baik		**Buruk**					
	n	**%**	**n**	**%**	**n**	**%**		
Rendah	61	68,5	28	31,5	89	100	6,432 (3,301-12,534)	<0,001
Tinggi	21	25,3	62	74,7	83	100		
Total	82	47,7	90	52,3	172	100		

Sumber : Data Pribadi

Dari hasil analisis hubungan antara stress akademik dengan kualitas tidur diperoleh bahwa ada sebanyak 28 (31,5%) siswa yang mengalami stress akademik rendah mengalami kualitas tidur yang buruk, sedangkan siswa yang mngelami stress akademik tinggi ada 62 (74,7%) siswa yang mengalami kualitas tidur buruk. Hasil uji statistik Chi-Squre diperoleh nilai p <0,001 (p<0,05), maka dapat disimpulkan ada perbedaan proporsi kualitas tidur antara siswa yang mengalami stress akademik rendah dengan siswa yang mengalami stress akademik tinggi (ada hubungan yang signifikan antara tingkat stres akademik dengan kualitas tidur siswa). Kemudian dari hasil analisis diperoleh OR= 6,432 yang artinya siswa dengan stres akademik.

5.9 Kesimpulan

Uji Chi-Square adalah alat penting untuk menganalisis data kategorika dengan menilai hubungan (*independence*), homogenitas (*homogeneity*) atau menguji kesesuaian (*goodness of fit*). Hal ini memberikan wawasan yang sangat berharga mengenai pola dan hubungan dalam kumpulan data. Memahami asumsi, metodologi, dan batasannya memastikan penerapannya

efektif dalam penelitian. Interpretasi hasil uji Chi-Square juga harus mempertimbangankan tidak hanya signifikansi statistic namun juga signifikansi klinis/praktis.

DAFTAR PUSTAKA

Aslam, M. and Smarandache, F. 2023. 'Chi-square test for imprecise data in consistency table', *Frontiers in Applied Mathematics and Statistics*, 9. Available at: https://doi.org/10.3389/fams.2023.1279638.

Franke, T.M., Ho, T. and Christie, C.A. 2012. 'The Chi-Square Test: Often Used and More Often Misinterpreted', *American Journal of Evaluation*, 33(3), pp. 448–458. Available at: https://doi.org/10.1177/1098214011426594.

Grove, S.K, and Gray, J.R. .2020. *Burns & Grove's The Practice of Nursing Research : Appraisal, Synthesis, and Generation of Evidence*. 9th edn. Texas: Elsevier.

Mchugh, M.L. .2013. 'The Chi-square test of independence Lessons in biostatistics', *Biochemia Medica*, 23(2), pp. 143–9. Available at: http://dx.doi.org/10.11613/BM.2013.018.

Nihan, S.T. 2020. 'Karl Pearsons chi-square tests', *Educational Research and Reviews*, 15(9), pp. 575–580. Available at: https://doi.org/10.5897/err2019.3817.

Pandis, N. 2016. 'The chi-square test', *American Journal of Orthodontics and Dentofacial Orthopedics*, 150(5), pp. 898–899. Available at: https://doi.org/10.1016/j.ajodo.2016.08.009.

Rana, R. and Singhal, R. 2015. 'Chi-square test and its application in hypothesis testing', *Journal of the Practice of Cardiovascular Sciences*, 1(1), p. 69. Available at: https://doi.org/10.4103/2395-5414.157577.

Schober, P. and Vetter, T.R. 2019. 'Statistical Minute', *International Anesthesia Research Society*, 129(2), p. 2019.

Shukla, D. 2023. 'A narrative review on types of data and scales of measurement: An initial step in the statistical analysis of medical data', *Cancer Research, Statistics, and Treatment*, 6(2), pp. 279–283. Available at: https://doi.org/10.4103/crst.crst_1_23.

Surury, I. 2020. *Buku Praktikum Statistical Program for Social Science*. Jakarta: Universitas Muhammadiyah Jakarta.

BAB 6
PENERAPAN BIOSTATISTIKA DALAM PENELITIAN KLINIS DAN KESEHATAN MASYARAKAT

Oleh Mardiani Gina Hartoyo

6.1 Pendahuluan

6.1.1 Pengantar Biostatistik

Biostatistik merupakan cabang statistika yang secara khusus berkaitan dengan aplikasi statistika dalam bidang ilmu kedokteran, biologi dan kesehatan masyarakat. Dalam pengantar ini, akan membahas definisi biostatistik, tujuan serta ruang lingkupnya dan peran biostatistik dalam penelitian klinis dan kesehatan masyarakat.

6.1.2 Definisi Biostatistik

Biostatistik merupakan ilmu yang mempelajari penggunaan teknik-teknik statistika untuk menganalisis data yang berkaitan dengan bidang biologi, kedokteran dan kesehatan masyarakat. Ini mencakup pengumpulan, penyajian, analisis dan interpretasi data untuk menghasilkan iformasi yang relevan dalam konteks kesehatan manusia, penyakit dan intervensi medis.

6.1.3 Tujuan dan Ruang Lingkup Biostatistik

Tujuan Utama Biostatistik adalah untuk menyediakan alat dan metode statistik yang diperlukan untuk mengevaluasi bukti-bukti empiris dalam ilmu kedokteran dan kesehatan masyarakat. Ini termasuk merancang studi penelitian, menganalisis data dan membuat kesimpulan yang didukung oleh bukti statistik.

Ruang lingkup biostatistik meliputi berbagai teknik statistika yang digunakan dalam penelitian klinis dab kesehatan masyarakat, seperti statistika deskriptif, inferensial, analisis regresi, survival

analysis, meta-analysis dll. Dan ini juga mencakup desain penelitian, pemilihan sampel, interpretasi hasil dan komunikasi hasil kepada pemangku kepentingan yang berbeda.

6.1.4 Peran Biostatistik dalam Penelitian Klinis dan Kesehatan Masyarakat.

Perancangan studi: biostatistik membantu dalam merancang studi penelitian yang kuat dengan memilih desain yang tepat, menentukan ukuran sampel yang diperlukan dan mengembangkan protokol penelitian ini.

Analisis data: biostatistik memungkinkan analisis data yang tepat menggunakan teknik statistika yang sesuai, sehingga memungkinkan peneliti untuk mengidentifikasi pola, hubungan dan tren yang relevan dalam data.

Interprestasi hasil: biostatistik membantu dalam menginterprestasikan hasil analisis data dengan benar, sehingga memungkikan para peneliti dan praktisi kesehatan untuk membuat kesimpulan yang didukung oleh bukti statistik.

Pengambilan keputusan: berdasarkan hasil analisis statistik, biostatistik membantu dalam mengambil keputusan yang informasional dalam konteks klinis dan kesehatan masyarakat termasuk diagnosis, pengobatan, pencegahan dan kebijakan kesehatan.

Dengan ini biostatistik memainkan peran dalam menghasilkan bukti-bukti ilmiah yang diperlukan untuk memahami dan mengatasi tantangan kesehatan masyarakat serta meningkatkan praktek klinis.

6.2 Statistik Deskriptif

Statistik deskriptif adalah cabang dari statistika yang berkaitan dengan penyajian dan penggambaran data secara ringkas serta informatif. Ini mencakup berbagai teknik untuk merangkum, mengatur dan menggambarkan data agar dapat dipahami dengan lebih baik.

6.2.1 Tujuan Statistik Deskriptif

Statistik deskriptif bertujuan untuk menyajikan dan menggambarkan karakteristik dasar dari suatu set data, baik itu populasi atau sampel, tanpa melakukan inferensi atau kesimpulan tentang populasi yang lebih besar.

6.2.2 Pemilihan dan Penyajian Data

Pemilihan data: langkah awal dalam statistik deskriptif adalah memilih data yang relevan untuk dianalisis, baik data numerik maupun kategorikal.

Penyajian data: data dapat disajikan dalam bentuk tabel, grafik atau diagram, tergantung pada jenis data dan tujuan analisisnya.

6.2.3 Ukuran Pemusatan Data

Mean (rata-rata): rata-rata adalah jumlah dari semua nilai dalam data yang dibagi dengan jumlah total nilai. Ini memberikan gambaran tentang pusat distribusi data.

Median (nilai tengah): median adalah nilai tengah dalam data saat diurutkan. Ini tidak dipengaruhi oleh nilai ekstrem dan memberikan gambaran tentang pusat distribusi data.

Modus (nilai yang paling sering muncul): modus adalah nilai yang paling sering muncul dalam data. Ini berguna untuk data kategorikal atau data yang memiliki banyak pengelompokan.

6.2.4 Ukuran Penyebaran Data

Rentang: adalah perbedaan antara nilai maksimum dan nilai minimum dalam data. Ini memberikan gambaran tentang sebaran keseluruhan data.

Varians: adalah rata-rata dari kuadrat perbedaan antara setiap nilai data dan rata-rata. Ini memberikan gambaran tentang sebaran data dari rata-rata.

Standar Deviasi: adalah akar kuadrat dari varians. Ini memberikan gambaran tentang nilai seberapa jauh nilai-nilai data tersebut dari rata-rata.

6.2.5 Grafik dan Diagram untuk Analisis Data

Histogram: adalah grafik yang digunakan untuk menampilkan distribusi frekuensi dari data numerik.

Diagram batang (Bar Chart): diagram batang digunakan untuk menampilkan data kategorikal dalam bentuk batang vertikal atau horizontal.

Diagram lingkaran (Pie Chart): diagram lingkaran digunakan untuk menampilkan proporsi relatif dari kategori dalam data.

Statistik deskriptif memberikan dasar yang penting untuk memahami karakteristik dasar dari suatu set data, menjadi langkah awal dalam analisis statistik yang lebih lanjut.

6.3 Statistik Inferensial

Statistik inferensial bertujuan untuk membuat generalisasi atau kesimpulan tentang populasi berdasarkan informasi yang diperoleh dari sampel data. Ini melibatkan penggunaan probabilitas untuk mengevaluasi kecocokan antara data sampel dan hipotesis yang diuji.

6.3.1 Uji Hipotesis

Uji hipotesis adalah prosedur statistik yang digunakan untuk menguji klaim atau asumsi tentang populasi dengan menggunakan sampel data. Melibatkan pembentukan hipotesis nol dan hipotesis alternatif serta penggunaan nilai p-nilai untuk menentukan apakah terdapat cukup bukti untuk menolak hipotesis nol.

6.3.2 Analisis Regresi dan Korelasi

Analisis regresi digunakan untuk mengevaluasi hubungan antara satu atau lebih variabel independen dengan variabel dependen. Ini melibatkan pembentukan model matematika untuk memprediksi nilai variabel dependen berdasarkan nilai variabel independen yang diberikan.

Analisis korelasi digunakan untuk mengevaluasi kekuatan dan arah hubungan antara dua variabel numerik. Ini memberikan gambaran tentang sejauh mana dua variabel bergerak bersama-sama.

6.3.3 Analisis Variansi (ANOVA)

Analisis variansi(Anova) digunakan untuk membandingkan rata-rata dari tiga atau lebih kelompok yang berbeda untuk menentukan apakah ada perbedaan yang siqnifikan di antara mereka. Ini bermanfaat ketika ingin menentukan apakah ada perbedaan yang siqnifikan antara kelompok perlakuan dalam suatu percobaan.

6.3.4 Pengujian Non Parametrik

Pengujian non parametrik adalah teknik statistik yang digunakan ketika data tidak memenuhi asumsi dari pengujian parametrik, seperti normalisasi atau homogenitas varians. Ini melibatkan penggunaan peringkat atau perbandinagn dari data asli, tanpa memperhitungkan distribusi tertentu.

6.3.5 Interpretasi Hasil Statistik Inferensial

Interpretasi hasil statistik inferensial melibatkan pengambilan keputusan tentang apakah terdapat cukup bukti untuk menolak hipotesis nol. Ini melibatkan penilaian terhadap nilai p-nilai, dimana nilai p yang ditentukan menunjukkan bahwa terdapat cukup bukti utnuk menolak hipotesis nol.

Statistik inferensial memberikan alat yang kuat untuk membuat generalisasi tentang populasi berdasarkan sampel data yang diamati. Namun interpretasi hasil harus dilakukan dengan hati-hati dan memperhatikan asumsi-asumsi yang mendasarinya.

6.4 Desain Penelitian

Desain penelitian merupakan rencana sistematis yang digunakan untuk mengumpulkan data yang relevan dan valid untuk menjawab pertanyaan penelitian atau menguji hipotesis tertentu. Desain penelitian mencakup berbagai elemen, termasuk pemilihan populasi atau sampel, pengumpulan data dan analisis statistik.

Desain penelitian adalah rencana atau kerangka kerja yang digunakan untuk merencanakan dan melaksanakan studi ilmiah dengan tujuan untuk mengumpulkan data yang valid dan relevan untuk menjawab penelitian atau menguji hipotesis tertentu.

6.4.1 Desain Studi Observasional

Desain ini melibatkan pengamatan terhadap individu atau kelompok tanpa intervensi yang direncanakan. Studi ini mencatat informasi tentang perilaku, karakteristik atau kejadian tanpa mengubah variabel yang diamati. Jenis-jenis desain studi observasional antara lain studi potong lintang (*cross sectional*), studi kohort dan studi kasus kontrol.

6.4.2 Desain Studi Eksperimental

Desain studi eksperimental melibatkan manipulasi variabel independen untuk menentukan efeknya terhadap variabel dependen. Studi ini memungkinkan peneliti untuk menarik kesimpulan tentang hubungan sebab-akibat antara variabel. Desain studi ini sering kali melibatkan pembagian subjek ke dalam kelompok perlakuan dan kelompok kontrol.

6.4.3 Perencanaan Sampling

Perencanaan sampling merupakan proses pemeliharaan elemen-elemen dari populasi yang akan diselidiki untuk menjadi bagian dari sampel. Hal ini melibatkan pemilihan teknik sampling yang tepat, seperti sampling acak sederhana, stratifikasi atau klaster serta menentukan ukuran sampel yang memadai untuk mencapai tingkat kepercayaan dan akurasi yang diinginkan.

6.4.4 Pengendalian Variabel pada Desain Penelitian

Pengendalian variabel adalah proses memastikan bahwa variabel-variabel yang tidak diinginkan atau tidak relevansi/tidak mempengaruhi hasil penelitian. Ini melibatkan desain eksperimen yang cermat, pemilihan kontrol yang tepat dan pemantauan terhadap faktor-faktor yang dapat mempengaruhi hasil.

Desain penelitian yang membaik merupakan kunci untuk menghasilkan data yang valid dan apat diandalkan, sehingga memungkinkan peneliti untuk membuat kesimpulan yang akurat dan dapat dipercaya tentang fenomena yang diamati. Dengan menggunakan desain penelitian yang sesuai, peneliti dapat mengurangi bias dan meningkatkan kepercayaan terhadap hasil penelitian mereka.

6.5 Analisis Spasial dan Temporal

Analisis spasial dan temporal merupakan pendekatan yang penting dalam penelitian kesehatan masyarakat untuk memahami sebaran geografis dan perubahan waktu dari berbagai faktor kesehatan. Ini melibatkan penggunaan teknik-teknuk khusus untuk memeriksa pola spasial dan temporal dari penyakit, determinan kesehatan dan intervensi.

6.5.1 Penggunaan GIS dalam Analisis Spasial

Sistem Informasi Geografis (GIS) adalah alat yang sangat berguna dalam analisis spasial dalam kesehatan masyarakat. GIS memungkinkan pengguna untuk memetakan data kesehatan, memvisualisasikan pola spasial dan mengidentifikasi area-area yang rentan terhadap masalah kesehatan tertentu. Hal ini dapat membantu dalam perencanaan intervensi kesehatan yang tepat sasaran dan alokasi sumber daya.

6.5.2 Analisis Temporal dan Time Series

Analisis temporal melibatkan pemahaman terhadap perubahan waktu dari variabel-variabel kesehatan seperti kejadian penyakit, tingkat kematian atau faktor risiko. Analisis ini menggunakan teknik time series untuk mengidentifikasi tren, musimanitas dan pola jangka panjang dalam data kesehatan.

6.5.3 Model Spasial Temporal dalam Penelitian Kesehatan

Model spasial temporal merupakan pendekatan yang kompleks menggabungkan analisis spasial dan temporal untuk memahami sebaran geografis dari fenomena kesehatan seiring waktu. Ini memungkinkan peneliti untuk mengevaluasi bagaimana pola penyakit berkembang sepanjang waktu dan ruang serta faktor apa yang mempengaruhi pola tersebut.

Analisis spasial dan temporal merupakan alat penting dalam penelitian kesehatan masyarakat karena memungkinkan peneliti untuk melihat lebih dari angka-angkah dan statistik. Dengan memperhatikan dimensi geografis dan waktu, peneliti dapat mengidentifikasi pola-pola yang mungkin terlewatkan dalam

analisis kesehatan konvensional, sehingga memungkinkan pengembangan intervensi yang lebih efektif dan tepat sasaran.

6.6 Analisis Survival

Analisis survival adalah teknik statistik yang digunakan untuk menganalisis data waktu kejadian, seperti waktu sampai terjadinya suatu peristiwa tertentu, misalnya waktu sampai kematian, waktu sampai kambuhnya penyakit atau waktu samapai suatu kejadian medis lainnya. Analisis survival sangat berguna dalam penelitian klinis, epidemiologi dan kesehatan masyarakat.

6.6.1 Konsep Dasar Analisis Survival

Konsep dasar dari analisis survival adalah memeriksa waktu antara peristiwa awal (seperti diagnosis penyakit atau pemberian pengobatan) hingga peristiwa akhir (seperti kematian atau kesembuhan). Data ini sering kali disensor artinya kita mungkin tidak tau dengan pasti kapan peristiwa akhir terjadi untuk beberapa subjek dalam studi.

6.6.2 Kurva Survival dan Tabel Risiko

Kurva Survival adalah grafik yang menunjukkan proporsi subjek yang masih ”bertahan hidup” (belum mengalami peristiwa akhir) pada titik waktu tertentu setelah peristiwa awal.

Tabel risiko adalah tabel yang menunjukkan jumlah subjek yang mengalami peristiwa akhir pada setiap titik waktu tertentu, bersama dengan jumlah subjek yang masih berisiko pada saat itu.

6.6.3 Metode Estimasi Survival (Kaplan-Meier)

Metode kaplan-meier adalah teknik yang umum digunakan untuk mengestimasi kurva survival dari data censored. Metode ini menghitung estimasi peluang bertahan hidup pada setiap titik waktu berdasarkan proporsi subjek yang masih bertahan hidup pada titik waktu sebelumnya.

6.6.4 Metode Regresi Cox

Metode ini juga dikenal sebagai model proporsional hazard cox adalah metode statistik yang digunakan untuk mengevaluasi

efek variabel-variabel independen terhadap waktu kejadian suatu peristiwa, sambil mengontrol variabel-variabel lainnya. Model ini menghasilkan perkiraan hazard ratio yang meunjukkan perubahan relatif dalam risiko peristiwa untuk setiap satuan perubahan dalam variabel independen.

Analisis survival memberikan pemahaman yang mendalam tentang waktu sampai terjadinya suatu peristiwa penting dalam konteks medis dan kesehatan masyarakat. Dengan menggunakan teknik-teknik ini, peneliti dapat mengidentifikasi faktor-faktor risiko, memprediksi prognosis dan mengevaluasi efektivitas intervensi dalam situasi di mana waktu adalah faktor kunci.

6.7 Meta-Analisis

Meta-analisis melibatkan pengumpulan data dari berbagai studi yang relevan tentang topik yang sama, kemudian menganalisis data secara bersama-sama untuk mendapatkan kesimpulan yang lebih kuat daripada yang dapat diberikan oleh setiap studi secara individual. Meta analisis sering digunakan untuk memperoleh estimasi efek yang lebih stabil dan dapat dipercaya serta untuk mengeksplorasi variasi antara studi.

6.7.1 Proses Pengumpulan Studi

Proses ini dimulai dengan pencarian studi-studi yang relevan menggunakan basis data ilmiah dan pustaka terkait. Studi-studi yang dipilih harus memenuhi kriteria inklusi yang ditetapkan sebelumnya, seperti desain studi, populasi yang diteliti dan hasil yang diukur. Setelah studi-studi yang relevan ditemukan , data dari setiap studi diekstraksi dan disusun dalam format yang sesuai untuk analisis meta-analisis.

6.7.2 Penggunaan Efek Ukuran dalam Meta-Analisis

Efek ukuran adalah ukuran statistik yang menggambarkan besarnyaefek dari suatu intervensi atau hubungan antara variabel dalam suatu studi. Dalam meta-analisis,efek ukuran dari setiap studi digunakan untuk menghitung bobot relatif dari masing-masing studi dalam analasis gabungan. Beberapa contoh efek ukuran yang umum

digunakan adalah risiko relatif, odds ratio dan perbedaan mean standar.

6.7.3 Interpretasi Hasil Meta-Analisis

Interpretasi hasil meta-analisis melibatkan evaluasi keseluruhan efek gabungan dari studi-studi yang telah dianalisis. Ini mencakup memperhitungkan kekuatan dan kelemahan dari setiap studi serta mengidentifikasi variasi antara studi. Kesimpulan dari meta-analisis harus disampaikan dengan hati-hati, mengakui ketidakpastian yang mungkin terkait dengan hasil tersebut. Interpretasi yang tepat memerlukan pertimbangan terhadap konteks klinis dan metodologis dari studi-studi yang terlibat.

Meta-analisis adalah alat yang kuat untuk menyintesis bukti ilmiah dari berbagai sumber, membantu dalam pengambilan keputusan klinis dan formulasi kebijakan kesehatan.

6.8 Analisis Kausalitas

Analisis kausalitas adalah upaya untuk menentukan apakah suatu hubungan sebab-akibat benar-benar ada antara dua variabel atau lebih dalam sutu konteks tertentu. Ini merupakan aspek penting dalam peneltian kesehatan untuk memahami hubungan antara faktor risiko, intervensi dan hasil kesehatan.

6.8.1 Konsep Dasar Kausalitas

Konsep ini melibatkan pengidentifikasian hubungan sebab-akibat antara dua variabel atau lebih. Untuk menyimpulkan ada hubungan kausal, beberapa kriteria harus dipenuhi,termasuk hubungan temporal,hubungan antara penyebab dan efek serta eliminasi alternatif.

6.8.2 Penggunaan Model Kausal dalam Penelitian Kesehatan

Model kausal adalah kerangka kerja yang digunakan untuk menggambarkan hubungan sebab-akibat antara variabel-variabel dalam suatu sistem. Dalam penelitian kesehatan, model kausal digunakan untuk mengidentifikasi faktor-faktor risiko dan mekanisme yang mendasari penyakit atau kondisi kesehatan tertentu serta merancang intervensi yang efektirf.

6.8.3 Analisis Propensity Score Matching

Analisis propensity score matching adalah metode statistik yang digunakan untuk mengurangi bias dalam penilaian efek suatu intervensi atau perlakuan dalam penelitian observasional. Ini melibatkan pencocokan subjek yang menerima perlakuan dengan subjek yang tidak menerima perlakuan berdasarkan skor kecenderungan untuk menerima perlakuan tersebut, sehingga menyeimbangkan karakteristik awal antara kedua kelompok.

6.8.4 Evaluasi Faktor Kausal pada Rancangan Penelitian

Evaluasi faktor kausal pada rancangan penelitian melibatkan desain studi yang dirancang dengan cermat untuk mengidentifikasi hubungan sebab-akibat antara variabel-variabel yang diamati.

Rancangan penelitian yang kuat seperti studi eksperimental dapat memberikan bukti yang lebih kuat tentang kausalitas daripada studi observasional yang lemah.

Analisis kausal adalah aspek penting dalam penelitian kesehatan untuk memahami hubungan antara variabel-variabel yang diamati dan menentukan implikasi praktir dari temuan tersebut.

6.9 Pengujian Sensitivitas dan Analisis Subgrup

Pengujian ini adalah teknik-teknik penting dalam penelitian klinis utnuk memahami seberapa baik suatu tes diagnostik atau intervensi dapat mengidentifikasi kondisi tertentu dan bagaimana efektifnya dalam kelompok tertentu.

6.9.1 Sensitivitas dan Specificity

Sensitivitas adalah kemampuan suatu tes diagnostik untuk mendeteksi keberadaan kondisi yang sesungguhnya. Sensitivitas dihitung sebagai proporsi subjek dengan kondisi yang unsubjek yang sebenarnya positif.

Specificity adalah kemampuan suatu tes diagnostik untuk mengecualikan keberadaan kondisi yang tidak ada. Specificity dihitung sebagai proporsi subjek tanpa kondisi yang diuji negatif oleh tes diagnostik dibandingkan dengan jumlah total subjek yang sebenarnya negatif.

6.9.2 Analisis Subgrup dalam Penelitian Klinis

Analisis subgrup adalah proses membagi sampel penelitian ke dalam subkelompok berdasarkan karakteristik tertentu, seperti usia, jenis kelamin atau penyakit penyerta dan kemudian membandingkan efek suatu intervensi di antara subkelompok tersebut. Ini membantu untuk mengevaluasi apakah intervensi tersebut efektif secara konsisten di seluruh populasi atau hanya di beberapa subkelompok.

6.9.3 Penggunaan Uji Sensitivitas dalam Penyaringan

Uji sensitivitas adalah teknik yang digunakan untuk mengevaluasi seberapa baik suatu tes diagnostik dapat mengidentifikasi keberadaan kondisi tertentu di populasi yang diteliti. Uji sensitivitas sering digunakan dalam penyaringan untuk mendeteksi penyakit atau kondisi kesehatan pada tahap awal, sehingga memungkinkan intervensi lebih lanjut untuk dilakukan.

Pengujian sensitivitas dan analisis subgrup merupakan alat yang penting dalam penekitian klinis untuk memahami efektivitas dan aplikabilitas dari tes diagnostik dan intervensi.

6.10 Aplikasi Statistik Terapan

Statsitik terapan memiliki berbagai aplikasi penting dalam penyajian hasil penelitian, interpretasi statistik dalam publikasi ilmiah dan pengambilan keputusan klinis.

6.10.1 Penyajian Hasil Penelitian

Grafik dan Tabel digunakan untuk menyajikan data secara visual dan terstruktur. Ini membantu pembaca untuk memahami pola, tren dan perbandingan dalam data dengan lebih baik.

Deskripsi statistik seperti mean, median dan standar deviasi, digunakan untuk merangkum dan menggambarkan karakteristik dasar dari data.

Kurva survival dan analisis temporal, untuk penelitian yang melibatkan waktu, kurva survival dan analisis temporal digunakan untuk menunjukkan perkembangan waktu dari peristiwa yang diamati.

6.10.2 Interpretasi Statistik dalam Publikasi Ilmiah

Penjelasan metode analisis digunakan dalam penelitian sangat penting untuk memastikan reproduktibilitas dan kepercayaan terhadap hasil.

Interpretasi hasil harus dilakukan dengan hati-hati, mengakui batasan dari studi dan kemungkinan sumber bias. Kesimpulan harus didasarkan pada bukti yang kuat dan disertai dengan interpretasi yang relevan secara klinis.

6.10.3 Peran Statistik dalam Pengambilan Keputusan Klinis

Evaluasi evidens, statistik membantu dokter dalam mengevaluasi evidens ilmiah yang mendukung pilihan pengobatan atau intervensi tertentu.

Analisis risiko dan manfaat, statistik digunakan untuk mengevaluasi risiko dan manfaat dari berbagai pilihan pengobatan atau tindakan medis, membantu dokter dan pasien dalam pengambilan keputusan yang informasional.

6.10.4 Kesalhan Umum dalam Interpretasi Statistik

Penafsiran kausalitas adalah mengasumsi kausalitas dari hubungan yang diamati tanpa mempertimbangkan faktor-faktor lain atau desain penelitian.

Pengabaian asumsi statistik, dari metode statistik tertentu dapat mengarah pada interpretasi yang tidak akurat dari hasil analisis.

Statistik terapan memiliki peran kunci dalam penelitian ilmiah dan pengambilan keputusan klinis. Namun penting untuk mmenggunakan statistik dengan hati-hati, menghindari kesalahan umum dan interpretasi dan memastikan bahwa hasilnya disajikan serta diinterpretasikan dengan benar.

6.11 Kesimpulan

Penerapan biostatistik dalam penelitian klinis dan kesehatan adalah kunci untuk menghasilkan bukti ilmiah yang kuat dan relevan dalam memahami masalah kesehatan, menganalisis intervensi dan membuat keputusan yang berdampak pada praktik

klinis dan kebijakan kesehatan masyarakat. Kesimpulan dalam pembahasan ini sebagai berikut :

1. Peran penting biostatistik, memberikan kerangka kerja yang kuat untuk merancang, menganalisis dan menginterpretasikan data dalam penelitian klinis dan kesehatan masyarakat. Ini membantu mengidentifikasi pola, tren dan hubungan antara variabel yang relevan.
2. Desain penelitian yang efektif, pemahaman yang baik tentang biostatistik memungkinkan peneliti untuk merancang studi yang tepat dan efisien, baik studi observasional maupun eksperimental serta memilih sampel yang representatif.
3. Analisis data yang akurat, melalui penggunaan teknik analisis statistik yang tepat seperti analisis regresi, uji hipotesis dan analisis survival, biostatistik membantu dalam memeriksa hubungan sebab-akibat, mengevaluasi efektivitas intervensi dan memahami prediktor risiko kesehatan.
4. Pengambilan keputusan yang informasional, interpretasi hasil statistik yang benar dan hati-hati memungkinkan para peneliti, praktisi klinis dan pembuat kebijakan untuk mengambil keputusan yang berdasarkan bukti, baik terkait dengan diagnosis, pengobatan atau tindakan pencegahan.
5. Komunikasi yang efektif, pengetahuan tentang biostatistik membantu dalam penyajian hasil penelitian secara jelas dan efektif kepada berbagai pemangku masyarakat umum dan pembuat kebijakan.
6. Pengembangan pengetahuan dan inovasi, melalui penerapan biostatistik yang cermat, penelitian klinis dan kesehatan masyarakat dapat terus mengembangkan pengetahuan tentang penyakit, faktor risiko dan intervensi yang dapat meningkatkan kesehatan populasi secara keseluruhan.

Dengan demikian penerapan bistatistik bukan hanya merupakan alat analisis, tetapi sebagai fondasi untuk pengambilan keputusan yang informasional, pembangunan pengetahuan dan perubahan positif dalam praktek klinis dan kebijakan kesehatan masyarakat.

DAFTAR PUSTAKA

Barlianto W, Andrajati R, Kurniawan AN. *Epidemiologi untuk Praktek Klinis*. Jakarta: Buku Kedokteran EGC;2018

Fasli R. *Statistika Kesehatan: Aplikasi dalam Riset Kesehatan*. Jakarta: Erlangga;2015

Irianto SK, Mangunnegoro H. *Biostatistik: Konsep dan Aplikasi untuk Penelitian Kesehatan*. Jakarta: Trans Info Media; 2019

Iskandarsyah A, Kusumaratna R, Yunihastuti E. *Pedoman Praktis Riset Epidemiologi*. Jakarta: Salemba Medika; 2016

Mardiati E, Setyohadi B. *Pengantar Biostatistik untuk Kedokteran dan Kesehatan Masyarakat*. Jakarta: Sagung Seto; 2017

Sastroasmoro S, Ismael S. *Dasar-Dasar Metodologi Penelitian Klinis.* 5th ed. Jakarta: Sagung Seto;2014

Soegianto D, Suryanto E, Sudoyo AW. *Epidemiologi Klinik Praktis: Panduan Penelitian Klinis*. Jakarta: EGC;2017

Sudoyo AW, Setiati S, Alwi I, et al., editors. *Buku Ajar Ilmu Penyakit Dalam.* 6th ed. Jakarta: Interna Publishing: 2014

Sunjaya DK, Pardosi JF, *Biostatistik Kedokteran dan Kesehatan Masyarakat.* Jakarta: Mitra Wacana Media; 2015

Yusuf S, Susanto T, Sancoyo GS,. *Statistik Kesehatan: Konsep dan Aplikasi untuk Praktek Kesehatan*. Jakarta: Salemba Medika; 2018.

BIODATA PENULIS

Nur Al-faida, SKM, M.Kes
Staf Pengajar Program Studi S1 Gizi
Sekolah Tinggi Kesehatan Persada Nabire, Papua Tengah

Penulis lahir di Sumulluk pada tanggal 25 September 1993. Penulis menjadi dosen tetap di Program Studi S1 Gizi STIKes Persada Nabire sejak tahun 2020. Pada tahun 2020-2025, penulis dipercayakan sebagai Ketua Lembaga Penelitian dan Pengabdian kepada Masyarakat (LPPM) di Sekolah Tinggi Ilmu Kesehatan Persada Nabire (STIKPEN).

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

Penulis adalah pengajar mata kuliah berikut ini: Matematika, Sosiologi Dasar, Fisiologi, Ilmu Pangan, Biostatistika, Gizi Ibu dan Anak 1, Antropologi Sosial Gizi, Epidemiologi Gizi, Teknologi Pangan Gizi, Metodologi Penelitian Gizi, dan Inovasi Ilmu Pangan Lokal.

BIODATA PENULIS

Dr. Fenita Purnama Sari Indah, SKM., M.Kes
Dosen Program Studi Kesehatan Masyarakat
STIKes Widya Dharma Husada Tangerang

Penulis lahir di Bandar Lampung, 12 Juni 1991. Pada tahun 2009-2013 menempuh S1 Kesehatan Masyarakat dengan Peminatan Biostatistika di Universitas Diponegoro, Semarang. Melanjutkan studi S2 Kesehatan Masyarakat pada 2013 hingga 2015 di perguruan tinggi yang sama. Sejak September 2020- Januari 2024 penulis menempuh Pendidikan doktoral di Fakultas Kesehatan Masyarakat, Universitas Indonesia.

Penulis aktif menulis artikel di beberapa jurnal ilmiah nasional, internasional, maupun buku. Saat ini tercatat sudah pernah mendapatkan beberapa hibah penelitian berupa Penelitian Fundamental, Penelitian Dasar Kompetitif Nasional, Penelitian Dosen Pemula dan Hibah BKKBN. Penulis juga merupakan *awardee* Beasiswa LPDP. Kepakaran Penulis adalah Kesehatan Masyarakat, Promosi Kesehatan, HIV/AIDS dan Statistik Fasilitas Pelayanan Kesehatan. Moto Penulis yaitu "*Work hard, Pray Hard and Thankful*". Penulis dapat dihubungi pada alamat email: fenita.purnama@masda.ac.id

BIODATA PENULIS

Andi Indrawati, S.Si., M.Kes.
Dosen Program Studi Sarjana Kebidanan
Institut Kesehatan dan Bisnis Kurnia Jaya Persada

Penulis lahir di Ujung Pandang tanggal 23 Maret 1973. Menyelesaikan Gelar Sarjana Biologi pada Program Studi Biologi Fakultas MIPA Universitas Hasanuddin tahun 1997 dan memperoleh gelar Magister pada Program Studi Biomedik Universitas Hasanuddin tahun 2009. Saat ini penulis adalah dosen tetap pada Program Studi Sarjana Kebidanan Fakultas Kesehatan Institut Kesehatan dan Bisnis Kurnia Jaya Persada di Kota Palopo, Provinsi Sulawesi Selatan. Penulis mengampuh matakuliah Metode Penelitian dan Biostatistik.

BIODATA PENULIS

Dr. Rahmawati, S. Si, M. Kes.
Dosen Tetap Program Studi D3 Teknologi Laboratorium Medis (TLM) Politeknik Kesehatan Muhammadiyah Makassar

Penulis lahir di Desa Kasiwiang Kecamatan Suli Kabupaten Luwu pada tanggal 12 April 1980, merupakan anak kedua dari dua bersaudara dari pasangan Abd. Muis dan Rapidah Hasan. Penulis adalah dosen tetap yayasan pada Program Studi D3 Teknologi Laboratorium Medis (TLM) Politeknik Kesehatan Muhammadiyah Makassar. Menyelesaikan pendidikan sekolah dasar di SD Negeri 20 Cimpu Kecamatan Suli Kabupaten Luwu pada tahun 1992. Pada tahun 1995 telah menyelesaikan pendidikan sekolah menengah pertama pada SMP Negeri 2 Kabupaten Takalar. Kemudian pada tahun 1998 menyelesaikan pendidikan sekolah menengah atas pada Sekolah Menengah Farmasi (SMF) Depkes Ujungpandang, Makassar.

Penulis menyelesaikan pendidikan S1 pada Jurusan Kimia Fakultas MIPA Universitas Hasanuddin pada tahun 2003. Pada tahun 2004 melanjutkan S2 pada Jurusan Kesehatan Lingkungan Fakultas Kesehatan Masyarakat (FKM) Universitas Hasanuddin dan berhasil menyelesaikan studi pada tahun 2006. Kemudian pada tahun 2013 melanjutkan S3 pada jurusan Ilmu Kimia Fakultas MIPA Universitas Hasanuddin dan berhasil meraih gelar Doktor pada tahun 2018. Penulis menekuni bidang ilmu kimia analitik dan biokimia serta toksikologi klinik.

BIODATA PENULIS

Ns.Puji Astuti Wiratmo.Skep.MN
Dosen Program Studi Ners
Fakultas Keperawatan & Kebidanan Universitas Binawan

Penulis adalah dosen tetap pada Program Studi Ners Fakultas Keperawatan dan Kebidanan Universitas Binawan. Penulis mengajar pada keilmuan Keperawatan Medikal Bedah dan Riset Keperawatan. Penulis menyelesaikan pendidikan sarjana keperawatan di Fakultas Ilmu Keperawatan Universitas Indonesia dan melanjutkan studi S2 pada keilmuan Profesional Studies di Faculty of Nursing Midwifery and Health dan Faculty of Education University of Technology Sydney, Australia. Saat ini penulis aktif melakukan penelitian dan bimbingan penelitian teoritis maupun penelitian lapangan pada mahasiswa keperawatan.

BIODATA PENULIS

Mardiani Gina Hartoyo,SKM.,MKes.
Dosen Program Studi Ilmu Kesehatan Masyarakat
Fakultas Ilmu Kesehatan

Penulis seorang akademisi dengan memperoleh gelar Magister Kesehatan Masyarakat dari Universitas Sam Ratulangi. Dengan pengalaman sebagai Dosen Tetap di Universitas Trinita, beliau menunjukkan komitmennya terhadap pengembangan ilmu kesehatan. Gelar Sarjana Kesehatan Masyarakat pada tahun 2009 dan kelanjutan studi hingga meraih gelar Magister pada tahun 2012, menciptakan landasan kuat bagi dedikasi Mardiani dalam dunia akademis.

Mardiani Gina Hartoyo juga menjabat sebagai Wakil Rektor III bidang SDM, Kemahasiswaan, dan Alumni di Universitas Trinita, memperkuat perannya dalam pembinaan mahasiswa. Pengalaman praktis di Rumah Sakit Polri Bhayangkara, Kota Manado, memberikan wawasan berharga terhadap realitas kesehatan di lapangan dan menjadi modal penting dalam peran pendidikannya.

Selain itu, Mardiani Gina Hartoyo juga membawa pengalaman bekerja di RS. Pancaran Kasih GMIM Kota Manado. Keterlibatannya di rumah sakit tersebut menambah dimensi praktis dalam pemahamannya, melibatkan diri dalam merawat pasien, mengelola kasus kesehatan, dan berkolaborasi dalam tim medis multidisiplin. Pengalaman ini semakin melengkapi wawasan dan

keahlian Mardiani, yang kini juga menjabat sebagai Direktur di Klinik Mata Trinita Amurang, Provinsi Sulawesi Utara. Gabungan pendidikan formal yang baik, pengalaman praktis di berbagai rumah sakit, dan tanggung jawab di klinik membuat Mardiani Gina Hartoyo menjadi figur yang inspiratif dan berperan penting dalam membentuk generasi muda di bidang kesehatan.