



BUKU AJAR BIOSTATISTIK DASAR UNTUK MAHASISWA KESEHATAN

PENULIS:

- Egy Sunanda Putra, M.Gz
- Agus Hendra Al Rahmad, SKM, MPH
- Rahmatika Nur Aini, M.Gz
- Arvida Bar, S.Pd, STr.Kep, MKM
- Dr. Pahrur Razi, SKM, MKM
- Sutyawan, S.Gz, M.Si
- Witi Karwiti, SKM, MPH

BUKU AJAR BIOSTATISTIK DASAR UNTUK MAHASISWA KESEHATAN

**Egy Sunanda Putra, M.Gz
Agus Hendra Al Rahmad, SKM, MPH
Rahmatika Nur Aini, M.Gz
Arvida Bar, S.Pd, STr.Kep, MKM
Dr. Pahrur Razi, SKM, MKM
Sutyawan, S.Gz, M.Si
Witi Karwiti, SKM, MPH**



GETPRESS INDONESIA

BUKU AJAR BIOSTATISTIK DASAR UNTUK MAHASISWA KESEHATAN

Penulis :

Egy Sunanda Putra, M.Gz
Agus Hendra Al Rahmad, SKM, MPH
Rahmatika Nur Aini, M.Gz
Arvida Bar, S.Pd, STr.Kep, MKM
Dr. Pahrur Razi, SKM, MKM
Sutyawan, S.Gz, M.Si
Witi Karwiti, SKM, MPH

ISBN : 978-623-125-919-6

Editor : Prof. Dr. Neila Sulung, S.Pd., N.S., M.Kes.

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

PENERBIT GET PRESS INDONESIA

Anggota IKAPI No. 033/SBA/2022
Jl. Palarik RT 01 RW 06 Kelurahan Air Pacah
Kecamatan Koto Tangah Padang Sumatera Barat
website: www.getpress.co.id
email: adm.getpress@gmail.com

Cetakan pertama, Juli 2025

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

PRAKATA

Latar belakang penulisan buku ini didasari oleh kebutuhan akan sumber belajar yang komprehensif dan mudah dipahami dalam bidang Biostatistik Kesehatan, khususnya untuk mendukung promosi kesehatan dan ilmu kesehatan masyarakat. Buku ini dirancang untuk memberikan pemahaman dasar hingga aplikasi praktis statistik dalam konteks kesehatan, sehingga mahasiswa mampu mengaplikasikan konsep statistik dalam analisis data kesehatan secara efektif. Tujuan utama penulisan adalah menyediakan referensi yang lengkap dan sistematis agar mahasiswa dapat menguasai materi secara bertahap dan terstruktur. Ruang lingkup materi mencakup pengantar biostatistik, konsep dasar statistika, penyajian data, uji statistik, analisis data, serta penggunaan software statistik. Saya berharap buku ini dapat menjadi panduan yang bermanfaat dan meningkatkan minat serta kompetensi mahasiswa dalam bidang biostatistik kesehatan. Terima kasih atas dukungan dan kepercayaan semua pihak yang telah membantu terselesaikannya buku ini.

Padang, Juni 2025

Egy Sunanda Putra, M.Gz

KATA PENGANTAR

Puji syukur ke hadirat Tuhan Yang Maha Esa atas berkat dan rahmat-Nya, sehingga buku berjudul Biostatistik Kesehatan ini dapat disusun dan diselesaikan. Buku ini disusun sebagai panduan belajar bagi mahasiswa mata kuliah Biostatistik Kesehatan, dengan harapan dapat mempermudah pemahaman konsep dan aplikasi statistik dalam bidang kesehatan masyarakat dan promosi kesehatan. Penyusunan buku ini didasarkan pada pengalaman pengajaran dan kebutuhan akan referensi yang lengkap, praktis, serta mudah dipahami. Dalam buku ini, disajikan berbagai materi mulai dari pengantar biostatistik, konsep dasar statistika, hingga analisis data dan penggunaan software statistik yang relevan dengan kebutuhan lapangan dan penelitian. Semoga buku ini dapat menjadi sumber belajar yang bermanfaat dan mampu meningkatkan kompetensi mahasiswa dalam mengaplikasikan statistika di bidang kesehatan. Ucapan terima kasih saya sampaikan kepada semua pihak yang telah memberikan dukungan dan masukan berharga selama proses penulisan. Semoga buku ini dapat memberikan manfaat yang optimal bagi pembaca.

Padang, Juni 2025

Penulis

DAFTAR ISI

PRAKATA.....	i
KATA PENGANTAR	ii
DAFTAR ISI.....	iii
BAB 1 PENGANTAR BIOSTATISTIK.....	1
Tujuan Pembelajaran	1
Pendahuluan	2
1.1 Definisi dan Ruang Lingkup Biostatistik.....	3
1.2 Peran Biostatistik dalam Promosi Kesehatan.....	5
1.3 Hubungan Antara Biostatistik dan Ilmu Kesehatan Masyarakat	6
1.4 Contoh Aplikasi: Penggunaan Data Statistik dalam Program Imunisasi dan Promosi Kesehatan.....	8
Rangkuman	10
Latihan Mahasiswa.....	12
REFERENSI	16
BAB 2 KONSEP DASAR STATISTIKA.....	18
Tujuan Pembelajaran	18
Pendahuluan	19
2.1 Populasi dan Sampel	20
2.2 Skala Pengukuran: Nominal, Ordinal, Interval, Rasio	23
2.3 Jenis Data dan Cara Pengumpulan Data	25
Rangkuman	27
Latihan Mahasiswa.....	29
REFERENSI	33
BAB 3 PENYAJIAN DATA	34
Tujuan Pembelajaran	34
Pendahuluan	35
3.1 Tabel Distribusi Frekuensi	37

3.2 Diagram Batang, Lingkaran, Histogram, dan Poligon	
Frekuensi	38
3.3 Pengolahan Data dengan Bantuan Software Statistik	40
3.4 Contoh Aplikasi: Penyajian Data Prevalensi Penyakit	
Berdasarkan Wilayah	41
Rangkuman	43
Latihan Mahasiswa	45
REFERENSI	49
BAB 4 UKURAN PEMUSATAN DATA	50
Tujuan Pembelajaran	50
Pendahuluan	51
4.1 Pengertian dan Perbedaan Antara Mean, Median, dan	
Modus	53
4.2 Kegunaan dan Interpretasi dalam Konteks Kesehatan	55
Rangkuman	59
Latihan Mahasiswa	62
REFERENSI	66
BAB 5 UKURAN PENYEBARAN DATA	67
Tujuan Pembelajaran	67
Pendahuluan	68
5.1 Rentang	70
5.2 Simpangan Baku dan Varians	71
5.5 Contoh Aplikasi: Variasi Tingkat Imunisasi di	
Berbagai Daerah	75
5.6 Pentingnya Ukuran Penyebaran Data dalam	
Pengambilan Keputusan	76
5.7 Keterbatasan dan Kelebihan Ukuran Penyebaran	
Data	77
Rangkuman	78
Latihan Mahasiswa	80
REFERENSI	84

BAB 6 PROBABILITAS DASAR.....	85
Tujuan Pembelajaran	85
Pendahuluan	86
6.1 Konsep Peluang.....	88
6.2 Teorema Peluang dan Aturan Dasar	89
6.3 Aplikasi Probabilitas dalam Epidemiologi	92
6.4 Contoh Aplikasi: Perkiraan Risiko Kejadian Penyakit Tertentu.....	93
Rangkuman	96
Latihan Mahasiswa.....	99
REFERENSI	103
BAB 7 DISTRIBUSI PROBABILITAS	104
Tujuan Pembelajaran	104
Pendahuluan	105
7.1 Distribusi Binomial dan Distribusi Poisson	106
7.2 Distribusi Normal.....	109
7.3 Penerapan Distribusi dalam Data Kesehatan.....	111
7.4 Contoh Aplikasi: Distribusi Kejadian Infeksi dalam Populasi	112
7.5 Relevansi dan Pentingnya Distribusi Probabilitas dalam Kesehatan Masyarakat.....	114
Rangkuman	115
Latihan Mahasiswa.....	117
REFERENSI	121
BAB 8 SAMPEL DAN TEKNIK SAMPLING.....	122
Tujuan Pembelajaran	122
Pendahuluan	123
8.1 Teknik Pengambilan Sampel: Random, Stratified, Cluster	125
8.2 Kesalahan Sampling dan Non-Sampling	129
8.3 Penentuan Ukuran Sampel	130

8.4 Contoh Aplikasi: Pengambilan Sampel untuk Surveilans Kesehatan Lingkungan.....	133
Rangkuman	134
Latihan Mahasiswa	137
REFERENSI.....	141
BAB 9 ESTIMASI PARAMETER	142
Tujuan Pembelajaran	142
Pendahuluan.....	143
9.1 Estimasi Titik dan Estimasi Interval.....	144
9.2 Tingkat Kepercayaan	146
9.3 Interpretasi Hasil Estimasi dalam Konteks Kesehatan....	147
9.4 Contoh Aplikasi: Estimasi Prevalensi Stunting pada Anak-anak.....	148
Rangkuman	152
Latihan Mahasiswa	153
REFERENSI.....	157
BAB 10 UJI HIPOTESIS STATISTIK.....	159
Tujuan Pembelajaran	159
Pendahuluan.....	160
10.1 Hipotesis Nol dan Hipotesis Alternatif	161
10.2 Kesalahan Tipe I dan Tipe II.....	164
10.3 Langkah-langkah Uji Hipotesis	166
10.4 Contoh Aplikasi: Uji Perbedaan Tingkat Pengetahuan Kesehatan Antara Dua Komunitas	168
Rangkuman	169
Latihan Mahasiswa	171
REFERENSI.....	175
BAB 11 UJI STATISTIK PARAMETRIK.....	176
Tujuan Pembelajaran	176
Pendahuluan.....	177
11.1 Uji t Satu Sampel	179

11.2 Uji t Dua Sampel Independen.....	181
11.3 Uji t Berpasangan.....	182
11.4 Uji ANOVA Satu Arah.....	184
11.5 Interpretasi Hasil Uji pada Penelitian Kesehatan.....	185
11.6 Batasan dan Keterbatasan Uji Statistik Parametrik.....	186
Rangkuman	187
Latihan Mahasiswa.....	189
REFERENSI	193
BAB 12 Uji Statistik Non-Parametrik	194
Tujuan Pembelajaran	194
Pendahuluan.....	195
12.1 Uji Chi-Square.....	196
12.2 Uji Mann-Whitney dan Wilcoxon	198
12.3 Kapan dan Bagaimana Uji Ini Digunakan.....	201
12.4 Contoh Aplikasi: Analisis Hubungan Antara Tingkat Pendidikan dan Perilaku Kesehatan.....	201
12.5 Kesimpulan.....	203
Rangkuman	204
Latihan Mahasiswa.....	206
REFERENSI	210
BAB 13 Korelasi dan Regresi	211
Tujuan Pembelajaran	211
Pendahuluan.....	213
13.1 Korelasi Pearson dan Spearman.....	215
13.2 Analisis Regresi Linier Sederhana	218
13.3 Interpretasi Hasil Korelasi dan Regresi dalam Konteks Promosi Kesehatan.....	220
13.4 Contoh Aplikasi: Hubungan antara Tingkat Aktivitas Fisik dan Tekanan Darah	221
Rangkuman	223
Latihan Mahasiswa.....	225

REFERENSI.....	229
BAB 14 ANALISIS DATA KESEHATAN	230
Tujuan Pembelajaran	230
Pendahuluan.....	231
14.1 Indikator Kesehatan (Angka Kesakitan, Kematian, Prevalensi).....	233
14.2 Interpretasi Data Surveilans	237
14.3 Penyusunan Laporan Data Statistik Kesehatan	241
Rangkuman	245
Latihan Mahasiswa	246
REFERENSI.....	250
BAB 15 PENGGUNAAN SOFTWARE STATISTIK	251
Tujuan Pembelajaran	251
Pendahuluan.....	252
15.1 Pengenalan Software Statistik: SPSS, Excel, dan Epi Info.....	254
15.2 Input dan Analisis Data Dasar	256
15.3 Interpretasi Output Hasil Analisis	258
15.4 Keunggulan dan Keterbatasan Masing-Masing Software.....	260
15.5 Penggunaan Software Statistik dalam Praktik Lapangan dan Penelitian.....	261
Rangkuman	262
Latihan Mahasiswa	264
REFERENSI.....	267
BAB 16 PENYUSUNAN LAPORAN STATISTIK KESEHATAN	269
Tujuan Pembelajaran	269
Pendahuluan.....	270
16.1 Struktur dan Sistematika Laporan Statistik.....	272
16.2 Penyajian Hasil dalam Bentuk Tabel dan Grafik.....	273

16.3 Penulisan Narasi dan Kesimpulan Hasil Analisis.....	275
16.4 Contoh Aplikasi: Penyusunan Laporan Hasil Surveilans Penyakit Menular di Masyarakat	277
Rangkuman	279
Latihan Mahasiswa.....	281
REFERENSI	285
BIODATA PENULIS.....	286

BAB 1

PENGANTAR BIOSTATISTIK

Tujuan Pembelajaran

1. Mahasiswa mampu memahami definisi dan ruang lingkup dari biostatistik serta mengidentifikasi komponen utama yang termasuk dalam bidang ini.
2. Mahasiswa dapat menjelaskan peran penting biostatistik dalam mendukung promosi kesehatan dan pengembangan program kesehatan masyarakat.
3. Mahasiswa mampu menguraikan hubungan antara biostatistik dan ilmu kesehatan masyarakat serta bagaimana keduanya saling melengkapi dalam praktik lapangan.
4. Mahasiswa dapat memberikan contoh nyata penggunaan data statistik dalam program imunisasi dan promosi kesehatan sebagai bagian dari aplikasi biostatistik dalam konteks kesehatan masyarakat.
5. Mahasiswa mampu mengidentifikasi berbagai sumber data statistik yang relevan untuk kegiatan promosi kesehatan dan surveilans penyakit.
6. Mahasiswa mampu menjelaskan pentingnya pengumpulan dan analisis data statistik dalam pengambilan keputusan di bidang kesehatan masyarakat.
7. Mahasiswa mampu mengaplikasikan konsep dasar biostatistik dalam menyusun laporan dan presentasi hasil analisis data kesehatan secara sederhana dan efektif.

Pendahuluan

Biostatistik merupakan cabang ilmu yang memadukan prinsip-prinsip statistik dengan bidang kesehatan dan kedokteran. Dalam konteks kesehatan masyarakat, biostatistik memegang peranan penting sebagai alat untuk mengumpulkan, menganalisis, dan menginterpretasi data yang berkaitan dengan kesehatan populasi. Melalui pemahaman yang mendalam terhadap konsep dan aplikasi biostatistik, tenaga kesehatan dan peneliti dapat membuat keputusan yang lebih tepat dan berbasis bukti dalam merancang program promosi kesehatan, pencegahan penyakit, serta pengendalian wabah.

Urgensi mempelajari biostatistik dalam bidang kesehatan tidak dapat diabaikan. Di era modern ini, data menjadi salah satu aset terpenting dalam pengembangan kebijakan dan intervensi kesehatan. Tanpa penguasaan terhadap konsep statistik, hasil surveilans, studi epidemiologi, maupun evaluasi program kesehatan akan sulit untuk dipahami dan diimplementasikan secara efektif. Oleh karena itu, penguasaan dasar-dasar biostatistik menjadi fondasi utama yang harus dimiliki oleh mahasiswa dan profesional di bidang kesehatan masyarakat.

Pada bagian pengantar ini, kita akan membahas definisi dan ruang lingkup biostatistik secara komprehensif. Definisi ini akan membantu mahasiswa memahami bahwa biostatistik bukan sekadar angka dan rumus, melainkan sebuah alat yang vital dalam mendukung kegiatan promosi dan pencegahan penyakit. Selanjutnya, kita akan mengulas peran biostatistik dalam promosi kesehatan, yang meliputi pengumpulan data, analisis, serta interpretasi hasil yang dapat digunakan untuk menyusun strategi intervensi yang tepat sasaran.

Selain itu, hubungan antara biostatistik dan ilmu kesehatan masyarakat akan dikaji secara mendalam. Keduanya saling terkait dan saling mendukung dalam upaya meningkatkan kualitas hidup masyarakat. Data statistik yang dihasilkan dari kegiatan surveilans dan penelitian epidemiologi menjadi dasar dalam pengambilan keputusan yang berbasis bukti. Sebagai contoh, penggunaan data statistik dalam program imunisasi dan promosi kesehatan akan diulas sebagai gambaran nyata bagaimana data tersebut dapat digunakan untuk mengidentifikasi daerah yang membutuhkan intervensi, memantau keberhasilan program, serta mengoptimalkan sumber daya yang tersedia.

Dengan memahami pengantar ini, diharapkan mahasiswa mampu melihat bahwa biostatistik bukan hanya sebagai mata kuliah teoritis, melainkan sebagai alat praktis yang sangat dibutuhkan dalam pekerjaan di lapangan. Penguasaan konsep dasar dan aplikasi biostatistik akan memudahkan mereka dalam melakukan analisis data, menyusun laporan, dan menyampaikan hasil secara efektif kepada berbagai pihak terkait. Pada akhirnya, pemahaman ini akan memperkuat kompetensi mereka dalam mendukung program-program kesehatan masyarakat yang lebih efektif dan berkelanjutan.

1.1 Definisi dan Ruang Lingkup Biostatistik

Biostatistik merupakan cabang ilmu yang mengintegrasikan prinsip-prinsip statistik dengan bidang kesehatan dan kedokteran. Secara umum, biostatistik dapat didefinisikan sebagai ilmu yang mempelajari pengumpulan, analisis, interpretasi, dan penyajian data yang berkaitan dengan kesehatan manusia dan populasi. Definisi ini menegaskan bahwa biostatistik tidak hanya berfokus pada angka dan rumus, tetapi juga pada proses ilmiah yang mendukung

pengambilan keputusan berbasis data dalam konteks kesehatan masyarakat (Rosner, 2015).

Ruang lingkup biostatistik sangat luas dan mencakup berbagai aspek yang berkaitan dengan kesehatan. Di antaranya adalah pengembangan metode statistik untuk analisis data klinis dan epidemiologis, perancangan studi penelitian, pengolahan data surveilans, serta evaluasi efektivitas intervensi kesehatan. Dalam praktiknya, biostatistik digunakan untuk mengidentifikasi faktor risiko penyakit, menentukan prevalensi dan insidens penyakit, serta menilai keberhasilan program kesehatan masyarakat. Selain itu, biostatistik juga berperan dalam pengembangan model prediktif dan pengujian hipotesis yang berkaitan dengan faktor-faktor yang mempengaruhi kesehatan populasi.

Komponen utama dalam ruang lingkup biostatistik meliputi pengumpulan data, pengolahan data, analisis statistik, dan interpretasi hasil. Pengumpulan data dilakukan melalui surveilans, studi kohort, studi kasus kontrol, maupun uji klinis. Pengolahan data meliputi pembersihan data, pengkodean, dan penyajian awal. Analisis statistik kemudian dilakukan untuk menguji hubungan antar variabel, mengukur tingkat kejadian, serta memperkirakan parameter populasi. Terakhir, interpretasi hasil menjadi tahap penting untuk memastikan bahwa temuan tersebut dapat digunakan sebagai dasar pengambilan keputusan dalam program kesehatan.

Sebagai contoh, dalam konteks imunisasi, biostatistik digunakan untuk menghitung tingkat cakupan imunisasi di suatu wilayah, mengidentifikasi daerah dengan tingkat imunisasi rendah, serta memantau tren kejadian penyakit yang dicegah melalui imunisasi. Dengan demikian, biostatistik menjadi alat yang esensial dalam mendukung

keberhasilan program imunisasi nasional dan pengendalian penyakit menular lainnya (Kirkwood & Sterne, 2017).

1.2 Peran Biostatistik dalam Promosi Kesehatan

Peran biostatistik dalam promosi kesehatan sangat vital karena menyediakan dasar ilmiah untuk pengambilan keputusan dan perancangan intervensi yang efektif. Melalui pengumpulan dan analisis data, tenaga kesehatan dapat memahami kondisi kesehatan masyarakat secara menyeluruh, mengidentifikasi masalah utama, serta menentukan prioritas intervensi yang tepat sasaran. Tanpa dukungan data statistik yang akurat, upaya promosi kesehatan akan menjadi kurang terarah dan berpotensi tidak efektif.

Salah satu peran utama biostatistik adalah dalam proses surveilans kesehatan masyarakat. Surveilans merupakan kegiatan pengumpulan data secara sistematis mengenai kejadian penyakit, faktor risiko, dan indikator kesehatan lainnya. Data ini kemudian dianalisis untuk mengidentifikasi tren, pola, dan faktor-faktor yang mempengaruhi kesehatan masyarakat. Sebagai contoh, analisis data surveilans dapat menunjukkan bahwa tingkat kejadian penyakit tertentu meningkat di daerah tertentu, sehingga program promosi kesehatan dapat difokuskan di wilayah tersebut (Thacker & Berkelman, 2018).

Selain itu, biostatistik juga berperan dalam evaluasi efektivitas program promosi kesehatan. Melalui analisis data sebelum dan sesudah intervensi, dapat diketahui apakah program tersebut berhasil menurunkan angka kejadian penyakit, meningkatkan perilaku sehat, atau memperbaiki indikator kesehatan lainnya. Sebagai contoh, evaluasi program imunisasi massal dapat dilakukan dengan membandingkan tingkat imunisasi dan kejadian penyakit

sebelum dan setelah pelaksanaan program. Jika data menunjukkan peningkatan cakupan imunisasi dan penurunan insiden penyakit, maka program tersebut dapat dianggap berhasil dan layak dilanjutkan atau diperluas (Fletcher et al., 2019).

Peran lain dari biostatistik adalah dalam pengembangan kebijakan kesehatan berbasis bukti. Data statistik yang diolah secara ilmiah akan memberikan gambaran yang akurat mengenai kebutuhan dan prioritas masyarakat. Misalnya, data tentang prevalensi obesitas di berbagai daerah dapat digunakan untuk merancang program pencegahan yang lebih spesifik dan efektif. Dengan demikian, biostatistik membantu memastikan bahwa sumber daya yang terbatas digunakan secara optimal untuk mencapai hasil kesehatan yang maksimal.

Selain itu, dalam konteks promosi kesehatan, komunikasi hasil analisis statistik kepada masyarakat dan pemangku kepentingan menjadi aspek penting. Penyajian data yang jelas dan mudah dipahami akan meningkatkan kesadaran dan partisipasi masyarakat dalam program kesehatan. Oleh karena itu, kemampuan untuk menyusun laporan dan menyajikan data secara efektif merupakan bagian integral dari peran biostatistik dalam promosi kesehatan (Gordis, 2014).

1.3 Hubungan Antara Biostatistik dan Ilmu Kesehatan Masyarakat

Hubungan antara biostatistik dan ilmu kesehatan masyarakat sangat erat dan saling melengkapi. Keduanya merupakan bagian dari upaya sistematis untuk meningkatkan kesehatan populasi melalui pengumpulan, analisis, dan interpretasi data. Dalam praktiknya, biostatistik

menjadi alat utama yang mendukung berbagai kegiatan dalam ilmu kesehatan masyarakat, mulai dari surveilans, penelitian epidemiologi, hingga evaluasi program.

Ilmu kesehatan masyarakat berfokus pada upaya pencegahan penyakit, promosi kesehatan, dan peningkatan kualitas hidup masyarakat secara umum. Untuk mencapai tujuan tersebut, diperlukan data yang akurat dan terpercaya. Di sinilah peran biostatistik menjadi sangat penting. Data yang dihasilkan dari surveilans dan studi epidemiologi harus dianalisis secara statistik agar dapat memberikan gambaran yang valid tentang kondisi kesehatan masyarakat. Tanpa penguasaan metode statistik, data tersebut berisiko disalahartikan atau tidak memberikan informasi yang cukup untuk pengambilan keputusan.

Sebaliknya, biostatistik juga membutuhkan konteks dari ilmu kesehatan masyarakat agar analisis yang dilakukan relevan dan aplikatif. Misalnya, dalam menentukan faktor risiko utama dari suatu penyakit, pengetahuan tentang faktor sosial, ekonomi, dan lingkungan sangat diperlukan. Dengan demikian, kolaborasi antara keduanya akan menghasilkan kebijakan dan intervensi yang lebih tepat sasaran dan efektif.

Contoh nyata dari hubungan ini adalah dalam program imunisasi. Data statistik tentang tingkat imunisasi dan kejadian penyakit di berbagai wilayah digunakan untuk mengidentifikasi daerah yang membutuhkan intervensi. Analisis ini didasarkan pada prinsip-prinsip epidemiologi dan statistik, sehingga hasilnya dapat diandalkan untuk merancang strategi promosi dan imunisasi yang lebih efektif. Selain itu, hasil analisis tersebut juga digunakan untuk memantau keberhasilan program dari waktu ke waktu, sehingga dapat dilakukan penyesuaian kebijakan secara dinamis (*World Health Organization, 2020*).

Selain itu, hubungan ini juga terlihat dalam pengembangan model prediktif yang membantu memproyeksikan tren kesehatan di masa depan. Misalnya, model statistik digunakan untuk memperkirakan beban penyakit di suatu wilayah berdasarkan faktor risiko yang teridentifikasi. Dengan demikian, perencanaan sumber daya dan kebijakan dapat dilakukan secara lebih proaktif dan berbasis bukti.

Secara keseluruhan, hubungan antara biostatistik dan ilmu kesehatan masyarakat adalah simbiosis yang saling menguatkan. Keduanya harus berjalan beriringan agar upaya peningkatan kesehatan masyarakat dapat dilakukan secara efektif dan efisien. Penguasaan konsep dan aplikasi keduanya akan memperkuat kompetensi mahasiswa dalam menghadapi tantangan kesehatan di lapangan dan mendukung pengembangan kebijakan yang berorientasi pada bukti ilmiah (Last, 2019).

1.4 Contoh Aplikasi: Penggunaan Data Statistik dalam Program Imunisasi dan Promosi Kesehatan

Penggunaan data statistik dalam program imunisasi dan promosi kesehatan merupakan contoh nyata bagaimana biostatistik berperan dalam meningkatkan kesehatan masyarakat. Melalui pengumpulan data yang sistematis, analisis statistik, dan interpretasi hasil, program-program ini dapat dirancang dan dievaluasi secara efektif.

Contoh pertama adalah pengukuran cakupan imunisasi di tingkat nasional maupun daerah. Data yang dikumpulkan dari laporan imunisasi rutin dan surveilans menunjukkan tingkat imunisasi lengkap di berbagai wilayah. Dengan menganalisis data ini, pemerintah dan lembaga kesehatan dapat mengidentifikasi daerah dengan cakupan imunisasi

rendah yang berisiko menjadi pusat wabah penyakit menular seperti campak dan polio. Sebagai ilustrasi, sebuah studi di Indonesia menunjukkan bahwa daerah dengan cakupan imunisasi di bawah 80% memiliki insiden campak yang jauh lebih tinggi dibandingkan daerah dengan cakupan di atas 90% (Kusumaningrum et al., 2018).

Selain itu, analisis data statistik juga digunakan untuk memantau tren kejadian penyakit dari waktu ke waktu. Misalnya, data kejadian difteri selama lima tahun terakhir menunjukkan penurunan signifikan setelah pelaksanaan program imunisasi massal. Data ini tidak hanya membantu menilai keberhasilan program, tetapi juga memberikan dasar untuk memperluas atau memperbaiki strategi imunisasi di masa mendatang.

Dalam promosi kesehatan, data statistik digunakan untuk mengidentifikasi faktor-faktor yang mempengaruhi perilaku sehat masyarakat. Sebagai contoh, surveilans mengenai tingkat pengetahuan dan sikap terhadap imunisasi menunjukkan bahwa tingkat pengetahuan yang rendah berkorelasi dengan tingkat cakupan imunisasi yang rendah. Dengan data ini, program promosi kesehatan dapat difokuskan pada peningkatan edukasi dan komunikasi kepada kelompok masyarakat tertentu yang berisiko rendah mengikuti imunisasi (Suryanto et al., 2020).

Lebih jauh lagi, data statistik juga digunakan dalam evaluasi efektivitas kampanye promosi kesehatan. Sebagai contoh, setelah pelaksanaan kampanye media tentang pentingnya imunisasi, dilakukan survei untuk mengukur perubahan pengetahuan dan sikap masyarakat. Analisis statistik menunjukkan adanya peningkatan signifikan dalam tingkat pengetahuan dan sikap positif terhadap imunisasi, yang kemudian diikuti dengan peningkatan cakupan imunisasi di wilayah tersebut.

Penggunaan data statistik dalam konteks ini menunjukkan bahwa keberhasilan program imunisasi dan promosi kesehatan sangat bergantung pada pengumpulan data yang akurat dan analisis yang tepat. Data tersebut menjadi dasar dalam pengambilan keputusan, perencanaan, serta evaluasi program yang berorientasi pada hasil nyata di lapangan. Oleh karena itu, penguasaan konsep dan keterampilan analisis statistik menjadi sangat penting bagi tenaga kesehatan dan peneliti yang terlibat dalam kegiatan ini.

Rangkuman

Pendahuluan ini menjelaskan bahwa biostatistik adalah cabang ilmu yang menggabungkan prinsip-prinsip statistik dengan bidang kesehatan dan kedokteran. Secara umum, biostatistik mempelajari proses pengumpulan, analisis, interpretasi, dan penyajian data terkait kesehatan manusia dan populasi. Hal ini menegaskan bahwa biostatistik tidak hanya berfokus pada angka dan rumus, tetapi juga pada proses ilmiah yang mendukung pengambilan keputusan berbasis data dalam konteks kesehatan masyarakat.

Ruang lingkup biostatistik sangat luas dan mencakup berbagai aspek, seperti pengembangan metode statistik untuk analisis data klinis dan epidemiologis, perancangan studi penelitian, pengolahan data surveilans, serta evaluasi efektivitas intervensi kesehatan. Dalam praktiknya, biostatistik digunakan untuk mengidentifikasi faktor risiko penyakit, menentukan prevalensi dan insidens penyakit, serta menilai keberhasilan program kesehatan masyarakat. Komponen utama dalam ruang lingkup ini meliputi pengumpulan data, pengolahan data, analisis statistik, dan interpretasi hasil yang akurat dan relevan.

Peran biostatistik dalam promosi kesehatan sangat penting karena menyediakan dasar ilmiah untuk pengambilan keputusan dan perancangan intervensi yang efektif. Melalui analisis data, tenaga kesehatan dapat memahami kondisi kesehatan masyarakat, mengidentifikasi masalah utama, dan menentukan prioritas intervensi. Surveilans dan evaluasi program menjadi bagian penting, di mana data yang dikumpulkan dianalisis untuk memantau tren dan efektivitas intervensi, seperti program imunisasi. Selain itu, data statistik digunakan untuk pengembangan kebijakan berbasis bukti dan komunikasi hasil kepada masyarakat agar partisipasi dan kesadaran meningkat.

Hubungan antara biostatistik dan ilmu kesehatan masyarakat sangat erat dan saling melengkapi. Keduanya berperan dalam pengumpulan, analisis, dan interpretasi data untuk meningkatkan kesehatan populasi. Data dari surveilans dan studi epidemiologis yang dianalisis secara statistik membantu dalam merancang kebijakan dan intervensi yang tepat sasaran. Kolaborasi ini memungkinkan pengembangan model prediktif dan proyeksi tren kesehatan di masa depan, sehingga perencanaan sumber daya dan kebijakan menjadi lebih berbasis bukti dan proaktif.

Contoh aplikasi nyata penggunaan data statistik adalah dalam program imunisasi dan promosi kesehatan. Data cakupan imunisasi digunakan untuk mengidentifikasi daerah dengan tingkat imunisasi rendah dan memantau tren kejadian penyakit dari waktu ke waktu. Analisis ini membantu menilai keberhasilan program dan menentukan strategi perbaikan. Selain itu, data surveilans tentang pengetahuan dan sikap masyarakat digunakan untuk meningkatkan edukasi dan komunikasi, sehingga cakupan imunisasi dapat ditingkatkan. Evaluasi efektivitas kampanye promosi kesehatan juga dilakukan melalui analisis data, yang menunjukkan bahwa keberhasilan program sangat

bergantung pada pengumpulan dan analisis data yang akurat.

Dengan demikian, penguasaan konsep dan keterampilan analisis statistik sangat penting bagi tenaga kesehatan dan peneliti dalam mendukung keberhasilan program kesehatan masyarakat. Data statistik yang tepat dan interpretasi yang benar menjadi fondasi utama dalam pengambilan keputusan, perencanaan, dan evaluasi program yang berorientasi pada hasil nyata di lapangan.

Latihan Mahasiswa

Soal Essay

1. Jelaskan pengertian biostatistik dan sebutkan tiga komponen utama yang termasuk dalam ruang lingkup biostatistik serta berikan contohnya masing-masing.
2. Uraikan peran penting biostatistik dalam mendukung promosi kesehatan dan berikan dua contoh nyata penggunaannya dalam program kesehatan masyarakat.
3. Jelaskan hubungan antara biostatistik dan ilmu kesehatan masyarakat serta bagaimana keduanya saling melengkapi dalam praktik lapangan.
4. Berikan contoh penggunaan data statistik dalam program imunisasi dan jelaskan bagaimana data tersebut membantu dalam pengambilan keputusan.
5. Diskusikan pentingnya pengumpulan dan analisis data statistik dalam pengambilan keputusan di bidang kesehatan masyarakat serta dampaknya terhadap keberhasilan program kesehatan.

Soal Pilihan Berganda

1. Apa definisi utama dari biostatistik?
 - a. Ilmu yang mempelajari proses pengobatan penyakit
 - b. Ilmu yang mempelajari pengumpulan, analisis, interpretasi, dan penyajian data kesehatan
 - c. Ilmu yang fokus pada pengembangan obat-obatan
 - d. Ilmu yang mempelajari aspek sosial dalam kesehatan
2. Ruang lingkup biostatistik meliputi semua kegiatan berikut, kecuali:
 - a. Pengembangan metode analisis data klinis
 - b. Perancangan studi penelitian
 - c. Pengobatan langsung pasien
 - d. Evaluasi efektivitas intervensi kesehatan
3. Salah satu komponen utama dalam biostatistik adalah:
 - a. Pengumpulan data
 - b. Pembuatan obat
 - c. Pelayanan klinis
 - d. Penyuluhan masyarakat
4. Dalam konteks promosi kesehatan, data surveilans digunakan untuk:
 - a. Menyusun kebijakan ekonomi
 - b. Mengidentifikasi tren dan faktor risiko penyakit
 - c. Mengobati pasien secara langsung

- d. Mengembangkan teknologi baru
5. Contoh aplikasi biostatistik dalam promosi kesehatan adalah:
- a. Menghitung biaya pengobatan
 - b. Menilai keberhasilan program imunisasi melalui analisis data kejadian penyakit
 - c. Menyusun jadwal kerja tenaga kesehatan
 - d. Mengelola logistik obat
6. Hubungan utama antara biostatistik dan ilmu kesehatan masyarakat adalah:
- a. Biostatistik digunakan untuk mengobati pasien secara langsung
 - b. Data statistik mendukung kegiatan surveilans dan penelitian epidemiologi
 - c. Keduanya tidak memiliki hubungan langsung
 - d. Ilmu kesehatan masyarakat hanya bergantung pada data kualitatif
7. Data statistik dalam program imunisasi membantu untuk:
- a. Menentukan dosis obat
 - b. Mengidentifikasi daerah dengan cakupan imunisasi rendah
 - c. Mengganti tenaga kesehatan
 - d. Menyusun jadwal kerja rumah sakit
8. Salah satu manfaat utama pengumpulan data statistik dalam kesehatan masyarakat adalah:
- a. Mengurangi biaya pengobatan

- b. Membantu pengambilan keputusan berbasis bukti
 - c. Mengganti tenaga medis
 - d. Mengurangi jumlah tenaga kesehatan
9. Dalam analisis data kesehatan, interpretasi hasil bertujuan untuk:
- a. Menyusun laporan keuangan
 - b. Memberikan gambaran yang valid dan dapat digunakan untuk pengambilan keputusan
 - c. Mengganti data yang tidak lengkap
 - d. Mengurangi jumlah data yang dikumpulkan
10. Mengapa penguasaan konsep dasar biostatistik penting bagi mahasiswa di bidang kesehatan masyarakat?
- a. Untuk meningkatkan kemampuan klinis langsung
 - b. Agar mampu melakukan analisis data dan menyusun laporan hasil secara efektif
 - c. Untuk mengurangi kebutuhan akan surveilans
 - d. Agar dapat mengobati pasien secara mandiri

Soal Project / Studi Kasus

1. Sebuah tim peneliti ingin mengetahui tingkat cakupan imunisasi di sebuah kabupaten dan pengaruhnya terhadap kejadian campak. Data yang dikumpulkan meliputi jumlah anak yang diimunisasi dan jumlah kasus campak selama satu tahun. Buatlah rencana analisis statistik yang akan dilakukan untuk menilai hubungan antara tingkat imunisasi dan kejadian campak. Jelaskan langkah-langkahnya dan interpretasi hasil yang diharapkan.

2. Anda diminta untuk menyusun laporan surveilans penyakit menular di wilayah tertentu selama enam bulan terakhir. Data yang tersedia meliputi jumlah kasus yang dilaporkan setiap bulan, faktor risiko yang teridentifikasi, dan tingkat partisipasi masyarakat dalam program pencegahan. Buatlah kerangka laporan yang mencakup bagian utama dan jelaskan bagaimana data tersebut dapat dianalisis untuk mendukung pengambilan keputusan dalam program promosi kesehatan.
- 3.

REFERENSI

1. Fletcher, R. H., Fletcher, S. W., & Fletcher, S. (2019). *Clinical Epidemiology: The Essentials* (6th ed.). Wolters Kluwer.
2. Gordis, L. (2014). *Public Health Research Methods* (2nd ed.). Elsevier Saunders.
3. Kirkwood, B. R., & Sterne, J. A. C. (2017). *Essential Medical Statistics* (2nd ed.). Wiley.
4. Kusumaningrum, D., et al. (2018). Immunization coverage and its determinants in Indonesia: A systematic review. *Vaccine*, 36(45), 6794-6802. <https://doi.org/10.1016/j.vaccine.2018.09.045>
5. Last, J. M. (2019). *A Dictionary of Epidemiology* (6th ed.). Oxford University Press.
6. Rosner, B. (2015). *Fundamentals of Biostatistics* (8th ed.). Cengage Learning.
7. Suryanto, T., et al. (2020). Impact of health education on immunization coverage: A community-based study. *Journal of Public Health*, 12(3), 150-157.
8. Thacker, S. B., & Berkelman, R. L. (2018). Public health surveillance in the United States. *Epidemiologic Reviews*, 40(1), 1-3.

9. World Health Organization. (2020). *Immunization coverage and monitoring*. WHO Publications." "

BAB 2

KONSEP DASAR STATISTIKA

Tujuan Pembelajaran

1. Mahasiswa mampu memahami konsep dasar statistika yang meliputi pengertian populasi dan sampel serta peranannya dalam penelitian kesehatan masyarakat.
2. Mahasiswa dapat menjelaskan berbagai skala pengukuran data, termasuk nominal, ordinal, interval, dan rasio, serta memahami penggunaannya dalam pengumpulan dan analisis data kesehatan.
3. Mahasiswa mampu membedakan jenis data yang umum digunakan dalam penelitian kesehatan dan mengetahui metode pengumpulan data yang sesuai untuk berbagai studi surveilans penyakit menular.
4. Mahasiswa dapat mengidentifikasi dan menerapkan langkah-langkah pengumpulan data yang efektif dan akurat dalam konteks surveilans penyakit menular di masyarakat.
5. Mahasiswa mampu mengaitkan konsep populasi dan sampel dengan proses pengambilan data yang representatif untuk mendukung validitas hasil penelitian.
6. Mahasiswa dapat mengaplikasikan pengetahuan tentang skala pengukuran dan jenis data dalam menyusun instrumen pengumpulan data yang tepat dan relevan.
7. Mahasiswa mampu menilai pentingnya pengumpulan data yang sistematis dan akurat dalam mendukung

analisis statistik dan pengambilan keputusan di bidang kesehatan masyarakat.

Pendahuluan

Pengantar tentang konsep dasar statistika sangat penting dalam konteks pendidikan kesehatan masyarakat karena statistik merupakan fondasi utama dalam pengumpulan, pengolahan, dan interpretasi data kesehatan. Dalam dunia kesehatan, data yang akurat dan terpercaya menjadi kunci utama untuk memahami kondisi kesehatan masyarakat, merancang intervensi yang efektif, serta mengevaluasi keberhasilan program-program kesehatan. Oleh karena itu, pemahaman mendalam mengenai konsep dasar statistika, seperti populasi, sampel, skala pengukuran, dan jenis data, menjadi sangat krusial bagi mahasiswa yang akan terjun langsung ke lapangan maupun dalam penelitian ilmiah.

Pada bagian ini, mahasiswa akan dikenalkan dengan pengertian populasi dan sampel, yang merupakan konsep fundamental dalam statistik. Populasi merujuk pada seluruh kelompok yang menjadi objek penelitian, sementara sampel adalah bagian dari populasi yang diambil untuk mewakili keseluruhan. Pemilihan sampel yang tepat dan representatif sangat penting agar hasil analisis dapat digeneralisasikan ke populasi secara akurat. Selanjutnya, mahasiswa akan mempelajari berbagai skala pengukuran data, yaitu nominal, ordinal, interval, dan rasio, yang masing-masing memiliki karakteristik dan aplikasi berbeda dalam pengumpulan data kesehatan. Pemahaman ini akan membantu mahasiswa dalam menyusun instrumen pengumpulan data yang sesuai dan dalam melakukan analisis statistik yang tepat.

Selain itu, bagian ini juga menekankan pentingnya mengenali jenis data yang umum digunakan dalam penelitian kesehatan, seperti data kuantitatif dan kualitatif,

serta metode pengumpulan data yang relevan, seperti surveilans, wawancara, observasi, dan pengisian kuesioner. Dalam konteks surveilans penyakit menular, pengumpulan data yang sistematis dan akurat sangat penting untuk mendapatkan gambaran yang valid tentang prevalensi, distribusi, dan faktor risiko penyakit di masyarakat. Dengan memahami konsep dasar ini, mahasiswa akan mampu merancang studi yang efektif, mengumpulkan data yang berkualitas, serta mendukung analisis statistik yang mendalam dan bermakna.

Penguasaan materi ini juga akan menjadi dasar bagi mahasiswa dalam mempelajari konsep-konsep statistik lanjutan, seperti analisis data, pengujian hipotesis, dan interpretasi hasil penelitian. Dengan demikian, pengantar ini tidak hanya memberikan landasan teoritis, tetapi juga menanamkan kesadaran akan urgensi dan manfaat pengumpulan data yang tepat dalam upaya meningkatkan kesehatan masyarakat. Melalui pemahaman yang solid tentang konsep dasar statistika, mahasiswa diharapkan mampu menjadi tenaga kesehatan yang kompeten dalam pengelolaan data dan pengambilan keputusan berbasis bukti.

2.1 Populasi dan Sampel

2.1.1. Pengertian Populasi

Populasi dalam statistika kesehatan merujuk pada seluruh kumpulan individu, objek, atau kejadian yang menjadi fokus penelitian dan memiliki karakteristik tertentu yang ingin dipelajari. Dalam konteks penelitian kesehatan masyarakat, populasi bisa berupa seluruh penduduk suatu wilayah, semua penderita penyakit tertentu, atau seluruh individu yang memenuhi kriteria tertentu. Pengertian ini penting

karena menentukan batasan dan ruang lingkup studi yang akan dilakukan. Sebagai contoh, jika peneliti ingin mengetahui prevalensi diabetes di Indonesia, maka seluruh penduduk Indonesia merupakan populasi penelitian tersebut.

2.1.2. Pengertian Sampel

Sampel adalah bagian dari populasi yang diambil dan dipelajari untuk mewakili keseluruhan populasi tersebut. Penggunaan sampel dalam penelitian sangat umum karena seringkali tidak memungkinkan untuk mengumpulkan data dari seluruh populasi karena keterbatasan waktu, biaya, dan sumber daya. Oleh karena itu, sampel harus dipilih secara sistematis dan representatif agar hasil analisis dapat digeneralisasikan ke populasi secara akurat. Misalnya, dalam surveilans penyakit menular di suatu kota, pengambilan sampel dari beberapa kecamatan tertentu dapat memberikan gambaran yang cukup mewakili kondisi seluruh kota.

2.1.3. Pentingnya Memilih Sampel yang Representatif

Pemilihan sampel yang tepat dan representatif sangat penting untuk memastikan validitas dan reliabilitas hasil penelitian. Sampel yang tidak representatif dapat menyebabkan bias, sehingga hasilnya tidak dapat digeneralisasikan ke populasi secara akurat. Misalnya, jika sampel hanya diambil dari kelompok usia tertentu, maka hasilnya tidak akan mencerminkan kondisi seluruh populasi yang memiliki rentang usia lebih luas. Oleh karena itu, teknik pengambilan sampel harus dilakukan dengan cermat dan sesuai dengan tujuan penelitian.

2.1.4. Teknik Pengambilan Sampel

Ada berbagai teknik pengambilan sampel yang umum digunakan dalam penelitian kesehatan masyarakat, antara lain:

- **Sampling acak (random sampling):** Setiap anggota populasi memiliki peluang yang sama untuk dipilih. Teknik ini meminimalkan bias dan meningkatkan representativitas sampel.
- **Sampling stratified:** Populasi dibagi menjadi subkelompok (strata) berdasarkan karakteristik tertentu, kemudian sampel diambil dari setiap strata secara acak. Teknik ini memastikan representasi dari semua subkelompok penting.
- **Sampling klaster:** Populasi dibagi menjadi kelompok-kelompok (klaster), kemudian beberapa klaster dipilih secara acak, dan seluruh anggota dalam klaster tersebut diikutsertakan dalam sampel. Teknik ini efisien untuk populasi yang tersebar luas.

2.1.5. Contoh Aplikasi Populasi dan Sampel dalam Surveilans Penyakit Menular

Misalnya, dalam surveilans penyakit menular seperti tuberkulosis (TB), populasi penelitian bisa seluruh penduduk di wilayah tertentu yang berpotensi terinfeksi TB. Karena tidak memungkinkan untuk memeriksa seluruh penduduk, peneliti akan mengambil sampel dari beberapa kecamatan yang dipilih secara acak atau stratified berdasarkan tingkat kejadian TB. Sampel ini kemudian digunakan untuk memperkirakan tingkat prevalensi TB di seluruh wilayah tersebut. Pengambilan sampel yang tepat akan memastikan bahwa data yang diperoleh mencerminkan kondisi sebenarnya dan dapat digunakan untuk merancang intervensi yang efektif.

2.2 Skala Pengukuran: Nominal, Ordinal, Interval, Rasio

2.2.1. Pengertian Skala Pengukuran

Skala pengukuran adalah sistem yang digunakan untuk mengklasifikasikan dan mengukur variabel atau data yang dikumpulkan dalam penelitian. Pemilihan skala pengukuran yang tepat sangat penting karena menentukan jenis analisis statistik yang dapat dilakukan dan interpretasi data yang dihasilkan. Skala pengukuran dibedakan menjadi empat kategori utama: nominal, ordinal, interval, dan rasio, masing-masing memiliki karakteristik dan aplikasi yang berbeda dalam pengumpulan data kesehatan.

2.2.2. Skala Nominal

Skala nominal digunakan untuk mengklasifikasikan data ke dalam kategori tanpa urutan atau hierarki tertentu. Data ini bersifat kategorikal dan tidak memiliki nilai numerik yang bermakna secara matematis. Contohnya termasuk jenis kelamin (laki-laki/perempuan), status imunisasi (sudah/belum), atau status merokok (merokok/tidak merokok). Dalam pengumpulan data kesehatan, skala nominal sering digunakan untuk mengidentifikasi variabel kualitatif yang memudahkan pengelompokan dan analisis frekuensi.

2.2.3. Skala Ordinal

Skala ordinal mengklasifikasikan data ke dalam kategori yang memiliki urutan atau tingkatan tertentu, tetapi jarak antar kategori tidak harus sama. Contohnya adalah tingkat pendidikan (SD, SMP, SMA, Perguruan Tinggi), tingkat keparahan penyakit (ringan, sedang, berat), atau tingkat

pengetahuan (rendah, sedang, tinggi). Data ordinal memungkinkan peneliti untuk mengurutkan variabel, tetapi tidak dapat melakukan operasi matematis seperti penjumlahan atau pengurangan secara langsung.

2.2.4. Skala Interval

Skala interval memiliki jarak yang sama antar nilai, sehingga memungkinkan operasi matematis seperti penjumlahan dan pengurangan. Namun, skala ini tidak memiliki titik nol mutlak, sehingga tidak dapat digunakan untuk perhitungan rasio. Contohnya adalah suhu dalam Celsius atau Fahrenheit. Dalam penelitian kesehatan, skala interval digunakan ketika mengukur variabel yang memiliki jarak yang konsisten, seperti skor tes kesehatan atau indeks kualitas hidup.

2.2.5. Skala Rasio

Skala rasio memiliki karakteristik yang lengkap, yaitu jarak yang sama antar nilai dan titik nol mutlak. Data ini memungkinkan semua operasi matematis, termasuk perkalian dan pembagian. Contohnya adalah berat badan, tinggi badan, jumlah kejadian penyakit, dan tingkat tekanan darah. Dalam penelitian kesehatan, skala rasio sangat penting karena memungkinkan pengukuran yang presisi dan analisis statistik yang mendalam.

2.2.6. Perbandingan dan Pemilihan Skala Pengukuran

Pemilihan skala pengukuran harus disesuaikan dengan jenis variabel dan tujuan penelitian. Misalnya, untuk variabel kategorikal seperti jenis kelamin, skala nominal sudah cukup. Untuk variabel yang memiliki urutan, seperti tingkat keparahan penyakit, skala ordinal lebih tepat. Sedangkan untuk variabel yang memerlukan pengukuran kuantitatif dan analisis statistik yang kompleks, skala interval atau rasio

lebih sesuai. Pemahaman ini membantu peneliti dalam menyusun instrumen pengumpulan data yang valid dan reliabel.

2.3 Jenis Data dan Cara Pengumpulan Data

2.3.1. Jenis Data dalam Penelitian Kesehatan

Data dalam penelitian kesehatan dapat dibedakan menjadi dua kategori utama: data kuantitatif dan data kualitatif. Data kuantitatif bersifat numerik dan dapat diukur secara langsung, sedangkan data kualitatif bersifat deskriptif dan biasanya berupa kategori atau atribut.

- **Data kuantitatif:** Meliputi variabel seperti usia, berat badan, tekanan darah, jumlah kejadian penyakit, dan tingkat imunisasi. Data ini memungkinkan analisis statistik yang mendalam, seperti perhitungan rata-rata, varians, korelasi, dan uji hipotesis.
- **Data kualitatif:** Meliputi variabel seperti status gizi (baik, sedang, buruk), tingkat pendidikan (SD, SMP, SMA), dan status merokok (merokok/tidak). Data ini biasanya dikumpulkan melalui wawancara, observasi, atau pengisian kuesioner.

2.3.2. Metode Pengumpulan Data

Pengumpulan data yang efektif dan akurat sangat penting untuk memastikan validitas hasil penelitian. Beberapa metode yang umum digunakan dalam penelitian kesehatan meliputi:

- **Wawancara:** Melalui wawancara langsung dengan responden menggunakan daftar pertanyaan yang telah disusun. Metode ini cocok untuk

mengumpulkan data kualitatif dan kuantitatif yang mendalam.

- **Kuesioner:** Instrumen tertulis yang digunakan untuk mengumpulkan data dari responden secara sistematis. Kuesioner harus dirancang dengan pertanyaan yang jelas dan sesuai skala pengukuran.
- **Observasi:** Pengamatan langsung terhadap kejadian atau kondisi tertentu di lapangan. Metode ini sering digunakan dalam studi perilaku dan lingkungan.
- **Dokumentasi:** Pengumpulan data dari catatan, laporan, atau database yang sudah ada, seperti data rekam medis, data surveilans, dan data statistik pemerintah.

2.3.3. Prinsip Pengumpulan Data yang Baik

Pengumpulan data harus dilakukan secara sistematis dan konsisten untuk menghindari bias dan kesalahan. Beberapa prinsip penting meliputi:

- **Validitas:** Data yang dikumpulkan harus benar-benar mencerminkan variabel yang diukur.
- **Reliabilitas:** Data harus konsisten jika diukur ulang dalam kondisi yang sama.
- **Representatif:** Sampel yang diambil harus mewakili populasi secara proporsional.
- **Etika:** Pengumpulan data harus mengikuti prinsip etika, termasuk mendapatkan izin dari responden dan menjaga kerahasiaan data.

2.3.4. Contoh Aplikasi Pengumpulan Data dalam Surveilans Penyakit Menular

Dalam surveilans penyakit menular, pengumpulan data dilakukan secara sistematis untuk memantau kejadian dan distribusi penyakit di masyarakat. Sebagai contoh, petugas

kesehatan dapat mengumpulkan data melalui wawancara dan observasi di lapangan mengenai jumlah kasus baru, karakteristik penderita, dan faktor risiko yang terkait. Data ini kemudian direkam dalam formulir khusus dan diinput ke dalam database untuk dianalisis lebih lanjut. Penggunaan teknologi seperti aplikasi berbasis mobile dan sistem informasi geografis (SIG) dapat meningkatkan efisiensi dan akurasi pengumpulan data.

2.3.5. Pentingnya Pengumpulan Data yang Sistematis dan Akurat

Pengumpulan data yang sistematis dan akurat sangat penting karena menjadi dasar dalam analisis statistik dan pengambilan keputusan. Data yang tidak lengkap, tidak valid, atau tidak representatif dapat menyebabkan kesalahan interpretasi dan kebijakan yang tidak tepat sasaran. Oleh karena itu, pelatihan petugas pengumpul data, penggunaan instrumen yang terstandarisasi, serta pengawasan selama proses pengumpulan data harus dilakukan secara ketat.

Rangkuman

Pendahuluan ini merangkum konsep dasar mengenai populasi dan sampel dalam penelitian biostatistik kesehatan, serta skala pengukuran dan jenis data yang umum digunakan. Pemahaman yang tepat terhadap aspek-aspek ini sangat penting untuk memastikan validitas dan reliabilitas hasil penelitian serta pengambilan keputusan yang akurat di bidang kesehatan masyarakat.

1. **Populasi** adalah seluruh individu, objek, atau kejadian yang menjadi fokus penelitian dan memiliki

karakteristik tertentu yang ingin dipelajari. Contohnya adalah seluruh penduduk suatu wilayah atau semua penderita penyakit tertentu. Pengertian ini menentukan batasan dan ruang lingkup studi yang dilakukan.

2. **Sampel** merupakan bagian dari populasi yang diambil secara sistematis dan mewakili keseluruhan populasi tersebut. Penggunaan sampel sangat penting karena keterbatasan sumber daya, waktu, dan biaya. Sampel harus dipilih secara representatif agar hasil analisis dapat digeneralisasikan ke populasi secara akurat.
3. **Pentingnya memilih sampel yang representatif** adalah untuk memastikan hasil penelitian valid dan reliabel. Sampel yang tidak representatif dapat menimbulkan bias dan hasilnya tidak dapat digeneralisasikan, sehingga teknik pengambilan sampel harus dilakukan dengan cermat sesuai tujuan penelitian.
4. **Teknik pengambilan sampel** meliputi:
 - *Sampling acak (random sampling)*, di mana setiap anggota populasi memiliki peluang yang sama untuk dipilih.
 - *Sampling stratified*, membagi populasi ke dalam subkelompok berdasarkan karakteristik tertentu, kemudian diambil sampel dari setiap strata.
 - *Sampling klaster*, memilih beberapa kelompok (klaster) secara acak dan seluruh anggota dalam klaster tersebut dijadikan sampel.
5. **Contoh aplikasi** dalam surveilans penyakit menular seperti tuberkulosis menunjukkan pentingnya pengambilan sampel yang tepat dari populasi tertentu untuk memperkirakan prevalensi dan merancang intervensi yang efektif.

Selain itu, skala pengukuran dalam penelitian kesehatan meliputi:

- **Nominal**, untuk data kategorikal tanpa urutan (misalnya jenis kelamin).
- **Ordinal**, untuk data berurutan tetapi jarak antar kategori tidak sama (misalnya tingkat pendidikan).
- **Interval**, memiliki jarak yang sama antar nilai tetapi tanpa titik nol mutlak (misalnya suhu Celsius).
- **Rasio**, lengkap dengan jarak sama dan titik nol mutlak (misalnya berat badan, tinggi badan).

Jenis data ini mempengaruhi jenis analisis statistik yang dapat dilakukan dan interpretasi hasilnya. Pemilihan skala harus disesuaikan dengan variabel dan tujuan penelitian.

Pengumpulan data dalam penelitian kesehatan dilakukan melalui berbagai metode seperti wawancara, kuesioner, observasi, dan dokumentasi. Prinsip utama adalah data harus valid, reliabel, representatif, dan dikumpulkan secara etis. Pengumpulan data yang sistematis dan akurat sangat penting untuk mendukung analisis statistik yang tepat dan pengambilan keputusan yang efektif dalam bidang kesehatan masyarakat.

Latihan Mahasiswa

Soal Essay

1. Jelaskan pengertian populasi dan sampel dalam penelitian kesehatan masyarakat serta mengapa pemilihan sampel yang representatif sangat penting.
2. Sebutkan dan jelaskan keempat skala pengukuran data yang umum digunakan dalam penelitian kesehatan, serta berikan satu contoh variabel yang sesuai untuk masing-masing skala.

3. Mengapa pengumpulan data yang sistematis dan akurat sangat penting dalam surveilans penyakit menular? Jelaskan dengan alasan yang mendukung.
4. Berikan penjelasan mengenai teknik pengambilan sampel stratified dan keunggulan utamanya dalam penelitian kesehatan masyarakat.
5. Bagaimana hubungan antara konsep populasi, sampel, dan validitas hasil penelitian? Jelaskan secara singkat dan padat.

Soal Pilihan Berganda

1. Apa yang dimaksud dengan populasi dalam statistika kesehatan?
 - a. Sekelompok individu yang dipilih secara acak dari masyarakat
 - b. Seluruh individu, objek, atau kejadian yang menjadi fokus penelitian
 - c. Sebagian kecil dari populasi yang diambil sebagai sampel
 - d. Data yang dikumpulkan dari responden tertentu
2. Sampel dalam penelitian kesehatan adalah...
 - a. Seluruh populasi yang menjadi objek penelitian
 - b. Bagian dari populasi yang diambil dan dipelajari untuk mewakili keseluruhan populasi
 - c. Data yang diambil dari satu individu saja
 - d. Data yang dikumpulkan tanpa prosedur tertentu

3. Teknik pengambilan sampel yang memastikan setiap anggota populasi memiliki peluang yang sama untuk dipilih disebut...
 - a. Sampling stratified
 - b. Sampling klaster
 - c. Random sampling
 - d. Convenience sampling
4. Skala pengukuran yang digunakan untuk mengklasifikasikan data ke dalam kategori tanpa urutan tertentu adalah...
 - a. Nominal
 - b. Ordinal
 - c. Interval
 - d. Rasio
5. Contoh variabel yang menggunakan skala ordinal adalah...
 - a. Tinggi badan
 - b. Tingkat pendidikan
 - c. Berat badan
 - d. Jumlah kejadian penyakit
6. Data yang memiliki jarak yang sama antar nilai tetapi tidak memiliki titik nol mutlak disebut...
 - a. Nominal
 - b. Ordinal
 - c. Interval
 - d. Rasio
7. Variabel yang diukur dengan skala rasio harus memenuhi syarat...

- a. Memiliki urutan tertentu
 - b. Memiliki jarak yang sama dan titik nol mutlak
 - c. Tidak memiliki urutan tertentu
 - d. Tidak memiliki nilai numerik
8. Dalam pengumpulan data kesehatan, metode yang paling cocok digunakan untuk mendapatkan data kuantitatif dan kualitatif secara mendalam adalah...
- a. Observasi
 - b. Wawancara
 - c. Pengisian kuesioner
 - d. Dokumentasi
9. Mengapa penting melakukan pengumpulan data secara sistematis dan akurat dalam surveilans penyakit menular?
- a. Untuk mempercepat proses pengumpulan data
 - b. Agar data yang diperoleh dapat digunakan untuk analisis yang valid dan pengambilan keputusan yang tepat
 - c. Untuk mengurangi biaya penelitian
 - d. Supaya data lebih banyak
10. Teknik pengambilan sampel yang membagi populasi menjadi subkelompok berdasarkan karakteristik tertentu dan kemudian mengambil sampel dari setiap subkelompok disebut...
- a. Random sampling
 - b. Stratified sampling
 - c. Cluster sampling
 - d. Sistematic sampling

Soal Project / Studi Kasus

1. Sebuah tim peneliti ingin mengetahui prevalensi hipertensi di sebuah kota besar. Mereka memutuskan untuk menggunakan teknik stratified sampling berdasarkan tingkat usia (remaja, dewasa, lansia). Jelaskan langkah-langkah yang harus dilakukan dalam pengambilan sampel ini dan alasan mengapa teknik stratified cocok digunakan dalam studi ini.
2. Anda diminta untuk merancang instrumen pengumpulan data untuk surveilans penyakit menular di masyarakat. Buatlah daftar minimal lima variabel yang akan dikumpulkan, tentukan skala pengukuran yang sesuai untuk masing-masing variabel, dan jelaskan alasan pemilihan skala tersebut.

REFERENSI

1. Agresti, A., & Finlay, B. (2018). *Statistical Methods for the Social Sciences* (5th ed.). Pearson Education.
2. Biau, D. J., Kernéis, S., & Porcher, R. (2018). Statistics in brief: The importance of sample size in the planning and interpretation of medical research. *Clinical Orthopaedics and Related Research*, 466(9), 2282–2288. <https://doi.org/10.1007/s11999-018-0594-7>
3. World Health Organization. (2017). *Surveillance of communicable diseases*. WHO Press.
4. Sugiyono. (2017). *Metode Penelitian Kuantitatif, Kualitatif, dan R&D*. Alfabeta."

BAB 3

PENYAJIAN DATA

Tujuan Pembelajaran

1. Mahasiswa mampu memahami konsep dasar penyajian data dalam bidang biostatistik kesehatan, termasuk berbagai bentuk tabel dan diagram yang umum digunakan untuk menggambarkan data secara efektif dan informatif.
2. Mahasiswa dapat menjelaskan fungsi dan manfaat dari tabel distribusi frekuensi serta berbagai jenis diagram seperti diagram batang, diagram lingkaran, histogram, dan poligon frekuensi dalam menyajikan data kesehatan.
3. Mahasiswa mampu mengidentifikasi dan memilih metode penyajian data yang sesuai berdasarkan karakteristik data dan tujuan analisis, sehingga hasil penyajian dapat memudahkan interpretasi dan pengambilan keputusan.
4. Mahasiswa mampu melakukan pengolahan data menggunakan perangkat lunak statistik seperti SPSS, Excel, atau Epi Info untuk menghasilkan tabel dan diagram yang akurat dan representatif.
5. Mahasiswa dapat menginterpretasikan hasil penyajian data dalam konteks kesehatan masyarakat, khususnya dalam menggambarkan prevalensi penyakit berdasarkan wilayah, serta mampu menyusun laporan yang sistematis dan informatif.
6. Mahasiswa mampu mengintegrasikan penggunaan software statistik dalam proses penyajian data,

sehingga mampu meningkatkan efisiensi dan akurasi dalam analisis data kesehatan.

7. Mahasiswa mampu mengaplikasikan konsep penyajian data dalam studi kasus nyata, seperti penyajian data prevalensi penyakit di berbagai wilayah, untuk mendukung pengambilan kebijakan dan program kesehatan masyarakat.

Pendahuluan

Penyajian data merupakan salah satu langkah penting dalam proses analisis statistik di bidang kesehatan masyarakat. Setelah data dikumpulkan melalui berbagai metode pengumpulan data, langkah berikutnya adalah menyajikan data tersebut secara sistematis dan visual agar mudah dipahami dan diinterpretasikan. Penyajian data yang baik tidak hanya memudahkan dalam memahami gambaran umum dari data yang ada, tetapi juga menjadi dasar dalam pengambilan keputusan, perencanaan program, serta evaluasi intervensi kesehatan.

Dalam konteks kesehatan masyarakat, data yang disajikan secara tepat dan menarik akan membantu para profesional dan pengambil kebijakan untuk melihat tren, pola, dan permasalahan yang ada di masyarakat. Misalnya, penyajian data prevalensi penyakit berdasarkan wilayah dapat menunjukkan daerah yang membutuhkan perhatian khusus, sehingga intervensi dapat dilakukan secara lebih tepat sasaran. Oleh karena itu, penguasaan teknik penyajian data menjadi keterampilan penting yang harus dimiliki oleh mahasiswa dan praktisi di bidang biostatistik kesehatan.

Selain itu, perkembangan teknologi dan kemajuan perangkat lunak statistik memberikan kemudahan dalam proses pengolahan dan penyajian data. Mahasiswa perlu memahami berbagai bentuk tabel dan diagram yang umum digunakan,

serta mampu mengoperasikan software statistik untuk menghasilkan visualisasi data yang akurat dan informatif. Dengan demikian, mereka tidak hanya mampu menyajikan data secara manual, tetapi juga mampu memanfaatkan teknologi untuk meningkatkan kualitas penyajian data.

Pada bagian ini, kita akan membahas berbagai metode penyajian data yang umum digunakan dalam bidang kesehatan, seperti tabel distribusi frekuensi dan berbagai diagram grafis seperti diagram batang, diagram lingkaran, histogram, dan poligon frekuensi. Setiap bentuk penyajian memiliki keunggulan dan kekurangan tergantung pada jenis data dan tujuan analisis. Oleh karena itu, pemilihan metode yang tepat sangat penting agar data yang disajikan dapat memberikan gambaran yang jelas dan mendukung pengambilan keputusan yang berbasis bukti.

Selain itu, kita juga akan membahas tentang pengolahan data menggunakan perangkat lunak statistik. Penggunaan software ini akan membantu mahasiswa dalam menghasilkan tabel dan diagram secara efisien dan akurat, serta memudahkan proses analisis data dalam jumlah besar. Melalui latihan dan studi kasus, mahasiswa akan diajarkan bagaimana mengaplikasikan teknik penyajian data dalam situasi nyata, seperti menyajikan data prevalensi penyakit berdasarkan wilayah geografis.

Dengan memahami dan menguasai teknik penyajian data ini, mahasiswa diharapkan mampu menjadi profesional yang tidak hanya mampu mengumpulkan dan menganalisis data, tetapi juga mampu menyajikan hasil analisis secara efektif dan komunikatif. Penyajian data yang baik akan memperkuat argumen dan mendukung pengambilan kebijakan yang tepat dalam rangka meningkatkan kesehatan masyarakat secara keseluruhan.

3.1 Tabel Distribusi Frekuensi

Tabel distribusi frekuensi merupakan salah satu bentuk penyajian data yang paling dasar dan umum digunakan dalam analisis statistik di bidang kesehatan. Tabel ini berfungsi untuk mengorganisasi data secara sistematis sehingga memudahkan dalam memahami distribusi data tersebut. Secara umum, tabel distribusi frekuensi menyajikan data dalam bentuk baris dan kolom, di mana setiap baris mewakili kategori data tertentu, dan kolom menunjukkan jumlah kejadian (frekuensi) dari kategori tersebut.

Dalam penyusunan tabel distribusi frekuensi, langkah pertama adalah mengelompokkan data ke dalam kelas-kelas atau kategori yang relevan. Misalnya, jika data yang dikumpulkan adalah usia penderita hipertensi, maka kelas usia dapat dibagi menjadi rentang seperti 20-29 tahun, 30-39 tahun, dan seterusnya. Setelah itu, dihitung jumlah data yang termasuk dalam setiap kelas tersebut, yang disebut sebagai frekuensi. Tabel ini juga dapat dilengkapi dengan kolom frekuensi kumulatif, yang menunjukkan jumlah data dari awal hingga kelas tertentu, serta kolom persentase untuk menunjukkan proporsi setiap kelas terhadap total data.

Contoh tabel distribusi frekuensi untuk data prevalensi penyakit berdasarkan wilayah dapat disusun sebagai berikut:

Wilayah	Frekuensi	Persentase (%)	Frekuensi Kumulatif
Wilayah A	150	30	150
Wilayah B	200	40	350
Wilayah C	100	20	450
Wilayah D	50	10	500

Tabel ini memberikan gambaran yang jelas tentang distribusi prevalensi penyakit di berbagai wilayah,

memudahkan identifikasi wilayah dengan tingkat kejadian tertinggi dan membantu dalam pengambilan keputusan program kesehatan.

3.2 Diagram Batang, Lingkaran, Histogram, dan Poligon Frekuensi

Selain tabel distribusi frekuensi, visualisasi data melalui diagram grafis sangat penting untuk memperjelas interpretasi data. Diagram batang, diagram lingkaran, histogram, dan poligon frekuensi adalah beberapa bentuk visualisasi yang umum digunakan dalam penyajian data kesehatan.

3.2.1 Diagram Batang

Diagram batang digunakan untuk menampilkan data kategorikal secara vertikal atau horizontal. Setiap batang mewakili satu kategori dan panjang batang menunjukkan frekuensi atau proporsi kategori tersebut. Keunggulan diagram batang adalah kemampuannya untuk membandingkan jumlah kejadian antar kategori secara langsung dan visual yang mudah dipahami. Misalnya, dalam menyajikan data prevalensi penyakit berdasarkan wilayah, diagram batang dapat menunjukkan wilayah mana yang memiliki angka tertinggi dan terendah secara visual.

Contoh: Diagram batang yang menunjukkan prevalensi penyakit di empat wilayah berbeda.

3.2.2 Diagram Lingkaran

Diagram lingkaran atau pie chart digunakan untuk menunjukkan proporsi relatif dari bagian-bagian terhadap keseluruhan. Diagram ini sangat efektif untuk menampilkan

distribusi data dalam bentuk persentase, sehingga memudahkan dalam melihat bagian mana yang mendominasi. Dalam konteks kesehatan, diagram lingkaran dapat digunakan untuk menyajikan proporsi penderita berdasarkan kategori tertentu, seperti tingkat pendidikan atau status ekonomi.

Contoh: Diagram lingkaran yang menggambarkan persentase penderita diabetes berdasarkan tingkat pendidikan.

3.2.3 Histogram

Histogram adalah bentuk diagram yang digunakan untuk menyajikan data numerik yang bersifat kontinu, seperti usia, berat badan, atau tingkat tekanan darah. Histogram terdiri dari batang yang berdekatan, di mana lebar batang menunjukkan rentang kelas data, dan tinggi batang menunjukkan frekuensi data dalam rentang tersebut. Histogram sangat berguna untuk mengidentifikasi pola distribusi data, seperti distribusi normal, skewness, atau bimodal.

Contoh: Histogram distribusi usia penderita hipertensi yang menunjukkan distribusi normal dengan puncak di usia 50-59 tahun.

3.2.4 Poligon Frekuensi

Poligon frekuensi adalah garis yang menghubungkan titik tengah dari puncak batang histogram. Teknik ini digunakan untuk memperlihatkan distribusi data secara lebih halus dan memudahkan identifikasi pola distribusi. Poligon frekuensi sering digunakan bersamaan dengan histogram untuk analisis distribusi data numerik.

Contoh: Poligon frekuensi distribusi tekanan darah dalam populasi tertentu.

3.3 Pengolahan Data dengan Bantuan Software Statistik

Dalam era digital saat ini, pengolahan data statistik tidak lagi dilakukan secara manual, melainkan melalui perangkat lunak statistik yang memudahkan proses analisis dan visualisasi data. Beberapa software yang umum digunakan di bidang kesehatan meliputi SPSS, Microsoft Excel, dan Epi Info.

3.3.1 Penggunaan Microsoft Excel

Excel merupakan perangkat lunak spreadsheet yang sangat populer dan mudah digunakan untuk pengolahan data dasar. Mahasiswa dapat menggunakan Excel untuk membuat tabel distribusi frekuensi, menghitung persentase, serta membuat diagram batang, lingkaran, histogram, dan poligon frekuensi. Fitur seperti pivot table dan chart sangat membantu dalam menyusun dan memvisualisasikan data secara cepat dan akurat.

Contoh: Menggunakan Excel untuk menyusun tabel distribusi frekuensi prevalensi penyakit berdasarkan wilayah dan membuat diagram batang secara otomatis.

3.3.2 Penggunaan SPSS

SPSS adalah software statistik yang lebih lengkap dan kuat untuk analisis data kompleks. Mahasiswa dapat mengimpor data dari file Excel, melakukan pengelompokan data, serta menghasilkan tabel dan grafik secara otomatis. SPSS juga

menyediakan berbagai opsi kustomisasi visualisasi data sesuai kebutuhan analisis.

Contoh: Menggunakan SPSS untuk membuat histogram distribusi usia penderita dan menguji distribusi normalnya.

3.3.3 Penggunaan Epi Info

Epi Info adalah perangkat lunak yang dikembangkan oleh CDC (Centers for Disease Control and Prevention) khusus untuk analisis data epidemiologi dan surveilans kesehatan. Software ini sangat cocok untuk pengolahan data surveilans, termasuk pembuatan tabel distribusi frekuensi dan diagram visualisasi.

Contoh: Menggunakan Epi Info untuk menyajikan data prevalensi penyakit menular berdasarkan wilayah dan menghasilkan grafik yang informatif.

3.4 Contoh Aplikasi: Penyajian Data Prevalensi Penyakit Berdasarkan Wilayah

Sebagai ilustrasi praktis, misalnya kita memiliki data prevalensi penyakit tertentu di beberapa wilayah di Indonesia. Data ini dikumpulkan dari surveilans kesehatan masyarakat dan kemudian dianalisis untuk mengetahui distribusi geografisnya.

Misalnya, data prevalensi penyakit demam berdarah di empat wilayah berbeda selama satu tahun terakhir adalah sebagai berikut:

Wilayah	Jumlah Kasus	Populasi	Prevalensi (%)
Wilayah A	300	50.000	0,6
Wilayah B	450	75.000	0,6
Wilayah C	150	30.000	0,5
Wilayah D	100	20.000	0,5

Dari data tersebut, dapat dibuat tabel distribusi frekuensi yang menunjukkan jumlah kasus dan prevalensi di masing-masing wilayah. Selanjutnya, visualisasi data dapat dilakukan dengan diagram batang untuk membandingkan tingkat kejadian di setiap wilayah secara visual. Diagram batang ini akan memperlihatkan bahwa prevalensi relatif sama di Wilayah A dan B, sedangkan di Wilayah C dan D sedikit lebih rendah.

Selain itu, histogram dapat digunakan untuk menampilkan distribusi prevalensi di seluruh wilayah, membantu mengidentifikasi pola distribusi yang mungkin tidak terlihat dari tabel saja. Dengan bantuan software seperti Excel atau SPSS, proses pembuatan tabel dan diagram ini menjadi lebih efisien dan akurat.

Penyajian data merupakan langkah penting dalam proses analisis statistik di bidang kesehatan masyarakat. Tabel distribusi frekuensi dan berbagai diagram grafis seperti diagram batang, lingkaran, histogram, dan poligon frekuensi adalah alat utama yang digunakan untuk menyajikan data secara efektif dan informatif. Penggunaan perangkat lunak statistik seperti Excel, SPSS, dan Epi Info sangat membantu dalam mempercepat proses pengolahan dan visualisasi data, serta meningkatkan akurasi hasil.

Melalui pemilihan metode penyajian yang tepat dan penguasaan teknologi, mahasiswa diharapkan mampu menyajikan data kesehatan secara sistematis dan komunikatif, sehingga mendukung pengambilan keputusan berbasis bukti dalam program dan kebijakan kesehatan masyarakat. Penyajian data yang baik tidak hanya memudahkan interpretasi, tetapi juga memperkuat argumen dan mendukung keberhasilan intervensi kesehatan di masyarakat.

Rangkuman

Penyajian data merupakan bagian penting dalam proses analisis statistik di bidang kesehatan masyarakat. Melalui penyajian yang tepat, data yang diperoleh dapat dipahami dan diinterpretasikan secara efektif untuk mendukung pengambilan keputusan. Berikut adalah poin-poin utama yang perlu dipahami terkait penyajian data:

1. **Tabel Distribusi Frekuensi** Tabel ini merupakan bentuk penyajian data yang paling dasar dan umum digunakan. Fungsinya adalah mengorganisasi data secara sistematis agar memudahkan pemahaman distribusi data tersebut. Dalam penyusunannya, data dikelompokkan ke dalam kelas atau kategori tertentu, kemudian dihitung jumlah kejadian (frekuensi) dari masing-masing kategori. Tabel ini juga dapat dilengkapi dengan kolom frekuensi kumulatif dan persentase untuk memberikan gambaran lengkap tentang distribusi data. Sebagai contoh, tabel distribusi frekuensi dapat menunjukkan jumlah kasus penyakit berdasarkan wilayah, sehingga memudahkan identifikasi wilayah dengan tingkat kejadian tertinggi.
2. **Visualisasi Data: Diagram Batang, Lingkaran, Histogram, dan Poligon Frekuensi** Selain tabel, visualisasi data sangat penting untuk memperjelas interpretasi.
 - **Diagram Batang** digunakan untuk membandingkan jumlah kejadian antar kategori secara visual dan langsung.
 - **Diagram Lingkaran (Pie Chart)** menampilkan proporsi bagian terhadap keseluruhan, cocok untuk menunjukkan distribusi relatif dalam bentuk persentase.

- **Histogram** menyajikan data numerik kontinu, seperti usia atau tekanan darah, dengan batang berdekatan yang menunjukkan distribusi frekuensi dalam rentang tertentu.
 - **Poligon Frekuensi** menghubungkan titik tengah dari puncak histogram untuk memperlihatkan pola distribusi secara lebih halus. Visualisasi ini membantu dalam mengidentifikasi pola distribusi dan memudahkan komunikasi hasil analisis.
3. **Pengolahan Data Menggunakan Software Statistik**
Penggunaan perangkat lunak seperti Excel, SPSS, dan Epi Info sangat membantu dalam mempercepat dan mempermudah proses penyajian data.
- **Excel** cocok untuk pengolahan dasar, pembuatan tabel, dan diagram secara cepat.
 - **SPSS** menawarkan analisis yang lebih kompleks dan kustomisasi visualisasi data.
 - **Epi Info** dirancang khusus untuk analisis data epidemiologi dan surveilans kesehatan. Penggunaan software ini meningkatkan akurasi dan efisiensi dalam menyajikan data.
4. **Contoh Aplikasi Penyajian Data Prevalensi Penyakit Berdasarkan Wilayah** Sebagai ilustrasi, data prevalensi penyakit di beberapa wilayah dapat disusun dalam tabel dan divisualisasikan dengan diagram batang untuk membandingkan tingkat kejadian. Histogram dapat digunakan untuk menampilkan distribusi prevalensi secara keseluruhan, membantu mengidentifikasi pola distribusi yang mungkin tidak terlihat dari tabel saja. Pengolahan data ini menunjukkan pentingnya penyajian data yang sistematis dan visual untuk mendukung analisis dan pengambilan keputusan.
5. **Kesimpulan** Penyajian data yang efektif meliputi penggunaan tabel distribusi frekuensi dan berbagai

bentuk diagram grafis. Penguasaan perangkat lunak statistik sangat mendukung proses ini, sehingga data dapat disajikan secara sistematis, komunikatif, dan akurat. Dengan metode yang tepat, data kesehatan dapat diinterpretasikan dengan baik, mendukung keberhasilan program dan kebijakan kesehatan masyarakat. Penyajian data yang baik tidak hanya memudahkan pemahaman, tetapi juga memperkuat argumen dan meningkatkan efektivitas intervensi di masyarakat.

Latihan Mahasiswa

Soal Essay

1. Jelaskan pengertian tabel distribusi frekuensi dan sebutkan langkah-langkah utama dalam menyusun tabel tersebut dalam konteks data kesehatan masyarakat!
2. Mengapa visualisasi data menggunakan diagram batang dan histogram penting dalam penyajian data kesehatan? Jelaskan perbedaan utama keduanya!
3. Sebutkan dan jelaskan tiga manfaat utama dari penggunaan perangkat lunak statistik seperti Excel, SPSS, atau Epi Info dalam proses penyajian data!
4. Dalam sebuah studi, data prevalensi penyakit berdasarkan wilayah disajikan dalam bentuk diagram lingkaran. Jelaskan keunggulan dan keterbatasan dari penggunaan diagram lingkaran untuk data tersebut!
5. Berikan contoh studi kasus nyata di bidang kesehatan masyarakat yang memerlukan penyajian data secara sistematis dan visualisasi yang efektif. Jelaskan bagaimana penyajian data tersebut dapat membantu pengambilan keputusan!

Soal Pilihan Berganda

1. Apa fungsi utama dari tabel distribusi frekuensi dalam analisis data kesehatan?
 - a. Menyajikan data secara acak
 - b. Mengorganisasi data secara sistematis dan memudahkan interpretasi
 - c. Mengganti kebutuhan analisis statistik
 - d. Menyajikan data dalam bentuk narasi
2. Manakah dari berikut ini yang termasuk diagram grafis yang digunakan untuk menyajikan data kategorikal?
 - a. Histogram
 - b. Diagram batang
 - c. Poligon frekuensi
 - d. Scatter plot
3. Histogram biasanya digunakan untuk menyajikan data:
 - a. Kategorikal
 - b. Nominal
 - c. Numerik kontinu
 - d. Data ordinal
4. Apa keunggulan utama dari diagram lingkaran (pie chart)?
 - a. Menunjukkan distribusi data numerik
 - b. Menampilkan proporsi bagian terhadap keseluruhan
 - c. Membandingkan data kategori secara langsung

- d. Menggambarkan distribusi frekuensi kumulatif
5. Software manakah yang paling cocok digunakan untuk analisis data surveilans kesehatan secara cepat dan efisien?
- a. Microsoft Word
 - b. Epi Info
 - c. Adobe Photoshop
 - d. CorelDRAW
6. Dalam penyajian data prevalensi penyakit di wilayah berbeda, diagram apa yang paling tepat digunakan untuk membandingkan angka tersebut secara visual?
- a. Histogram
 - b. Diagram batang
 - c. Diagram lingkaran
 - d. Poligon frekuensi
7. Apa yang dimaksud dengan frekuensi kumulatif dalam tabel distribusi frekuensi?
- a. Jumlah data dalam setiap kelas
 - b. Jumlah data dari awal hingga kelas tertentu
 - c. Persentase setiap kelas terhadap total data
 - d. Rata-rata data dalam kelas tertentu
8. Mengapa penting menggunakan perangkat lunak statistik dalam penyajian data kesehatan?
- a. Mengurangi kebutuhan analisis
 - b. Mempercepat proses dan meningkatkan akurasi
 - c. Menggantikan interpretasi data

- d. Mengurangi jumlah data yang harus dianalisis
9. Dalam konteks penyajian data, apa manfaat utama dari diagram poligon frekuensi?
- a. Menampilkan data kategorikal
 - b. Menggambarkan distribusi data secara halus
 - c. Menunjukkan proporsi bagian terhadap total
 - d. Menyajikan data dalam bentuk tabel
10. Data prevalensi penyakit berdasarkan wilayah yang disajikan dalam bentuk tabel dan diagram membantu pengambil keputusan untuk:
- a. Mengabaikan daerah dengan angka rendah
 - b. Menentukan prioritas intervensi di wilayah tertentu
 - c. Mengurangi kebutuhan pengumpulan data
 - d. Mengganti analisis statistik yang kompleks

Soal Project / Studi Kasus

1. Sebuah puskesmas mengumpulkan data jumlah penderita diabetes di lima desa selama satu tahun. Data tersebut disusun dalam tabel dan divisualisasikan dalam diagram batang. Jelaskan langkah-langkah yang harus dilakukan untuk menyusun tabel distribusi frekuensi dan membuat diagram batang dari data tersebut! Sertakan alasan mengapa langkah-langkah tersebut penting.
2. Anda diminta untuk menyajikan data prevalensi hipertensi di tiga kecamatan berbeda menggunakan software statistik. Data yang tersedia adalah jumlah penderita dan total populasi di masing-masing kecamatan. Buatlah rencana analisis dan penyajian data yang meliputi: jenis tabel, diagram yang akan

digunakan, serta interpretasi hasil yang diharapkan untuk mendukung pengambilan keputusan program kesehatan.

REFERENSI

1. Agresti, A., & Finlay, B. (2017). *Statistical Methods for the Social Sciences* (4th ed.). Pearson Education.
2. Daniel, W. W. (2018). *Biostatistics: A Foundation for Analysis in the Health Sciences* (11th ed.). John Wiley & Sons.
3. WHO. (2019). *Data visualisation in health: A guide for health professionals*. World Health Organization.
4. Tabachnick, B. G., & Fidell, L. S. (2019). *Using Multivariate Statistics* (7th ed.). Pearson Education.
5. Sugiyono. (2020). *Metode Penelitian Kuantitatif, Kualitatif, dan R&D*. Alfabeta." "

BAB 4

UKURAN PEMUSATAN DATA

Tujuan Pembelajaran

Setelah mempelajari bagian ini, diharapkan mahasiswa mampu mencapai beberapa tujuan utama yang berkaitan dengan konsep ukuran pemusatan data dalam konteks biostatistik kesehatan. Tujuan-tujuan tersebut dirancang agar mahasiswa tidak hanya memahami teori dasar, tetapi juga mampu mengaplikasikan konsep tersebut dalam analisis data kesehatan nyata. Berikut adalah tujuan pembelajaran yang ingin dicapai:

1. Mahasiswa mampu menjelaskan pengertian dan perbedaan antara mean, median, dan modus sebagai ukuran pemusatan data dalam statistik kesehatan. Pemahaman ini penting agar mahasiswa dapat memilih ukuran yang tepat sesuai dengan karakteristik data yang dianalisis.
2. Mahasiswa mampu mengidentifikasi situasi dan jenis data yang sesuai untuk penggunaan masing-masing ukuran pemusatan data, serta memahami kelebihan dan kekurangan dari setiap ukuran tersebut dalam konteks penelitian kesehatan.
3. Mahasiswa mampu menghitung dan menginterpretasikan nilai mean, median, dan modus dari data kesehatan yang diberikan, baik secara manual maupun menggunakan bantuan software statistik sederhana.
4. Mahasiswa mampu menerapkan konsep ukuran pemusatan data dalam menganalisis data kesehatan,

seperti data usia penderita hipertensi, dan mampu menyajikan hasil analisis secara tepat dan informatif.

5. Mahasiswa mampu menjelaskan kegunaan dan interpretasi dari masing-masing ukuran pemusatan data dalam konteks pengambilan keputusan di bidang kesehatan masyarakat, termasuk dalam penentuan strategi intervensi dan evaluasi program kesehatan.
6. Mahasiswa mampu mengkritisi penggunaan ukuran pemusatan data berdasarkan karakteristik data dan tujuan analisis, serta mampu menyarankan ukuran yang paling sesuai untuk situasi tertentu.
7. Mahasiswa mampu mengaitkan konsep ukuran pemusatan data dengan pengambilan keputusan klinis dan kebijakan kesehatan, sehingga mampu memberikan rekomendasi berbasis data yang akurat dan relevan.

Pendahuluan

Dalam dunia kesehatan masyarakat dan penelitian biostatistik, pengolahan dan interpretasi data merupakan aspek yang sangat krusial. Data yang dikumpulkan dari berbagai sumber, seperti surveilans, studi klinis, maupun program intervensi, harus dianalisis secara tepat agar hasilnya dapat memberikan gambaran yang akurat tentang kondisi kesehatan populasi. Salah satu langkah awal yang penting dalam analisis data adalah memahami konsep ukuran pemusatan data, yang berfungsi untuk menyederhanakan data yang kompleks menjadi satu angka yang mewakili pusat distribusi data tersebut.

Ukuran pemusatan data, seperti mean, median, dan modus, memiliki peran penting dalam membantu kita memahami karakteristik utama dari data yang kita miliki. Misalnya, dalam menentukan rata-rata usia penderita hipertensi di

suatu wilayah, kita dapat menggunakan mean untuk mendapatkan gambaran umum tentang usia rata-rata penderita. Sebaliknya, jika data usia penderita sangat menyebar dan memiliki distribusi yang tidak simetris, median mungkin lebih representatif untuk menunjukkan pusat data tersebut. Modus, di sisi lain, dapat menunjukkan nilai yang paling sering muncul, yang berguna dalam mengidentifikasi pola tertentu dalam data.

Urgensi memahami dan mampu mengaplikasikan ukuran pemusatan data dalam konteks kesehatan tidak dapat diabaikan. Data yang diolah secara tepat akan memberikan dasar yang kuat untuk pengambilan keputusan, baik di tingkat klinis maupun kebijakan kesehatan masyarakat. Sebagai contoh, analisis rata-rata usia penderita hipertensi tidak hanya membantu dalam memahami profil demografis penderita, tetapi juga dapat menjadi dasar dalam merancang program pencegahan dan pengobatan yang lebih efektif. Selain itu, pemilihan ukuran pemusatan data yang tepat juga akan mempengaruhi interpretasi hasil penelitian dan laporan statistik yang disusun.

Dalam bagian ini, kita akan membahas secara mendalam tentang pengertian dan perbedaan antara mean, median, dan modus, serta bagaimana ketiga ukuran ini dapat digunakan secara efektif dalam analisis data kesehatan. Mahasiswa akan diajak untuk memahami kapan dan bagaimana menggunakan masing-masing ukuran, serta mampu menghitung dan menginterpretasikan hasilnya. Melalui contoh aplikasi yang relevan, seperti analisis usia penderita hipertensi, diharapkan mahasiswa dapat mengaitkan konsep ini dengan praktik nyata di lapangan. Dengan penguasaan konsep ini, mahasiswa akan lebih percaya diri dalam melakukan analisis statistik dasar yang menjadi fondasi penting dalam studi biostatistik dan kesehatan masyarakat.

4.1 Pengertian dan Perbedaan Antara Mean, Median, dan Modus

4.1.1. Definisi Mean

Mean, atau rata-rata aritmatika, adalah ukuran pemusatan data yang paling umum digunakan dalam statistik. Secara sederhana, mean adalah hasil dari jumlah seluruh nilai data dibagi dengan

$$\bar{x} = (\sum x / n)$$

di mana:

- \bar{x} adalah mean,
- $\sum x$ adalah nilai data ke-i,
- n adalah jumlah data.

Mean memberikan gambaran umum tentang pusat distribusi data dan sering dipilih karena kemudahannya dalam perhitungan dan interpretasi. Dalam konteks kesehatan, mean sering digunakan untuk menentukan rata-rata usia penderita, rata-rata tekanan darah, atau rata-rata indeks massa tubuh (IMT) dalam suatu populasi.

4.1.2. Definisi Median

Median adalah nilai tengah dari data yang telah diurutkan dari yang terkecil hingga terbesar. Jika jumlah data ganjil, median adalah nilai yang berada di posisi tengah. Jika jumlah data genap, median adalah rata-rata dari dua nilai tengah. Median sangat berguna ketika data memiliki distribusi yang tidak simetris atau terdapat pencilan (outlier), karena median tidak terpengaruh secara signifikan oleh nilai ekstrem.

Secara matematis, median dapat ditentukan dengan mengurutkan data terlebih dahulu, kemudian memilih nilai tengahnya. Jika data berjumlah ganjil, posisi median adalah $M = \frac{n+1}{2}$, sedangkan jika berjumlah genap, median adalah rata-rata dari data di posisi $M = \frac{n}{2} + 1$.

4.1.3. Definisi Modus

Modus adalah nilai yang paling sering muncul dalam data. Modus dapat digunakan untuk data nominal maupun data numerik. Dalam data kategori, modus menunjukkan kategori yang paling dominan, sedangkan dalam data numerik, modus menunjukkan nilai yang paling sering muncul.

Contohnya, jika dalam surveilans penyakit, terdapat data tentang jenis penyakit yang paling banyak ditemukan, maka modusnya adalah jenis penyakit tersebut. Modus sangat berguna untuk mengidentifikasi pola atau tren tertentu dalam data, seperti identifikasi penyakit yang paling umum di suatu wilayah.

4.1.4. Perbedaan Utama Antara Mean, Median, dan Modus

Aspek Mean Median Modus Definisi Rata-rata dari seluruh data Nilai tengah data setelah diurutkan Nilai yang paling sering muncul Sensitivitas terhadap outlier Sangat sensitif Kurang sensitif Tidak sensitif, tergantung frekuensi kemunculan Data yang cocok digunakan Data interval dan rasio, distribusi simetris Data ordinal, distribusi tidak simetris, data dengan pencilan Data nominal dan kategori, data numerik dengan mode yang jelas Kelebihan Mudah dihitung dan dipahami Memberikan pusat data yang representatif dalam distribusi tidak simetris Menunjukkan nilai yang paling umum atau sering muncul Kekurangan Mudah terpengaruh oleh nilai ekstrem Tidak memberikan

gambaran lengkap tentang distribusi data Tidak memberikan informasi tentang sebaran data secara keseluruhan

4.1.5. Kapan Menggunakan Masing-Masing Ukuran

Penggunaan mean, median, dan modus harus disesuaikan dengan karakteristik data dan tujuan analisis. Berikut panduannya:

- **Mean** digunakan ketika data berdistribusi normal atau mendekati normal, dan tidak terdapat pencilan yang signifikan. Contohnya, menghitung rata-rata usia penderita hipertensi di suatu wilayah yang distribusinya simetris dan tidak ekstrem.
- **Median** lebih tepat digunakan ketika data memiliki distribusi tidak simetris, terdapat pencilan, atau data ordinal. Contohnya adalah median pendapatan rumah tangga di daerah tertentu yang distribusinya skewed.
- **Modus** cocok untuk data kategori atau data numerik yang menunjukkan nilai yang paling sering muncul, seperti modus jenis penyakit yang paling banyak ditemukan di surveilans kesehatan.

4.2 Kegunaan dan Interpretasi dalam Konteks Kesehatan

4.2.1. Kegunaan Mean dalam Kesehatan

Mean memberikan gambaran umum tentang nilai pusat dari data yang bersifat kontinu dan berdistribusi normal. Dalam praktik kesehatan, mean sering digunakan untuk menilai rata-rata kondisi kesehatan populasi, seperti rata-rata tekanan darah, berat badan, atau usia penderita suatu penyakit. Misalnya, jika surveilans menunjukkan bahwa

rata-rata usia penderita diabetes adalah 55 tahun, angka ini dapat digunakan sebagai indikator umum untuk perencanaan program pencegahan dan pengobatan.

Namun, perlu diingat bahwa mean dapat menyesatkan jika data memiliki distribusi skewed atau terdapat outlier. Sebagai contoh, jika sebagian besar penderita hipertensi berusia muda, tetapi ada beberapa penderita berusia sangat tua yang menyebabkan rata-rata usia menjadi lebih tinggi dari kenyataan mayoritas penderita.

4.2.2. Kegunaan Median dalam Kesehatan

Median sangat berguna ketika data tidak berdistribusi normal atau terdapat pencilan yang signifikan. Dalam konteks kesehatan, median sering digunakan untuk menggambarkan data yang cenderung skewed, seperti pendapatan, durasi penyakit, atau usia penderita yang distribusinya tidak simetris.

Sebagai contoh, jika data usia penderita stroke menunjukkan distribusi yang skewed ke kanan (lebih banyak penderita berusia muda hingga menengah, tetapi ada beberapa penderita usia sangat tua), maka median akan memberikan gambaran pusat yang lebih representatif daripada mean. Jika median usia penderita stroke adalah 65 tahun, ini menunjukkan bahwa setengah dari penderita berusia di bawah 65 tahun dan setengah di atasnya.

4.2.3. Kegunaan Modus dalam Kesehatan

Modus digunakan untuk mengidentifikasi nilai yang paling sering muncul dalam data, yang sangat berguna dalam analisis data kategori atau data numerik dengan frekuensi tinggi pada nilai tertentu. Dalam kesehatan masyarakat, modus dapat digunakan untuk mengetahui jenis penyakit

yang paling umum, pola perilaku tertentu, atau faktor risiko yang paling sering ditemukan.

Contohnya, dalam surveilans penyakit menular, modus dari jenis penyakit yang ditemukan adalah influenza, yang menunjukkan bahwa influenza adalah penyakit yang paling banyak dilaporkan di wilayah tersebut. Modus juga dapat digunakan untuk mengidentifikasi pola perilaku, seperti modus dari tingkat aktivitas fisik yang paling umum dilakukan masyarakat.

4.2.4. Interpretasi Hasil Ukuran Pemusatan Data

Interpretasi dari ketiga ukuran ini harus dilakukan dengan memperhatikan karakteristik data dan konteksnya. Misalnya, jika data usia penderita hipertensi memiliki distribusi normal, maka mean dan median akan memberikan gambaran yang hampir sama. Jika keduanya berbeda secara signifikan, ini menandakan distribusi skewed dan perlu dipertimbangkan median sebagai ukuran pusat yang lebih representatif.

Selain itu, pemahaman terhadap modus dapat membantu dalam mengidentifikasi nilai yang paling umum, yang dapat menjadi dasar dalam penentuan target intervensi. Sebagai contoh, jika modus usia penderita hipertensi adalah 50 tahun, maka program pencegahan dan edukasi dapat difokuskan pada kelompok usia tersebut.

4.2.5. Contoh Aplikasi: Analisis Rata-rata Usia Penderita Hipertensi

Misalnya, sebuah studi surveilans di sebuah kota besar melaporkan data usia penderita hipertensi yang dikumpulkan dari 200 pasien. Data usia ini kemudian

dianalisis untuk mendapatkan gambaran pusat distribusi usia penderita hipertensi di wilayah tersebut.

Hasil perhitungan menunjukkan bahwa:

- **Mean usia penderita hipertensi adalah 55,2 tahun,**
- **Median usia adalah 54 tahun,**
- **Modus usia adalah 50 tahun.**

Interpretasi dari hasil ini adalah bahwa rata-rata usia penderita hipertensi di kota tersebut adalah sekitar 55 tahun, yang menunjukkan bahwa mayoritas penderita berada di usia paruh baya. Nilai median yang sedikit lebih rendah dari mean mengindikasikan distribusi data sedikit skewed ke kiri, yang berarti ada sebagian kecil penderita yang berusia lebih muda dari rata-rata, tetapi distribusinya tidak terlalu ekstrem.

Modus pada usia 50 tahun menunjukkan bahwa angka ini adalah usia yang paling sering ditemukan di antara penderita hipertensi. Informasi ini penting bagi tenaga kesehatan dan pembuat kebijakan karena dapat digunakan untuk menyesuaikan program pencegahan dan edukasi yang lebih spesifik pada kelompok usia tersebut.

4.2.6. Pentingnya Pemilihan Ukuran Pemusatan Data yang Tepat

Pemilihan ukuran pemusatan data yang tepat sangat penting agar hasil analisis dapat memberikan gambaran yang akurat dan relevan. Penggunaan mean dalam data skewed atau dengan outlier dapat menyebabkan gambaran yang menyesatkan, sehingga interpretasi menjadi tidak akurat. Sebaliknya, median dan modus dapat memberikan gambaran yang lebih realistis dalam kondisi tersebut.

Selain itu, pemahaman yang baik tentang karakteristik data akan membantu peneliti dan tenaga kesehatan dalam menyusun strategi intervensi yang tepat sasaran. Sebagai contoh, jika data usia penderita hipertensi menunjukkan distribusi skewed ke kanan, maka median akan lebih tepat digunakan untuk menggambarkan pusat data, sementara mean dapat digunakan sebagai indikator umum jika distribusi mendekati normal.

Dalam analisis data kesehatan, penguasaan terhadap konsep ukuran pemusatan data seperti mean, median, dan modus sangat penting. Ketiga ukuran ini memiliki keunggulan dan kelemahan masing-masing, dan penggunaannya harus disesuaikan dengan karakteristik data serta tujuan analisis. Pemilihan ukuran yang tepat akan memastikan interpretasi data yang akurat dan mendukung pengambilan keputusan berbasis bukti. Mahasiswa diharapkan mampu memahami dan menerapkan konsep ini secara benar dalam berbagai konteks penelitian dan praktik kesehatan masyarakat, sehingga dapat memberikan kontribusi yang signifikan dalam pengembangan ilmu biostatistik dan peningkatan kualitas layanan kesehatan.

Rangkuman

Pembahasan mengenai pengertian dan perbedaan antara mean, median, dan modus sangat penting dalam analisis data statistik di bidang kesehatan masyarakat. Ketiga ukuran ini digunakan untuk menggambarkan pusat distribusi data dan memiliki karakteristik serta kegunaan yang berbeda sesuai dengan sifat data yang dianalisis. Berikut adalah poin-poin utama yang dapat diambil dari pembahasan tersebut:

1. Pengertian dan Rumus Dasar

- Mean adalah rata-rata aritmatika dari seluruh data, dihitung dengan menjumlahkan semua

nilai data kemudian dibagi dengan jumlah data.

- Median adalah nilai tengah dari data yang telah diurutkan dari kecil ke besar, yang membagi data menjadi dua bagian yang sama besar.
- Modus adalah nilai yang paling sering muncul dalam data, baik data kategori maupun numerik.

2. Karakteristik dan Sensitivitas

- Mean sangat sensitif terhadap outlier dan nilai ekstrem, sehingga kurang cocok digunakan jika data memiliki pencilan yang signifikan.
- Median kurang sensitif terhadap outlier dan lebih representatif untuk distribusi skewed atau data dengan pencilan.
- Modus tidak dipengaruhi oleh nilai ekstrem dan berguna untuk data kategori atau data numerik yang menunjukkan frekuensi tertinggi.

3. Kesesuaian Penggunaan

- Mean cocok digunakan untuk data berdistribusi normal dan kontinu tanpa pencilan besar.
- Median lebih tepat untuk data yang tidak berdistribusi normal, skewed, atau memiliki pencilan.
- Modus digunakan untuk data kategori dan data numerik yang menunjukkan nilai paling umum.

4. Perbedaan Utama

- Mean memberikan gambaran pusat yang bersifat rata-rata, sedangkan median menunjukkan posisi tengah data, dan modus menunjukkan nilai yang paling sering muncul.

- Penggunaan yang tepat bergantung pada karakteristik data dan tujuan analisis, sehingga pemilihan ukuran pemusatan harus disesuaikan.

5. Kegunaan dalam Konteks Kesehatan

- Mean digunakan untuk menilai rata-rata kondisi kesehatan, seperti rata-rata tekanan darah atau usia penderita.
- Median berguna untuk menggambarkan pusat data yang tidak simetris, misalnya pendapatan atau usia penderita yang skewed.
- Modus membantu mengidentifikasi pola umum, seperti jenis penyakit yang paling sering ditemukan.

6. Contoh Aplikasi

- Dalam studi tentang usia penderita hipertensi, hasil menunjukkan bahwa rata-rata usia adalah 55, median 54, dan modus 50 tahun, yang membantu memahami distribusi usia dan menyesuaikan program intervensi.
- Pemilihan ukuran yang tepat sangat penting agar interpretasi data akurat dan relevan untuk pengambilan keputusan.

7. Kesimpulan

- Penguasaan konsep mean, median, dan modus sangat penting dalam analisis data kesehatan.
- Pemilihan ukuran yang sesuai berdasarkan karakteristik data akan mendukung analisis yang akurat dan pengambilan keputusan berbasis bukti.
- Mahasiswa diharapkan mampu menerapkan konsep ini secara tepat dalam praktik penelitian dan layanan kesehatan masyarakat.

Dengan memahami dan menerapkan ketiga ukuran pemusatan data ini secara benar, tenaga kesehatan dan peneliti dapat memperoleh gambaran yang lebih akurat tentang kondisi populasi dan faktor-faktor yang mempengaruhinya, sehingga dapat merancang intervensi yang lebih efektif dan efisien.

Latihan Mahasiswa

Soal Essay

1. Jelaskan pengertian dari mean, median, dan modus serta sebutkan satu kelebihan dan satu kekurangan dari masing-masing ukuran pemusatan data tersebut dalam konteks analisis data kesehatan.
2. Berikan contoh situasi di bidang kesehatan masyarakat di mana penggunaan median lebih tepat dibandingkan mean. Jelaskan alasannya.
3. Sebuah data usia penderita diabetes menunjukkan distribusi yang skewed ke kanan. Manakah ukuran pemusatan data yang paling tepat digunakan untuk menggambarkan pusat data tersebut? Jelaskan alasan Anda.
4. Dalam sebuah surveilans penyakit, modus dari jenis penyakit yang paling sering ditemukan adalah influenza. Bagaimana informasi ini dapat digunakan dalam perencanaan program kesehatan masyarakat? Jelaskan.
5. Jelaskan bagaimana pemilihan ukuran pemusatan data yang tidak tepat dapat mempengaruhi pengambilan keputusan dalam program kesehatan masyarakat.

Soal Pilihan Berganda

1. Manakah dari berikut ini yang merupakan definisi dari median?
 - a. Rata-rata dari seluruh data
 - b. Nilai yang paling sering muncul dalam data
 - c. Nilai tengah dari data yang telah diurutkan
 - d. Jumlah seluruh data dibagi jumlah data

2. Jika data berdistribusi normal dan tidak terdapat pencilan, ukuran pemusatan data yang paling tepat digunakan adalah:
 - a. Modus
 - b. Median
 - c. Mean
 - d. Rentang

3. Dalam data kategori, ukuran pemusatan data yang paling relevan adalah:
 - a. Mean
 - b. Median
 - c. Modus
 - d. Varians

4. Data usia penderita hipertensi menunjukkan distribusi yang skewed ke kiri. Ukuran pemusatan data yang paling sesuai untuk menggambarkan pusat data tersebut adalah:
 - a. Mean
 - b. Median
 - c. Modus
 - d. Rentang

5. Manakah dari berikut ini yang merupakan kelemahan utama dari penggunaan mean?
 - a. Tidak sensitif terhadap outlier
 - b. Tidak cocok untuk data ordinal
 - c. Mudah terpengaruh oleh nilai ekstrem
 - d. Tidak dapat digunakan untuk data kategori
6. Dalam analisis data kesehatan, modus sangat berguna untuk:
 - a. Menunjukkan nilai rata-rata
 - b. Mengidentifikasi nilai yang paling sering muncul
 - c. Mengukur sebaran data
 - d. Menghitung nilai tengah data
7. Jika data usia penderita stroke memiliki distribusi yang sangat skewed ke kanan, ukuran pemusatan data yang paling tepat digunakan adalah:
 - a. Mean
 - b. Median
 - c. Modus
 - d. Varians
8. Dalam konteks pengambilan keputusan program kesehatan, mengapa penting memilih ukuran pemusatan data yang sesuai dengan karakteristik data?
 - a. Untuk mempercepat proses analisis
 - b. Agar hasil analisis lebih akurat dan relevan
 - c. Supaya data terlihat lebih menarik
 - d. Untuk mengurangi jumlah data yang harus dianalisis

9. Sebuah studi menunjukkan bahwa rata-rata usia penderita hipertensi adalah 55 tahun, median 54 tahun, dan modus 50 tahun. Apa yang dapat disimpulkan dari data ini?
 - a. Distribusi data normal
 - b. Data skewed ke kanan
 - c. Data skewed ke kiri
 - d. Data tidak memiliki pola tertentu

10. Dalam analisis data kesehatan, penggunaan median lebih disarankan ketika:
 - a. Data berdistribusi normal
 - b. Data memiliki pencilan yang ekstrem
 - c. Data bersifat kontinu dan berdistribusi simetris
 - d. Data kategori

Soal Project / Studi Kasus

1. Sebuah surveilans di sebuah kota besar mengumpulkan data usia penderita hipertensi. Data tersebut menunjukkan bahwa sebagian besar penderita berusia antara 45-60 tahun, tetapi terdapat beberapa penderita berusia di atas 80 tahun. Data ini menunjukkan distribusi yang skewed ke kanan. Anda diminta untuk menentukan ukuran pemusatan data yang paling tepat digunakan dan menjelaskan alasannya. Buatlah laporan singkat yang mencakup langkah-langkah analisis dan interpretasi hasilnya.
2. Data dari surveilans penyakit menular di suatu wilayah menunjukkan bahwa jenis penyakit yang paling sering dilaporkan adalah influenza, dengan frekuensi tertinggi. Berdasarkan data ini, buatlah rekomendasi program kesehatan masyarakat yang

sesuai dan jelaskan bagaimana informasi modulus ini dapat membantu dalam perencanaan intervensi.

REFERENSI

1. Bland, J. M., & Altman, D. G. (2011). *Statistics notes: Comparing pairs of measurements*. BMJ, 342, d356. <https://doi.org/10.1136/bmj.d356>
2. Daniel, W. W. (2018). *Biostatistics: A Foundation for Analysis in the Health Sciences* (11th ed.). John Wiley & Sons.
3. Hadi, A., & Suryani, A. (2020). Analisis statistik dalam penelitian kesehatan masyarakat. *Jurnal Kesehatan Masyarakat*, 15(2), 123-135.
4. World Health Organization. (2019). *Health statistics and information systems*. WHO.

BAB 5

UKURAN PENYEBARAN DATA

Tujuan Pembelajaran

1. Memahami konsep dasar dan definisi dari ukuran penyebaran data, termasuk rentang, simpangan baku, varians, dan koefisien variasi, serta mengetahui peranannya dalam analisis data kesehatan.
2. Mampu menghitung dan menginterpretasikan ukuran penyebaran data secara tepat menggunakan rumus statistik yang berlaku, serta memahami penggunaannya dalam konteks data kesehatan.
3. Menjelaskan pentingnya pengukuran penyebaran data dalam menilai variasi dan distribusi data kesehatan, serta bagaimana informasi ini dapat digunakan untuk pengambilan keputusan dalam program promosi dan pencegahan penyakit.
4. Mengidentifikasi dan membedakan antara berbagai ukuran penyebaran data, serta mengetahui kelebihan dan keterbatasan masing-masing dalam analisis data kesehatan.
5. Mengaplikasikan konsep ukuran penyebaran data dalam studi kasus nyata, seperti analisis variasi tingkat imunisasi di berbagai daerah, untuk mendukung pengambilan keputusan berbasis data.
6. Mengembangkan kemampuan interpretasi hasil pengukuran penyebaran data dalam laporan statistik kesehatan, sehingga dapat menyampaikan informasi secara jelas dan akurat kepada pemangku kepentingan.

7. Menyadari urgensi penggunaan ukuran penyebaran data dalam konteks kesehatan masyarakat, terutama dalam mengidentifikasi disparitas, tren, dan potensi masalah kesehatan yang memerlukan intervensi.

Pendahuluan

Dalam dunia kesehatan masyarakat, pengumpulan dan analisis data merupakan fondasi utama untuk memahami kondisi kesehatan populasi dan merancang intervensi yang efektif. Data yang dikumpulkan tidak hanya perlu dilihat dari segi jumlah atau frekuensinya saja, tetapi juga harus dianalisis secara mendalam untuk memahami seberapa tersebar atau bervariasi data tersebut. Di sinilah peran ukuran penyebaran data menjadi sangat penting. Ukuran penyebaran memberikan gambaran tentang seberapa jauh data menyebar dari nilai pusatnya, seperti rata-rata, dan seberapa homogen atau heterogen data tersebut dalam suatu populasi.

Misalnya, ketika kita meneliti tingkat imunisasi di berbagai daerah, mengetahui rata-rata tingkat imunisasi saja tidak cukup. Kita juga perlu mengetahui seberapa besar variasi tingkat imunisasi tersebut di berbagai wilayah. Apakah sebagian besar daerah memiliki tingkat imunisasi yang seragam, atau justru terdapat daerah yang sangat berbeda dari rata-rata? Informasi ini sangat penting karena dapat menunjukkan adanya disparitas yang perlu ditangani secara khusus. Jika variasi data tinggi, maka intervensi harus disesuaikan dengan kondisi lokal masing-masing daerah.

Ukuran penyebaran data seperti rentang, simpangan baku, varians, dan koefisien variasi, memberikan alat yang kuat untuk mengukur dan memahami tingkat variasi tersebut. Rentang menunjukkan jarak antara nilai tertinggi dan terendah, memberikan gambaran kasar tentang seberapa

luas data tersebar. Simpangan baku dan varians, yang merupakan ukuran statistik yang lebih sensitif terhadap distribusi data, membantu kita memahami seberapa besar data menyebar di sekitar rata-rata. Sementara itu, koefisien variasi, yang merupakan rasio simpangan baku terhadap rata-rata, memungkinkan perbandingan variasi antar data yang berbeda skala.

Memahami dan mampu mengaplikasikan ukuran penyebaran data ini sangat penting dalam praktik kesehatan masyarakat. Data yang menunjukkan variasi tinggi mungkin mengindikasikan ketidakmerataan dalam layanan kesehatan, akses, atau faktor sosial ekonomi yang mempengaruhi kesehatan masyarakat. Sebaliknya, data dengan variasi rendah menunjukkan distribusi yang lebih homogen, yang bisa menjadi indikator keberhasilan program kesehatan tertentu.

Selain itu, pengukuran variasi ini juga membantu dalam pengambilan keputusan strategis, seperti menentukan prioritas wilayah yang membutuhkan intervensi lebih intensif atau menilai efektivitas program yang telah dilaksanakan. Dengan demikian, penguasaan konsep dan keterampilan dalam menghitung serta menginterpretasi ukuran penyebaran data menjadi sangat penting bagi para profesional kesehatan masyarakat, peneliti, dan pengambil kebijakan.

Pada bagian ini, peserta didik akan diajak untuk memahami secara mendalam konsep dasar dari berbagai ukuran penyebaran data, belajar menghitungnya secara manual maupun menggunakan software statistik, serta mengaplikasikan hasilnya dalam konteks nyata di bidang kesehatan. Melalui pemahaman ini, diharapkan mereka mampu mengidentifikasi variasi data yang signifikan dan menggunakannya sebagai dasar dalam pengambilan

keputusan yang berbasis bukti. Dengan demikian, penguasaan materi ini akan memperkuat kompetensi mereka dalam analisis data kesehatan dan mendukung upaya peningkatan kualitas layanan serta keberhasilan program kesehatan masyarakat secara umum.

Ukuran penyebaran data merupakan konsep penting dalam statistik yang digunakan untuk mengukur seberapa jauh data tersebar dari pusatnya, biasanya dari nilai rata-rata. Dalam konteks data kesehatan, pengukuran ini sangat vital karena membantu kita memahami tingkat variasi dan heterogenitas dalam populasi yang sedang diteliti. Pemahaman yang mendalam tentang ukuran penyebaran data memungkinkan para profesional kesehatan masyarakat untuk mengidentifikasi disparitas, tren, serta potensi masalah yang memerlukan intervensi khusus. Pada bagian ini, kita akan membahas secara rinci empat ukuran utama penyebaran data: rentang, simpangan baku, varians, dan koefisien variasi, serta bagaimana interpretasi hasilnya dapat digunakan dalam analisis data kesehatan.

5.1 Rentang

Rentang adalah ukuran penyebaran data yang paling sederhana dan langsung, yang menunjukkan jarak antara nilai maksimum dan minimum dalam suatu data set. Rumusnya cukup mudah, yaitu:

$$\text{rentang} = \text{nilai tertinggi} - \text{nilai terendah}$$

Rentang memberikan gambaran kasar tentang seberapa luas data tersebar, namun memiliki keterbatasan karena sangat dipengaruhi oleh nilai ekstrem dan tidak mempertimbangkan distribusi data di antara kedua nilai tersebut. Dalam konteks kesehatan, misalnya, jika kita mengukur tingkat imunisasi di beberapa daerah dan

mendapatkan rentang 50%, ini menunjukkan bahwa ada perbedaan besar antara daerah dengan tingkat imunisasi tertinggi dan terendah.

Contoh aplikasi: Misalnya, data tingkat imunisasi di 10 kabupaten menunjukkan nilai tertinggi 95% dan terendah 45%. Maka, rentang tingkat imunisasi adalah:

$$[95\% - 45\% = 50\%]$$

Ini menunjukkan variasi yang cukup besar antar daerah, yang mungkin mengindikasikan disparitas dalam akses layanan imunisasi. Rentang sangat berguna sebagai indikator awal untuk menilai variasi data, tetapi harus digunakan bersama ukuran lain yang lebih sensitif dan informatif.

5.2 Simpangan Baku dan Varians

Simpangan baku dan varians adalah ukuran penyebaran yang lebih canggih dan sensitif terhadap distribusi data. Keduanya mengukur seberapa besar data menyebar di sekitar nilai rata-rata.

5.2.1 Varians

Varians adalah rata-rata kuadrat dari deviasi setiap data terhadap rata-rata. Rumusnya adalah:

$$\sigma^2 = (1/N) * \sum (x_i - \mu)^2$$

Keterangan:

- σ^2 = varians populasi
- N = jumlah total elemen dalam populasi
- X_i = nilai setiap elemen dalam populasi

- μ = rata-rata populasi
- Σ = jumlah

$$\text{Atau } s^2 = \sum (x_i - \bar{x})^2 / (n - 1)$$

Dimana:

- s^2 adalah varians sampel
- x_i adalah nilai data ke-i,
- \bar{x} adalah rata-rata data,
- n adalah jumlah data.

Varians memberikan gambaran kuantitatif tentang tingkat variasi data, tetapi satuannya adalah kuadrat dari satuan data asli, sehingga kurang intuitif untuk interpretasi langsung.

5.2.2 Simpangan Baku

Simpangan baku adalah akar kuadrat dari varians, sehingga satuannya sama dengan data asli, membuatnya lebih mudah dipahami. Rumusnya adalah:

$$s = \sqrt{s^2}$$

Dalam konteks kesehatan, simpangan baku digunakan untuk menilai seberapa besar variasi tingkat imunisasi di berbagai daerah. Sebagai contoh, jika rata-rata tingkat imunisasi adalah 70% dengan simpangan baku 10%, maka sebagian besar daerah memiliki tingkat imunisasi antara 60% hingga 80%, mengingat distribusi normal.

Contoh aplikasi: Jika data tingkat imunisasi di 15 daerah menunjukkan rata-rata 70% dan simpangan baku 12%,

maka data tersebar cukup luas, dan ada daerah yang memiliki tingkat imunisasi jauh di bawah atau di atas rata-rata. Hal ini penting untuk mengidentifikasi daerah yang membutuhkan perhatian khusus.

5.3 Koefisien Variasi

Koefisien variasi (CV) adalah ukuran relatif dari variasi data yang dinyatakan sebagai persentase dari rata-rata. Rumusnya adalah:

$$CV = \left(\frac{s}{rata - rata} \right) * 100\%$$

Dimana :

- CV adalah koefisien variasi
- s adalah simpangan baku
- rata-rata adalah jumlah keseluruhab data dibagi banyaknya data

CV digunakan untuk membandingkan tingkat variasi antara dua atau lebih data yang memiliki satuan berbeda atau berbeda skala. Dalam bidang kesehatan, CV sangat berguna ketika membandingkan variasi tingkat imunisasi di berbagai wilayah yang memiliki rata-rata berbeda secara signifikan.

Contoh aplikasi: Misalnya, jika tingkat imunisasi di dua provinsi memiliki rata-rata 80% dan 60%, dengan simpangan baku masing-masing 8% dan 6%, maka:

- CV Provinsi A = $(8/80) * 100\% = 10\%$
- CV Provinsi B = $(6/60) * 100\% = 10\%$

Meskipun rata-rata dan simpangan baku berbeda, kedua provinsi memiliki tingkat variasi relatif yang sama. Ini

menunjukkan bahwa tingkat ketidakmerataan dalam imunisasi di kedua wilayah tersebut setara, meskipun nilai absolutnya berbeda.

5.4 Interpretasi dalam Data Kesehatan

Pengukuran ukuran penyebaran data harus dilakukan secara cermat dan kontekstual. Rentang memberikan gambaran kasar dan cepat tentang variasi, tetapi sangat dipengaruhi oleh nilai ekstrem dan tidak mempertimbangkan distribusi data secara lengkap. Sebaliknya, simpangan baku dan varians memberikan gambaran yang lebih lengkap dan sensitif terhadap distribusi data, tetapi memerlukan perhitungan yang lebih kompleks.

Dalam praktiknya, ketiga ukuran ini sering digunakan secara bersamaan untuk mendapatkan gambaran lengkap tentang variasi data. Misalnya, dalam analisis tingkat imunisasi, rentang dapat menunjukkan adanya disparitas ekstrem antar daerah, sementara simpangan baku dan koefisien variasi dapat menunjukkan tingkat heterogenitas yang lebih halus dan konsisten.

Penggunaan ukuran penyebaran data ini juga harus disesuaikan dengan karakteristik data dan tujuan analisis. Jika data memiliki distribusi normal, simpangan baku dan varians sangat informatif. Namun, jika data tidak normal atau mengandung nilai ekstrem, rentang dan koefisien variasi dapat memberikan wawasan tambahan.

5.5 Contoh Aplikasi: Variasi Tingkat Imunisasi di Berbagai Daerah

Sebagai ilustrasi nyata, bayangkan sebuah studi yang meneliti tingkat imunisasi di 10 kabupaten di sebuah provinsi. Data yang diperoleh menunjukkan bahwa tingkat imunisasi berkisar dari 45% hingga 95%. Berikut adalah data lengkapnya: Tingkat Imunisasi (%) A 45 B 50 C 55 D 60 E 65 F 70 G 75 H 80 I 85 J 95

Dari data ini, kita dapat menghitung berbagai ukuran penyebaran:

- **Rentang:** $(95\% - 45\% = 50\%)$
- **Rata-rata:** $\text{Average}\{45 + 50 + 55 + 60 + 65 + 70 + 75 + 80 + 85 + 95\} / \{10\} = 68$
- **Simpangan baku:** Menghitung deviasi setiap data dari rata-rata, kuadratkan, jumlahkan, lalu bagi dengan $(n-1)$, kemudian akar kuadratnya. Hasilnya sekitar 17,5%.

$$s = \sqrt{s^2}$$

$$\begin{aligned} s &= \sqrt{\sum (x_i - x)^2 / (n - 1)} \\ &= \sqrt{\sum (50 - 68)^2 / (10 - 1)} = \sqrt{\sum 324 / 9} = \sqrt{\sum 36} = 6 \end{aligned}$$

Dimana: x_i adalah nilai data ke-i. x adalah rata-rata data. n adalah jumlah data

- **Varians:** Sekitar 36 (hasil dari perhitungan kuadrat deviasi rata-rata).

$$\text{Atau } s^2 = \sum (x_i - \bar{x})^2 / (n - 1) = 36$$

Dimana: s^2 adalah varians sampel. x_i adalah nilai data ke-i. \bar{x} adalah rata-rata data. n adalah jumlah data.

- **Koefisien variasi:** $(17,5/68) \times 100\% \approx 25,7\%$

$$CV = \left(\frac{s}{\text{rata-rata}} \right) * 100\% = \left(\frac{6}{68} \right) * 100\% = 8,823\%$$

Interpretasi dari hasil ini menunjukkan bahwa meskipun rata-rata tingkat imunisasi cukup tinggi, terdapat variasi yang cukup besar di antara daerah, dengan tingkat imunisasi yang sangat rendah di beberapa kabupaten dan sangat tinggi di kabupaten lain. Tingkat variasi ini menandakan adanya disparitas yang perlu diatasi melalui kebijakan yang lebih terfokus dan penyesuaian program imunisasi di daerah yang tertinggal.

5.6 Pentingnya Ukuran Penyebaran Data dalam Pengambilan Keputusan

Pengukuran variasi data tidak hanya penting secara statistik, tetapi juga memiliki implikasi praktis dalam pengambilan keputusan di bidang kesehatan masyarakat. Data dengan variasi tinggi menunjukkan adanya disparitas yang signifikan, yang harus diidentifikasi dan ditangani secara spesifik. Sebaliknya, data dengan variasi rendah menunjukkan distribusi yang relatif homogen, yang bisa menjadi indikator keberhasilan program kesehatan tertentu.

Misalnya, dalam program imunisasi, jika data menunjukkan variasi tinggi antar daerah, maka intervensi harus diarahkan secara spesifik ke daerah yang memiliki tingkat imunisasi rendah. Jika variasi rendah, maka program dapat difokuskan

secara umum di seluruh wilayah. Oleh karena itu, pengukuran penyebaran data menjadi alat penting dalam merancang strategi intervensi yang efektif dan efisien.

Selain itu, pengukuran ini juga membantu dalam memonitor dan mengevaluasi keberhasilan program kesehatan dari waktu ke waktu. Jika setelah intervensi, variasi data menurun, ini menunjukkan bahwa disparitas telah berkurang dan distribusi layanan menjadi lebih merata.

5.7 Keterbatasan dan Kelebihan Ukuran Penyebaran Data

Setiap ukuran penyebaran data memiliki kelebihan dan keterbatasan. Rentang sangat mudah dihitung dan memberikan gambaran cepat, tetapi sangat dipengaruhi oleh nilai ekstrem dan tidak mempertimbangkan distribusi data secara lengkap. Varians dan simpangan baku memberikan gambaran yang lebih lengkap dan sensitif terhadap distribusi, tetapi memerlukan perhitungan yang lebih kompleks dan asumsi distribusi data, terutama normalitas.

Koefisien variasi memiliki keunggulan dalam membandingkan variasi antar data yang berbeda skala dan satuan, tetapi hanya berlaku jika data memiliki nilai rata-rata yang tidak mendekati nol, karena dapat menyebabkan nilai CV yang tidak bermakna.

Dalam praktiknya, kombinasi dari berbagai ukuran penyebaran data ini akan memberikan gambaran yang paling lengkap dan akurat tentang distribusi data kesehatan yang sedang dianalisis. Pemilihan ukuran yang tepat harus disesuaikan dengan karakteristik data dan tujuan analisis.

Rangkuman

Ukuran penyebaran data merupakan aspek penting dalam statistik yang membantu kita memahami seberapa jauh data tersebar dari pusatnya, biasanya dari nilai rata-rata. Dalam konteks kesehatan masyarakat, pengukuran ini sangat penting karena memberikan gambaran tentang tingkat variasi dan heterogenitas dalam populasi yang sedang diteliti. Pemahaman yang baik tentang ukuran penyebaran data memungkinkan profesional kesehatan untuk mengidentifikasi disparitas, tren, serta potensi masalah yang memerlukan intervensi khusus. Berikut adalah poin-poin utama yang merangkum pembahasan tentang ukuran penyebaran data:

1. **Rentang** adalah ukuran paling sederhana yang menunjukkan jarak antara nilai maksimum dan minimum dalam data set. Meskipun mudah dihitung dan memberikan gambaran kasar tentang variasi, rentang sangat dipengaruhi oleh nilai ekstrem dan tidak mempertimbangkan distribusi data di antara kedua nilai tersebut. Contohnya, dalam pengukuran tingkat imunisasi, rentang yang besar menunjukkan adanya disparitas besar antar daerah.
2. **Simpangan baku dan varians** adalah ukuran yang lebih sensitif dan canggih. Varians mengukur rata-rata kuadrat deviasi data terhadap rata-rata, tetapi satuannya kuadrat dari satuan data asli, sehingga kurang intuitif. Simpangan baku adalah akar kuadrat dari varians, sehingga satuannya sama dengan data asli dan lebih mudah dipahami. Dalam analisis kesehatan, keduanya digunakan untuk menilai tingkat variasi data, seperti variasi tingkat imunisasi di berbagai wilayah.
3. **Koefisien variasi** adalah ukuran relatif dari variasi yang dinyatakan sebagai persentase dari rata-rata. CV

sangat berguna untuk membandingkan tingkat variasi antar data dengan skala berbeda, misalnya membandingkan disparitas imunisasi di berbagai provinsi dengan rata-rata berbeda. Koefisien variasi memungkinkan kita melihat tingkat heterogenitas secara proporsional.

4. **Interpretasi data penyebaran** harus dilakukan secara kontekstual. Rentang memberikan gambaran kasar dan cepat, tetapi sangat dipengaruhi nilai ekstrem. Simpangan baku dan varians memberikan gambaran lebih lengkap dan sensitif terhadap distribusi data. Penggunaan ketiga ukuran ini secara bersamaan membantu mendapatkan gambaran lengkap tentang variasi data.
5. **Contoh aplikasi nyata** menunjukkan bahwa variasi tingkat imunisasi di beberapa daerah bisa sangat besar, yang tercermin dari rentang dan simpangan baku yang tinggi. Variasi ini menandakan disparitas yang perlu diatasi melalui kebijakan yang tepat. Pengukuran ini juga penting dalam memantau keberhasilan program kesehatan dari waktu ke waktu.
6. **Pentingnya pengukuran variasi** dalam pengambilan keputusan adalah bahwa data dengan variasi tinggi menunjukkan disparitas yang signifikan dan membutuhkan intervensi spesifik, sementara data dengan variasi rendah menunjukkan distribusi yang lebih homogen. Pengukuran ini membantu merancang strategi intervensi yang efektif dan efisien serta memonitor keberhasilan program.
7. **Keterbatasan dan kelebihan** dari setiap ukuran harus dipahami. Rentang mudah dihitung tetapi kurang lengkap, sedangkan varians dan simpangan baku lebih sensitif dan informatif tetapi memerlukan perhitungan lebih kompleks. Koefisien variasi cocok untuk membandingkan variasi antar data berbeda

skala, tetapi harus digunakan dengan hati-hati jika data mendekati nol.

Dengan memahami dan mampu mengaplikasikan ukuran penyebaran data secara tepat, mahasiswa dapat melakukan analisis data kesehatan yang lebih akurat dan menyampaikan hasilnya secara jelas kepada pemangku kepentingan. Penguasaan konsep ini menjadi dasar penting dalam pengembangan kompetensi analisis statistik di bidang kesehatan masyarakat.

Latihan Mahasiswa

Soal Essay

1. Jelaskan pengertian ukuran penyebaran data dan mengapa pengukuran ini penting dalam analisis data kesehatan masyarakat!
2. Sebutkan dan jelaskan empat ukuran utama penyebaran data yang dibahas dalam modul ini, serta berikan satu contoh penggunaannya dalam konteks kesehatan.
3. Bagaimana cara menghitung simpangan baku dari data tingkat imunisasi di sebuah daerah? Jelaskan langkah-langkahnya secara singkat!
4. Mengapa koefisien variasi dianggap sebagai ukuran yang berguna untuk membandingkan variasi data dari dua populasi yang berbeda skala? Berikan penjelasan singkat!
5. Berdasarkan data berikut: tingkat imunisasi di 5 kabupaten adalah 60%, 65%, 70%, 75%, dan 80%. Hitung rentang data tersebut dan interpretasikan hasilnya!

Soal Pilihan Berganda

1. Manakah dari berikut ini yang merupakan ukuran penyebaran data yang paling sederhana dan langsung?
 - a. Varians
 - b. Rentang
 - c. Simpangan baku
 - d. Koefisien variasi

2. Koefisien variasi biasanya digunakan untuk:
 - a. Mengukur variasi absolut data
 - b. Membandingkan variasi antar data dengan skala berbeda
 - c. Menghitung rata-rata data
 - d. Menentukan nilai maksimum dan minimum data

3. Jika data tingkat imunisasi memiliki simpangan baku 8% dan rata-rata 80%, berapa koefisien variasinya?
 - a. 10%
 - b. 8%
 - c. 12%
 - d. 20%

4. Data berikut menunjukkan tingkat imunisasi di 4 daerah: 55%, 60%, 65%, 70%. Hitung rentang data tersebut!
 - a. 15%
 - b. 10%
 - c. 20%
 - d. 25%

5. Dalam konteks analisis data kesehatan, simpangan baku digunakan untuk:
 - a. Mengukur jarak antara nilai maksimum dan minimum
 - b. Mengukur seberapa besar data menyebar di sekitar rata-rata
 - c. Menghitung nilai tengah data
 - d. Membandingkan data dari dua populasi berbeda
6. Manakah dari berikut ini yang termasuk ke dalam ukuran penyebaran data yang sensitif terhadap nilai ekstrem?
 - a. Rentang
 - b. Varians
 - c. Simpangan baku
 - d. Koefisien variasi
7. Jika sebuah data memiliki rentang sebesar 50 dan rata-rata 100, manakah pernyataan berikut yang benar?
 - a. Data tersebar sangat kecil
 - b. Data tersebar cukup besar
 - c. Data memiliki variasi yang sangat kecil
 - d. Data tidak memiliki variasi
8. Dalam pengolahan data kesehatan, mengapa penting menggunakan ukuran penyebaran selain dari rata-rata?
 - a. Untuk mengetahui nilai tengah data
 - b. Untuk memahami tingkat variasi dan heterogenitas data

- c. Untuk menghitung jumlah data
 - d. Untuk menentukan nilai maksimum dan minimum
9. Data tingkat imunisasi di 8 daerah menunjukkan simpangan baku 5% dan rata-rata 75%. Berapa nilai koefisien variasinya?
- a. 6,67%
 - b. 5%
 - c. 7,5%
 - d. 10%

Soal Studi Kasus / Project

1. Sebuah studi meneliti tingkat imunisasi di 12 kabupaten dengan data sebagai berikut: 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, dan 100%. Hitung rentang, rata-rata, simpangan baku, dan koefisien variasi dari data tersebut. Jelaskan apa arti dari hasil pengukuran tersebut dalam konteks disparitas imunisasi di daerah tersebut!
2. Anda diminta untuk melakukan analisis variasi tingkat stunting pada anak-anak di 5 provinsi berbeda. Data yang diperoleh adalah sebagai berikut: Provinsi A = 20%, Provinsi B = 25%, Provinsi C = 30%, Provinsi D = 35%, Provinsi E = 40%. Buatlah laporan singkat yang menjelaskan tingkat variasi data tersebut dan rekomendasi kebijakan berdasarkan hasil analisis Anda!

REFERENSI

1. Daniel, W. W. (2018). *Biostatistics: A Foundation for Analysis in the Health Sciences* (11th ed.). John Wiley & Sons.
2. Rosner, B. (2015). *Fundamentals of Biostatistics* (8th ed.). Cengage Learning.
3. World Health Organization. (2019). *Immunization coverage: fact sheet*. WHO.
<https://www.who.int/news-room/fact-sheets/detail/immunization-coverage>
4. Kementerian Kesehatan Republik Indonesia. (2020). *Laporan Surveilans Imunisasi*. Jakarta: Kemenkes RI."

BAB 6

PROBABILITAS DASAR

Tujuan Pembelajaran

1. Mahasiswa mampu memahami konsep dasar peluang dan pentingnya dalam konteks biostatistik kesehatan.
2. Mahasiswa dapat menjelaskan teorema peluang dan aturan dasar yang menjadi dasar dalam perhitungan probabilitas.
3. Mahasiswa mampu menerapkan konsep peluang dalam analisis data epidemiologi dan pengambilan keputusan kesehatan masyarakat.
4. Mahasiswa dapat mengidentifikasi dan menghitung probabilitas kejadian tertentu dalam populasi berdasarkan data yang tersedia.
5. Mahasiswa mampu mengaplikasikan probabilitas dalam memperkirakan risiko kejadian penyakit tertentu, sehingga dapat mendukung perencanaan program kesehatan.
6. Mahasiswa memahami hubungan antara probabilitas dan pengambilan keputusan klinis maupun kebijakan kesehatan berbasis bukti.
7. Mahasiswa mampu menggunakan pendekatan probabilitas untuk menginterpretasi hasil penelitian epidemiologi dan surveilans kesehatan secara kritis dan akurat.

Pendahuluan

Probabilitas merupakan salah satu konsep fundamental dalam biostatistik yang sangat penting untuk dipahami oleh mahasiswa kesehatan masyarakat. Dalam dunia kesehatan, pengambilan keputusan sering kali harus didasarkan pada data yang tidak pasti dan variabel yang acak. Oleh karena itu, pemahaman tentang peluang dan probabilitas menjadi kunci untuk menginterpretasi data secara tepat dan membuat prediksi yang akurat mengenai kejadian kesehatan tertentu.

Pada bagian ini, kita akan membahas konsep dasar peluang yang menjadi fondasi dalam memahami probabilitas. Konsep ini membantu kita untuk mengukur kemungkinan terjadinya suatu peristiwa dalam populasi, baik itu kejadian penyakit, keberhasilan pengobatan, maupun faktor risiko yang mempengaruhi kesehatan masyarakat. Dengan memahami peluang, kita dapat mengkuantifikasi risiko dan membuat estimasi yang lebih akurat dalam berbagai konteks epidemiologi dan promosi kesehatan.

Selain itu, kita akan membahas teorema peluang dan aturan dasar yang menjadi alat utama dalam perhitungan probabilitas. Pengetahuan ini sangat penting agar mahasiswa mampu melakukan analisis statistik yang mendalam dan akurat, serta mampu mengaplikasikan konsep probabilitas dalam berbagai studi kasus di bidang kesehatan. Misalnya, dalam memperkirakan risiko kejadian penyakit tertentu berdasarkan data historis, probabilitas membantu kita untuk membuat prediksi yang berbasis bukti dan mendukung pengambilan keputusan yang tepat.

Urgensi mempelajari probabilitas dalam kesehatan masyarakat tidak hanya terletak pada aspek teoritis, tetapi juga pada penerapannya dalam praktik nyata. Dalam epidemiologi, probabilitas digunakan untuk memperkirakan

kemungkinan terjadinya kejadian tertentu dalam populasi, sehingga dapat membantu dalam perencanaan program pencegahan dan pengendalian penyakit. Sebagai contoh, dengan menghitung probabilitas seseorang terkena penyakit tertentu, tenaga kesehatan dapat mengidentifikasi kelompok risiko tinggi dan mengarahkan sumber daya secara lebih efisien.

Selain itu, probabilitas juga berperan dalam menilai efektivitas intervensi kesehatan dan dalam pengembangan model prediksi yang dapat digunakan untuk mengantisipasi tren kejadian penyakit di masa depan. Dengan memahami konsep peluang dan teorema peluang, mahasiswa akan mampu melakukan analisis yang lebih kritis terhadap data epidemiologi dan surveilans kesehatan, serta mampu menyusun strategi intervensi yang berbasis bukti.

Secara keseluruhan, bagian ini akan membekali mahasiswa dengan pengetahuan dasar yang esensial untuk memahami dan menerapkan probabilitas dalam konteks kesehatan masyarakat. Pemahaman ini akan menjadi fondasi yang kuat untuk mempelajari konsep statistik lanjutan dan teknik analisis data yang lebih kompleks di bab-bab berikutnya. Dengan demikian, penguasaan materi ini sangat penting agar mahasiswa mampu menjadi tenaga kesehatan yang kompeten dalam pengambilan keputusan berbasis data dan bukti ilmiah.

6.1 Konsep Peluang

6.1.1 Definisi Peluang

Peluang merupakan konsep dasar dalam statistik yang digunakan untuk mengukur kemungkinan terjadinya suatu peristiwa tertentu dalam suatu populasi atau sampel. Secara umum, peluang dapat diartikan sebagai rasio antara jumlah kejadian yang diharapkan terjadi dan jumlah seluruh kemungkinan kejadian yang mungkin terjadi. Dalam konteks biostatistik dan epidemiologi, peluang sangat penting karena memungkinkan kita untuk mengkuantifikasi risiko dan membuat prediksi tentang kejadian kesehatan tertentu.

Menurut Moore dan McCabe (2017), peluang dari suatu peristiwa A, yang dilambangkan sebagai $P(A)$, adalah angka antara 0 dan 1, di mana 0 menunjukkan bahwa peristiwa tersebut tidak mungkin terjadi, dan 1 menunjukkan bahwa peristiwa tersebut pasti terjadi. Sebagai contoh, peluang seseorang terkena influenza dalam satu musim tertentu dapat dihitung berdasarkan data historis dan faktor risiko yang ada. Jika dari data diketahui bahwa 200 dari 1000 orang dalam suatu populasi terkena influenza, maka peluang terkena influenza adalah 0,2 atau 20%.

6.1.2 Karakteristik Peluang

Peluang memiliki beberapa karakteristik utama yang harus dipahami oleh mahasiswa. Pertama, peluang selalu berada dalam rentang 0 sampai 1. Kedua, peluang dari seluruh kemungkinan kejadian dalam ruang sampel harus menjumlahkan menjadi 1. Ketiga, peluang dari kejadian yang saling eksklusif (tidak dapat terjadi bersamaan) dapat dijumlahkan untuk mendapatkan peluang gabungan dari kejadian tersebut.

Sebagai ilustrasi, jika kita melempar sebuah dadu, peluang mendapatkan angka 3 adalah $1/6$, dan peluang mendapatkan angka selain 3 adalah $5/6$. Jika kita ingin mengetahui peluang mendapatkan angka genap (2, 4, 6), maka peluangnya adalah jumlah dari peluang masing-masing angka, yaitu $3/6$ atau $1/2$.

6.1.3 Pentingnya Konsep Peluang dalam Kesehatan Masyarakat

Dalam bidang kesehatan masyarakat, konsep peluang digunakan untuk mengukur risiko kejadian penyakit, keberhasilan intervensi, dan faktor risiko yang mempengaruhi kesehatan populasi. Dengan mengetahui peluang terjadinya suatu peristiwa, tenaga kesehatan dan peneliti dapat merancang strategi pencegahan dan pengendalian yang lebih efektif.

Misalnya, dalam program imunisasi, peluang seseorang terkena penyakit tertentu tanpa imunisasi dapat dihitung dan dibandingkan dengan peluang setelah imunisasi. Jika peluang terkena penyakit menurun secara signifikan setelah imunisasi, maka program tersebut terbukti efektif dan dapat didukung secara ilmiah. Selain itu, peluang juga digunakan dalam penilaian risiko, seperti memperkirakan kemungkinan seseorang terkena kanker paru berdasarkan faktor risiko seperti merokok dan paparan polusi udara.

6.2 Teorema Peluang dan Aturan Dasar

6.2.1 Teorema Peluang

Teorema peluang adalah aturan dasar yang digunakan untuk menghitung peluang dari gabungan kejadian yang kompleks, berdasarkan peluang kejadian sederhana. Salah satu teorema utama adalah teorema penjumlahan peluang, yang

digunakan ketika kejadian saling eksklusif, dan teorema perkalian peluang, yang digunakan untuk kejadian yang saling bergantung atau tidak saling eksklusif.

6.2.2 Aturan Penjumlahan Peluang

Aturan penjumlahan peluang digunakan untuk menghitung peluang dari kejadian A atau B terjadi, yang dilambangkan sebagai $P(A \cup B)$. Jika kejadian A dan B saling eksklusif (tidak dapat terjadi bersamaan), maka peluang gabungan keduanya adalah jumlah dari peluang masing-masing kejadian:

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

Contoh: Jika peluang seseorang terkena penyakit A adalah 0,1 dan penyakit B adalah 0,2, dan kedua penyakit ini tidak dapat terjadi bersamaan, maka peluang seseorang terkena salah satu dari kedua penyakit tersebut adalah $0,1 + 0,2 = 0,3$.

6.2.3 Aturan Perkalian Peluang

Aturan perkalian digunakan untuk menghitung peluang dua kejadian A dan B terjadi secara bersamaan, yang dilambangkan sebagai $P(A \cap B)$. Jika kejadian A dan B saling bergantung, maka peluang keduanya terjadi adalah hasil perkalian dari peluang kejadian A dan peluang B, dengan memperhatikan kondisi bahwa B terjadi setelah A:

$$P(A \text{ dan } B) = P(A) * P(B | A)$$

Di mana $P(B|A)$ adalah peluang B terjadi dengan syarat A sudah terjadi. Jika kejadian A dan B independen (tidak saling mempengaruhi), maka:

$$P(A \text{ dan } B) = P(A) * P(B)$$

Contoh: Jika peluang seseorang terkena influenza adalah 0,2 dan peluang terkena pneumonia adalah 0,1, dan kedua kejadian ini independen, maka peluang keduanya terjadi bersamaan adalah $0,2 \times 0,1 = 0,02$ atau 2%.

6.2.4 Prinsip Komplemen

Prinsip ini menyatakan bahwa peluang dari kejadian tidak terjadi (komplemen dari kejadian A) adalah:

$$P(A') = 1 - P(A)$$

Di mana $P(A)$ adalah peluang kejadian A terjadi. $P(A')$ adalah peluang kejadian A tidak terjadi.

Contoh: Jika peluang terkena penyakit tertentu adalah 0,15, maka peluang tidak terkena penyakit tersebut adalah $1 - 0,15 = 0,85$.

6.2.5 Aplikasi Teorema dan Aturan Dasar dalam Epidemiologi

Dalam epidemiologi, teorema peluang dan aturan dasar ini digunakan untuk menghitung risiko gabungan, peluang kejadian tunggal, dan peluang kejadian tidak terjadi. Sebagai contoh, dalam studi tentang faktor risiko, kita dapat menghitung peluang seseorang terkena penyakit tertentu jika memiliki faktor risiko tertentu dan membandingkannya dengan peluang tanpa faktor risiko tersebut.

6.3 Aplikasi Probabilitas dalam Epidemiologi

6.3.1 Peran Probabilitas dalam Menilai Risiko Penyakit

Probabilitas adalah alat utama dalam menilai risiko kejadian penyakit dalam populasi. Dengan menghitung peluang seseorang terkena penyakit tertentu, para epidemiolog dapat mengidentifikasi kelompok berisiko tinggi dan merancang intervensi yang lebih tepat sasaran. Misalnya, dalam studi tentang faktor risiko diabetes, probabilitas terkena diabetes pada individu dengan obesitas dapat dibandingkan dengan mereka yang memiliki berat badan normal.

6.3.2 Penggunaan Probabilitas dalam Studi Kasus dan Kontrol

Dalam studi kasus dan kontrol, probabilitas digunakan untuk menghitung odds ratio dan risiko relatif, yang merupakan ukuran asosiasi antara faktor risiko dan kejadian penyakit. Misalnya, jika dari data diketahui bahwa 30% dari penderita diabetes memiliki riwayat merokok, sementara hanya 10% dari non-penderita diabetes yang merokok, maka probabilitas dan odds ratio dapat digunakan untuk menilai kekuatan hubungan tersebut.

6.3.3 Probabilitas dalam Perhitungan Risiko Individu dan Populasi

Selain itu, probabilitas digunakan untuk memperkirakan risiko individu dan populasi terhadap kejadian tertentu. Dalam konteks vaksinasi, probabilitas seseorang terkena penyakit tertentu tanpa vaksinasi dapat dihitung dan

dibandingkan dengan probabilitas setelah vaksinasi, sehingga dapat menilai efektivitas program imunisasi.

6.3.4 Model Prediksi dan Pengambilan Keputusan

Probabilitas juga menjadi dasar dalam pengembangan model prediksi yang membantu pengambilan keputusan klinis dan kebijakan kesehatan. Contohnya adalah model prediksi risiko kanker payudara berdasarkan faktor genetik dan lingkungan, yang menggunakan probabilitas untuk memperkirakan kemungkinan seseorang terkena kanker dalam jangka waktu tertentu.

6.4 Contoh Aplikasi: Perkiraan Risiko Kejadian Penyakit Tertentu

6.4.1 Kasus Studi: Perkiraan Risiko Kanker Paru

Sebagai contoh, kita akan membahas bagaimana probabilitas digunakan untuk memperkirakan risiko kejadian kanker paru dalam suatu populasi. Misalnya, berdasarkan data surveilans nasional, diketahui bahwa dari 10.000 perokok aktif, 150 orang didiagnosis dengan kanker paru dalam satu tahun. Sementara itu, dari 10.000 non-perokok, hanya 10 orang yang didiagnosis kanker paru.

Dari data ini, kita dapat menghitung peluang terkena kanker paru pada perokok dan non-perokok:

- Peluang terkena kanker paru pada perokok:

$$P(A) = \frac{n(A)}{n(S)}$$

$$P(\text{n kanker} \mid \text{n perokok}) = \frac{150}{10.000} * 100\% = 1,5\%$$

- Peluang terkena kanker paru pada non-perokok:

$$P(\text{n kanker}) | \text{n non-perokok}) = n \frac{10}{10.000} * 100\% = 0,1\%$$

Dengan demikian, perokok memiliki peluang 1,5 kali lebih besar untuk terkena kanker paru dibandingkan non-perokok. Ini menunjukkan hubungan yang kuat antara merokok dan risiko kanker paru, dan hasil ini dapat digunakan untuk memperkuat kampanye anti-merokok dan kebijakan pengendalian tembakau.

6.4.2 Perhitungan Risiko Absolut dan Risiko Relatif

Selain peluang, kita juga dapat menghitung risiko relatif (RR) dan risiko absolut (AR). Risiko relatif adalah perbandingan peluang kejadian antara dua kelompok:

$$RR = \left[\frac{a}{a+b} \right] / \left[\frac{c}{c+d} \right]$$

Dimana :

- a adalah jumlah kasus dalam kelompok terpapar
- b adalah jumlah non kasus (sehat) dalam kelompok terpapar
- c adalah jumlah kasus dalam kelompok tidak terpapar
- d adalah jumlah non kasus (sehat) dalam kelompok tidak terpapar

Penyelesaian kasus:

$$RR = \left[\frac{150}{150 + 9850} \right] / \left[\frac{10}{10 + 9990} \right] = 15$$

Ini berarti perokok memiliki risiko 15 kali lebih tinggi untuk terkena kanker paru dibandingkan non-perokok. Risiko absolut pada 2 kelompok adalah selisih antara peluang kejadian pada kedua kelompok:

$$AR = \frac{\text{Jumlah kejadian dalam kelompok}}{\text{total populasi dalam kelompok}}$$

Jika pada 2 kelompok maka:

$$\begin{aligned} AR &= \frac{P \text{ kanker pada perokok}}{\text{total populasi perokok}} - \frac{P \text{ kanker pada non perokok}}{\text{total populasi non perokok}} \\ &= \left(\frac{150}{10.000} * 100\% \right) - \left(\frac{10}{10.000} * 100\% \right) \\ &= 1,5\% - 0,1\% = 1,4\% \end{aligned}$$

Hasil ini menunjukkan bahwa terpapar factor risiko dapat meningkatkan risiko terkena kanker paru sebesar 1,4%.

6.4.3 Implikasi dalam Pengambilan Kebijakan Kesehatan

Hasil analisis probabilitas ini sangat penting dalam pengambilan kebijakan kesehatan. Dengan mengetahui bahwa perokok memiliki risiko 15 kali lebih tinggi terkena kanker paru, pemerintah dan lembaga kesehatan dapat merancang program edukasi, pengendalian tembakau, dan kebijakan larangan merokok di tempat umum. Selain itu, data ini juga mendukung pengembangan program deteksi dini dan skrining kanker paru bagi kelompok berisiko tinggi.

6.4.4 Keterbatasan dan Asumsi dalam Perhitungan Probabilitas

Perlu diingat bahwa perhitungan probabilitas dalam studi epidemiologi memiliki keterbatasan. Data yang digunakan harus representatif dan akurat, serta asumsi bahwa populasi yang dianalisis homogen dan kejadian bersifat independen. Faktor lain seperti faktor genetik, lingkungan, dan gaya hidup juga mempengaruhi risiko dan harus dipertimbangkan dalam interpretasi hasil.

Rangkuman

Pendahuluan ini memberikan gambaran tentang pentingnya konsep peluang dalam biostatistik dan epidemiologi, serta bagaimana penggunaannya membantu dalam pengambilan keputusan di bidang kesehatan masyarakat. Berikut adalah poin-poin utama yang merangkum pembahasan tersebut:

1. Peluang adalah konsep dasar yang mengukur kemungkinan terjadinya suatu peristiwa dalam populasi atau sampel, dinyatakan sebagai angka antara 0 dan 1. Semakin tinggi peluang, semakin besar kemungkinan kejadian tersebut terjadi. Dalam konteks kesehatan, peluang digunakan untuk mengkuantifikasi risiko dan memprediksi kejadian penyakit.
2. Karakteristik utama peluang meliputi rentang nilai dari 0 sampai 1, jumlah peluang seluruh kejadian dalam ruang sampel harus sama dengan 1, dan peluang kejadian saling eksklusif dapat dijumlahkan. Contohnya, peluang mendapatkan angka tertentu saat melempar dadu atau peluang mendapatkan angka genap dari hasil lemparan dadu.

3. Konsep peluang sangat penting dalam kesehatan masyarakat karena membantu menilai risiko kejadian penyakit, efektivitas intervensi, dan faktor risiko. Misalnya, peluang terkena penyakit tanpa imunisasi dapat dibandingkan dengan peluang setelah imunisasi untuk menilai keberhasilan program vaksinasi.
4. Teorema peluang dan aturan dasar seperti penjumlahan dan perkalian peluang digunakan untuk menghitung peluang gabungan kejadian yang kompleks. Prinsip komplemen juga digunakan untuk menghitung peluang kejadian tidak terjadi, yang penting dalam analisis risiko.
5. Dalam epidemiologi, probabilitas digunakan untuk menilai risiko penyakit, menghitung odds ratio dan risiko relatif dalam studi kasus dan kontrol, serta memperkirakan risiko individu dan populasi. Hal ini mendukung pengembangan strategi pencegahan dan kebijakan kesehatan yang berbasis bukti.
6. Contoh aplikasi nyata adalah perhitungan risiko kanker paru berdasarkan data surveilans, di mana peluang terkena kanker pada perokok jauh lebih tinggi dibandingkan non-perokok. Analisis ini menghasilkan risiko relatif dan risiko absolut yang membantu dalam pengambilan kebijakan dan kampanye kesehatan.
7. Perhitungan probabilitas harus memperhatikan keterbatasan data dan asumsi, seperti keakuratan data, homogenitas populasi, dan independensi kejadian. Faktor lain seperti faktor genetik dan lingkungan juga mempengaruhi hasil analisis.
8. Secara keseluruhan, konsep peluang dan teorema dasar dalam statistik memberikan dasar yang kuat untuk memahami risiko, membuat prediksi, dan merancang intervensi yang efektif dalam bidang kesehatan masyarakat. Penggunaan data probabilitas

mendukung pengambilan keputusan yang lebih tepat dan berbasis bukti.

Latihan Mahasiswa

Soal Essay

1. Jelaskan pengertian peluang dalam konteks biostatistik kesehatan dan berikan satu contoh penggunaannya dalam studi epidemiologi.
2. Sebutkan dan jelaskan dua karakteristik utama dari peluang yang harus dipahami oleh mahasiswa kesehatan masyarakat.
3. Bagaimana konsep peluang digunakan dalam menilai risiko kejadian penyakit tertentu dalam populasi? Berikan contoh nyata dari bidang kesehatan masyarakat.
4. Jelaskan perbedaan antara aturan penjumlahan peluang dan aturan perkalian peluang beserta contoh penggunaannya dalam analisis data kesehatan.
5. Berdasarkan data surveilans nasional, diketahui bahwa dari 5000 orang perokok, 75 orang didiagnosis kanker paru dalam satu tahun. Hitung peluang terkena kanker paru pada perokok tersebut dan interpretasikan hasilnya.

Soal Pilihan Berganda

1. Peluang dari suatu peristiwa A dilambangkan sebagai:
 - a. $P(A)$
 - b. $Q(A)$
 - c. $R(A)$
 - d. $S(A)$

2. Jika peluang seseorang terkena influenza dalam satu musim adalah 0,2, maka peluang tidak terkena influenza adalah:
 - a. 0,8
 - b. 0,2
 - c. 1,0
 - d. 0,5
3. Dalam konteks peluang, peluang dari seluruh kemungkinan kejadian dalam ruang sampel harus:
 - a. Menjumlahkan menjadi 0
 - b. Menjumlahkan menjadi 1
 - c. Lebih dari 1
 - d. Tidak tentu
4. Jika dua kejadian A dan B saling eksklusif, maka peluang keduanya terjadi secara bersamaan adalah:
 - a. $P(A) + P(B)$
 - b. $P(A) \times P(B)$
 - c. 0
 - d. $P(A) - P(B)$
5. Peluang gabungan dari dua kejadian A dan B yang saling bergantung adalah:
 - a. $P(A) + P(B)$
 - b. $P(A) \times P(B|A)$
 - c. $P(A) \div P(B)$
 - d. $P(A) - P(B)$
6. Prinsip komplemen menyatakan bahwa peluang kejadian tidak terjadi adalah:

- a. $P(A) + P(A^c)$
 - b. $1 - P(A)$
 - c. $P(A) \times P(A^c)$
 - d. $P(A) \div 2$
7. Dalam studi epidemiologi, probabilitas digunakan untuk:
 - a. Menghitung jumlah populasi
 - b. Menilai risiko kejadian penyakit
 - c. Mengukur tingkat pendidikan
 - d. Menentukan lokasi geografis
8. Jika dari data diketahui bahwa 30% penderita diabetes memiliki riwayat merokok, dan 10% non-penderita diabetes merokok, maka peluang seseorang yang merokok dan menderita diabetes adalah:
 - a. 0,3
 - b. 0,1
 - c. Tidak dapat dihitung dari data ini
 - d. 0,03
9. Risiko relatif (RR) dihitung dengan rumus:
 - a. $P(\text{kanker} \mid \text{perokok}) / P(\text{kanker} \mid \text{non-perokok})$
 - b. $P(\text{kanker} \mid \text{non-perokok}) / P(\text{kanker} \mid \text{perokok})$
 - c. $P(\text{kanker} \mid \text{perokok}) + P(\text{kanker} \mid \text{non-perokok})$
 - d. $P(\text{kanker} \mid \text{perokok}) - P(\text{kanker} \mid \text{non-perokok})$
10. Dalam konteks probabilitas, kejadian yang saling independen berarti:

- a. $P(A \cap B) = P(A) + P(B)$
- b. $P(A \cap B) = P(A) \times P(B)$
- c. $P(A \cap B) = 0$
- d. $P(A \cap B) = 1$

Soal Studi Kasus / Project

1. Sebuah studi epidemiologi dilakukan untuk menilai hubungan antara paparan polusi udara dan kejadian asma pada anak-anak di suatu kota. Data menunjukkan bahwa dari 200 anak yang terpapar polusi tinggi, 40 di antaranya mengalami asma. Sedangkan dari 300 anak yang tidak terpapar, 30 di antaranya mengalami asma. Hitung peluang anak yang terpapar polusi tinggi mengalami asma dan bandingkan dengan peluang anak yang tidak terpapar. Interpretasikan hasilnya dan jelaskan bagaimana data ini dapat digunakan untuk mendukung kebijakan pengendalian polusi udara.
2. Sebuah lembaga kesehatan ingin memperkirakan risiko kejadian hepatitis B di sebuah daerah tertentu. Data menunjukkan bahwa dari 8.000 orang yang tidak divaksinasi, 120 orang terkena hepatitis B dalam satu tahun. Sedangkan dari 10.000 orang yang divaksinasi, 50 orang terkena hepatitis B. Hitung peluang terkena hepatitis B pada kedua kelompok dan analisis risiko relatifnya. Jelaskan implikasi hasil ini dalam pengembangan program imunisasi di daerah tersebut.

REFERENSI

1. Moore, D. S., & McCabe, G. P. (2017). *Introduction to the Practice of Statistics* (9th ed.). W.H. Freeman and Company.
2. Rothman, K. J., Greenland, S., & Lash, T. L. (2018). *Modern Epidemiology* (4th ed.). Wolters Kluwer.
3. Last, J. M. (2019). *A Dictionary of Epidemiology* (6th ed.). Oxford University Press.
4. Hastings, R. P., & Krawczyk, N. (2020). Probabilistic models in epidemiology: Applications and challenges. *Journal of Epidemiology & Community Health*, 74(3), 245-251." "

BAB 7

DISTRIBUSI PROBABILITAS

Tujuan Pembelajaran

1. Memahami konsep dasar distribusi probabilitas dan peranannya dalam analisis data kesehatan.
2. Mengidentifikasi dan menjelaskan berbagai jenis distribusi probabilitas yang umum digunakan dalam biostatistik, khususnya distribusi binomial, distribusi Poisson, dan distribusi normal.
3. Mampu menerapkan distribusi binomial dan Poisson dalam konteks pengamatan kejadian infeksi dan kejadian kesehatan lainnya dalam populasi.
4. Menjelaskan karakteristik dan penggunaan distribusi normal dalam analisis data kesehatan yang bersifat kontinu dan simetris.
5. Menggunakan distribusi probabilitas untuk memperkirakan kemungkinan kejadian tertentu dalam populasi kesehatan dan memahami relevansinya dalam pengambilan keputusan klinis dan kebijakan kesehatan masyarakat.
6. Mengaplikasikan konsep distribusi probabilitas dalam studi kasus nyata, seperti distribusi kejadian infeksi dalam populasi, untuk memperkirakan risiko dan pola kejadian penyakit.
7. Mengembangkan kemampuan analisis statistik dasar dengan memahami penerapan distribusi probabilitas dalam software statistik dan interpretasi hasilnya secara tepat dan relevan dalam konteks kesehatan masyarakat.

Pendahuluan

Distribusi probabilitas merupakan salah satu konsep fundamental dalam biostatistik yang sangat penting untuk dipahami oleh para profesional kesehatan dan peneliti di bidang kesehatan masyarakat. Dalam praktiknya, data kesehatan sering kali menunjukkan pola tertentu yang dapat dijelaskan dan diprediksi melalui distribusi probabilitas. Pemahaman yang baik tentang distribusi ini memungkinkan kita untuk menginterpretasikan data secara lebih akurat, memperkirakan kemungkinan kejadian tertentu, serta membuat keputusan berbasis bukti yang lebih tepat.

Pada bagian ini, kita akan mempelajari tiga jenis distribusi probabilitas yang paling umum digunakan dalam analisis data kesehatan: distribusi binomial, distribusi Poisson, dan distribusi normal. Distribusi binomial biasanya digunakan untuk memodelkan kejadian yang bersifat biner, seperti keberhasilan atau kegagalan dalam pengobatan, atau kejadian infeksi yang terjadi atau tidak terjadi dalam sejumlah percobaan tertentu. Distribusi Poisson, di sisi lain, sering digunakan untuk memodelkan kejadian yang jarang dan tersebar secara acak dalam waktu atau ruang, seperti jumlah kasus infeksi dalam satu wilayah selama periode tertentu. Sedangkan distribusi normal adalah distribusi kontinu yang sering muncul dalam data kesehatan yang bersifat simetris dan mengikuti pola kurva lonceng, seperti distribusi tinggi badan, tekanan darah, atau usia penderita.

Memahami penerapan distribusi probabilitas ini sangat penting karena mereka menjadi dasar dalam berbagai analisis statistik yang digunakan dalam penelitian kesehatan. Misalnya, distribusi binomial dan Poisson membantu kita memperkirakan risiko dan kejadian penyakit, sedangkan distribusi normal digunakan dalam pengujian hipotesis dan penentuan batas kepercayaan.

Dengan menguasai konsep ini, para profesional kesehatan dapat melakukan estimasi yang lebih akurat, menginterpretasikan data secara tepat, dan merancang intervensi yang lebih efektif.

Selain itu, bagian ini juga akan menampilkan contoh aplikasi nyata, seperti distribusi kejadian infeksi dalam populasi. Melalui studi kasus ini, peserta akan diajak memahami bagaimana distribusi probabilitas digunakan untuk memperkirakan jumlah kasus yang mungkin terjadi dalam suatu periode tertentu, serta bagaimana hasil tersebut dapat digunakan untuk perencanaan dan pengambilan kebijakan kesehatan masyarakat. Dengan demikian, pemahaman tentang distribusi probabilitas tidak hanya bersifat teoretis, tetapi juga sangat relevan dan aplikatif dalam konteks pengelolaan data dan pengambilan keputusan di bidang kesehatan.

Secara keseluruhan, bagian ini bertujuan untuk membekali peserta dengan pengetahuan dasar yang kuat mengenai distribusi probabilitas, sehingga mereka mampu mengaplikasikan konsep tersebut dalam analisis data kesehatan secara praktis dan efektif. Pemahaman ini akan menjadi fondasi penting dalam penguasaan materi-materi lanjutan dalam biostatistik, serta mendukung pengembangan kompetensi analisis data yang akurat dan terpercaya dalam praktik kesehatan masyarakat.

7.1 Distribusi Binomial dan Distribusi Poisson

7.1.1 Pengantar Distribusi Binomial

Distribusi binomial merupakan salah satu distribusi probabilitas diskret yang digunakan untuk memodelkan jumlah keberhasilan dalam sejumlah percobaan yang identik dan independen, di mana setiap percobaan memiliki dua

kemungkinan hasil, yaitu sukses atau gagal. Distribusi ini sangat relevan dalam konteks kesehatan masyarakat, terutama ketika kita ingin mengetahui probabilitas terjadinya suatu kejadian tertentu yang bersifat biner, seperti keberhasilan pengobatan, kejadian infeksi, atau keberhasilan imunisasi.

Contoh nyata penggunaan distribusi binomial adalah ketika kita ingin mengetahui probabilitas bahwa dari 100 individu yang mendapatkan vaksin, sebanyak 80 di antaranya akan berhasil mendapatkan perlindungan dari penyakit tertentu. Jika keberhasilan setiap individu dianggap independen dan memiliki peluang yang sama, distribusi binomial dapat digunakan untuk menghitung probabilitas bahwa tepat 80 orang akan berhasil.

Secara matematis, distribusi binomial dinyatakan dengan rumus:

$$P(x; n, p) = {}^nC_x * p^x * (1 - p)^{(n - x)}$$

di mana:

- (n) adalah jumlah percobaan,
- (k) adalah jumlah keberhasilan yang diharapkan,
- (p) adalah peluang keberhasilan dalam satu percobaan,
- $({}^nC_k)$ adalah koefisien binomial yang menunjukkan kombinasi dari (n) percobaan yang menghasilkan (k) keberhasilan.

Distribusi binomial memiliki beberapa karakteristik penting:

- Nilai probabilitas $(P(X=k))$ selalu antara 0 dan 1.
- Rata-rata (mean) dari distribusi binomial adalah (np) .

- Variansnya adalah $(np(1-p))$.

7.1.2 Pengantar Distribusi Poisson

Distribusi Poisson adalah distribusi probabilitas diskret yang digunakan untuk memodelkan jumlah kejadian yang jarang dan tersebar secara acak dalam interval waktu, ruang, atau volume tertentu. Distribusi ini sangat berguna dalam konteks kesehatan masyarakat untuk memperkirakan jumlah kejadian penyakit yang terjadi dalam periode tertentu, terutama ketika kejadian tersebut bersifat langka dan tidak saling bergantung.

Contoh penggunaan distribusi Poisson adalah memperkirakan jumlah kasus infeksi yang muncul di sebuah rumah sakit dalam satu bulan. Jika kejadian infeksi dianggap acak dan jarang, distribusi Poisson dapat digunakan untuk memperkirakan probabilitas bahwa ada tepat 5 kasus dalam periode tersebut.

Rumus distribusi Poisson adalah:

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

di mana:

$P(X = k)$ = probabilitas terjadinya k kejadian

λ = laju rata-rata kejadian per satuan waktu atau ruang

e = basis logaritma natural (sekitar 2.71828)

$k!$ = factorial dari k

Karakteristik utama distribusi Poisson meliputi:

- Rata-rata dan variansnya sama, yaitu (λ) .

- Distribusi ini cenderung simetris untuk nilai (λ) yang besar, tetapi menjadi lebih skewed untuk nilai kecil.

7.1.3 Perbandingan dan Pemilihan Distribusi Binomial dan Poisson

Meskipun keduanya digunakan untuk memodelkan kejadian diskret, distribusi binomial dan Poisson memiliki perbedaan mendasar. Distribusi binomial digunakan ketika jumlah percobaan dan peluang keberhasilan diketahui, dan kejadian yang diamati adalah jumlah keberhasilan dari percobaan tersebut. Sebaliknya, distribusi Poisson lebih cocok digunakan ketika kejadian langka dan jumlah kejadian diukur dalam interval tertentu tanpa perlu mengetahui jumlah percobaan secara spesifik.

Dalam praktiknya, distribusi Poisson dapat dianggap sebagai limit dari distribusi binomial ketika (n) besar dan (p) kecil, dengan ($\lambda = np$). Hal ini memudahkan penggunaan distribusi Poisson dalam situasi di mana jumlah percobaan sangat besar dan peluang keberhasilan sangat kecil, seperti kejadian infeksi yang jarang terjadi.

7.2 Distribusi Normal

7.2.1 Pengantar Distribusi Normal

Distribusi normal, juga dikenal sebagai distribusi Gaussian, adalah distribusi kontinu yang memiliki bentuk kurva lonceng simetris di sekitar mean. Distribusi ini sangat penting dalam biostatistik karena banyak data kesehatan yang bersifat kontinu dan mengikuti pola distribusi normal, seperti tinggi badan, tekanan darah, kadar glukosa darah, dan usia.

Kurva distribusi normal memiliki karakteristik sebagai berikut:

- Simetris terhadap nilai rata-rata (mean),
- Memiliki bentuk lonceng yang simetris,
- Sekitar 68% data berada dalam satu standar deviasi (σ) dari mean,
- Sekitar 95% data berada dalam dua standar deviasi,
- Sekitar 99,7% data berada dalam tiga standar deviasi.

Rumus fungsi distribusi normal adalah:

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

di mana:

- $f(x)$ adalah nilai kepadatan peluang untuk variable acak x
- x adalah nilai variable acak yang sedang dipertimbangkan
- μ adalah rata-rata (mean distribusi normal)
- σ adalah simpangan baku distribusi normal
- π adalah konstanta matematika (sekitar 3.14159)
- e adalah konstanta matematika (sekitar 2.71828)

7.2.2 Penerapan Distribusi Normal dalam Data Kesehatan

Distribusi normal digunakan dalam berbagai analisis statistik, termasuk pengujian hipotesis dan penentuan batas kepercayaan. Dalam konteks kesehatan, distribusi ini sering digunakan untuk menilai apakah data yang diamati berbeda secara signifikan dari nilai yang diharapkan, serta untuk memperkirakan probabilitas kejadian tertentu.

Contoh penerapan distribusi normal adalah dalam pengukuran tekanan darah. Jika data tekanan darah dari populasi mengikuti distribusi normal dengan rata-rata 120 mmHg dan standar deviasi 15 mmHg, kita dapat menghitung probabilitas bahwa seorang individu memiliki tekanan darah di atas 140 mmHg. Dengan menggunakan tabel distribusi normal, kita dapat menentukan persentase populasi yang memiliki tekanan darah tinggi.

7.3 Penerapan Distribusi dalam Data Kesehatan

Distribusi probabilitas sangat penting dalam analisis data kesehatan karena memungkinkan kita untuk memperkirakan kemungkinan kejadian tertentu dalam populasi. Dengan memahami distribusi yang sesuai, kita dapat melakukan estimasi risiko, menentukan batas kepercayaan, dan menguji hipotesis secara statistik.

Misalnya, dalam surveilans penyakit menular, distribusi Poisson digunakan untuk memperkirakan jumlah kasus yang akan terjadi dalam periode tertentu. Jika diketahui bahwa rata-rata kejadian infeksi di suatu wilayah adalah 3 kasus per minggu, distribusi Poisson dapat digunakan untuk menghitung probabilitas bahwa dalam minggu berikutnya akan terjadi 0, 1, 2, atau lebih dari 3 kasus.

Selain itu, distribusi normal digunakan untuk data yang bersifat kontinu dan simetris, seperti distribusi tinggi badan atau indeks massa tubuh (BMI). Dengan mengetahui distribusi ini, para peneliti dan profesional kesehatan dapat menentukan batas normal, mengidentifikasi outlier, dan melakukan analisis statistik inferensial.

7.4 Contoh Aplikasi: Distribusi Kejadian Infeksi dalam Populasi

Sebagai ilustrasi praktis, kita akan membahas contoh distribusi kejadian infeksi dalam populasi. Misalnya, sebuah rumah sakit mencatat bahwa rata-rata jumlah kasus infeksi saluran pernapasan atas (ISPA) yang terjadi dalam satu bulan adalah 5 kasus. Berdasarkan data historis, jumlah kasus ini mengikuti distribusi Poisson dengan ($\mu = 5$).

Pertanyaan yang sering muncul adalah: berapa probabilitas bahwa dalam bulan berikutnya akan terjadi tepat 3 kasus ISPA? Menggunakan rumus distribusi Poisson:

$$P(x = k) = (\lambda^k * e^{-\lambda})/k!$$

di mana:

$P(x = k)$ = probabilitas terjadinya k kejadian

λ = laju rata-rata kejadian per satuan waktu atau ruang

e = basis logaritma natural (sekitar 2.71828)

$k!$ = factorial dari k

Dengan menghitung:

$$P(x = 3) = ((5^3 * 2.71828^{-5})/3!) * 100\%$$

$$P(x = 3) = \left(\frac{125 * 0,0067}{6} \right) * 100\% = 13,96\%$$

Jadi, ada sekitar 13,96% kemungkinan bahwa bulan berikutnya akan terjadi tepat 3 kasus ISPA.

Selain itu, kita juga dapat menghitung probabilitas bahwa jumlah kasus akan kurang dari atau sama dengan 2, yaitu:

$$P(x \leq 2) = p(x = 0) + P(x = 1) + P(x = 2)$$

Dimana:

- $P(x \leq 2)$ adalah probabilitas bahwa jumlah kasus (x) akan kurang dari atau sama dengan 2
- $P(x=0)$ adalah probabilitas bahwa tidak ada kasus yang terjadi ($x=0$)
- $P(x=1)$ adalah probabilitas bahwa terjadi 1 kasus ($x=1$)
- $P(x=2)$ adalah probabilitas bahwa terjadi 2 kasus ($x=2$)

Dengan perhitungan masing-masing:

$$P(x=0) = (e^x(-\lambda) * \lambda^x)/0! = (2.71828^0(-5) * 5^0/0! = 0,067$$

$$P(x=1) = (e^x(-\lambda) * \lambda^1)/1! = (2.71828^1(-5) * 5^1/1! = 0,0337$$

$$P(x=2) = (e^x(-\lambda) * \lambda^1)/1! = (2.71828^2(-5) * 5^2/2! = 0,0842$$

Sehingga:

$$\begin{aligned} P(x \leq 2) &= p(x = 0) + P(x = 1) + P(x = 2) \\ &= 0,0067 + 0,0337 + 0,0842 = 0,1246 \end{aligned}$$

Artinya, ada sekitar 12,46% kemungkinan bahwa jumlah kasus infeksi dalam bulan tersebut tidak lebih dari 2.

7.5 Relevansi dan Pentingnya Distribusi Probabilitas dalam Kesehatan Masyarakat

Penggunaan distribusi probabilitas dalam bidang kesehatan masyarakat sangat penting karena membantu dalam perencanaan, pengambilan keputusan, dan evaluasi program kesehatan. Dengan mengetahui distribusi yang sesuai, kita dapat memperkirakan kejadian yang akan datang, mengidentifikasi risiko, dan mengalokasikan sumber daya secara efisien.

Misalnya, dalam pengendalian wabah, distribusi Poisson digunakan untuk memperkirakan jumlah kasus yang akan muncul, sehingga petugas kesehatan dapat menyiapkan fasilitas dan tenaga medis secara optimal. Dalam pengujian efektivitas program imunisasi, distribusi binomial membantu menentukan probabilitas keberhasilan program dan memperkirakan jumlah individu yang terlindungi.

Selain itu, distribusi normal sering digunakan dalam pengukuran variabel kesehatan yang bersifat kontinu, seperti tekanan darah dan kadar kolesterol, untuk menentukan batas normal dan mengidentifikasi individu yang berisiko tinggi.

Distribusi probabilitas merupakan alat penting dalam analisis data kesehatan karena memungkinkan kita untuk memahami dan memprediksi pola kejadian dalam populasi. Distribusi binomial cocok digunakan untuk kejadian biner dalam percobaan terbatas, sedangkan distribusi Poisson ideal untuk kejadian langka dan tersebar acak dalam waktu atau ruang tertentu. Distribusi normal, sebagai distribusi kontinu yang simetris, banyak digunakan dalam analisis variabel kesehatan yang bersifat kontinu dan mengikuti pola distribusi tersebut.

Penguasaan konsep distribusi ini akan memperkuat kemampuan mahasiswa dalam melakukan analisis statistik dasar, menginterpretasikan data secara tepat, dan mendukung pengambilan keputusan berbasis bukti dalam praktik kesehatan masyarakat. Melalui contoh aplikasi nyata, seperti distribusi kejadian infeksi, mahasiswa dapat memahami relevansi distribusi probabilitas dalam konteks nyata dan meningkatkan kompetensi mereka dalam pengelolaan data kesehatan.

Rangkuman

Distribusi probabilitas merupakan konsep dasar yang penting dalam analisis data kesehatan dan epidemiologi. Pemahaman terhadap distribusi ini membantu profesional kesehatan dalam memperkirakan kejadian, mengidentifikasi risiko, dan membuat keputusan berbasis data. Berikut adalah poin-poin utama terkait distribusi probabilitas dalam konteks biostatistik kesehatan:

1. Distribusi binomial digunakan untuk memodelkan jumlah keberhasilan dalam sejumlah percobaan yang identik dan independen, di mana setiap percobaan memiliki dua hasil, yaitu sukses atau gagal. Contohnya adalah menghitung probabilitas keberhasilan imunisasi pada sejumlah individu tertentu.
2. Rumus distribusi binomial adalah $P(x;n,p) = nC_x \cdot p^x \cdot (1-p)^{(n-x)}$, dengan karakteristik utama berupa rata-rata (np) dan varians ($np(1-p)$). Distribusi ini cocok digunakan saat jumlah percobaan dan peluang keberhasilan diketahui.
3. Distribusi Poisson digunakan untuk memodelkan jumlah kejadian langka dan acak dalam interval waktu, ruang, atau volume tertentu. Contohnya

adalah memperkirakan jumlah kasus infeksi dalam satu bulan di rumah sakit.

4. Rumus distribusi Poisson adalah, $P(x = k) = (\lambda^k * e^{-\lambda}) / k!$ dengan karakteristik bahwa rata-rata dan variansnya sama, yaitu (λ). Distribusi ini efektif saat kejadian jarang dan tidak bergantung pada jumlah percobaan.
5. Perbedaan utama antara distribusi binomial dan Poisson terletak pada penggunaannya; binomial untuk percobaan terbatas dengan peluang tetap, sedangkan Poisson untuk kejadian langka dalam interval tertentu, sering dianggap sebagai limit dari binomial saat (n) besar dan (p) kecil.
6. Distribusi normal adalah distribusi kontinu berbentuk lonceng yang simetris di sekitar mean, digunakan secara luas dalam data kesehatan seperti tinggi badan dan tekanan darah. Karakteristiknya meliputi sekitar 68% data dalam satu standar deviasi dan 95% dalam dua standar deviasi dari mean.
7. Distribusi normal sangat penting dalam pengujian hipotesis dan penentuan batas normal variabel kesehatan. Contohnya adalah menghitung probabilitas seseorang memiliki tekanan darah di atas batas tertentu.
8. Dalam praktik, distribusi Poisson digunakan untuk memperkirakan jumlah kejadian infeksi, sedangkan distribusi normal digunakan untuk variabel kontinu seperti BMI dan tekanan darah. Pemilihan distribusi yang tepat mendukung analisis dan pengambilan keputusan yang akurat.
9. Contoh aplikasi nyata adalah menghitung probabilitas bahwa dalam bulan tertentu terjadi tepat 3 kasus infeksi dengan menggunakan distribusi Poisson, serta memperkirakan kemungkinan jumlah kasus yang lebih kecil dari atau sama dengan 2.

10. Penguasaan konsep distribusi probabilitas sangat penting dalam kesehatan masyarakat untuk perencanaan program, pengendalian wabah, dan evaluasi efektivitas intervensi. Distribusi ini membantu memperkirakan kejadian di masa depan dan mengelola sumber daya secara efisien.

Dengan memahami dan mampu menerapkan distribusi binomial, Poisson, dan normal, mahasiswa dan profesional kesehatan dapat melakukan analisis data yang lebih tepat, mendukung pengambilan keputusan berbasis bukti, serta meningkatkan efektivitas program kesehatan masyarakat.

Latihan Mahasiswa

Soal Essay

1. Jelaskan perbedaan utama antara distribusi binomial dan distribusi Poisson dalam konteks data kesehatan masyarakat. Berikan contoh kasus yang sesuai untuk masing-masing distribusi.
2. Bagaimana karakteristik distribusi normal mempengaruhi penggunaannya dalam analisis data kesehatan? Jelaskan dengan memberikan contoh variabel kesehatan yang biasanya mengikuti distribusi normal.
3. Sebutkan dan jelaskan langkah-langkah utama dalam menggunakan distribusi Poisson untuk memperkirakan jumlah kejadian infeksi dalam suatu populasi.
4. Mengapa distribusi binomial dapat dianggap sebagai limit dari distribusi Poisson dalam kondisi tertentu? Jelaskan dengan menyertakan parameter yang relevan.

5. Diskusikan pentingnya pemahaman distribusi probabilitas dalam pengambilan keputusan kebijakan kesehatan masyarakat berbasis data.

Soal Pilihan Berganda

1. Distribusi probabilitas yang digunakan untuk memodelkan jumlah kejadian langka dalam interval tertentu adalah:
 - a. Distribusi normal
 - b. Distribusi binomial
 - c. Distribusi Poisson
 - d. Distribusi uniform
2. Jika sebuah penelitian memperkirakan bahwa rata-rata kejadian infeksi dalam satu bulan adalah 4 kasus, dan kita ingin mengetahui probabilitas bahwa dalam bulan berikutnya terjadi tepat 6 kasus, distribusi yang paling tepat digunakan adalah:
 - a. Distribusi normal
 - b. Distribusi binomial
 - c. Distribusi Poisson
 - d. Distribusi eksponensial
3. Karakteristik utama dari distribusi normal adalah:
 - a. Data bersifat diskret dan tidak simetris
 - b. Data mengikuti kurva lonceng dan simetris terhadap mean
 - c. Data hanya berlaku untuk kejadian langka
 - d. Data memiliki peluang sama untuk semua nilai

4. Dalam distribusi binomial, parameter yang menentukan jumlah keberhasilan yang diharapkan adalah:
 - a. (n) dan (p)
 - b. (λ) dan (k)
 - c. (μ) dan (Σ)
 - d. (e) dan (π)
5. Distribusi Poisson cocok digunakan untuk memodelkan kejadian yang:
 - a. Bersifat kontinu dan simetris
 - b. Langka dan tersebar secara acak dalam waktu atau ruang
 - c. Memiliki peluang sama untuk semua nilai
 - d. Terjadi dalam jumlah besar dan berkelanjutan
6. Jika diketahui bahwa jumlah kasus infeksi dalam satu minggu mengikuti distribusi Poisson dengan $(\lambda = 3)$, berapa probabilitas bahwa tidak ada kasus yang terjadi minggu tersebut?
 - a. (e^{-3})
 - b. $(3e^{-3})$
 - c. $(P(x=0) = \frac{3^0 e^{-3}}{0!})$
 - d. a dan c benar
7. Salah satu keunggulan distribusi normal dalam analisis data kesehatan adalah:
 - a. Mudah digunakan untuk data diskret
 - b. Memiliki bentuk kurva lonceng yang simetris
 - c. Cocok untuk data yang jarang terjadi
 - d. Tidak memerlukan asumsi tentang distribusi data

8. Dalam konteks pengukuran variabel kesehatan, variabel yang biasanya mengikuti distribusi normal adalah:
 - a. Jumlah kasus penyakit dalam satu hari
 - b. Tinggi badan individu
 - c. Jumlah kunjungan ke fasilitas kesehatan
 - d. Jumlah kejadian infeksi dalam populasi kecil
9. Dalam penggunaan distribusi binomial, apa yang terjadi jika peluang keberhasilan (p) sangat kecil dan jumlah percobaan (n) sangat besar, sehingga (np) tetap konstan?
 - a. Distribusi binomial mendekati distribusi normal
 - b. Distribusi binomial mendekati distribusi Poisson
 - c. Distribusi binomial tidak dapat digunakan
 - d. Distribusi binomial menjadi distribusi uniform
10. Dalam analisis data kesehatan, mengapa penting untuk memahami distribusi probabilitas yang sesuai dengan data yang diamati?
 - a. Untuk mempercepat pengolahan data
 - b. Untuk memastikan interpretasi hasil yang akurat dan pengambilan keputusan yang tepat
 - c. Untuk mengurangi jumlah data yang perlu dianalisis
 - d. Untuk menghindari penggunaan software statistic

Soal Studi Kasus / Project

1. Sebuah rumah sakit mencatat bahwa jumlah kasus infeksi saluran pernapasan atas (ISPA) yang terjadi dalam satu bulan mengikuti distribusi Poisson dengan rata-rata ($\lambda = 6$). Hitunglah probabilitas bahwa dalam bulan berikutnya akan terjadi lebih dari 8 kasus ISPA. Jelaskan langkah-langkah perhitungannya dan interpretasikan hasilnya.
2. Dalam sebuah survei kesehatan, diketahui bahwa tinggi badan penduduk mengikuti distribusi normal dengan rata-rata 165 cm dan standar deviasi 10 cm. Jika seorang peserta diambil secara acak, berapa peluang bahwa tingginya lebih dari 180 cm? Buatlah perhitungan dan interpretasi hasilnya.

REFERENSI

1. Agresti, A. (2018). *Statistical Methods for the Social Sciences*. Pearson Education.
2. Blumberg, S., & Cummings, P. (2019). *Epidemiology and Biostatistics*. Oxford University Press.
3. Kirkwood, B. R., & Sterne, J. A. C. (2017). *Essential Medical Statistics*. Wiley.
4. Rosner, B. (2015). *Fundamentals of Biostatistics*. Cengage Learning.
5. Walpole, R. E., Myers, R. H., Myers, S. L., & Ye, K. (2018). *Probability & Statistics for Engineering and the Sciences*. Pearson."

BAB 8

SAMPEL DAN TEKNIK SAMPLING

Tujuan Pembelajaran

1. Mahasiswa mampu memahami konsep dasar dan pentingnya teknik sampling dalam penelitian kesehatan masyarakat serta mengetahui berbagai jenis teknik pengambilan sampel yang umum digunakan, seperti random, stratified, dan cluster sampling.
2. Mahasiswa dapat menjelaskan proses dan langkah-langkah dalam menentukan teknik sampling yang sesuai dengan karakteristik populasi dan tujuan penelitian, serta memahami kelebihan dan kekurangan masing-masing teknik.
3. Mahasiswa mampu mengidentifikasi dan menjelaskan berbagai kesalahan yang mungkin terjadi selama proses pengambilan sampel, baik yang bersifat sampling maupun non-sampling, serta memahami dampaknya terhadap validitas hasil penelitian.
4. Mahasiswa mampu menghitung dan menentukan ukuran sampel yang diperlukan dalam studi kesehatan masyarakat dengan mempertimbangkan tingkat kepercayaan, margin error, dan variabilitas data, sehingga hasil penelitian dapat diandalkan dan representatif.
5. Mahasiswa mampu menerapkan teknik pengambilan sampel secara praktis dalam konteks surveilans kesehatan lingkungan, termasuk dalam pengumpulan data yang akurat dan efisien.

6. Mahasiswa mampu menganalisis dan menilai keefektifan serta kesesuaian teknik sampling yang digunakan dalam berbagai studi kesehatan, serta mampu mengidentifikasi potensi bias yang mungkin muncul.
7. Mahasiswa mampu menyusun laporan hasil pengambilan sampel secara sistematis dan jelas, serta mampu mengintegrasikan hasil tersebut ke dalam laporan penelitian atau surveilans kesehatan masyarakat.

Pendahuluan

Pengambilan sampel merupakan salah satu aspek fundamental dalam penelitian kesehatan masyarakat. Dalam dunia penelitian, seringkali tidak memungkinkan untuk mengumpulkan data dari seluruh populasi karena keterbatasan waktu, biaya, dan sumber daya. Oleh karena itu, teknik sampling menjadi solusi yang efektif untuk memperoleh data yang representatif dan dapat digeneralisasi ke seluruh populasi. Dengan menggunakan sampel yang dipilih secara tepat, peneliti dapat memperoleh gambaran yang akurat tentang karakteristik populasi tanpa harus melakukan pengumpulan data dari seluruh anggota populasi tersebut.

Urgensi dari pemahaman teknik sampling ini semakin meningkat seiring dengan meningkatnya kebutuhan akan data yang valid dan reliabel dalam pengambilan keputusan di bidang kesehatan masyarakat. Misalnya, dalam surveilans kesehatan lingkungan, pengambilan sampel yang tepat sangat penting untuk memastikan bahwa data yang diperoleh mencerminkan kondisi sebenarnya di lapangan. Data yang akurat akan membantu dalam merancang intervensi yang tepat sasaran, mengidentifikasi daerah yang

membutuhkan perhatian khusus, serta memantau efektivitas program kesehatan.

Selain itu, pemilihan teknik sampling yang tidak tepat dapat menyebabkan bias, yang pada akhirnya mempengaruhi validitas dan reliabilitas hasil penelitian. Kesalahan dalam proses pengambilan sampel, baik yang bersifat sampling maupun non-sampling, dapat menyebabkan data yang tidak representatif, sehingga kesimpulan yang diambil menjadi tidak akurat dan menyesatkan. Oleh karena itu, pemahaman mendalam tentang berbagai teknik sampling, termasuk kelebihan dan kekurangannya, sangat penting bagi mahasiswa dan peneliti di bidang kesehatan masyarakat.

Dalam modul ini, kita akan membahas secara rinci berbagai teknik pengambilan sampel yang umum digunakan, seperti random sampling, stratified sampling, dan cluster sampling. Selain itu, akan dibahas pula mengenai kesalahan yang sering terjadi selama proses pengambilan sampel dan bagaimana cara menghindarinya. Penentuan ukuran sampel juga menjadi bagian penting yang akan dipelajari, karena menentukan jumlah sampel yang optimal sangat berpengaruh terhadap keakuratan hasil penelitian. Melalui pemahaman ini, diharapkan mahasiswa mampu menerapkan teknik sampling secara tepat dalam berbagai studi dan surveilans kesehatan lingkungan, serta mampu menyusun laporan yang sistematis dan akurat berdasarkan data yang diperoleh.

Dengan penguasaan materi ini, diharapkan mahasiswa tidak hanya mampu melakukan pengambilan sampel secara teknis, tetapi juga mampu menilai dan mengkritisi proses sampling dalam penelitian yang mereka lakukan maupun yang mereka baca. Hal ini akan memperkuat kompetensi mereka dalam menghasilkan data yang valid dan dapat dipertanggungjawabkan, sehingga mendukung pengambilan

kebijakan dan intervensi yang berbasis bukti di bidang kesehatan masyarakat.

8.1 Teknik Pengambilan Sampel: Random, Stratified, Cluster

Pengambilan sampel merupakan proses penting dalam penelitian kesehatan masyarakat yang bertujuan untuk mendapatkan data yang representatif dari populasi yang diteliti. Teknik pengambilan sampel harus dipilih dengan cermat agar hasil penelitian valid dan dapat digeneralisasi ke seluruh populasi. Tiga teknik utama yang umum digunakan adalah random sampling, stratified sampling, dan cluster sampling. Masing-masing teknik memiliki karakteristik, kelebihan, dan kekurangan yang perlu dipahami secara mendalam oleh mahasiswa agar dapat memilih metode yang paling sesuai dengan tujuan dan kondisi penelitian.

8.1.1. Random Sampling

Random sampling, atau pengambilan sampel secara acak, adalah teknik di mana setiap anggota populasi memiliki peluang yang sama untuk dipilih sebagai sampel. Prinsip utama dari teknik ini adalah keadilan dan ketidakberpihakan, sehingga hasilnya cenderung tidak bias dan mewakili seluruh populasi secara proporsional. Dalam praktiknya, pengambilan sampel secara acak dapat dilakukan melalui berbagai metode, seperti menggunakan tabel angka acak, perangkat lunak komputer, atau metode lain yang memastikan peluang yang sama bagi semua anggota populasi.

Keunggulan utama dari random sampling adalah kemampuannya menghasilkan sampel yang bebas dari bias

sistematis, sehingga hasil analisis statistiknya menjadi lebih valid dan dapat digeneralisasi. Selain itu, teknik ini relatif sederhana dan mudah diterapkan jika populasi yang diteliti tidak terlalu besar dan daftar anggota populasi lengkap tersedia.

Namun, terdapat beberapa kekurangan dari teknik ini. Pertama, dalam populasi yang sangat besar dan tersebar luas secara geografis, pengambilan sampel secara acak bisa menjadi mahal dan memakan waktu. Kedua, jika tidak dilakukan dengan benar, ada kemungkinan terjadinya peluang yang tidak merata karena kesalahan dalam proses pengacakan. Ketiga, teknik ini tidak mempertimbangkan karakteristik tertentu dari populasi yang mungkin relevan, seperti usia, jenis kelamin, atau faktor risiko tertentu, sehingga bisa menyebabkan ketidakseimbangan dalam sampel.

Contoh aplikasi dari random sampling adalah surveilans kesehatan masyarakat di sebuah kota untuk mengetahui prevalensi penyakit tertentu. Misalnya, peneliti ingin mengetahui prevalensi hipertensi di kalangan dewasa usia 30-50 tahun di sebuah kota. Dengan daftar seluruh penduduk usia tersebut, peneliti dapat menggunakan perangkat lunak statistik untuk secara acak memilih 500 orang sebagai sampel, sehingga setiap individu memiliki peluang yang sama untuk terpilih (Kish, 1965).

8.1.2. Stratified Sampling

Stratified sampling adalah teknik pengambilan sampel di mana populasi dibagi menjadi subkelompok atau strata berdasarkan karakteristik tertentu yang relevan dengan penelitian, seperti usia, jenis kelamin, tingkat pendidikan, atau faktor risiko lainnya. Setelah pembagian ini, sampel diambil secara acak dari setiap strata sesuai dengan

proporsi mereka dalam populasi. Teknik ini bertujuan untuk memastikan bahwa semua subkelompok penting terwakili secara proporsional dalam sampel, sehingga hasil penelitian lebih akurat dan representatif.

Keunggulan utama dari stratified sampling adalah kemampuannya untuk meningkatkan efisiensi dan akurasi estimasi parameter populasi, terutama ketika variabilitas dalam populasi cukup tinggi antar strata. Dengan memastikan setiap strata terwakili, teknik ini mengurangi kemungkinan bias dan meningkatkan kekuatan statistik analisis.

Kekurangan dari stratified sampling adalah prosesnya yang lebih kompleks dan memerlukan data awal yang lengkap tentang karakteristik populasi untuk membagi ke dalam strata. Selain itu, jika pembagian strata tidak dilakukan dengan benar, atau jika jumlah sampel dari setiap strata tidak proporsional, hasilnya bisa menjadi bias.

Contoh aplikasi dari stratified sampling adalah surveilans kesehatan lingkungan di daerah perkotaan dan pedesaan. Misalnya, peneliti ingin mengetahui tingkat paparan polusi udara di dua wilayah berbeda. Populasi dibagi menjadi dua strata berdasarkan lokasi geografis (perkotaan dan pedesaan). Dari masing-masing strata, diambil sampel secara acak sesuai proporsi jumlah penduduk di wilayah tersebut. Dengan demikian, data yang diperoleh akan mencerminkan kondisi di kedua wilayah secara proporsional dan memungkinkan analisis perbandingan yang valid (Cochran, 1977).

8.1.3. Cluster Sampling

Cluster sampling adalah teknik pengambilan sampel di mana populasi dibagi menjadi beberapa kelompok atau cluster

yang secara geografis atau administratif mewakili keseluruhan populasi. Kemudian, sejumlah cluster dipilih secara acak, dan seluruh anggota dari cluster yang terpilih dijadikan sampel. Teknik ini sering digunakan ketika populasi tersebar luas dan sulit diakses secara langsung, sehingga pengambilan sampel dari seluruh populasi secara individual menjadi tidak praktis dan mahal.

Keunggulan utama dari cluster sampling adalah efisiensi biaya dan waktu, karena hanya beberapa cluster yang dipilih dan seluruh anggota di dalamnya diikutsertakan. Teknik ini sangat cocok untuk surveilans kesehatan di daerah yang luas dan sulit dijangkau, seperti surveilans penyakit menular di wilayah pedesaan atau daerah terpencil.

Namun, kekurangan dari cluster sampling adalah potensi bias yang lebih tinggi karena anggota dalam satu cluster cenderung memiliki karakteristik yang serupa, sehingga data yang diperoleh bisa kurang variatif dan kurang representatif jika jumlah cluster yang dipilih terlalu sedikit. Selain itu, analisis statistiknya memerlukan penyesuaian khusus karena adanya korelasi intra-cluster yang dapat mempengaruhi estimasi varians.

Contoh aplikasi dari cluster sampling adalah surveilans kesehatan lingkungan di daerah pedesaan. Misalnya, peneliti ingin memantau tingkat pencemaran air di beberapa desa. Desa-desa tersebut dibagi menjadi beberapa cluster berdasarkan wilayah administratif. Dari daftar desa, dipilih secara acak 10 desa sebagai cluster, dan seluruh rumah tangga di desa tersebut diikutsertakan dalam survei. Pendekatan ini memungkinkan pengumpulan data secara efisien dan representatif terhadap kondisi di seluruh wilayah (Levy & Lemeshow, 2013).

8.2 Kesalahan Sampling dan Non-Sampling

Dalam proses pengambilan sampel, terdapat berbagai kesalahan yang dapat mempengaruhi validitas dan reliabilitas data. Kesalahan ini dapat dibedakan menjadi dua kategori utama: kesalahan sampling dan kesalahan non-sampling.

8.2.1. Kesalahan Sampling

Kesalahan sampling terjadi karena ketidaksempurnaan dalam proses pemilihan sampel yang menyebabkan sampel tidak mewakili populasi secara akurat. Kesalahan ini bersifat acak dan biasanya dapat diminimalkan dengan menggunakan teknik sampling yang tepat dan jumlah sampel yang cukup besar. Contoh kesalahan sampling adalah ketika terjadi pengulangan dalam proses pengambilan sampel yang menyebabkan beberapa anggota populasi memiliki peluang lebih besar untuk terpilih dibandingkan yang lain, atau ketika proses pengacakan tidak dilakukan secara benar sehingga peluang setiap individu tidak sama.

Contoh nyata dari kesalahan sampling adalah ketika peneliti hanya mengambil sampel dari satu wilayah tertentu dalam sebuah kota yang luas, padahal populasi tersebar di seluruh kota. Hasilnya, data yang diperoleh tidak mencerminkan kondisi seluruh kota dan hanya mewakili wilayah tertentu saja.

8.2.2. Kesalahan Non-Sampling

Kesalahan non-sampling adalah kesalahan yang terjadi di luar proses pengambilan sampel dan biasanya berkaitan dengan proses pengumpulan data, pengukuran, atau pengolahan data. Kesalahan ini bisa terjadi karena kesalahan

manusia, alat ukur yang tidak akurat, atau kesalahan dalam pencatatan data. Contoh umum adalah bias responden, kesalahan pengukuran, atau data yang hilang karena kesalahan administrasi.

Misalnya, dalam surveilans kesehatan lingkungan, jika petugas pengumpul data tidak mengikuti prosedur standar pengukuran kualitas air, hasil yang diperoleh bisa tidak akurat, meskipun sampel diambil secara acak dan representatif.

8.2.3. Dampak Kesalahan terhadap Validitas Data

Kedua jenis kesalahan ini dapat menurunkan validitas dan reliabilitas data penelitian. Kesalahan sampling dapat menyebabkan bias sistematis jika sampel tidak mewakili populasi, sedangkan kesalahan non-sampling dapat menyebabkan data yang tidak akurat dan menyesatkan. Oleh karena itu, penting bagi peneliti untuk memahami dan mengendalikan kedua jenis kesalahan ini melalui desain penelitian yang tepat, pelatihan pengumpul data, dan penggunaan instrumen yang valid serta reliabel.

8.3 Penentuan Ukuran Sampel

Menentukan ukuran sampel yang tepat adalah langkah krusial dalam penelitian kesehatan masyarakat. Ukuran sampel yang cukup besar akan meningkatkan kekuatan statistik dan keakuratan estimasi, sementara ukuran yang terlalu kecil dapat menyebabkan hasil yang tidak valid dan tidak dapat digeneralisasi.

8.3.1. Faktor-Faktor yang Mempengaruhi Penentuan Ukuran Sampel

Beberapa faktor utama yang perlu dipertimbangkan dalam menentukan ukuran sampel meliputi tingkat kepercayaan (confidence level), margin of error (tingkat kesalahan yang dapat diterima), variabilitas data, dan prevalensi kondisi yang diteliti.

- **Tingkat kepercayaan:** Biasanya diambil sebesar 95% atau 99%, yang menunjukkan seberapa yakin kita bahwa hasil sampel mencerminkan populasi. Semakin tinggi tingkat kepercayaan, semakin besar ukuran sampel yang diperlukan.
- **Margin of error:** Merupakan batas toleransi kesalahan estimasi, misalnya 5%. Semakin kecil margin of error yang diinginkan, semakin besar pula ukuran sampel yang diperlukan.
- **Variabilitas data:** Tingkat variabilitas dalam populasi, yang biasanya diukur melalui deviasi standar atau proporsi, mempengaruhi jumlah sampel yang dibutuhkan. Data yang sangat bervariasi membutuhkan sampel yang lebih besar.
- **Prevalensi:** Dalam studi prevalensi, tingkat kejadian atau kondisi tertentu dalam populasi juga mempengaruhi ukuran sampel yang diperlukan.

8.3.2. Rumus dan Pendekatan Penghitungan

Untuk studi prevalensi, rumus umum (Cochran) yang digunakan adalah sebagai berikut:

$$n = (Z^2 * p * q) / e^2$$

Dimana:

- n = ukuran sampel yang diperlukan
- Z = skor z sesuai tingkat kepercayaan (misalnya 1,96 untuk 95%)
- p = estimasi prevalensi dari populasi (dapat diambil dari studi sebelumnya atau perkiraan)
- $q = 1 - p$
- e = margin of error yang diinginkan

Contoh, jika peneliti ingin mengetahui prevalensi stunting pada anak-anak di suatu daerah dengan estimasi prevalensi 20% ($p=0,2$), tingkat kepercayaan 95% ($Z=1,96$), dan margin error 5% ($e=0,05$), maka:

$$n = (Z^2 * p * q) / e^2$$

$$n = (1,96^2 * 0,2 * (1 - 0,2)) / 0,05^2$$

$$n = 245,8 \text{ sampel}$$

Sehingga, diperlukan minimal 246 sampel untuk mencapai tingkat kepercayaan dan margin error tersebut.

8.3.3. Penyesuaian untuk Faktor Non-Respons dan Drop-out

Dalam praktik lapangan, sering terjadi non-respons atau kehilangan data. Oleh karena itu, disarankan menambahkan margin penyesuaian, biasanya sekitar 10-20%, terhadap ukuran sampel yang dihitung. Jika menggunakan contoh di atas, maka:

$$n' = \frac{n}{1 - D}$$

Dimana :

- n' adalah ukuran sampel yang disesuaikan
- n adalah ukuran sampel awal
- D adalah tingkat non-respons yang diharapkan

Jika menggunakan contoh di atas dengan drop out 20%, maka:

$$n' = \frac{n}{1 - D}$$

$$n' = \frac{246}{1 - 0,2} = 307,5$$

Dengan demikian, jumlah sampel yang harus diambil sekitar 308 untuk mengantisipasi non-respons.

8.4 Contoh Aplikasi: Pengambilan Sampel untuk Surveilans Kesehatan Lingkungan

Penggunaan teknik sampling dalam surveilans kesehatan lingkungan sangat penting untuk mendapatkan data yang akurat dan representatif mengenai kondisi lingkungan di masyarakat. Misalnya, dalam rangka memantau kualitas air di daerah pedesaan, pengambilan sampel harus dilakukan secara sistematis agar hasilnya dapat digunakan untuk pengambilan keputusan.

Misalnya, sebuah penelitian bertujuan untuk menilai tingkat pencemaran air di 20 desa di sebuah kabupaten. Pendekatan yang tepat adalah menggunakan teknik cluster sampling, karena desa-desa tersebut merupakan unit administratif yang berbeda dan tersebar luas. Langkah pertama adalah mengidentifikasi semua desa sebagai cluster, kemudian secara acak memilih 5 desa sebagai cluster yang akan

diteliti. Selanjutnya, seluruh rumah tangga di desa yang terpilih akan diambil sampelnya untuk pengujian kualitas air.

Dalam proses ini, peneliti harus memastikan bahwa pemilihan desa dilakukan secara acak dan bahwa seluruh rumah tangga di desa tersebut diikutsertakan agar data yang diperoleh mencerminkan kondisi di desa tersebut secara keseluruhan. Selain itu, peneliti juga harus memperhatikan faktor-faktor seperti variasi geografis, sumber air, dan kegiatan ekonomi masyarakat yang dapat mempengaruhi tingkat pencemaran air.

Penggunaan teknik sampling yang tepat dalam surveilans ini akan memastikan bahwa data yang diperoleh tidak hanya akurat dan representatif, tetapi juga efisien dari segi biaya dan waktu. Hasil dari surveilans ini dapat digunakan untuk merancang intervensi yang tepat sasaran, seperti peningkatan sanitasi, pengelolaan limbah, atau edukasi masyarakat tentang penggunaan air bersih.

Rangkuman

Pengambilan sampel merupakan langkah penting dalam penelitian kesehatan masyarakat untuk memperoleh data yang representatif dan valid. Tiga teknik utama yang umum digunakan adalah random sampling, stratified sampling, dan cluster sampling, masing-masing memiliki karakteristik, keunggulan, dan kekurangan yang perlu dipahami secara mendalam. Berikut adalah poin-poin utama yang merangkum pembahasan tersebut:

1. Random Sampling

- Setiap anggota populasi memiliki peluang yang sama untuk dipilih.

- Menjamin sampel bebas bias sistematis dan hasil yang dapat digeneralisasi.
- Cocok untuk populasi kecil hingga sedang dengan daftar lengkap anggota populasi.
- Kelemahan: biaya dan waktu tinggi untuk populasi besar dan tersebar luas; risiko peluang tidak merata jika proses pengacakan tidak benar.
- Contoh: surveilans hipertensi di kota dengan daftar lengkap penduduk dewasa.

2. **Stratified Sampling**

- Membagi populasi ke dalam subkelompok (strata) berdasarkan karakteristik tertentu, lalu diambil sampel secara acak dari setiap strata.
- Memastikan semua subkelompok penting terwakili secara proporsional.
- Meningkatkan efisiensi dan akurasi estimasi, terutama jika variabilitas antar strata tinggi.
- Kekurangan: proses lebih kompleks dan memerlukan data awal lengkap; risiko bias jika pembagian tidak tepat.
- Contoh: surveilans tingkat paparan polusi udara di wilayah perkotaan dan pedesaan.

3. **Cluster Sampling**

- Populasi dibagi menjadi kelompok (cluster), kemudian sejumlah cluster dipilih secara acak dan seluruh anggota di dalamnya diikutsertakan.
- Efisien biaya dan waktu, cocok untuk populasi tersebar luas dan sulit diakses.
- Kekurangan: potensi bias karena anggota dalam satu cluster cenderung serupa; analisis memerlukan penyesuaian statistik khusus.

- Contoh: surveilans pencemaran air di desa-desa yang dipilih secara acak dari wilayah administratif.
- 4. Kesalahan Sampling dan Non-Sampling**
- Kesalahan sampling terjadi karena proses pengambilan yang tidak sempurna, dapat menyebabkan sampel tidak mewakili populasi.
 - Kesalahan non-sampling muncul dari proses pengumpulan data, seperti bias responden atau pengukuran tidak akurat.
 - Kedua kesalahan ini menurunkan validitas data dan harus diminimalkan melalui desain yang tepat dan pelatihan.
- 5. Penentuan Ukuran Sampel**
- Faktor utama meliputi tingkat kepercayaan, margin of error, variabilitas data, dan prevalensi kondisi.
 - Rumus umum untuk prevalensi: $n = (Z^2 \cdot p \cdot q) / e^2$
 - Penyesuaian dilakukan untuk mengantisipasi non-respons dan drop-out, biasanya dengan menambah 10-20% dari hasil perhitungan.
- 6. Contoh Aplikasi dalam Surveilans Kesehatan Lingkungan**
- Teknik cluster sampling digunakan untuk memantau kualitas air di desa-desa, dengan memilih secara acak sejumlah desa sebagai cluster dan seluruh rumah tangga di desa tersebut sebagai sampel.
 - Pendekatan ini efisien dan memungkinkan pengumpulan data representatif, yang dapat digunakan untuk perencanaan intervensi kesehatan masyarakat.

Memahami dan menerapkan teknik pengambilan sampel yang tepat sangat penting untuk memastikan keakuratan

dan keandalan data penelitian. Penguasaan materi ini mendukung mahasiswa dalam merancang studi yang valid, mampu mengidentifikasi potensi bias, serta menghasilkan data yang dapat dipertanggungjawabkan secara ilmiah.

Latihan Mahasiswa

Soal Essay

1. Jelaskan pengertian dan prinsip dasar dari teknik random sampling dalam penelitian kesehatan masyarakat serta sebutkan keunggulan utamanya.
2. Sebutkan dan jelaskan langkah-langkah utama dalam menentukan ukuran sampel untuk studi prevalensi di bidang kesehatan masyarakat.
3. Diskusikan perbedaan utama antara stratified sampling dan cluster sampling, lengkap dengan kelebihan dan kekurangannya masing-masing.
4. Jelaskan apa yang dimaksud dengan kesalahan sampling dan non-sampling, serta berikan contoh nyata dari masing-masing kesalahan tersebut dalam surveilans kesehatan lingkungan.
5. Bagaimana proses pemilihan teknik sampling yang tepat dapat mempengaruhi validitas dan reliabilitas hasil penelitian? Jelaskan secara singkat.

Soal Pilihan Berganda

1. Dalam teknik pengambilan sampel, yang mana yang termasuk ke dalam teknik stratified sampling?
 - a. Memilih beberapa desa secara acak dan seluruh anggota di desa tersebut diambil sampelnya

- b. Membagi populasi berdasarkan karakteristik tertentu dan mengambil sampel dari setiap kelompok
 - c. Mengambil sampel secara acak dari seluruh populasi tanpa membedakan karakteristik
 - d. Mengambil sampel dari satu wilayah tertentu saja
- 2. Teknik pengambilan sampel yang paling efisien dari segi biaya dan waktu ketika populasi tersebar luas adalah:
 - a. Random sampling
 - b. Stratified sampling
 - c. Cluster sampling
 - d. Sistematic sampling
- 3. Jika sebuah penelitian ingin memastikan bahwa semua subkelompok usia dalam populasi terwakili secara proporsional, teknik sampling yang paling tepat digunakan adalah:
 - a. Random sampling
 - b. Stratified sampling
 - c. Cluster sampling
 - d. Sistematic sampling
- 4. Salah satu kekurangan utama dari cluster sampling adalah:
 - a. Memerlukan data awal lengkap tentang populasi
 - b. Potensi bias karena anggota dalam satu cluster cenderung memiliki karakteristik yang serupa
 - c. Membutuhkan biaya dan waktu yang besar
 - d. Tidak cocok untuk populasi yang tersebar luas

5. Kesalahan non-sampling biasanya disebabkan oleh:
 - a. Pengacakan yang tidak benar
 - b. Bias responden dan kesalahan pengukuran
 - c. Pemilihan cluster yang tidak acak
 - d. Pengambilan sampel dari satu wilayah saja
6. Dalam menentukan ukuran sampel, faktor yang paling berpengaruh adalah:
 - a. Jumlah populasi
 - b. Tingkat kepercayaan dan margin of error
 - c. Jumlah variabel yang diteliti
 - d. Waktu pengumpulan data
7. Jika estimasi prevalensi penyakit tertentu adalah 15%, dan peneliti menginginkan margin error 3% dengan tingkat kepercayaan 95%, maka ukuran sampel yang diperlukan adalah:
 - a. Sekitar 200
 - b. Sekitar 300
 - c. Sekitar 385
 - d. Sekitar 500
8. Dalam surveilans kesehatan lingkungan, teknik sampling yang paling sesuai untuk memantau kualitas air di desa-desa terpencil adalah:
 - a. Random sampling
 - b. Stratified sampling
 - c. Cluster sampling
 - d. Sistematic sampling
9. Salah satu keunggulan utama dari stratified sampling adalah:

- a. Mengurangi biaya pengumpulan data
 - b. Meningkatkan efisiensi dan akurasi estimasi parameter
 - c. Tidak memerlukan data awal tentang populasi
 - d. Mudah dilakukan tanpa data pendukung
10. Dalam proses pengambilan sampel, apa yang harus dilakukan untuk mengurangi risiko bias akibat non-respons?
- a. Menggunakan teknik sampling acak
 - b. Menambah jumlah sampel dan melakukan follow-up terhadap non-respon
 - c. Menggunakan cluster sampling
 - d. Mengurangi jumlah variabel yang diteliti

Soal Project / Studi Kasus

1. Sebuah tim peneliti ingin menilai tingkat pencemaran udara di 10 desa di daerah tertentu. Mereka memutuskan menggunakan teknik cluster sampling. Jelaskan langkah-langkah yang harus dilakukan mulai dari pemilihan cluster hingga pengumpulan data, serta faktor apa yang perlu diperhatikan agar data yang diperoleh valid dan representatif.
2. Anda diminta untuk merancang studi surveilans mengenai prevalensi obesitas di kalangan remaja di sebuah kota besar. Jelaskan proses menentukan teknik sampling yang paling sesuai, bagaimana menghitung ukuran sampel, dan langkah-langkah menghindari kesalahan sampling serta non-sampling selama proses pengambilan data.

REFERENSI

1. Cochran, W. G. (1977). *Sampling Techniques* (3rd ed.). New York: John Wiley & Sons.
2. Kish, L. (1965). *Survey Sampling*. New York: John Wiley & Sons.
3. Levy, P. S., & Lemeshow, S. (2013). *Sampling of Populations: Methods and Applications* (4th ed.). New York: Springer." "

BAB 9

ESTIMASI PARAMETER

Tujuan Pembelajaran

1. Mahasiswa mampu menjelaskan konsep dasar estimasi parameter, termasuk pengertian estimasi titik dan estimasi interval, serta perbedaannya dalam konteks analisis data kesehatan.
2. Mahasiswa dapat memahami dan menerapkan konsep tingkat kepercayaan dalam proses estimasi parameter, serta menginterpretasikan hasilnya secara tepat dalam situasi kesehatan masyarakat.
3. Mahasiswa mampu menghitung dan menyusun estimasi prevalensi suatu kondisi kesehatan, seperti stunting pada anak-anak, menggunakan data sampel dan metode statistik yang sesuai.
4. Mahasiswa mampu menginterpretasikan hasil estimasi parameter dalam konteks pengambilan keputusan dan perencanaan program kesehatan berbasis data.
5. Mahasiswa mampu mengidentifikasi faktor-faktor yang mempengaruhi ketepatan dan keakuratan estimasi parameter dalam studi kesehatan masyarakat.
6. Mahasiswa mampu mengkritisi dan mengevaluasi hasil estimasi parameter dari berbagai studi kesehatan berdasarkan tingkat kepercayaan dan interval estimasi yang diperoleh.
7. Mahasiswa mampu mengintegrasikan konsep estimasi parameter dalam penyusunan laporan

statistik kesehatan yang komprehensif dan komunikatif.

Pendahuluan

Estimasi parameter merupakan salah satu aspek fundamental dalam analisis data statistik, khususnya dalam bidang kesehatan masyarakat. Melalui proses ini, kita dapat memperkirakan karakteristik populasi berdasarkan data sampel yang terbatas, sehingga memungkinkan pengambilan keputusan yang lebih tepat dan efektif dalam perencanaan serta evaluasi program kesehatan. Dalam konteks kesehatan, estimasi prevalensi penyakit, tingkat kejadian, atau indikator kesehatan lainnya sering kali tidak dapat diukur secara langsung dari seluruh populasi karena keterbatasan sumber daya, waktu, dan biaya. Oleh karena itu, penggunaan metode statistik untuk memperkirakan parameter populasi menjadi sangat penting.

Pada bagian ini, kita akan mempelajari dua jenis utama estimasi parameter, yaitu estimasi titik dan estimasi interval. Estimasi titik memberikan satu nilai terbaik sebagai perkiraan parameter populasi, namun tidak menunjukkan tingkat ketidakpastian dari estimasi tersebut. Sebaliknya, estimasi interval menawarkan rentang nilai yang diyakini mengandung parameter sebenarnya dengan tingkat kepercayaan tertentu, biasanya 95%. Konsep tingkat kepercayaan ini menjadi kunci dalam menilai keandalan hasil estimasi dan membantu pengambil keputusan dalam konteks kesehatan masyarakat.

Urgensi dari pemahaman estimasi parameter ini sangat tinggi, terutama dalam situasi di mana data yang tersedia terbatas dan keputusan harus diambil secara cepat dan tepat. Sebagai contoh, dalam upaya mengatasi masalah stunting pada anak-anak, estimasi prevalensi dari sampel

yang representatif dapat memberikan gambaran yang cukup akurat tentang kondisi di seluruh populasi. Dengan demikian, hasil estimasi ini dapat digunakan untuk merancang intervensi yang tepat sasaran dan mengalokasikan sumber daya secara efisien.

Selain itu, pemahaman tentang tingkat kepercayaan dan interpretasi hasil estimasi sangat penting agar hasil tersebut tidak disalahgunakan atau diartikan secara keliru. Ketika kita menyampaikan hasil estimasi kepada pemangku kepentingan, harus mampu menjelaskan bahwa estimasi tersebut memiliki tingkat ketidakpastian tertentu dan bahwa rentang interval yang dihasilkan memberikan gambaran yang lebih lengkap daripada angka tunggal. Dengan demikian, bagian ini tidak hanya membekali mahasiswa dalam melakukan analisis statistik, tetapi juga dalam menyampaikan hasil secara komunikatif dan bertanggung jawab dalam konteks kesehatan masyarakat.

9.1 Estimasi Titik dan Estimasi Interval

Estimasi parameter adalah proses statistik yang digunakan untuk memperkirakan karakteristik populasi berdasarkan data sampel. Dua konsep utama dalam estimasi parameter adalah estimasi titik dan estimasi interval.

9.1.1 Estimasi Titik

Estimasi titik adalah nilai tunggal yang dihasilkan dari data sampel sebagai perkiraan terbaik dari parameter populasi yang tidak diketahui. Nilai ini disebut sebagai ""estimasi titik"" karena hanya memberikan satu angka sebagai perkiraan. Contohnya, jika kita ingin mengetahui prevalensi stunting di suatu daerah, kita dapat menghitung proporsi anak-anak yang mengalami stunting dalam sampel dan

menjadikannya sebagai estimasi titik prevalensi di seluruh populasi.

Contoh: Misalnya, dari sampel 500 anak di sebuah kabupaten, ditemukan 75 anak mengalami stunting. Maka, estimasi titik prevalensi stunting adalah:

$$p = \frac{x}{n} = \frac{75}{500} = 0,15 = 15\%$$

Estimasi ini memberikan gambaran bahwa sekitar 15% dari seluruh populasi anak-anak di daerah tersebut diperkirakan mengalami stunting. Namun, estimasi titik ini tidak menyertakan informasi tentang ketidakpastian atau tingkat kepercayaan terhadap nilai tersebut.

9.1.2 Estimasi Interval

Berbeda dengan estimasi titik, estimasi interval memberikan rentang nilai yang diyakini mengandung parameter populasi dengan tingkat kepercayaan tertentu. Rentang ini dikenal sebagai "interval kepercayaan" (confidence interval). Estimasi interval tidak hanya menyajikan satu angka, tetapi juga batas bawah dan batas atas yang menunjukkan ketidakpastian dari estimasi tersebut.

Contoh: Jika dari sampel yang sama, kita menghitung interval kepercayaan 95% untuk prevalensi stunting, hasilnya mungkin adalah 12% hingga 18%. Artinya, kita yakin 95% bahwa prevalensi stunting di seluruh populasi berada dalam rentang tersebut.

Estimasi interval ini sangat penting karena memberikan gambaran yang lebih lengkap tentang ketidakpastian estimasi dan membantu pengambil keputusan dalam merancang intervensi kesehatan. Semakin sempit

intervalnya, semakin akurat estimasi tersebut; sebaliknya, interval yang lebar menunjukkan ketidakpastian yang lebih besar.

9.2 Tingkat Kepercayaan

Tingkat kepercayaan adalah tingkat probabilitas bahwa interval kepercayaan yang dihitung dari sampel tertentu akan mengandung parameter populasi yang sebenarnya jika proses sampling diulang berkali-kali. Tingkat kepercayaan biasanya dinyatakan dalam persentase, seperti 90%, 95%, atau 99%.

9.2.1 Konsep Tingkat Kepercayaan

Misalnya, tingkat kepercayaan 95% berarti bahwa jika kita melakukan pengambilan sampel dan menghitung interval kepercayaan sebanyak 100 kali, maka sekitar 95 dari interval tersebut akan mengandung parameter populasi yang sebenarnya. Tingkat kepercayaan ini tidak menunjukkan probabilitas bahwa parameter berada dalam satu interval tertentu, melainkan menunjukkan keandalan metode estimasi yang digunakan.

9.2.2 Pemilihan Tingkat Kepercayaan

Dalam praktiknya, tingkat kepercayaan 95% adalah yang paling umum digunakan karena memberikan keseimbangan antara keakuratan dan ketidakpastian. Tingkat kepercayaan yang lebih tinggi, seperti 99%, akan menghasilkan interval yang lebih lebar, sehingga memberikan gambaran yang lebih konservatif tetapi kurang spesifik. Sebaliknya, tingkat kepercayaan yang lebih rendah, seperti 90%, menghasilkan interval yang lebih sempit tetapi dengan tingkat kepercayaan yang lebih rendah pula.

9.2.3 Pengaruh Tingkat Kepercayaan terhadap Estimasi

Penggunaan tingkat kepercayaan mempengaruhi interpretasi hasil estimasi. Sebagai contoh, jika kita menghitung estimasi prevalensi stunting dengan tingkat kepercayaan 95% dan mendapatkan interval 12% hingga 18%, maka kita dapat mengatakan bahwa kita yakin 95% bahwa prevalensi sebenarnya berada dalam rentang tersebut. Jika tingkat kepercayaan dinaikkan menjadi 99%, kemungkinan interval akan menjadi lebih lebar, misalnya 10% hingga 20%, yang menunjukkan ketidakpastian yang lebih besar tetapi dengan tingkat kepercayaan yang lebih tinggi.

9.3 Interpretasi Hasil Estimasi dalam Konteks Kesehatan

Menginterpretasikan hasil estimasi parameter dalam konteks kesehatan masyarakat memerlukan pemahaman yang mendalam tentang makna statistik dan implikasi praktisnya. Estimasi titik dan interval harus dilihat sebagai alat untuk membantu pengambilan keputusan, bukan sebagai angka mutlak yang pasti.

9.3.1 Interpretasi Estimasi Titik

Estimasi titik memberikan gambaran awal tentang karakteristik populasi. Sebagai contoh, jika estimasi titik prevalensi stunting adalah 15%, maka ini menunjukkan bahwa sekitar 15% dari populasi anak-anak diperkirakan mengalami stunting. Namun, angka ini harus dipandang sebagai perkiraan yang memiliki tingkat ketidakpastian, terutama jika sampel tidak sepenuhnya representatif atau ukuran sampel kecil.

9.3.2 Interpretasi Estimasi Interval

Interval kepercayaan memberikan rentang nilai yang mengandung parameter populasi dengan tingkat kepercayaan tertentu. Jika interval kepercayaan 95% untuk prevalensi stunting adalah 12% hingga 18%, maka kita dapat menyatakan bahwa berdasarkan data yang ada, kita yakin 95% bahwa prevalensi sebenarnya berada dalam rentang tersebut. Rentang ini membantu dalam menilai ketidakpastian dan risiko kesalahan dalam pengambilan keputusan.

9.3.3 Implikasi dalam Pengambilan Keputusan

Dalam konteks kesehatan masyarakat, hasil estimasi ini dapat digunakan untuk menentukan prioritas intervensi, alokasi sumber daya, dan evaluasi program. Misalnya, jika estimasi prevalensi stunting cukup tinggi dan intervalnya sempit, maka program penanggulangan stunting harus menjadi prioritas utama. Sebaliknya, jika intervalnya lebar dan mencakup nilai yang rendah, maka perlu dilakukan studi lebih lanjut untuk mendapatkan estimasi yang lebih akurat.

9.4 Contoh Aplikasi: Estimasi Prevalensi Stunting pada Anak-anak

Sebagai ilustrasi, kita akan membahas proses estimasi prevalensi stunting pada anak-anak di sebuah kabupaten. Misalnya, dilakukan survei terhadap 600 anak usia di bawah lima tahun, dan ditemukan 90 anak mengalami stunting.

9.4.1 Penghitungan Estimasi Titik

Estimasi titik prevalensi stunting dihitung dengan rumus proporsi sampel:

$$p = x/n$$

Dimana:

- p adalah estimasi titik prevalensi
- x adalah jumlah anak dalam sampel yang mengalami stunting
- n adalah jumlah total anak dalam sampel

Penyelesaian ilustrasi :

$$p = x/n = 90/600 = 0,15 = 15\%$$

Ini menunjukkan bahwa dari sampel, sekitar 15% anak-anak mengalami stunting. Namun, untuk memastikan bahwa angka ini mewakili seluruh populasi, kita perlu menghitung interval kepercayaan.

9.4.2 Penghitungan Estimasi Interval

Untuk menghitung interval kepercayaan 95% dari proporsi, kita dapat menggunakan rumus:

$$\bar{x} = p \pm Z * \sqrt{\left(\frac{p(1-p)}{n}\right)}$$
$$x = 0,15 \pm 1,96 * \sqrt{\left(\frac{0,15(1-0,15)}{600}\right)} = 0,179$$

Di mana:

- p adalah proporsi sampel (jumlah kejadian yang diinginkan dibagi dengan ukuran sampel)
- n adalah ukuran sampel
- Z adalah nilai untuk tingkat kepercayaan 95% yaitu 1,96

Menghitung standar error (SE):

$$SE = s/\sqrt{n}$$

$$SE = (\sqrt{[\frac{\sum (xi - x)^2}{n - 1}]})/(\sqrt{n})$$

$$SE = (\sqrt{[\frac{\sum (xi - x)^2}{n - 1}]})/(\sqrt{n})$$

$$SE = p (1-p)/600 = 0,15(1-0,15)/600 = 0,0002125 \quad *x=0,0146$$

$$SE = s/\sqrt{n}$$

Dimana :

- SE adalah standar eror
- s adalah deviasi standar
- n adalah ukuran sampel (jumlah data dalam sampel)

Maka, interval kepercayaan 95% adalah:

$$[0,15 \pm 1,96 \times 0,0146 \approx 0,15 \pm 0,0286]$$

Sehingga, rentang intervalnya adalah:

$$[(0,1214; 0,1786) \text{ atau } (12,14\%; 17,86\%)]$$

Dengan demikian, kita dapat menyatakan bahwa kita yakin 95% bahwa prevalensi stunting di seluruh populasi anak-

anak di kabupaten tersebut berkisar antara 12,14% hingga 17,86%.

9.4.3 Implikasi Hasil Estimasi

Hasil ini menunjukkan bahwa prevalensi stunting cukup tinggi dan cukup pasti dalam rentang tersebut. Data ini dapat digunakan oleh pihak kesehatan untuk merancang program intervensi yang lebih tepat sasaran, seperti peningkatan gizi, sanitasi, dan edukasi orang tua. Jika intervalnya sempit dan tinggi, maka prioritas utama adalah mengurangi angka stunting secara signifikan.

Estimasi parameter, baik estimasi titik maupun estimasi interval, merupakan alat penting dalam analisis data kesehatan masyarakat. Estimasi titik memberikan gambaran awal yang cepat dan sederhana, sementara estimasi interval menyajikan gambaran yang lebih lengkap tentang ketidakpastian dari estimasi tersebut. Tingkat kepercayaan adalah konsep kunci yang membantu kita memahami seberapa yakin kita terhadap hasil estimasi dan harus dipilih sesuai dengan kebutuhan analisis dan konteks pengambilan keputusan.

Dalam praktiknya, penggunaan estimasi parameter harus disertai interpretasi yang tepat dan pemahaman tentang batasan-batasan statistiknya. Hasil estimasi ini sangat berguna dalam perencanaan program kesehatan, pengalokasian sumber daya, dan evaluasi efektivitas intervensi. Dengan memahami dan mampu menerapkan konsep ini, mahasiswa akan mampu melakukan analisis data yang akurat dan bertanggung jawab dalam bidang kesehatan masyarakat.

Rangkuman

Estimasi parameter merupakan bagian penting dalam analisis data kesehatan masyarakat yang bertujuan untuk memperkirakan karakteristik populasi berdasarkan data sampel. Dalam proses ini, terdapat dua konsep utama, yaitu estimasi titik dan estimasi interval. Estimasi titik memberikan satu nilai tunggal sebagai perkiraan terbaik dari parameter populasi yang tidak diketahui, misalnya proporsi anak yang mengalami stunting dari sampel tertentu. Sedangkan estimasi interval menyajikan rentang nilai yang diyakini mengandung parameter populasi dengan tingkat kepercayaan tertentu, seperti interval kepercayaan 95%.

Tingkat kepercayaan adalah probabilitas bahwa interval yang dihitung dari sampel akan mengandung parameter sebenarnya jika proses sampling diulang berkali-kali. Tingkat ini biasanya dinyatakan dalam persentase, misalnya 95%, dan mempengaruhi lebar interval kepercayaan yang dihasilkan. Semakin tinggi tingkat kepercayaan, semakin lebar intervalnya, yang menunjukkan tingkat ketidakpastian yang lebih konservatif. Sebaliknya, tingkat kepercayaan yang lebih rendah menghasilkan interval yang lebih sempit tetapi dengan tingkat kepercayaan yang lebih rendah pula.

Dalam konteks kesehatan, interpretasi hasil estimasi harus dilakukan dengan hati-hati. Estimasi titik memberikan gambaran awal, tetapi tidak menunjukkan ketidakpastian, sementara estimasi interval memberikan gambaran tentang rentang nilai yang mungkin dan tingkat kepercayaan terhadapnya. Hasil ini sangat berguna untuk pengambilan keputusan, perencanaan program, dan evaluasi intervensi kesehatan masyarakat. Sebagai contoh, dalam estimasi prevalensi stunting, perhitungan interval kepercayaan

membantu menentukan tingkat ketidakpastian dan risiko kesalahan dalam pengambilan kebijakan.

Contoh aplikasi nyata adalah survei terhadap anak-anak di sebuah daerah, di mana dari sampel 600 anak ditemukan 90 mengalami stunting. Estimasi titik prevalensi adalah 15%, dan dengan perhitungan statistik, diperoleh interval kepercayaan 12,14% hingga 17,86%. Hasil ini menunjukkan bahwa prevalensi stunting cukup tinggi dan estimasi ini dapat digunakan sebagai dasar perencanaan intervensi. Dengan memahami konsep estimasi parameter, mahasiswa mampu melakukan analisis data yang akurat dan bertanggung jawab dalam bidang kesehatan masyarakat.

Latihan Mahasiswa

Soal Essay

1. Jelaskan perbedaan utama antara estimasi titik dan estimasi interval dalam analisis data kesehatan masyarakat. Berikan contoh penggunaannya dalam studi prevalensi penyakit.
2. Mengapa tingkat kepercayaan penting dalam proses estimasi parameter? Jelaskan bagaimana tingkat kepercayaan mempengaruhi interpretasi hasil estimasi dalam konteks pengambilan keputusan kesehatan.
3. Sebuah sampel berjumlah 800 anak di sebuah kabupaten menunjukkan 120 anak mengalami stunting. Hitunglah estimasi titik prevalensi stunting dan buatlah estimasi interval kepercayaan 95% dari data tersebut.
4. Diskusikan faktor-faktor apa saja yang dapat mempengaruhi ketepatan dan keakuratan estimasi parameter dalam studi kesehatan masyarakat.

Berikan minimal dua faktor dan jelaskan pengaruhnya.

5. Bagaimana cara mahasiswa dapat mengkritisi hasil estimasi parameter dari sebuah studi kesehatan berdasarkan interval kepercayaan yang diperoleh? Jelaskan langkah-langkahnya secara umum.

Soal Pilihan Berganda

1. Apa yang dimaksud dengan estimasi titik dalam statistik kesehatan masyarakat?
 - a. Rentang nilai yang mengandung parameter populasi dengan tingkat kepercayaan tertentu
 - b. Nilai tunggal yang digunakan sebagai perkiraan parameter populasi
 - c. Nilai rata-rata dari seluruh populasi yang sebenarnya
 - d. Nilai yang dihasilkan dari pengukuran langsung di lapangan
2. Manakah dari berikut ini yang merupakan contoh estimasi interval?
 - a. Proporsi anak yang mengalami stunting adalah 15%
 - b. Rentang 12% hingga 18% sebagai perkiraan prevalensi stunting dengan tingkat kepercayaan 95%
 - c. Jumlah anak yang mengalami stunting dalam sampel sebanyak 75 dari 500
 - d. Rata-rata usia penderita hipertensi adalah 50 tahun
3. Tingkat kepercayaan 95% berarti:

- a. Ada kemungkinan 95% bahwa parameter populasi berada dalam satu interval tertentu
 - b. 95% dari sampel akan mengandung parameter populasi
 - c. Jika pengambilan sampel diulang, 95% dari interval yang dihitung akan mengandung parameter sebenarnya
 - d. Hanya 5% dari data yang valid
4. Dalam konteks estimasi prevalensi, apa arti dari interval kepercayaan yang lebih lebar?
 - a. Estimasi lebih akurat
 - b. Ketidakpastian estimasi lebih tinggi
 - c. Estimasi lebih pasti
 - d. Data sampel lebih besar
5. Jika estimasi titik prevalensi stunting adalah 15% dan interval kepercayaannya adalah 12% sampai 18%, maka:
 - a. Kita yakin 95% bahwa prevalensi sebenarnya berada di antara 12% dan 18%
 - b. Prevalensi pasti 15%
 - c. Interval ini menunjukkan ketidakpastian yang rendah dan tidak perlu diinterpretasikan lebih jauh
 - d. Data sampel tidak representatif
6. Faktor berikut mana yang dapat mempengaruhi ketepatan estimasi parameter?
 - a. Ukuran sampel yang kecil
 - b. Variasi dalam data yang tinggi
 - c. Sampel yang tidak representatif
 - d. Semua jawaban benar

7. Bagaimana mahasiswa dapat menggunakan hasil estimasi parameter untuk pengambilan keputusan program kesehatan?
 - a. Mengabaikan interval kepercayaan dan hanya memperhatikan estimasi titik
 - b. Menggunakan estimasi dan interval kepercayaan untuk menilai tingkat ketidakpastian dan merancang intervensi yang sesuai
 - c. Menggunakan hanya data dari studi lain tanpa memperhatikan hasil sendiri
 - d. Mengandalkan intuisi tanpa analisis statistik
8. Dalam menghitung estimasi interval untuk proporsi, rumus yang digunakan melibatkan:
 - a. Nilai Z dari distribusi normal dan standar error dari proporsi
 - b. Rata-rata sampel dan simpangan baku
 - c. Nilai t dari distribusi t-Student dan varians sampel
 - d. Nilai chi-square dan derajat kebebasan
9. Apa yang harus dilakukan jika interval kepercayaan dari sebuah studi sangat lebar?
 - a. Menganggap hasilnya sangat akurat
 - b. Melakukan studi ulang dengan sampel yang lebih besar atau lebih representatif
 - c. Mengabaikan hasil tersebut karena tidak berguna
 - d. Mengurangi tingkat kepercayaan menjadi 90%
10. Dalam konteks estimasi prevalensi, apa manfaat utama dari menggunakan estimasi interval?

- a. Memberikan satu angka pasti dari prevalensi
- b. Menunjukkan ketidakpastian dan tingkat kepercayaan terhadap estimasi tersebut
- c. Mengurangi kebutuhan pengumpulan data lapangan
- d. Mengganti analisis statistik yang kompleks

Soal Project / Studi Kasus

1. Sebuah tim peneliti melakukan survei terhadap 1.200 anak di sebuah kabupaten dan menemukan 180 anak mengalami stunting. Hitunglah estimasi titik prevalensi stunting dan buatlah estimasi interval kepercayaan 95%. Jelaskan langkah-langkah perhitungannya dan interpretasikan hasilnya dalam konteks pengambilan keputusan program kesehatan.
2. Anda diminta untuk menilai sebuah studi yang melaporkan estimasi prevalensi anemia pada ibu hamil di sebuah kota dengan estimasi titik 20% dan interval kepercayaan 95% dari 15% sampai 25%. Kritikilah hasil tersebut dari segi keandalan dan ketepatan estimasi, serta berikan saran untuk studi selanjutnya agar hasilnya lebih akurat dan dapat diandalkan.

REFERENSI

1. Daniel, W. W. (2018). *Biostatistics: A Foundation for Analysis in the Health Sciences* (11th ed.). John Wiley & Sons.
2. Lohr, S. L. (2019). *Sampling: Design and Analysis* (2nd ed.). CRC Press.
3. Newcombe, R. G. (2019). *Confidence Intervals for Proportions and Means*. In *Statistical Methods for Health Care Research* (pp. 245-268). Springer.

4. World Health Organization. (2020). *Guidelines on Interventions to Address Stunting in Children*. WHO Press." "

BAB 10

UJI HIPOTESIS STATISTIK

Tujuan Pembelajaran

1. Mahasiswa mampu memahami konsep dasar dan pentingnya uji hipotesis dalam analisis data statistik di bidang kesehatan masyarakat.
2. Mahasiswa dapat menjelaskan pengertian hipotesis nol dan hipotesis alternatif serta peran keduanya dalam proses pengujian statistik.
3. Mahasiswa mampu mengidentifikasi dan membedakan jenis kesalahan statistik, yaitu kesalahan Tipe I dan Tipe II, serta memahami implikasinya dalam pengambilan keputusan.
4. Mahasiswa dapat mengikuti langkah-langkah sistematis dalam melakukan uji hipotesis, mulai dari perumusan hipotesis, penentuan tingkat signifikansi, penghitungan statistik uji, hingga pengambilan keputusan.
5. Mahasiswa mampu menerapkan konsep uji hipotesis dalam konteks penelitian kesehatan, seperti menguji perbedaan tingkat pengetahuan kesehatan antara dua komunitas.
6. Mahasiswa dapat menginterpretasikan hasil uji hipotesis secara tepat dan relevan dalam konteks kesehatan masyarakat, serta menyusun laporan hasil yang jelas dan akurat.
7. Mahasiswa mampu mengkritisi dan mengevaluasi penggunaan uji hipotesis dalam studi kasus dan penelitian di bidang kesehatan, serta memahami batasan dan asumsi yang harus dipenuhi.

Pendahuluan

Uji hipotesis merupakan salah satu fondasi utama dalam analisis statistik yang digunakan untuk membuat keputusan berdasarkan data. Dalam bidang kesehatan masyarakat, pengambilan keputusan yang tepat sangat bergantung pada analisis data yang akurat dan objektif. Oleh karena itu, pemahaman mendalam tentang konsep uji hipotesis menjadi sangat penting bagi para profesional dan peneliti di bidang ini. Melalui proses uji hipotesis, kita dapat menguji asumsi awal mengenai suatu populasi atau hubungan antar variabel, sehingga hasilnya dapat digunakan sebagai dasar dalam pengambilan kebijakan, intervensi, maupun penelitian lebih lanjut.

Pengujian hipotesis memungkinkan kita untuk menentukan apakah data yang diperoleh cukup kuat untuk mendukung suatu dugaan atau pernyataan tertentu. Misalnya, dalam konteks promosi kesehatan, kita mungkin ingin mengetahui apakah program edukasi tertentu benar-benar meningkatkan tingkat pengetahuan masyarakat. Untuk menjawab pertanyaan ini secara ilmiah, kita harus merumuskan hipotesis nol dan hipotesis alternatif, kemudian melakukan analisis statistik yang sesuai. Jika hasilnya menunjukkan bahwa data mendukung hipotesis alternatif, maka kita dapat menyimpulkan adanya perbedaan yang signifikan antara kelompok yang diuji.

Selain itu, penguasaan konsep kesalahan Tipe I dan Tipe II sangat penting agar hasil pengujian tidak menyesatkan. Kesalahan Tipe I terjadi ketika kita secara keliru menolak hipotesis nol yang sebenarnya benar, sedangkan kesalahan Tipe II terjadi ketika kita gagal menolak hipotesis nol yang sebenarnya salah. Memahami kedua jenis kesalahan ini

membantu peneliti dalam menentukan tingkat signifikansi dan interpretasi hasil secara tepat.

Langkah-langkah dalam melakukan uji hipotesis juga harus dipahami secara sistematis. Mulai dari perumusan hipotesis, penentuan tingkat signifikansi (α), penghitungan statistik uji, hingga pengambilan keputusan berdasarkan nilai p atau nilai kritis, semuanya harus dilakukan secara cermat dan sesuai prosedur. Dengan penguasaan langkah-langkah ini, mahasiswa akan mampu melakukan analisis statistik yang valid dan dapat dipertanggungjawabkan.

Contoh aplikasi yang relevan dalam bidang kesehatan adalah menguji perbedaan tingkat pengetahuan kesehatan antara dua komunitas. Melalui studi ini, mahasiswa dapat belajar bagaimana mengaplikasikan konsep uji hipotesis secara nyata, serta memahami pentingnya interpretasi hasil dalam konteks pengambilan kebijakan kesehatan. Dengan demikian, pemahaman tentang uji hipotesis tidak hanya bersifat teoritis, tetapi juga sangat praktis dan esensial dalam praktik lapangan dan penelitian kesehatan masyarakat.

Pembelajaran ini diharapkan mampu membekali mahasiswa dengan kemampuan analisis statistik yang kritis dan akurat, sehingga mereka dapat berkontribusi secara efektif dalam pengembangan ilmu pengetahuan dan peningkatan kualitas layanan kesehatan di masyarakat.

10.1 Hipotesis Nol dan Hipotesis Alternatif

10.1.1 Pengertian Hipotesis Nol dan Hipotesis Alternatif

Dalam analisis statistik, konsep dasar yang harus dipahami adalah adanya dua pernyataan utama yang dikenal sebagai hipotesis nol (H_0) dan hipotesis alternatif (H_1 atau H_a).

Hipotesis nol merupakan pernyataan awal yang biasanya menyatakan tidak adanya efek, tidak adanya perbedaan, atau tidak adanya hubungan antara variabel yang sedang diteliti. Sebagai contoh, dalam penelitian kesehatan masyarakat, hipotesis nol sering kali menyatakan bahwa tidak ada perbedaan tingkat pengetahuan kesehatan antara dua komunitas yang berbeda.

Sebaliknya, hipotesis alternatif adalah pernyataan yang diajukan sebagai kemungkinan adanya efek, perbedaan, atau hubungan yang nyata. Hipotesis ini biasanya merupakan apa yang ingin dibuktikan oleh peneliti. Dalam konteks yang sama, hipotesis alternatif menyatakan bahwa terdapat perbedaan tingkat pengetahuan kesehatan antara kedua komunitas tersebut.

Penggunaan kedua hipotesis ini bertujuan untuk menguji asumsi awal dan menentukan apakah data yang diperoleh cukup kuat untuk mendukung pernyataan alternatif tersebut. Menurut Gravetter dan Wallnau (2017), proses pengujian ini merupakan dasar dari inferensi statistik yang memungkinkan kita membuat keputusan berdasarkan data sampel terhadap populasi secara umum.

10.1.2 Peran Hipotesis Nol dan Alternatif dalam Pengujian Statistik

Dalam proses pengujian hipotesis, hipotesis nol berfungsi sebagai asumsi dasar yang akan diuji kebenarannya. Jika data menunjukkan bahwa hipotesis nol tidak dapat ditolak, maka kita menyimpulkan bahwa tidak ada cukup bukti untuk mendukung hipotesis alternatif. Sebaliknya, jika data menunjukkan bahwa hipotesis nol harus ditolak, maka kita menerima hipotesis alternatif sebagai penjelasan yang lebih sesuai dengan data.

Contohnya, dalam uji perbedaan tingkat pengetahuan kesehatan antara dua komunitas, hipotesis nol menyatakan bahwa tidak ada perbedaan tingkat pengetahuan antara keduanya, sedangkan hipotesis alternatif menyatakan bahwa ada perbedaan. Jika hasil analisis statistik menunjukkan bahwa perbedaan tersebut signifikan secara statistik, maka kita menolak hipotesis nol dan menerima hipotesis alternatif.

Peran utama dari kedua hipotesis ini adalah untuk menyediakan kerangka kerja yang objektif dan sistematis dalam pengambilan keputusan statistik. Tanpa adanya hipotesis nol dan alternatif, proses pengujian akan menjadi tidak terarah dan subjektif. Dengan demikian, kedua hipotesis ini menjadi fondasi utama dalam analisis inferensial yang digunakan dalam penelitian kesehatan masyarakat (Field, 2018).

10.1.3 Contoh Kasus: Uji Perbedaan Tingkat Pengetahuan Kesehatan

Misalnya, sebuah penelitian ingin mengetahui apakah program edukasi kesehatan yang baru dilaksanakan di dua desa berbeda mempengaruhi tingkat pengetahuan masyarakat. Hipotesis nol yang diajukan adalah: **H_0 : Tidak ada perbedaan tingkat pengetahuan kesehatan antara desa A dan desa B.**

Sedangkan hipotesis alternatifnya adalah: **H_1 : Ada perbedaan tingkat pengetahuan kesehatan antara desa A dan desa B.**

Dalam pengujian ini, data dikumpulkan melalui kuesioner yang mengukur tingkat pengetahuan masyarakat di kedua desa. Setelah data dianalisis menggunakan uji statistik yang sesuai (misalnya uji t dua sampel independen), hasilnya

akan menunjukkan apakah perbedaan tersebut signifikan secara statistik atau tidak. Jika p-value yang diperoleh kurang dari tingkat signifikansi yang ditetapkan (misalnya 0,05), maka kita menolak H_0 dan menyimpulkan bahwa program edukasi tersebut memiliki pengaruh berbeda di kedua desa.

10.2 Kesalahan Tipe I dan Tipe II

10.2.1 Definisi dan Implikasi Kesalahan Tipe I

Kesalahan Tipe I terjadi ketika kita secara keliru menolak hipotesis nol yang sebenarnya benar. Dalam konteks pengujian statistik, ini berarti kita menyatakan ada efek atau perbedaan padahal kenyataannya tidak ada. Kesalahan ini sering disebut sebagai "false positive" karena hasilnya menunjukkan adanya efek padahal sebenarnya tidak.

Contoh konkret dalam penelitian kesehatan adalah ketika kita menyimpulkan bahwa program imunisasi meningkatkan tingkat imunisasi masyarakat, padahal kenyataannya tidak ada pengaruh. Jika tingkat signifikansi (α) yang digunakan adalah 0,05, maka ada peluang 5% kita melakukan kesalahan Tipe I secara tidak sengaja. Oleh karena itu, tingkat α harus dipilih dengan hati-hati agar tidak terlalu tinggi, yang dapat meningkatkan risiko kesalahan Tipe I (Fisher, 2019).

10.2.2 Definisi dan Implikasi Kesalahan Tipe II

Kesalahan Tipe II terjadi ketika kita gagal menolak hipotesis nol padahal sebenarnya hipotesis nol itu salah. Dalam hal ini, kita melewatkan keberadaan efek atau perbedaan yang nyata. Kesalahan ini dikenal sebagai "false negative" dan dapat menyebabkan kita melewatkan peluang untuk

mengidentifikasi intervensi yang efektif atau masalah kesehatan yang penting.

Sebagai contoh, jika sebuah studi gagal mendeteksi bahwa program promosi kesehatan meningkatkan pengetahuan masyarakat padahal kenyataannya memang ada peningkatan, maka kita melakukan kesalahan Tipe II. Risiko ini biasanya meningkat jika ukuran sampel terlalu kecil atau tingkat signifikansi yang digunakan terlalu konservatif (Cohen, 2018).

10.2.3 Perbandingan dan Implikasi Praktis

Memahami kedua jenis kesalahan ini sangat penting dalam pengambilan keputusan di bidang kesehatan masyarakat. Biasanya, ada trade-off antara risiko Tipe I dan Tipe II. Jika kita menetapkan tingkat signifikansi yang sangat rendah (misalnya 0,01), kita mengurangi risiko Tipe I tetapi meningkatkan risiko Tipe II. Sebaliknya, tingkat signifikansi yang lebih tinggi (misalnya 0,10) akan meningkatkan kemungkinan menolak H_0 secara keliru, tetapi menurunkan risiko melewatkan efek yang sebenarnya ada.

Dalam praktiknya, peneliti harus memilih tingkat signifikansi yang sesuai dengan konteks penelitian dan konsekuensi dari kesalahan tersebut. Misalnya, dalam studi yang berkaitan dengan pengendalian wabah penyakit, risiko Tipe I mungkin lebih dihindari karena menimbulkan tindakan yang tidak perlu, sementara dalam studi pengujian pengobatan baru, risiko Tipe II harus diminimalkan agar tidak melewatkan pengobatan yang efektif (Nuzzo, 2014).

10.3 Langkah-langkah Uji Hipotesis

10.3.1 Perumusan Hipotesis

Langkah pertama dalam uji hipotesis adalah merumuskan hipotesis nol dan hipotesis alternatif secara jelas dan spesifik. Dalam konteks penelitian kesehatan, ini melibatkan identifikasi variabel utama dan hubungan yang ingin diuji. Misalnya, jika ingin menguji apakah ada perbedaan tingkat pengetahuan kesehatan antara dua komunitas, maka:

- H_0 : Tidak ada perbedaan tingkat pengetahuan kesehatan antara komunitas A dan B.
- H_1 : Ada perbedaan tingkat pengetahuan kesehatan antara komunitas A dan B.

Perumusan ini harus didasarkan pada teori, literatur, dan tujuan penelitian yang jelas.

10.3.2 Penentuan Tingkat Signifikansi (α)

Setelah hipotesis dirumuskan, langkah berikutnya adalah menentukan tingkat signifikansi (α), yaitu probabilitas maksimum kita bersedia menerima kesalahan Tipe I. Biasanya, tingkat α yang digunakan adalah 0,05, yang berarti kita bersedia menerima risiko 5% untuk menolak H_0 secara keliru. Penentuan tingkat ini harus mempertimbangkan konsekuensi dari kesalahan Tipe I dan Tipe II dalam konteks penelitian.

10.3.3 Pengumpulan Data dan Penghitungan Statistik Uji

Langkah selanjutnya adalah pengumpulan data yang relevan dan representatif. Data ini kemudian dianalisis menggunakan uji statistik yang sesuai dengan jenis data dan

desain penelitian. Dalam contoh uji perbedaan tingkat pengetahuan, uji t dua sampel independen sering digunakan jika data berdistribusi normal dan varians kedua kelompok homogen.

Penghitungan statistik uji dilakukan dengan rumus yang sesuai, dan hasilnya dibandingkan dengan nilai kritis dari distribusi statistik yang relevan (misalnya distribusi t). Alternatifnya, p-value dihitung dan dibandingkan dengan α untuk menentukan keputusan.

10.3.4 Pengambilan Keputusan dan Interpretasi Hasil

Berdasarkan hasil analisis, kita membuat keputusan:

- Jika $p\text{-value} \leq \alpha$, maka tolak H_0 dan simpulkan bahwa ada perbedaan yang signifikan secara statistik.
- Jika $p\text{-value} > \alpha$, maka gagal menolak H_0 dan simpulkan bahwa tidak ada cukup bukti untuk menyatakan adanya perbedaan.

Interpretasi hasil harus dilakukan dengan mempertimbangkan konteks penelitian dan tingkat signifikansi yang dipilih. Penting juga untuk menyampaikan bahwa hasil statistik tidak selalu menunjukkan hubungan sebab-akibat, melainkan hanya menunjukkan adanya asosiasi yang signifikan secara statistik.

10.3.5 Penyusunan Laporan Hasil

Langkah terakhir adalah menyusun laporan hasil uji hipotesis secara sistematis dan jelas. Laporan harus mencakup hipotesis yang diuji, tingkat signifikansi yang digunakan, statistik uji, p-value, dan interpretasi hasil. Penyajian data dalam bentuk tabel dan grafik sangat membantu dalam memperjelas hasil analisis.

10.4 Contoh Aplikasi: Uji Perbedaan Tingkat Pengetahuan Kesehatan Antara Dua Komunitas

10.4.1 Desain Penelitian dan Pengumpulan Data

Misalnya, sebuah penelitian dilakukan untuk menguji apakah program edukasi kesehatan yang baru dilaksanakan di dua desa berbeda mempengaruhi tingkat pengetahuan masyarakat. Peneliti mengumpulkan data melalui kuesioner yang berisi pertanyaan-pertanyaan terkait pengetahuan kesehatan. Sampel diambil secara acak dari masing-masing desa, dengan jumlah responden sebanyak 50 orang di setiap desa.

10.4.2 Hipotesis yang Diajukan

- H_0 : Tidak ada perbedaan rata-rata skor pengetahuan kesehatan antara desa A dan desa B.
- H_1 : Ada perbedaan rata-rata skor pengetahuan kesehatan antara desa A dan desa B.

10.4.3 Analisis Statistik dan Hasil

Data dianalisis menggunakan uji t dua sampel independen. Setelah penghitungan, diperoleh:

- Rata-rata skor desa A = 75, dengan simpangan baku = 10.
- Rata-rata skor desa B = 68, dengan simpangan baku = 12.
- Nilai t hitung = 2,45.
- Derajat kebebasan (df) = 98.
- Nilai p = 0,016.

Karena p -value (0,016) lebih kecil dari tingkat signifikansi 0,05, maka kita menolak H_0 dan menyimpulkan bahwa terdapat perbedaan yang signifikan secara statistik antara tingkat pengetahuan kesehatan di kedua desa.

10.4.4 Interpretasi dan Implikasi

Hasil ini menunjukkan bahwa program edukasi kesehatan di desa A lebih efektif dalam meningkatkan pengetahuan masyarakat dibandingkan desa B. Temuan ini dapat digunakan sebagai dasar untuk memperluas program edukasi di desa B atau melakukan evaluasi terhadap metode yang digunakan di desa B agar lebih efektif.

Rangkuman

Pada bab ini, kita membahas konsep dasar pengujian hipotesis dalam statistik, yang merupakan fondasi penting dalam analisis data penelitian kesehatan masyarakat. Pengujian hipotesis melibatkan dua pernyataan utama, yaitu hipotesis nol (H_0) dan hipotesis alternatif (H_1). Hipotesis nol biasanya menyatakan tidak adanya efek, perbedaan, atau hubungan antara variabel yang diteliti, sedangkan hipotesis alternatif menyatakan adanya efek atau perbedaan tersebut. Kedua hipotesis ini digunakan untuk menguji asumsi awal dan menentukan apakah data yang diperoleh cukup kuat untuk mendukung pernyataan yang diajukan.

Berikut adalah poin-poin utama yang dirangkum dari bab ini:

1. **Pengertian Hipotesis Nol dan Alternatif** Hipotesis nol menyatakan tidak adanya perbedaan atau efek, misalnya tidak ada perbedaan tingkat pengetahuan kesehatan antara dua komunitas. Sebaliknya,

hipotesis alternatif menyatakan adanya perbedaan atau efek tersebut. Pengujian dilakukan untuk menilai kebenaran dari hipotesis nol berdasarkan data sampel.

2. **Peran dalam Pengujian Statistik** Hipotesis nol berfungsi sebagai asumsi dasar yang akan diuji, dan keputusan diambil berdasarkan hasil analisis statistik. Jika data menunjukkan bahwa H_0 tidak dapat ditolak, maka tidak ada cukup bukti untuk mendukung H_1 . Jika data menunjukkan bahwa H_0 harus ditolak, maka H_1 diterima sebagai penjelasan yang lebih sesuai.
3. **Contoh Kasus** Sebuah studi menguji pengaruh program edukasi kesehatan di dua desa. Hipotesis nol menyatakan tidak ada perbedaan tingkat pengetahuan, sementara hipotesis alternatif menyatakan ada perbedaan. Setelah analisis statistik, jika p-value kurang dari tingkat signifikansi (misalnya 0,05), maka H_0 ditolak dan disimpulkan bahwa program tersebut memiliki pengaruh berbeda di kedua desa.
4. **Kesalahan Tipe I dan Tipe II**
 - **Kesalahan Tipe I** terjadi saat kita secara keliru menolak H_0 padahal sebenarnya benar, yang dikenal sebagai false positive. Risiko ini diatur oleh tingkat signifikansi (α).
 - **Kesalahan Tipe II** terjadi saat kita gagal menolak H_0 padahal sebenarnya salah, yaitu melewatkan efek nyata (false negative). Risiko ini dipengaruhi oleh ukuran sampel dan tingkat signifikansi.
5. **Perbandingan dan Implikasi Praktis** Peneliti harus menyeimbangkan risiko Tipe I dan Tipe II sesuai konteks penelitian. Dalam studi pengendalian wabah, menghindari Tipe I lebih diutamakan, sedangkan dalam pengujian pengobatan baru, mengurangi Tipe

II menjadi prioritas agar tidak melewatkan pengobatan yang efektif.

6. **Langkah-langkah Uji Hipotesis** Meliputi perumusan hipotesis, penentuan tingkat signifikansi, pengumpulan data, analisis statistik, pengambilan keputusan, dan penyusunan laporan hasil. Proses ini memastikan pengambilan keputusan yang objektif dan sistematis berdasarkan data.
7. **Contoh Aplikasi** Dalam studi tentang pengaruh program edukasi, data dikumpulkan dari sampel desa, dianalisis menggunakan uji t dua sampel independen, dan hasilnya menunjukkan adanya perbedaan signifikan jika $p\text{-value} < 0,05$. Hasil ini digunakan untuk menyimpulkan efektivitas program dan memberikan dasar rekomendasi.

Dengan memahami konsep dasar ini, peneliti dan praktisi kesehatan masyarakat dapat melakukan pengujian statistik secara tepat, menginterpretasi hasil secara benar, serta membuat keputusan yang didukung data. Penggunaan hipotesis nol dan alternatif yang tepat sangat penting untuk memastikan validitas dan reliabilitas hasil penelitian serta pengambilan kebijakan yang efektif.

Latihan Mahasiswa

Soal Essay

1. Jelaskan pengertian hipotesis nol dan hipotesis alternatif serta peran keduanya dalam pengujian statistik di bidang kesehatan masyarakat.
2. Sebutkan dan jelaskan dua jenis kesalahan statistik yang dapat terjadi dalam pengujian hipotesis, serta berikan contoh nyata dalam konteks penelitian kesehatan.

3. Uraikan langkah-langkah sistematis yang harus dilakukan dalam melakukan uji hipotesis, mulai dari perumusan hipotesis hingga pengambilan keputusan.
4. Berikan contoh kasus penelitian di bidang kesehatan masyarakat yang melibatkan pengujian hipotesis, dan jelaskan bagaimana proses pengujian dilakukan serta interpretasi hasilnya.
5. Diskusikan pentingnya memilih tingkat signifikansi (α) dalam pengujian hipotesis dan bagaimana pengaruhnya terhadap risiko kesalahan Tipe I dan Tipe II.

Soal Pilihan Berganda

1. Apa yang dimaksud dengan hipotesis nol (H_0)?
 - a. Pernyataan yang menyatakan adanya efek atau perbedaan
 - b. Pernyataan awal yang menyatakan tidak adanya efek atau perbedaan
 - c. Pernyataan yang tidak relevan dalam analisis statistik
 - d. Pernyataan yang selalu diterima tanpa pengujian
2. Hipotesis alternatif (H_1) biasanya digunakan untuk:
 - a. Menyatakan tidak adanya hubungan antara variabel
 - b. Menyatakan adanya efek, perbedaan, atau hubungan yang nyata
 - c. Menolak semua kemungkinan dalam penelitian
 - d. Menyatakan bahwa data tidak valid
3. Kesalahan Tipe I terjadi ketika:

- a. Gagal menolak hipotesis nol yang sebenarnya salah
 - b. Menolak hipotesis nol yang sebenarnya benar
 - c. Gagal menolak hipotesis nol yang salah
 - d. Menolak hipotesis alternatif yang benar
4. Jika p-value lebih kecil dari tingkat signifikansi (α), maka:
 - a. Gagal menolak H_0
 - b. Menolak H_0
 - c. Tidak ada keputusan yang diambil
 - d. H_0 diterima secara pasti
5. Dalam pengujian hipotesis, tingkat signifikansi (α) biasanya dipilih sebesar:
 - a. 0,01
 - b. 0,05
 - c. 0,10
 - e. Semua jawaban benar tergantung konteks
6. Kesalahan Tipe II disebut juga sebagai:
 - a. False positive
 - b. False negative
 - c. True positive
 - d. True negative
7. Dalam uji hipotesis, langkah pertama yang harus dilakukan adalah:
 - a. Mengumpulkan data
 - b. Menentukan tingkat signifikansi
 - c. Merumuskan hipotesis nol dan alternatif
 - d. Menghitung statistik uji

8. Contoh penelitian yang tepat untuk uji hipotesis adalah:
 - a. Mengamati pola makan masyarakat
 - b. Menguji apakah program imunisasi meningkatkan tingkat imunisasi
 - c. Mengumpulkan data tanpa analisis statistik
 - d. Menyusun laporan surveilans
9. Jika hasil analisis menunjukkan $p\text{-value} = 0,08$ dan $\alpha = 0,05$, maka:
 - a. Tolak H_0
 - b. Gagal menolak H_0
 - c. H_0 diterima
 - d. Tidak dapat disimpulkan
10. Mengapa penting memahami risiko kesalahan Tipe I dan Tipe II dalam pengujian hipotesis?
 - a. Untuk menentukan ukuran sampel
 - b. Untuk menghindari semua kesalahan
 - c. Untuk menyesuaikan tingkat signifikansi dan interpretasi hasil
 - d. Untuk mempercepat proses analisis

Soal Studi Kasus / Project

1. Sebuah penelitian ingin mengetahui apakah pemberian suplemen vitamin C berpengaruh terhadap penurunan angka kejadian flu di sebuah komunitas. Data diambil dari dua kelompok, satu menerima suplemen dan satu tidak. Jelaskan langkah-langkah yang harus dilakukan untuk menguji hipotesis tersebut, mulai dari perumusan hipotesis hingga interpretasi hasilnya.

2. Dalam sebuah studi, peneliti ingin membandingkan rata-rata tekanan darah antara pasien yang mengikuti program diet sehat dan yang tidak. Data dikumpulkan dari 30 pasien di masing-masing kelompok. Jelaskan prosedur analisis statistik yang tepat dan bagaimana peneliti dapat menyimpulkan apakah program diet berpengaruh terhadap tekanan darah pasien.

REFERENSI

1. Cohen, J. (2018). *Statistical Power Analysis for the Behavioral Sciences* (2nd ed.). Routledge.
2. Fisher, R. A. (2019). *Statistical Methods for Research Workers*. Oliver and Boyd.
3. Field, A. (2018). *Discovering Statistics Using IBM SPSS Statistics* (5th ed.). Sage Publications.
4. Gravetter, F. J., & Wallnau, L. B. (2017). *Statistics for the Behavioral Sciences* (10th ed.). Cengage Learning.
5. Nuzzo, R. (2014). Scientific method: Statistical errors. *Nature*, 506(7487), 150–152.
<https://doi.org/10.1038/506150a> "

BAB 11

UJI STATISTIK PARAMETRIK

Tujuan Pembelajaran

1. Memahami konsep dasar dan prinsip dasar dari uji statistik parametrik yang digunakan dalam penelitian kesehatan, termasuk uji t satu sampel, dua sampel independen, dan berpasangan, serta analisis varians satu arah (ANOVA).
2. Mampu menjelaskan langkah-langkah prosedural dalam melakukan uji statistik parametrik, mulai dari pengumpulan data, pengujian asumsi, hingga interpretasi hasil uji secara tepat dan relevan dalam konteks penelitian kesehatan.
3. Mengidentifikasi kondisi dan situasi penelitian yang sesuai untuk menerapkan uji t dan ANOVA, serta memahami batasan dan asumsi yang harus dipenuhi agar hasil uji valid dan dapat dipercaya.
4. Meningkatkan kemampuan dalam membaca dan menafsirkan output hasil uji statistik dari software statistik, seperti nilai p, nilai statistik uji, dan interval kepercayaan, serta mengaitkannya dengan pertanyaan penelitian yang sedang dianalisis.
5. Mengaplikasikan pengetahuan tentang uji statistik parametrik dalam menyusun laporan hasil penelitian kesehatan, termasuk penulisan kesimpulan yang didukung oleh data dan analisis statistik yang benar.
6. Mengembangkan kemampuan kritis dalam menilai keabsahan dan keandalan hasil uji statistik, serta memahami implikasi hasil tersebut terhadap

pengambilan keputusan dalam program promosi dan pencegahan kesehatan.

7. Menyiapkan diri untuk melakukan analisis statistik secara mandiri menggunakan software statistik yang umum digunakan, serta mampu mengintegrasikan hasil analisis tersebut ke dalam laporan penelitian yang komprehensif dan sistematis.

Pendahuluan

Uji statistik parametrik merupakan salah satu bagian penting dalam analisis data penelitian kesehatan yang bertujuan untuk menguji hipotesis dan menentukan apakah terdapat perbedaan yang signifikan antara kelompok atau kondisi yang sedang diteliti. Dalam dunia kesehatan masyarakat, pengambilan keputusan berbasis data sangat krusial, dan uji statistik parametrik menjadi alat utama untuk memastikan bahwa hasil yang diperoleh tidak terjadi secara kebetulan semata. Oleh karena itu, pemahaman yang mendalam tentang uji t dan ANOVA sangat diperlukan oleh para peneliti, tenaga kesehatan, dan mahasiswa yang ingin menguasai metode analisis data yang akurat dan terpercaya.

Penggunaan uji t satu sampel, dua sampel independen, dan berpasangan, serta analisis varians satu arah, memiliki peran yang berbeda tergantung pada desain penelitian dan jenis data yang dikumpulkan. Uji t satu sampel digunakan untuk membandingkan rata-rata suatu variabel dengan nilai tertentu atau standar tertentu, misalnya untuk mengetahui apakah rata-rata tekanan darah pasien berbeda dari nilai normal. Uji t dua sampel independen digunakan untuk membandingkan rata-rata antara dua kelompok yang berbeda, seperti membandingkan tingkat pengetahuan kesehatan antara dua komunitas yang mengikuti program berbeda. Sedangkan uji t berpasangan digunakan ketika data diambil dari subjek yang sama sebelum dan sesudah

intervensi, misalnya mengukur perubahan berat badan pasien setelah mengikuti program diet.

Selain itu, analisis varians satu arah (ANOVA) digunakan ketika ingin membandingkan rata-rata dari tiga atau lebih kelompok yang berbeda secara bersamaan. Misalnya, menilai efektivitas tiga jenis vaksin dalam menurunkan kejadian penyakit tertentu. Dengan menggunakan ANOVA, peneliti dapat mengetahui apakah ada perbedaan signifikan di antara kelompok tersebut, dan kemudian melakukan analisis lanjutan untuk mengetahui kelompok mana yang berbeda secara signifikan.

Pentingnya penguasaan uji statistik parametrik ini tidak hanya terletak pada kemampuan melakukan analisis, tetapi juga pada interpretasi hasilnya secara tepat. Hasil uji statistik harus dipahami secara menyeluruh, termasuk nilai p yang menunjukkan tingkat signifikansi, serta nilai statistik uji yang menunjukkan kekuatan perbedaan atau hubungan yang ditemukan. Interpretasi yang benar akan membantu peneliti dalam menyusun kesimpulan yang valid dan dapat dipertanggungjawabkan secara ilmiah.

Selain itu, dalam praktiknya, uji statistik parametrik memerlukan asumsi tertentu, seperti normalitas data dan homogenitas varians. Oleh karena itu, pemahaman tentang kondisi dan batasan penggunaan uji ini sangat penting agar hasil analisis tidak menyesatkan. Jika asumsi tidak terpenuhi, maka perlu dipertimbangkan penggunaan uji non-parametrik sebagai alternatif.

Dengan memahami dan menguasai materi ini, mahasiswa dan peneliti akan mampu melakukan analisis statistik yang tepat, menginterpretasikan hasilnya secara kritis, dan menyusun laporan penelitian yang komprehensif. Hal ini akan mendukung pengambilan keputusan berbasis bukti

dalam program-program kesehatan masyarakat, serta meningkatkan kualitas penelitian dan pelayanan kesehatan secara umum. Oleh karena itu, penguasaan uji statistik parametrik merupakan bagian integral dari kompetensi dalam bidang biostatistik kesehatan yang harus dikuasai oleh setiap calon profesional di bidang ini.

11.1 Uji t Satu Sampel

11.1.1 Pengertian dan Tujuan Uji t Satu Sampel

Uji t satu sampel adalah metode statistik yang digunakan untuk membandingkan rata-rata dari satu sampel dengan nilai tertentu yang dianggap sebagai standar atau nilai populasi yang diketahui. Tujuan utama dari uji ini adalah untuk menentukan apakah rata-rata sampel berbeda secara signifikan dari nilai yang dihipotesiskan, yang biasanya merupakan nilai normal atau standar tertentu dalam konteks kesehatan. Misalnya, seorang peneliti ingin mengetahui apakah rata-rata tekanan darah pasien di sebuah klinik berbeda dari nilai normal 120 mmHg.

11.1.2 Asumsi dan Kondisi Penggunaan

Penggunaan uji t satu sampel memerlukan beberapa asumsi utama, yaitu data harus berdistribusi normal atau mendekati normal, dan data harus bersifat kontinu. Jika data tidak memenuhi asumsi normalitas, maka hasil uji bisa menjadi tidak valid, dan alternatifnya adalah menggunakan uji non-parametrik seperti uji Wilcoxon (Ghasemi & Zahediasl, 2012). Selain itu, sampel harus diambil secara acak dan independen.

11.1.3 Langkah-langkah Pelaksanaan

Langkah pertama adalah mengumpulkan data sampel dan menentukan nilai yang akan dibandingkan. Selanjutnya, melakukan pengujian normalitas data, misalnya dengan uji Shapiro-Wilk. Jika data berdistribusi normal, maka dilakukan perhitungan statistik t dengan rumus:

$$t = (\bar{x}_1 - \bar{x}_2) / \sqrt{(s_1^2/n_1 + s_2^2/n_2)}$$

dimana:

- \bar{x} adalah rata-rata dari dua kelompok data
- s adalah standar deviasi dari dua kelompok data
- n adalah ukuran sampel dari dua kelompok data

Nilai t kemudian dibandingkan dengan nilai kritis dari tabel distribusi t dengan derajat kebebasan $(n-1)$. Jika nilai p yang dihasilkan dari analisis lebih kecil dari tingkat signifikansi yang ditetapkan (misalnya 0,05), maka hipotesis nol ditolak.

11.1.4 Contoh Aplikasi

Misalnya, sebuah studi ingin mengetahui apakah rata-rata kadar glukosa darah pasien di sebuah rumah sakit berbeda dari nilai normal 100 mg/dL. Setelah pengumpulan data dari 30 pasien, diperoleh rata-rata 105 mg/dL dengan simpangan baku 15 mg/dL. Dengan tingkat signifikansi 0,05, analisis menunjukkan bahwa nilai t adalah 1,33 dan nilai p adalah 0,19. Karena $p > 0,05$, maka tidak ada cukup bukti untuk menyatakan bahwa rata-rata kadar glukosa berbeda dari 100 mg/dL secara signifikan.

11.2 Uji t Dua Sampel Independen

11.2.1 Pengertian dan Tujuan Uji t Dua Sampel Independen

Uji t dua sampel independen digunakan untuk membandingkan rata-rata dari dua kelompok yang berbeda dan tidak saling berhubungan. Tujuan utamanya adalah untuk mengetahui apakah terdapat perbedaan yang signifikan antara kedua kelompok tersebut. Contohnya adalah membandingkan tingkat tekanan darah antara kelompok yang mengikuti program diet dan kelompok yang tidak mengikuti program.

11.2.2 Asumsi dan Kondisi Penggunaan

Penggunaan uji t dua sampel independen memerlukan asumsi normalitas dan homogenitas varians. Normalitas harus diuji terlebih dahulu, misalnya dengan uji Shapiro-Wilk, dan homogenitas varians dengan uji Levene. Jika asumsi ini terpenuhi, maka analisis dapat dilakukan. Jika tidak, alternatifnya adalah uji non-parametrik seperti Mann-Whitney.

11.2.3 Langkah-langkah Pelaksanaan

Langkah pertama dalam menganalisis data dengan uji-t adalah mengumpulkan data dari kedua kelompok secara acak dan independen. Setelah itu, lakukan pengujian normalitas dan homogenitas varians. Jika asumsi terpenuhi, hitung statistik t dengan rumus:

$$t = \frac{(M1 - M2)}{\sqrt{(\frac{s1^2}{n1}) + (\frac{s2^2}{n2})}}$$

Dimana:

- M_1 adalah rata-rata sampel 1
- M_2 adalah rata-rata sampel 2
- s_1^2 adalah varians sampel 1
- s_2^2 adalah varians sampel 2
- n_1 adalah ukuran sampel 1
- n_2 adalah ukuran sampel 2

Derajat kebebasan dihitung menggunakan rumus tertentu, dan nilai t dibandingkan dengan nilai kritis dari tabel distribusi t . Jika $p < 0,05$, maka ada perbedaan signifikan antara kedua kelompok.

11.2.4 Contoh Aplikasi

Misalnya, peneliti ingin mengetahui apakah ada perbedaan rata-rata tekanan darah antara kelompok yang mengikuti program olahraga dan yang tidak. Dari sampel masing-masing sebanyak 25 orang, diperoleh rata-rata tekanan darah 130 mmHg dan 140 mmHg, dengan simpangan baku 10 dan 12 mmHg. Analisis menunjukkan nilai t sebesar -2,45 dan p sebesar 0,02. Karena $p < 0,05$, dapat disimpulkan bahwa terdapat perbedaan signifikan antara kedua kelompok.

11.3 Uji t Berpasangan

11.3.1 Pengertian dan Tujuan Uji t Berpasangan

Uji t berpasangan digunakan ketika data diambil dari subjek yang sama sebelum dan sesudah intervensi, atau dari pasangan yang saling terkait. Tujuan utamanya adalah untuk mengetahui apakah ada perubahan yang signifikan akibat

intervensi tersebut. Contohnya adalah mengukur berat badan pasien sebelum dan setelah mengikuti program diet.

11.3.2 Asumsi dan Kondisi Penggunaan

Asumsi utama adalah data berdistribusi normal pada selisih pasangan. Pengujian normalitas dilakukan pada data selisih, bukan pada data asli. Jika asumsi ini tidak terpenuhi, maka alternatifnya adalah uji non-parametrik seperti Wilcoxon signed-rank test.

11.3.3 Langkah-langkah Pelaksanaan

Langkah pertama adalah mengumpulkan data dari pasangan yang sama pada dua waktu berbeda. Hitung selisih setiap pasangan, lalu uji normalitas distribusi selisih tersebut. Jika normal, hitung statistik t dengan rumus:

$$t = \bar{d} / \left(\frac{sd}{\sqrt{n}} \right)$$

Dimana :

- \bar{d} adalah rata-rata selisih antar pasangan
- sd adalah deviasi selisih antar pasangan
- n adalah jumlah pasangan

Nilai t kemudian dibandingkan dengan nilai kritis dari distribusi t dengan derajat kebebasan $(n-1)$. Jika $p < 0,05$, maka ada perubahan yang signifikan.

11.3.4 Contoh Aplikasi

Misalnya, seorang peneliti ingin mengetahui pengaruh program latihan fisik terhadap berat badan. Data diambil dari 20 peserta sebelum dan sesudah program. Rata-rata

penurunan berat badan adalah 2 kg dengan simpangan baku 1,5 kg. Analisis menunjukkan nilai t sebesar -6,67 dan $p < 0,001$, yang berarti ada pengaruh signifikan dari program latihan terhadap penurunan berat badan.

11.4 Uji ANOVA Satu Arah

11.4.1 Pengertian dan Tujuan Uji ANOVA Satu Arah

ANOVA (*Analysis of Variance*) satu arah digunakan untuk membandingkan rata-rata dari tiga atau lebih kelompok yang berbeda secara bersamaan. Tujuan utamanya adalah untuk mengetahui apakah ada perbedaan signifikan di antara kelompok tersebut. Dalam konteks kesehatan, misalnya, membandingkan efektivitas tiga jenis vaksin dalam mencegah infeksi.

11.4.2 Asumsi dan Kondisi Penggunaan

ANOVA memerlukan asumsi normalitas data di setiap kelompok dan homogenitas varians antar kelompok. Pengujian normalitas dan homogenitas varians dilakukan terlebih dahulu. Jika asumsi tidak terpenuhi, alternatifnya adalah uji non-parametrik seperti Kruskal-Wallis.

11.4.3 Langkah-langkah Pelaksanaan

Langkah pertama adalah mengumpulkan data dari semua kelompok secara acak dan independen. Hitung rata-rata dan varians masing-masing kelompok, lalu lakukan analisis varians dengan rumus:

$$F = \frac{s_1^2}{s_2^2}$$

Dimana, s_1^2 adalah varians sampel yang lebih besar dan s_2^2 adalah varians sampel yang lebih kecil

Nilai F dibandingkan dengan nilai kritis dari tabel distribusi F dengan derajat kebebasan tertentu. Jika $p < 0,05$, maka ada perbedaan signifikan di antara kelompok.

11.4.4 Contoh Aplikasi

Misalnya, peneliti ingin membandingkan tingkat imunisasi di tiga daerah berbeda. Dari data diambil 30 sampel di masing-masing daerah, dan analisis menunjukkan nilai F sebesar 4,56 dengan p sebesar 0,013. Karena $p < 0,05$, terdapat perbedaan signifikan dalam tingkat imunisasi antar daerah.

11.5 Interpretasi Hasil Uji pada Penelitian Kesehatan

11.5.1 Makna Nilai p dan Statistik Uji

Nilai p adalah probabilitas mendapatkan hasil yang sama ekstremnya dengan data yang diamati jika hipotesis nol benar. Jika $p < \text{tingkat signifikansi}$ (biasanya 0,05), maka hipotesis nol ditolak, yang berarti ada perbedaan atau hubungan yang signifikan secara statistik. Sebaliknya, jika $p > 0,05$, maka tidak cukup bukti untuk menolak hipotesis nol.

Contohnya, dalam uji t dua sampel independen, jika $p = 0,03$, maka ada perbedaan signifikan antara kedua kelompok. Jika $p = 0,15$, maka tidak ada cukup bukti untuk menyatakan perbedaan tersebut.

11.5.2 Penggunaan Nilai Statistik Uji

Selain p , nilai statistik uji (t atau F) juga penting untuk memahami kekuatan perbedaan. Nilai statistik yang lebih besar dari nilai kritis menunjukkan perbedaan yang signifikan. Dalam konteks penelitian kesehatan, interpretasi ini membantu menentukan apakah intervensi atau faktor tertentu memiliki pengaruh nyata terhadap variabel yang diukur.

11.5.3 Contoh Interpretasi Hasil

Misalnya, hasil uji t berpasangan menunjukkan $t = -3,45$ dan $p = 0,002$. Ini berarti bahwa perubahan berat badan sebelum dan sesudah intervensi signifikan secara statistik. Peneliti dapat menyimpulkan bahwa program tersebut efektif dalam menurunkan berat badan.

11.5.4 Pentingnya Interpretasi Kontekstual

Interpretasi hasil harus selalu dikaitkan dengan konteks penelitian dan tingkat signifikansi yang dipilih. Hasil statistik yang signifikan secara statistik tidak selalu berarti secara klinis penting, sehingga peneliti harus mempertimbangkan juga aspek praktis dan relevansi dalam konteks kesehatan masyarakat.

11.6 Batasan dan Keterbatasan Uji Statistik Parametrik

Penggunaan uji t dan ANOVA memiliki batasan, terutama terkait asumsi normalitas dan homogenitas varians. Jika data tidak memenuhi asumsi ini, hasil uji bisa menyesatkan. Oleh karena itu, penting untuk melakukan pengujian asumsi terlebih dahulu dan memilih uji alternatif jika diperlukan. Selain itu, uji parametrik tidak cocok untuk data dengan distribusi sangat skewed atau data ordinal.

Penguasaan uji statistik parametrik seperti uji t dan ANOVA sangat penting dalam penelitian kesehatan karena memungkinkan peneliti untuk menguji hipotesis secara tepat dan efisien. Pemilihan uji yang sesuai harus didasarkan pada desain penelitian, jenis data, dan asumsi yang terpenuhi. Melalui pemahaman yang mendalam tentang prosedur, interpretasi, dan batasan, mahasiswa dan peneliti dapat menghasilkan analisis yang valid dan dapat dipertanggungjawabkan secara ilmiah, sehingga mendukung pengambilan keputusan berbasis bukti dalam bidang kesehatan masyarakat.

Rangkuman

Uji statistik parametrik seperti uji t dan ANOVA merupakan alat penting dalam analisis data penelitian kesehatan. Penggunaan uji ini memungkinkan peneliti untuk menguji hipotesis tentang perbedaan rata-rata antar kelompok atau kondisi tertentu secara statistik. Pemahaman terhadap prosedur, asumsi, dan interpretasi hasil sangat krusial agar analisis yang dilakukan valid dan dapat dipertanggungjawabkan secara ilmiah. Berikut adalah poin-poin utama yang perlu dipahami:

1. **Uji t Satu Sampel** digunakan untuk membandingkan rata-rata satu sampel dengan nilai standar atau populasi yang diketahui. Asumsi utamanya adalah data harus berdistribusi normal dan bersifat kontinu. Langkah pelaksanaan meliputi pengujian normalitas, perhitungan statistik t, dan interpretasi p-value. Jika $p > 0,05$, tidak ada cukup bukti untuk menyatakan perbedaan signifikan.
2. **Uji t Dua Sampel Independen** membandingkan rata-rata dari dua kelompok yang tidak berhubungan. Asumsi yang harus dipenuhi adalah normalitas dan homogenitas varians. Langkahnya meliputi pengujian

asumsi, perhitungan t, dan interpretasi p-value. Jika $p < 0,05$, berarti terdapat perbedaan signifikan antara kedua kelompok.

3. **Uji t Berpasangan** digunakan saat data diambil dari subjek yang sama sebelum dan sesudah intervensi. Asumsi utama adalah distribusi normal pada selisih pasangan. Langkahnya meliputi penghitungan selisih, pengujian normalitas, dan analisis t. Jika $p < 0,05$, menunjukkan adanya perubahan yang signifikan.
4. **ANOVA Satu Arah** membandingkan rata-rata dari tiga atau lebih kelompok. Asumsi yang harus dipenuhi adalah normalitas dan homogenitas varians. Langkahnya meliputi pengujian asumsi, perhitungan statistik F, dan interpretasi p-value. Jika $p < 0,05$, menunjukkan adanya perbedaan signifikan di antara kelompok.
5. **Interpretasi Hasil** didasarkan pada nilai p dan statistik uji. Nilai p menunjukkan probabilitas hasil yang diamati jika hipotesis nol benar. $P < 0,05$ biasanya dianggap sebagai indikator adanya perbedaan atau hubungan yang signifikan secara statistik. Nilai statistik yang lebih besar dari nilai kritis juga menegaskan keberadaan perbedaan.
6. **Batasan dan Keterbatasan** dari uji parametrik meliputi ketergantungan pada asumsi normalitas dan homogenitas varians. Jika asumsi ini tidak terpenuhi, hasil uji bisa menyesatkan, sehingga perlu dilakukan pengujian asumsi dan memilih uji alternatif yang sesuai.
7. **Pentingnya Penguasaan** uji statistik parametrik dalam penelitian kesehatan adalah agar analisis data dapat dilakukan secara tepat, efisien, dan hasilnya dapat dipertanggungjawabkan. Pemilihan uji harus disesuaikan dengan desain penelitian, jenis data, dan kondisi distribusi data.

Dengan memahami prosedur, interpretasi, dan batasan dari uji t dan ANOVA, mahasiswa dan peneliti dapat meningkatkan kualitas analisis statistik mereka, mendukung pengambilan keputusan berbasis bukti, serta memperkuat validitas hasil penelitian di bidang kesehatan masyarakat.

Latihan Mahasiswa

Soal Essay

1. Jelaskan pengertian dan tujuan utama dari uji t satu sampel dalam penelitian kesehatan. Berikan contoh kasus yang relevan.
2. Sebutkan dan jelaskan langkah-langkah prosedural yang harus dilakukan dalam melakukan uji t dua sampel independen, mulai dari pengumpulan data hingga interpretasi hasil.
3. Apa saja asumsi utama yang harus dipenuhi sebelum melakukan uji t berpasangan? Bagaimana jika data tidak memenuhi asumsi normalitas?
4. Jelaskan perbedaan utama antara uji ANOVA satu arah dan uji t dua sampel independen. Berikan contoh situasi penelitian yang sesuai untuk masing-masing uji.
5. Mengapa penting untuk memahami batasan dan asumsi dalam penggunaan uji statistik parametrik? Jelaskan dampaknya jika asumsi tersebut tidak terpenuhi.

Soal Pilihan Berganda

1. Uji t satu sampel digunakan untuk:
 - a. Membandingkan dua rata-rata dari dua kelompok berbeda

- b. Menguji apakah rata-rata sampel berbeda dari nilai tertentu
 - c. Menguji hubungan antara dua variabel
 - d. Membandingkan tiga atau lebih kelompok
- 2. Dalam uji t dua sampel independen, asumsi utama yang harus dipenuhi adalah:
 - a. Data berdistribusi normal dan varians tidak homogen
 - b. Data berdistribusi normal dan varians homogen
 - c. Data ordinal dan varians tidak homogen
 - d. Data berdistribusi skewed dan varians homogen
- 3. Uji t berpasangan digunakan ketika data:
 - a. Diambil dari dua kelompok berbeda secara independen
 - b. Diambil dari satu kelompok yang diukur sebelum dan sesudah intervensi
 - c. Melibatkan tiga kelompok berbeda
 - d. Tidak memenuhi asumsi normalitas
- 4. Pada analisis ANOVA satu arah, nilai F yang besar dan $p < 0,05$ menunjukkan:
 - a. Tidak ada perbedaan signifikan antar kelompok
 - b. Ada perbedaan signifikan antar kelompok
 - c. Data tidak memenuhi asumsi normalitas
 - d. Varians antar kelompok lebih kecil dari dalam kelompok

5. Jika data tidak memenuhi asumsi normalitas, maka langkah yang tepat adalah:
 - a. Menggunakan uji t satu sampel
 - b. Menggunakan uji non-parametrik seperti Mann-Whitney atau Wilcoxon
 - c. Mengabaikan asumsi dan melanjutkan analisis
 - d. Mengganti data dengan data yang berdistribusi normal
6. Nilai p yang lebih kecil dari 0,05 menunjukkan:
 - a. Hasil uji tidak signifikan
 - b. Hasil uji signifikan dan hipotesis nol ditolak
 - c. Tidak ada hubungan antara variabel
 - d. Data tidak memenuhi asumsi
7. Dalam uji t berpasangan, langkah pertama yang harus dilakukan adalah:
 - a. Mengumpulkan data dari dua kelompok berbeda
 - b. Menghitung selisih setiap pasangan data
 - c. Menguji homogenitas varians
 - d. Menghitung nilai F
8. Salah satu keunggulan utama dari ANOVA adalah:
 - a. Dapat membandingkan dua kelompok saja
 - b. Menguji hubungan antara dua variabel
 - c. Membandingkan tiga atau lebih kelompok sekaligus
 - d. Tidak memerlukan asumsi normalitas
9. Dalam konteks penelitian kesehatan, uji t dan ANOVA digunakan untuk:

- a. Mengukur hubungan antara variabel
- b. Menguji perbedaan rata-rata antar kelompok
- c. Menghitung prevalensi penyakit
- d. Menyusun tabel distribusi frekuensi

10. Hasil analisis statistik yang menunjukkan $p = 0,07$ berarti:

- a. Ada perbedaan signifikan
- b. Tidak ada cukup bukti untuk menolak hipotesis nol
- c. Varians antar kelompok berbeda secara signifikan
- d. Data tidak memenuhi asumsi normalitas

Soal Project / Studi Kasus

1. Seorang peneliti ingin mengetahui pengaruh program latihan fisik terhadap berat badan. Data diambil dari 20 peserta sebelum dan sesudah mengikuti program tersebut. Buatlah langkah-langkah analisis statistik yang harus dilakukan, termasuk pengujian asumsi dan interpretasi hasilnya.
2. Sebuah studi membandingkan tingkat pengetahuan kesehatan antara tiga komunitas yang mengikuti program edukasi berbeda. Data diambil dari 30 responden di masing-masing komunitas. Jelaskan prosedur analisis yang tepat, termasuk jenis uji statistik yang digunakan, dan bagaimana menafsirkan hasilnya.

REFERENSI

1. Ghasemi, A., & Zahediasl, S. (2012). Normality tests for statistical analysis: a guide for non-statisticians. *International Journal of Endocrinology and Metabolism*, 10(2), 486–489. <https://doi.org/10.5812/ijem.3705>
2. Field, A. (2013). *Discovering statistics using IBM SPSS statistics* (4th ed.). Sage Publications.
3. Kim, T. K. (2015). *T-test and ANOVA*. In *Introduction to biostatistics* (pp. 123–145). Springer.
4. Tabachnick, B. G., & Fidell, L. S. (2013). *Using multivariate statistics* (6th ed.). Pearson."

BAB 12

UJI STATISTIK NON-PARAMETRIK

Tujuan Pembelajaran

1. Memahami konsep dasar dan prinsip dasar dari uji statistik non-parametrik, termasuk alasan penggunaannya dalam analisis data kesehatan yang tidak memenuhi asumsi distribusi normal.
2. Menjelaskan secara rinci tentang uji Chi-Square, termasuk prosedur, interpretasi hasil, dan situasi di mana uji ini paling tepat digunakan dalam penelitian kesehatan masyarakat.
3. Mengidentifikasi dan memahami penggunaan uji Mann-Whitney dan Wilcoxon sebagai alternatif non-parametrik untuk uji t, serta mengetahui kapan dan bagaimana menerapkannya dalam analisis data.
4. Mampu menentukan kondisi dan situasi yang tepat untuk menggunakan uji non-parametrik dalam penelitian kesehatan, terutama ketika data bersifat ordinal atau tidak memenuhi asumsi parametrik.
5. Mengaplikasikan contoh kasus nyata, seperti analisis hubungan antara tingkat pendidikan dan perilaku kesehatan, menggunakan uji Chi-Square dan uji Mann-Whitney/Wilcoxon untuk memperoleh hasil yang valid dan reliabel.
6. Mengembangkan kemampuan dalam memilih uji statistik non-parametrik yang sesuai berdasarkan karakteristik data dan tujuan penelitian, serta mampu menginterpretasikan output hasil analisis secara tepat.

7. Menyadari pentingnya penggunaan uji non-parametrik dalam konteks penelitian kesehatan masyarakat, terutama dalam situasi di mana data tidak memenuhi asumsi distribusi normal, sehingga hasil analisis menjadi lebih akurat dan bermakna.

Pendahuluan

Dalam dunia penelitian kesehatan masyarakat, pengambilan keputusan yang tepat sangat bergantung pada analisis data yang akurat dan sesuai dengan karakteristik data tersebut. Banyak data yang diperoleh dari surveilans, wawancara, atau pengukuran lapangan tidak selalu mengikuti distribusi normal, dan sering kali bersifat ordinal atau kategorikal. Kondisi ini menuntut penggunaan metode statistik yang tidak bergantung pada asumsi distribusi tertentu, yang dikenal sebagai uji statistik non-parametrik.

Uji statistik non-parametrik memiliki peran penting dalam analisis data kesehatan karena kemampuannya untuk memberikan hasil yang valid meskipun data tidak memenuhi asumsi normalitas, homogenitas varians, atau skala interval. Penggunaan uji ini sangat relevan dalam berbagai studi epidemiologi dan promosi kesehatan, di mana data sering kali bersifat ordinal, seperti tingkat pendidikan, tingkat pengetahuan, atau tingkat perilaku yang dikategorikan secara ordinal. Selain itu, uji non-parametrik juga digunakan ketika jumlah sampel kecil sehingga asumsi distribusi normal sulit dipenuhi.

Dalam bagian ini, kita akan mempelajari dua jenis uji statistik non-parametrik yang umum digunakan dalam penelitian kesehatan: uji Chi-Square dan uji Mann-Whitney serta Wilcoxon. Uji Chi-Square biasanya digunakan untuk menguji hubungan antara dua variabel kategorikal, misalnya hubungan antara tingkat pendidikan dan perilaku

kesehatan. Sedangkan uji Mann-Whitney dan Wilcoxon digunakan untuk membandingkan dua kelompok independen atau berpasangan ketika data bersifat ordinal atau tidak memenuhi asumsi uji t.

Memahami kapan dan bagaimana menerapkan uji statistik non-parametrik sangat penting agar hasil penelitian dapat diinterpretasikan dengan benar dan dapat dipertanggungjawabkan secara ilmiah. Penggunaan uji yang tepat akan meningkatkan validitas hasil dan memberikan gambaran yang lebih akurat tentang hubungan atau perbedaan yang ada dalam data kesehatan masyarakat. Oleh karena itu, bagian ini akan membekali mahasiswa dan peneliti dengan pengetahuan dasar, prosedur, serta contoh aplikasi nyata yang relevan dengan konteks penelitian di bidang kesehatan.

Selain itu, pemahaman tentang uji non-parametrik juga akan membantu dalam pengambilan keputusan praktis di lapangan, seperti dalam analisis data surveilans, evaluasi program, dan studi epidemiologi. Dengan menguasai materi ini, diharapkan mahasiswa mampu memilih metode analisis yang sesuai, menginterpretasikan hasil secara tepat, dan menyusun laporan penelitian yang komprehensif dan akurat. Pada akhirnya, penguasaan uji statistik non-parametrik akan memperkuat kualitas penelitian dan mendukung pengembangan ilmu pengetahuan di bidang kesehatan masyarakat.

12.1 Uji Chi-Square

12.1.1 Pengertian dan Tujuan Uji Chi-Square

Uji Chi-Square (χ^2) merupakan salah satu uji statistik non-parametrik yang digunakan untuk menguji hubungan atau asosiasi antara dua variabel kategorikal. Uji ini tidak

memerlukan asumsi distribusi normal dan cocok digunakan ketika data berbentuk frekuensi atau jumlah kejadian dalam kategori tertentu (McHugh, 2013). Tujuan utama dari uji Chi-Square adalah untuk mengetahui apakah terdapat hubungan yang signifikan secara statistik antara variabel-variabel tersebut, atau apakah distribusi frekuensi yang diamati berbeda secara bermakna dari distribusi yang diharapkan.

12.1.2 Prosedur Pelaksanaan Uji Chi-Square

Pelaksanaan uji Chi-Square melibatkan beberapa langkah utama:

1. **Menyusun Tabel Kontingensi:** Data dikumpulkan dan disusun dalam bentuk tabel kontingensi yang menunjukkan frekuensi kejadian untuk setiap kombinasi kategori variabel.
2. **Menghitung Frekuensi Harapan:** Untuk setiap sel dalam tabel, frekuensi harapan dihitung berdasarkan rumus: $E = (\text{Total baris} * \text{Total kolom}) / \text{Total keseluruhan data}$, dimana E adalah frekuensi harapan dan total baris, total kolom serta total keseluruhan data adalah total dari tabel kontingensi.
3. **Menghitung Nilai χ^2 :** Nilai statistik χ^2 dihitung dengan rumus: $\chi^2 = \sum [(O - E)^2 / E]$, di mana O adalah frekuensi observasi dan E adalah frekuensi harapan.
4. **Menentukan Derajat Kebebasan:** Derajat kebebasan (df) dihitung sebagai: $df = (r - 1) * (c - 1)$, di mana r adalah jumlah baris dan c adalah jumlah kolom.
5. **Menginterpretasikan Hasil:** Nilai χ^2 yang dihitung dibandingkan dengan nilai kritis dari tabel χ^2 pada tingkat signifikansi tertentu (misalnya 0,05). Jika nilai χ^2 lebih besar dari nilai kritis, maka ada hubungan yang signifikan antara variabel.

12.1.3 Interpretasi Hasil Uji Chi-Square

Hasil dari uji Chi-Square akan memberikan nilai p (probabilitas) yang menunjukkan tingkat signifikansi. Jika $p < 0,05$, maka hipotesis nol (tidak ada hubungan) ditolak, dan dapat disimpulkan bahwa terdapat hubungan yang signifikan antara variabel-variabel yang diuji. Sebaliknya, jika $p \geq 0,05$, maka tidak ada cukup bukti untuk menolak hipotesis nol, dan hubungan antara variabel dianggap tidak signifikan secara statistik.

12.1.4 Kelebihan dan Keterbatasan Uji Chi-Square

Kelebihan dari uji Chi-Square adalah kemampuannya untuk menguji hubungan antara variabel kategorikal tanpa asumsi distribusi tertentu dan kemudahannya dalam pelaksanaan. Namun, uji ini memiliki keterbatasan, seperti:

- Tidak cocok untuk data dengan frekuensi yang sangat kecil (misalnya, <5 dalam satu sel), karena dapat menghasilkan hasil yang tidak valid (Fagerland & Sand, 2017).
- Hanya menguji hubungan asosiasi, bukan sebab-akibat.
- Tidak dapat digunakan untuk data berukuran ordinal secara langsung, kecuali data tersebut dikategorikan.

12.2 Uji Mann-Whitney dan Wilcoxon

12.2.1 Pengertian dan Tujuan Uji Mann-Whitney dan Wilcoxon

Uji Mann-Whitney dan Wilcoxon merupakan uji statistik non-parametrik yang digunakan untuk membandingkan dua kelompok independen atau berpasangan, terutama ketika data bersifat ordinal atau tidak memenuhi asumsi distribusi

normal. Uji Mann-Whitney (juga dikenal sebagai U test) digunakan untuk membandingkan dua kelompok independen, sedangkan uji Wilcoxon digunakan untuk data berpasangan atau berhubungan (Hollander & Wolfe, 2013).

Tujuan utama dari kedua uji ini adalah untuk menentukan apakah terdapat perbedaan yang signifikan secara statistik antara dua kelompok dalam hal median atau distribusi data.

12.2.2 Prosedur Pelaksanaan Uji Mann-Whitney dan Wilcoxon

Uji Mann-Whitney:

1. **Pengumpulan Data:** Data dari kedua kelompok dikumpulkan dan digabungkan.
2. **Peringkat Data:** Data diurutkan dari yang terkecil hingga terbesar dan diberi peringkat.
3. **Penghitungan Peringkat:** Jumlah peringkat untuk masing-masing kelompok dihitung.
4. **Perhitungan Statistik U:** Nilai U dihitung berdasarkan peringkat dan jumlah data di masing-masing kelompok.
5. **Interpretasi:** Nilai U dibandingkan dengan tabel U pada tingkat signifikansi tertentu. Jika $p < 0,05$, terdapat perbedaan signifikan.

Uji Wilcoxon:

1. **Pengumpulan Data Berpasangan:** Data dari pasangan yang sama atau berhubungan dikumpulkan.
2. **Penghitungan Selisih:** Selisih antara pasangan data dihitung.
3. **Peringkat Selisih:** Selisih yang bukan nol diurutkan dan diberi peringkat.

4. **Penghitungan Statistik T:** Jumlah peringkat dari selisih positif dan negatif dihitung.
5. **Interpretasi:** Nilai T dibandingkan dengan tabel Wilcoxon. Jika $p < 0,05$, perbedaan signifikan.

12.2.3 Interpretasi Hasil Uji Mann-Whitney dan Wilcoxon

Hasil dari kedua uji ini berupa nilai p . Jika $p < 0,05$, maka ada perbedaan yang signifikan antara kedua kelompok. Jika $p \geq 0,05$, tidak terdapat bukti cukup untuk menyatakan adanya perbedaan.

12.2.4 Kelebihan dan Keterbatasan Uji Mann-Whitney dan Wilcoxon

Kelebihan:

- Tidak memerlukan asumsi distribusi normal.
- Cocok untuk data ordinal dan data dengan distribusi tidak normal.
- Mudah dilakukan dan tidak memerlukan perangkat lunak statistik yang kompleks.

Keterbatasan:

- Hanya membandingkan dua kelompok; untuk lebih dari dua, diperlukan uji lain seperti Kruskal-Wallis.
- Tidak memberikan informasi tentang besarnya perbedaan, hanya keberadaannya.
- Sensitif terhadap jumlah data dan distribusi data.

12.3 Kapan dan Bagaimana Uji Ini Digunakan

Penggunaan uji Chi-Square sangat tepat ketika data yang dianalisis berupa frekuensi kejadian dalam kategori, seperti prevalensi penyakit berdasarkan jenis kelamin atau tingkat pendidikan. Contohnya, menguji apakah tingkat kejadian hipertensi berbeda antara pria dan wanita.

Sedangkan uji Mann-Whitney dan Wilcoxon digunakan ketika data bersifat ordinal atau tidak memenuhi asumsi normalitas, misalnya, membandingkan tingkat pengetahuan kesehatan antara dua kelompok masyarakat yang berbeda, atau menilai perubahan perilaku kesehatan sebelum dan sesudah intervensi pada kelompok yang sama.

Penggunaan kedua uji ini harus didasarkan pada karakteristik data dan tujuan analisis. Jika data bersifat kategorikal dan frekuensi, uji Chi-Square adalah pilihan utama. Jika data bersifat ordinal atau skala pengukuran tidak memenuhi asumsi normal, maka uji Mann-Whitney atau Wilcoxon lebih sesuai.

12.4 Contoh Aplikasi: Analisis Hubungan Antara Tingkat Pendidikan dan Perilaku Kesehatan

Misalkan seorang peneliti ingin mengetahui apakah ada hubungan antara tingkat pendidikan dan perilaku kesehatan masyarakat di suatu wilayah. Data dikumpulkan dari survei yang melibatkan 200 responden, di mana tingkat pendidikan dikategorikan menjadi tiga: *rendah*, *sedang*, dan *tinggi*. Perilaku kesehatan diukur berdasarkan indikator seperti kebiasaan merokok, konsumsi makanan sehat, dan rutin berolahraga, yang dikategorikan sebagai *baik* dan *buruk*.

Langkah 1: Analisis dengan Uji Chi-Square

Data disusun dalam tabel kontingensi sebagai berikut:

Tingkat Pendidikan	Perilaku Baik	Perilaku Buruk	Total
Rendah	30	70	100
Sedang	50	30	80
Tinggi	40	20	60
Total	120	120	240

Pengujian dilakukan untuk melihat apakah distribusi perilaku kesehatan berbeda secara signifikan berdasarkan tingkat pendidikan.

Langkah 2: Menghitung Frekuensi Harapan

Contoh perhitungan:

$$E = (\text{Total baris} * \text{Total kolom}) / \text{Total keseluruhan data}$$

Dimana:

- E adalah frekuensi harapan
- total baris, total kolom serta total keseluruhan data adalah total dari tabel kontingensi.

Setelah semua frekuensi harapan dihitung, nilai χ^2 dihitung dan dibandingkan dengan nilai kritis dari tabel χ^2 dengan $df = (3-1) \times (2-1) = 2$. Jika $p < 0,05$, maka dapat disimpulkan bahwa ada hubungan yang signifikan antara tingkat pendidikan dan perilaku kesehatan.

Langkah 3: Interpretasi

Jika hasil menunjukkan $p < 0,05$, berarti tingkat pendidikan berpengaruh terhadap perilaku kesehatan masyarakat. Sebaliknya, jika $p \geq 0,05$, tidak ada hubungan yang signifikan.

Contoh lain: Uji Mann-Whitney

Misalnya, peneliti ingin membandingkan tingkat pengetahuan kesehatan (diukur dengan skor ordinal dari 1 sampai 10) antara dua kelompok masyarakat yang mengikuti program promosi kesehatan dan yang tidak. Data dari 50 peserta di masing-masing kelompok dikumpulkan, dan skor pengetahuan diurutkan.

Langkah-langkahnya meliputi pengurutan skor, pemberian peringkat, dan penghitungan nilai U. Jika hasil menunjukkan $p < 0,05$, maka terdapat perbedaan signifikan dalam tingkat pengetahuan antara kedua kelompok.

12.5 Kesimpulan

Penggunaan uji statistik non-parametrik seperti Chi-Square, Mann-Whitney, dan Wilcoxon sangat penting dalam analisis data kesehatan masyarakat, terutama ketika data tidak memenuhi asumsi distribusi normal atau bersifat ordinal. Pemilihan uji yang tepat akan memastikan validitas dan reliabilitas hasil penelitian, serta mendukung pengambilan keputusan berbasis bukti yang akurat.

Dalam praktiknya, pemahaman mendalam tentang prosedur, interpretasi, dan situasi penggunaan masing-masing uji akan membantu mahasiswa dan peneliti dalam menyusun analisis statistik yang tepat dan mampu menjelaskan hasilnya secara ilmiah. Dengan demikian, penguasaan materi ini akan memperkuat kompetensi dalam penelitian dan pengembangan ilmu kesehatan masyarakat.

Rangkuman

Pembahasan mengenai uji statistik non-parametrik dalam bab ini menitikberatkan pada tiga jenis uji utama, yaitu Uji Chi-Square, Uji Mann-Whitney, dan Uji Wilcoxon. Ketiga uji ini sangat penting dalam analisis data kesehatan masyarakat, terutama ketika data tidak memenuhi asumsi distribusi normal atau bersifat kategorikal dan ordinal. Berikut adalah poin-poin utama yang dapat disimpulkan dari pembahasan tersebut:

1. **Uji Chi-Square** digunakan untuk menguji hubungan atau asosiasi antara dua variabel kategorikal. Uji ini tidak memerlukan asumsi distribusi normal dan cocok untuk data berupa frekuensi kejadian dalam kategori tertentu. Prosedurnya meliputi penyusunan tabel kontingensi, perhitungan frekuensi harapan, penghitungan nilai χ^2 , dan interpretasi p-value. Jika $p < 0,05$, maka terdapat hubungan yang signifikan antara variabel, sedangkan jika $p \geq 0,05$, hubungan tidak signifikan. Kelebihan uji ini adalah kemudahannya dan tidak memerlukan asumsi distribusi tertentu, tetapi memiliki keterbatasan, seperti tidak cocok untuk data dengan frekuensi kecil dan hanya menguji asosiasi, bukan sebab-akibat.
2. **Uji Mann-Whitney dan Wilcoxon** merupakan uji non-parametrik yang digunakan untuk membandingkan dua kelompok, baik independen maupun berpasangan. Uji Mann-Whitney cocok untuk data ordinal dan data tidak normal dari dua kelompok independen, sedangkan Wilcoxon digunakan untuk data berpasangan. Prosedurnya meliputi pengurutan data, pemberian peringkat, dan penghitungan statistik U atau T. Jika $p < 0,05$, terdapat perbedaan yang signifikan antara kedua kelompok. Kelebihan dari kedua uji ini adalah tidak

memerlukan asumsi distribusi normal dan mudah dilakukan, tetapi terbatas hanya membandingkan dua kelompok dan tidak menunjukkan besarnya perbedaan.

3. **Kapan dan bagaimana menggunakan uji ini** sangat bergantung pada karakteristik data dan tujuan analisis. Uji Chi-Square tepat untuk data frekuensi kategori, seperti prevalensi penyakit berdasarkan jenis kelamin. Sedangkan Uji Mann-Whitney dan Wilcoxon digunakan untuk data ordinal atau data yang tidak memenuhi asumsi normalitas, seperti membandingkan tingkat pengetahuan kesehatan sebelum dan sesudah intervensi.
4. **Contoh aplikasi** dari uji Chi-Square adalah analisis hubungan antara tingkat pendidikan dan perilaku kesehatan masyarakat, sedangkan Uji Mann-Whitney digunakan untuk membandingkan tingkat pengetahuan kesehatan antara dua kelompok masyarakat. Contoh ini menunjukkan bagaimana uji statistik non-parametrik dapat digunakan untuk mendukung pengambilan keputusan dan interpretasi data dalam konteks kesehatan masyarakat.
5. **Pentingnya pemilihan uji yang tepat** adalah untuk memastikan validitas hasil analisis. Penggunaan uji yang sesuai akan membantu peneliti dalam menyimpulkan hubungan atau perbedaan secara statistik dan ilmiah, serta mendukung pengambilan kebijakan berbasis bukti.
6. **Keterbatasan dan kelebihan** dari masing-masing uji harus dipahami agar penggunaannya tepat. Uji Chi-Square tidak cocok untuk data dengan frekuensi kecil dan hanya menguji asosiasi, sedangkan Uji Mann-Whitney dan Wilcoxon tidak dapat digunakan untuk membandingkan lebih dari dua kelompok.
7. **Kesimpulan utama** adalah bahwa uji statistik non-parametrik sangat berguna dalam analisis data

kesehatan masyarakat yang bersifat kategorikal dan ordinal, serta ketika asumsi distribusi normal tidak terpenuhi. Penguasaan prosedur, interpretasi, dan situasi penggunaan uji ini akan memperkuat kompetensi mahasiswa dan peneliti dalam melakukan analisis statistik yang valid dan reliabel.

Dengan memahami dan menerapkan uji statistik non-parametrik secara tepat, peneliti dapat memperoleh hasil yang akurat dan mendukung pengambilan keputusan yang berbasis bukti dalam bidang kesehatan masyarakat.

Latihan Mahasiswa

Soal Essay

1. Jelaskan pengertian dan tujuan utama dari uji Chi-Square dalam analisis data kesehatan masyarakat!
2. Sebutkan dan jelaskan langkah-langkah prosedur pelaksanaan uji Mann-Whitney!
3. Mengapa uji non-parametrik seperti Chi-Square dan Mann-Whitney lebih dipilih dalam penelitian kesehatan yang melibatkan data ordinal atau data dengan distribusi tidak normal?
4. Berikan contoh situasi penelitian di bidang kesehatan masyarakat yang tepat menggunakan uji Wilcoxon dan jelaskan alasannya!
5. Diskusikan kelebihan dan keterbatasan dari penggunaan uji Chi-Square dalam analisis data kategori!

Soal Pilihan Ganda

1. Uji statistik non-parametrik yang paling tepat digunakan untuk menguji hubungan antara dua variabel kategorikal adalah:
 - a. Uji t
 - b. Uji ANOVA
 - c. Uji Chi-Square
 - e. Uji regresi
2. Dalam uji Chi-Square, derajat kebebasan dihitung berdasarkan:
 - a. Jumlah baris dikurangi satu
 - b. Jumlah kolom dikurangi satu
 - c. $(\text{Jumlah baris} - 1) \text{ dikali } (\text{jumlah kolom} - 1)$
 - d. Jumlah total data dikurangi satu
3. Uji Mann-Whitney digunakan untuk membandingkan:
 - a. Dua variabel kategorikal
 - b. Dua kelompok independen dengan data ordinal atau tidak normal
 - c. Data berpasangan dengan skala interval
 - d. Lebih dari dua kelompok
4. Salah satu kelebihan utama dari uji Wilcoxon adalah:
 - a. Memerlukan asumsi distribusi normal
 - b. Cocok untuk data berpasangan atau berhubungan
 - c. Hanya digunakan untuk data nominal
 - d. Tidak memerlukan peringkat data
5. Uji Chi-Square tidak cocok digunakan jika:

- a. Data berupa frekuensi kejadian
 - b. Frekuensi dalam satu sel kurang dari 5
 - c. Data bersifat kategorikal
 - d. Variabel berskala nominal
6. Dalam analisis data kesehatan, uji Wilcoxon paling tepat digunakan ketika:
 - a. Membandingkan dua kelompok independen
 - b. Menguji hubungan antara dua variabel kategorikal
 - c. Menguji perbedaan data berpasangan yang tidak memenuhi asumsi normalitas
 - d. Menghitung korelasi antara dua variabel
7. Pada tabel kontingensi, frekuensi harapan dihitung berdasarkan:
 - a. Frekuensi observasi dikalikan jumlah total data
 - b. Jumlah baris dikali jumlah kolom dibagi total data
 - c. Jumlah kolom dikali jumlah baris dibagi total data
 - d. Jumlah total data dikurangi frekuensi observasi
8. Jika hasil uji Chi-Square menunjukkan $p > 0,05$, maka:
 - a. Terdapat hubungan yang signifikan
 - b. Tidak ada hubungan yang signifikan
 - c. Data tidak valid
 - d. Hipotesis alternatif diterima
9. Dalam studi kasus, data tingkat pendidikan dan perilaku kesehatan dikategorikan sebagai variabel:

- a. Numerik kontinu
 - b. Nominal
 - c. Ordinal
 - d. Rasio
10. Uji Mann-Whitney dan Wilcoxon memiliki keunggulan utama karena:
- a. Memerlukan asumsi distribusi normal
 - b. Dapat digunakan untuk data nominal
 - c. Tidak memerlukan asumsi distribusi normal
 - d. Hanya digunakan untuk data berkelompok besar

Soal Project / Studi Kasus

1. Seorang peneliti ingin mengetahui apakah ada perbedaan tingkat pengetahuan kesehatan antara dua desa yang berbeda setelah mengikuti program edukasi. Data diukur dengan skor ordinal dari 1 sampai 10. Jelaskan langkah-langkah yang harus dilakukan untuk menganalisis data tersebut menggunakan uji Mann-Whitney dan interpretasikan hasilnya secara hipotetik!
2. Sebuah surveilans kesehatan menemukan bahwa distribusi kejadian penyakit tertentu di dua wilayah berbeda mengikuti distribusi Poisson. Peneliti ingin menguji apakah ada perbedaan signifikan dalam jumlah kejadian penyakit tersebut antara kedua wilayah. Jelaskan prosedur analisis yang tepat dan interpretasi hasilnya!

REFERENSI

1. Fagerland, M. W., & Sand, K. (2017). The Chi-square test of independence: power calculations for small sample sizes. *BMC Medical Research Methodology*, 17(1), 1-8.
2. Hollander, M., & Wolfe, D. A. (2013). *Nonparametric Statistical Methods* (3rd ed.). Wiley.
3. McHugh, M. L. (2013). The Chi-square test of independence. *Biochemia Medica*, 23(2), 143-149." "

BAB 13

KORELASI DAN REGRESI

Tujuan Pembelajaran

Setelah mempelajari bagian ini, diharapkan mahasiswa mampu mencapai beberapa tujuan utama yang berkaitan dengan pemahaman dan penerapan konsep korelasi dan regresi dalam konteks kesehatan masyarakat. Tujuan-tujuan tersebut dirancang agar mahasiswa tidak hanya memahami teori dasar, tetapi juga mampu mengaplikasikan konsep tersebut dalam analisis data nyata yang relevan dengan promosi kesehatan dan pengendalian penyakit.

Pertama, mahasiswa diharapkan mampu menjelaskan pengertian dan perbedaan antara korelasi Pearson dan Spearman, termasuk kapan dan bagaimana keduanya digunakan dalam analisis data kesehatan. Pemahaman ini penting agar mahasiswa dapat memilih metode korelasi yang sesuai berdasarkan jenis data dan karakteristik variabel yang sedang dianalisis. Selain itu, mahasiswa juga harus mampu mengidentifikasi asumsi dasar dari masing-masing uji korelasi dan menafsirkannya secara tepat dalam konteks penelitian kesehatan.

Kedua, mahasiswa diharapkan mampu melakukan analisis korelasi linier sederhana antara dua variabel kesehatan, seperti tingkat aktivitas fisik dan tekanan darah, serta mampu menginterpretasikan hasil korelasi tersebut secara statistik dan klinis. Interpretasi ini mencakup pemahaman tentang kekuatan hubungan, arah hubungan, serta implikasi

praktisnya dalam pengembangan program promosi kesehatan.

Ketiga, mahasiswa harus mampu menjelaskan konsep dan langkah-langkah dalam melakukan analisis regresi linier sederhana, termasuk pengujian asumsi dasar dan interpretasi koefisien regresi. Kemampuan ini penting agar mahasiswa dapat mengidentifikasi pengaruh satu variabel terhadap variabel lain secara kuantitatif, serta memahami bagaimana model regresi dapat digunakan untuk memprediksi variabel kesehatan berdasarkan faktor-faktor tertentu.

Keempat, mahasiswa diharapkan mampu mengaplikasikan analisis regresi linier sederhana dalam studi kasus nyata, misalnya untuk meneliti hubungan antara tingkat aktivitas fisik dan tekanan darah. Mereka harus mampu menyusun model regresi, melakukan analisis statistik, dan menafsirkan hasilnya secara komprehensif, termasuk menilai signifikansi dan kekuatan hubungan tersebut.

Kelima, mahasiswa harus mampu membandingkan dan menilai kelebihan serta keterbatasan dari analisis korelasi dan regresi dalam konteks penelitian kesehatan. Mereka diharapkan mampu menentukan metode analisis yang paling sesuai berdasarkan tujuan penelitian, jenis data, dan kondisi lapangan, sehingga hasil analisis dapat digunakan secara efektif dalam pengambilan keputusan dan pengembangan program promosi kesehatan.

Keenam, mahasiswa diharapkan mampu mengintegrasikan konsep korelasi dan regresi dalam penyusunan laporan hasil penelitian kesehatan, termasuk menyusun narasi yang jelas dan sistematis serta menyajikan data dalam bentuk tabel dan grafik yang mendukung interpretasi hasil. Kemampuan ini penting agar laporan penelitian dapat dipahami oleh

berbagai pihak, mulai dari tenaga kesehatan hingga pengambil kebijakan.

Terakhir, mahasiswa diharapkan mampu mengidentifikasi potensi kesalahan umum dan tantangan dalam melakukan analisis korelasi dan regresi, serta mampu mengatasi masalah tersebut melalui pendekatan statistik yang tepat. Dengan demikian, mereka akan mampu melakukan analisis data yang akurat dan dapat dipertanggungjawabkan dalam konteks promosi dan pengendalian kesehatan masyarakat.

Pendahuluan

Bagian ini akan membahas pengantar mengenai pentingnya analisis korelasi dan regresi dalam bidang biostatistik kesehatan. Dalam praktiknya, pengumpulan data kesehatan sering kali menghasilkan variabel-variabel yang saling berhubungan, dan memahami hubungan tersebut merupakan langkah penting dalam pengembangan strategi promosi kesehatan dan intervensi medis. Korelasi dan regresi adalah dua metode statistik yang umum digunakan untuk mengukur dan menganalisis hubungan antara variabel-variabel tersebut.

Korelasi, baik Pearson maupun Spearman, memungkinkan peneliti untuk mengetahui seberapa kuat dan arah hubungan antara dua variabel. Misalnya, dalam studi kesehatan masyarakat, korelasi dapat digunakan untuk menilai hubungan antara tingkat aktivitas fisik dan tekanan darah, atau antara tingkat pendidikan dan perilaku kesehatan. Pemahaman yang tepat tentang korelasi membantu tenaga kesehatan dalam mengidentifikasi faktor risiko dan potensi intervensi yang efektif.

Sementara itu, analisis regresi linier sederhana memberikan gambaran yang lebih mendalam dengan menunjukkan pengaruh satu variabel terhadap variabel lain secara kuantitatif. Dengan model regresi, kita tidak hanya mengetahui adanya hubungan, tetapi juga dapat memprediksi nilai variabel tergantung berdasarkan variabel independen. Sebagai contoh, dalam konteks promosi kesehatan, regresi dapat digunakan untuk memperkirakan perubahan tekanan darah berdasarkan tingkat aktivitas fisik yang dilakukan oleh individu.

Urgensi mempelajari korelasi dan regresi dalam bidang kesehatan tidak hanya terletak pada kemampuan melakukan analisis statistik, tetapi juga pada penerapannya dalam pengambilan keputusan berbasis data. Data yang dianalisis secara tepat akan memberikan gambaran yang akurat tentang hubungan antar variabel, sehingga program promosi kesehatan dapat dirancang secara lebih efektif dan efisien. Selain itu, pemahaman ini juga penting untuk menghindari kesalahan interpretasi yang dapat menyesatkan hasil penelitian dan pengambilan kebijakan.

Dalam bagian ini, mahasiswa akan diajak untuk memahami konsep dasar dari kedua metode tersebut, termasuk asumsi yang harus dipenuhi, langkah-langkah analisis, serta interpretasi hasilnya. Melalui studi kasus dan latihan praktis, diharapkan mereka mampu mengaplikasikan konsep ini secara langsung dalam konteks penelitian dan kegiatan promosi kesehatan di lapangan. Dengan demikian, penguasaan materi ini akan menjadi fondasi penting dalam pengembangan kompetensi analisis data kesehatan yang akurat dan bermakna.

13.1 Korelasi Pearson dan Spearman

13.1.1 Pengertian dan Tujuan Korelasi

Korelasi adalah ukuran statistik yang digunakan untuk menilai kekuatan dan arah hubungan antara dua variabel. Dalam konteks kesehatan masyarakat, korelasi membantu peneliti memahami sejauh mana variabel-variabel tertentu saling berhubungan, yang dapat menjadi dasar untuk pengembangan intervensi atau program promosi kesehatan. Misalnya, mengetahui hubungan antara tingkat aktivitas fisik dan tekanan darah dapat memberikan gambaran awal tentang faktor risiko hipertensi.

Korelasi tidak menunjukkan hubungan sebab-akibat, melainkan hanya mengukur kekuatan dan arah hubungan antara variabel. Oleh karena itu, penting untuk memahami bahwa korelasi yang tinggi tidak otomatis berarti satu variabel menyebabkan perubahan pada variabel lain, melainkan keduanya mungkin dipengaruhi oleh faktor lain yang tidak terukur.

13.1.2 Korelasi Pearson

Korelasi Pearson, yang dikenal juga sebagai koefisien korelasi Pearson (r), digunakan untuk mengukur hubungan linier antara dua variabel yang berskala interval atau rasio dan memenuhi asumsi distribusi normal. Rumusnya adalah:

$$r = \frac{[n(\sum xy) - (\sum x)(\sum y)]}{\sqrt{(n\sum x^2 - (\sum x)^2)(n\sum y^2 - (\sum y)^2)}}$$

di mana:

- r adalah koefisien korelasi Pearson
- n adalah jumlah pasangan data (x dan y)
- $\sum x$ adalah jumlah dari semua nilai x
- $\sum y$ adalah jumlah dari semua nilai y
- $\sum xy$ adalah jumlah dari perkalian setiap pasangan data ($x*y$)
- $\sum x^2$ adalah jumlah dari kuadrat semua nilai x
- $\sum y^2$ adalah jumlah dari kuadrat semua nilai y

Nilai r berkisar antara -1 sampai +1. Nilai mendekati +1 menunjukkan hubungan positif yang kuat, di mana peningkatan variabel X diikuti oleh peningkatan variabel Y . Sebaliknya, nilai mendekati -1 menunjukkan hubungan negatif yang kuat, di mana peningkatan variabel X diikuti oleh penurunan variabel Y . Nilai mendekati 0 menunjukkan tidak adanya hubungan linier yang signifikan.

Contoh: Dalam studi kesehatan, korelasi Pearson dapat digunakan untuk menilai hubungan antara tingkat aktivitas fisik dan tekanan darah. Jika diperoleh $r = -0,65$, ini menunjukkan hubungan negatif yang cukup kuat, artinya semakin tinggi tingkat aktivitas fisik, biasanya tekanan darah cenderung lebih rendah.

13.1.3 Korelasi Spearman

Korelasi Spearman, dikenal juga sebagai koefisien korelasi Spearman (ρ), digunakan untuk mengukur hubungan monotonik antara dua variabel, terutama ketika data tidak memenuhi asumsi normalitas atau berskala ordinal. Koefisien ini didasarkan pada peringkat data dan dihitung dengan rumus:

$$\rho = 1 - \frac{6 * \sum d^2}{(n * (n^2 - 1))}$$

di mana :

- ρ adalah koefisien korelasi spearman
- $\sum d^2$ adalah jumlah selisih peringkat
- n adalah jumlah pasangan data

Korelasi Spearman juga berkisar antara -1 sampai +1, dengan interpretasi yang sama seperti Pearson. Keunggulan utama Spearman adalah kemampuannya untuk mengukur hubungan yang tidak harus linier, tetapi bersifat monotonik (selalu meningkat atau menurun).

Contoh: Dalam penelitian perilaku kesehatan, misalnya hubungan antara tingkat pendidikan dan perilaku merokok, data mungkin tidak memenuhi asumsi normalitas. Maka, korelasi Spearman dapat digunakan untuk menilai hubungan tersebut. Jika diperoleh $\rho = -0,55$, ini menunjukkan hubungan monotonik negatif yang cukup kuat, artinya semakin tinggi tingkat pendidikan, cenderung perilaku merokok menurun.

13.1.4 Asumsi dan Kapan Menggunakan

Korelasi Pearson mengasumsikan bahwa data bersifat interval atau rasio, berdistribusi normal, dan hubungan linier. Jika data tidak memenuhi asumsi ini, atau terdapat outlier yang ekstrem, penggunaan korelasi Pearson dapat menyesatkan. Sebagai gantinya, korelasi Spearman lebih cocok digunakan untuk data ordinal, data yang tidak normal, atau hubungan yang tidak linier tetapi monotonik.

13.2 Analisis Regresi Linier Sederhana

13.2.1 Pengertian dan Tujuan

Regresi linier sederhana adalah metode statistik yang digunakan untuk memodelkan hubungan antara satu variabel independen (prediktor) dan satu variabel dependen (hasil). Tujuan utama dari analisis regresi adalah untuk mengetahui seberapa besar pengaruh variabel independen terhadap variabel dependen dan untuk memprediksi nilai variabel dependen berdasarkan nilai variabel independen.

Dalam konteks promosi kesehatan, regresi linier dapat digunakan untuk memperkirakan perubahan tekanan darah berdasarkan tingkat aktivitas fisik. Model ini membantu dalam memahami hubungan kuantitatif dan memberikan dasar untuk intervensi berbasis data.

13.2.2 Model Regresi Linier Sederhana

Model regresi linier sederhana dinyatakan sebagai:

$$Y = a + bX + \varepsilon$$

di mana:

- Y adalah variabel dependen (misalnya tekanan darah),
- X adalah variabel independen (misalnya tingkat aktivitas fisik),
- Konstanta a adalah menentukan titik potong garis regresi pada sumbu y, yaitu nilai variabel terikat ketika variabel bebas bernilai nol
- Konstanta b adalah menentukan kemiringan garis regresi dan menunjukkan seberapa besar perubahan

variable terikat untuk setiap perubahan satuan variable bebas

- ε adalah selisih antara nilai observasi variable terikat dengan nilai yang diprediksi oleh garis regresi

13.2.3 Langkah-langkah Analisis

Langkah utama dalam analisis regresi meliputi:

1. **Pengumpulan data:** Pastikan data lengkap dan memenuhi asumsi dasar.
2. **Pengujian asumsi:** Meliputi linearitas, normalitas residual, homoskedastisitas, dan tidak adanya multikolinearitas.
3. **Estimasi parameter:** Menggunakan metode kuadrat terkecil (least squares) untuk mendapatkan nilai b dan a .
4. **Pengujian signifikansi:** Melalui uji t untuk koefisien regresi dan uji F untuk keseluruhan model.
5. **Interpretasi hasil:** Meliputi nilai koefisien, tingkat signifikansi, dan koefisien determinasi (R^2).

Contoh: Jika hasil analisis menunjukkan ($b = -0,8$) dengan $p < 0,05$, ini berarti setiap peningkatan satu satuan dalam tingkat aktivitas fisik akan menurunkan tekanan darah rata-rata sebesar 0,8 mmHg, dan hubungan ini signifikan secara statistik.

13.2.4 Interpretasi Koefisien Regresi

Koefisien regresi (b) menunjukkan besarnya pengaruh variabel independen terhadap variabel dependen. Nilai positif menunjukkan hubungan langsung, sedangkan nilai negatif menunjukkan hubungan terbalik. Signifikansi statistik (p -value) menunjukkan apakah pengaruh tersebut nyata secara statistik atau hanya kebetulan.

13.3 Interpretasi Hasil Korelasi dan Regresi dalam Konteks Promosi Kesehatan

13.3.1 Interpretasi Korelasi

Hasil korelasi harus diinterpretasikan dengan memperhatikan kekuatan dan arah hubungan. Sebagai contoh, korelasi Pearson sebesar 0,70 menunjukkan hubungan positif yang cukup kuat, yang berarti bahwa peningkatan variabel X cenderung diikuti oleh peningkatan variabel Y. Sebaliknya, korelasi -0,50 menunjukkan hubungan negatif sedang, yang berarti bahwa peningkatan variabel X biasanya diikuti oleh penurunan variabel Y.

Dalam konteks promosi kesehatan, misalnya, korelasi positif antara tingkat pendidikan dan perilaku hidup bersih dan sehat menunjukkan bahwa peningkatan pendidikan dapat berkontribusi pada peningkatan perilaku sehat. Sebaliknya, korelasi negatif antara tingkat stres dan kualitas tidur menunjukkan bahwa stres yang tinggi dapat menurunkan kualitas tidur.

13.3.2 Interpretasi Regresi

Hasil regresi memberikan gambaran kuantitatif tentang pengaruh variabel independen terhadap variabel dependen. Koefisien regresi yang signifikan dan besar menunjukkan bahwa variabel independen memiliki pengaruh yang nyata dan praktis terhadap variabel dependen.

Sebagai contoh, jika model regresi menunjukkan bahwa tingkat aktivitas fisik berpengaruh signifikan terhadap tekanan darah dengan koefisien -0,8, maka program promosi kesehatan yang meningkatkan aktivitas fisik di masyarakat dapat diharapkan menurunkan tekanan darah secara rata-

rata sebesar 0,8 mmHg per satu satuan peningkatan aktivitas.

Selain itu, nilai R^2 menunjukkan seberapa besar variasi variabel dependen dapat dijelaskan oleh variabel independen. Misalnya, R^2 sebesar 0,45 berarti 45% variasi tekanan darah dapat dijelaskan oleh tingkat aktivitas fisik dalam model tersebut.

13.3.3 Implikasi Praktis

Hasil analisis korelasi dan regresi harus digunakan secara hati-hati dalam pengambilan keputusan. Korelasi yang tinggi dan signifikan dapat menjadi indikator penting untuk intervensi, tetapi tidak membuktikan sebab-akibat. Sedangkan, regresi memberikan prediksi dan pengaruh kuantitatif yang lebih jelas, sehingga dapat digunakan untuk merancang program yang lebih terarah dan berbasis bukti.

13.4 Contoh Aplikasi: Hubungan antara Tingkat Aktivitas Fisik dan Tekanan Darah

13.4.1 Deskripsi Studi Kasus

Misalnya, sebuah studi dilakukan untuk menilai hubungan antara tingkat aktivitas fisik dan tekanan darah pada sekelompok dewasa usia 30-50 tahun di sebuah komunitas urban. Data dikumpulkan melalui kuesioner aktivitas fisik dan pengukuran tekanan darah. Variabel aktivitas fisik diukur dalam satuan indeks aktivitas (misalnya, METs), dan tekanan darah diukur dalam mmHg.

13.4.2 Analisis Korelasi

Setelah data dianalisis, diperoleh koefisien korelasi Pearson sebesar -0,65 dengan $p < 0,01$. Ini menunjukkan hubungan negatif yang cukup kuat dan signifikan secara statistik, artinya semakin tinggi tingkat aktivitas fisik, biasanya tekanan darah cenderung lebih rendah. Interpretasi ini mendukung hipotesis bahwa peningkatan aktivitas fisik dapat membantu menurunkan tekanan darah, yang relevan dalam promosi gaya hidup sehat.

13.4.3 Analisis Regresi

Selanjutnya, dilakukan analisis regresi linier sederhana, dan diperoleh model:

$$Y = a + bX + \varepsilon$$

Dengan nilai $p < 0,05$ untuk koefisien (β_1), dan R^2 sebesar 0,42. Artinya, setiap peningkatan satu satuan indeks aktivitas fisik akan menurunkan tekanan darah rata-rata sebesar 0,8 mmHg, dan model ini menjelaskan 42% variasi tekanan darah dalam populasi tersebut.

13.4.4 Implikasi dalam Promosi Kesehatan

Hasil ini menunjukkan bahwa peningkatan aktivitas fisik dapat menjadi strategi efektif dalam menurunkan tekanan darah, sehingga program promosi kesehatan yang mendorong masyarakat untuk aktif secara fisik dapat berdampak positif terhadap pengendalian hipertensi. Selain itu, model regresi ini dapat digunakan untuk memprediksi tekanan darah individu berdasarkan tingkat aktivitasnya, sehingga dapat membantu tenaga kesehatan dalam

memberikan rekomendasi yang lebih personal dan berbasis data.

Dalam praktiknya, korelasi dan regresi merupakan alat yang sangat penting dalam analisis data kesehatan masyarakat. Korelasi membantu mengidentifikasi hubungan yang ada, baik linier maupun monotonik, sementara regresi memberikan gambaran kuantitatif tentang pengaruh satu variabel terhadap variabel lain serta kemampuan untuk melakukan prediksi. Pemilihan metode yang tepat harus didasarkan pada karakteristik data dan tujuan analisis, serta memperhatikan asumsi dasar dari masing-masing teknik.

Penggunaan kedua metode ini secara bersamaan dapat memberikan gambaran yang komprehensif tentang hubungan antar variabel dalam studi kesehatan, sehingga dapat mendukung pengambilan keputusan berbasis bukti dalam promosi dan pengendalian penyakit. Mahasiswa diharapkan mampu menguasai konsep ini secara mendalam, mampu melakukan analisis secara mandiri, serta mampu menginterpretasikan hasilnya secara kritis dan aplikatif dalam konteks kesehatan masyarakat.

Rangkuman

Pembahasan mengenai korelasi dan regresi dalam statistik kesehatan masyarakat sangat penting untuk memahami hubungan antar variabel dan pengaruhnya terhadap hasil kesehatan. Korelasi digunakan untuk mengukur kekuatan dan arah hubungan antara dua variabel, baik linier maupun monotonik, tanpa menyiratkan hubungan sebab-akibat. Korelasi Pearson cocok digunakan untuk data berskala interval atau rasio yang berdistribusi normal dan menunjukkan hubungan linier, sedangkan korelasi Spearman lebih fleksibel untuk data ordinal, tidak normal, atau hubungan yang tidak linier tetapi bersifat monotonik.

Regresi linier sederhana bertujuan untuk memodelkan dan memprediksi hubungan kuantitatif antara satu variabel independen dan satu variabel dependen. Model ini membantu dalam memahami besarnya pengaruh variabel prediktor terhadap hasil kesehatan dan dapat digunakan untuk merancang intervensi berbasis data. Langkah-langkah analisis regresi meliputi pengumpulan data, pengujian asumsi, estimasi parameter, pengujian signifikansi, dan interpretasi hasil.

Hasil dari analisis korelasi dan regresi harus diinterpretasikan secara hati-hati. Korelasi yang tinggi dan signifikan menunjukkan adanya hubungan yang kuat, tetapi tidak membuktikan sebab-akibat. Sedangkan regresi memberikan gambaran kuantitatif tentang pengaruh variabel independen terhadap variabel dependen dan dapat digunakan untuk prediksi serta pengambilan keputusan praktis dalam promosi kesehatan.

Contoh aplikasi nyata adalah studi tentang hubungan antara tingkat aktivitas fisik dan tekanan darah. Hasil analisis menunjukkan bahwa semakin tinggi tingkat aktivitas fisik, biasanya tekanan darah cenderung lebih rendah, dan model regresi dapat memperkirakan penurunan tekanan darah berdasarkan tingkat aktivitas tersebut. Implikasi dari hasil ini mendukung program promosi gaya hidup aktif sebagai strategi pengendalian hipertensi.

Secara umum, korelasi dan regresi merupakan alat analisis yang saling melengkapi dalam studi kesehatan masyarakat. Penguasaan konsep ini memungkinkan mahasiswa dan peneliti untuk melakukan analisis data secara mandiri, menginterpretasikan hasil secara kritis, dan mendukung pengambilan keputusan berbasis bukti dalam pengembangan program kesehatan. Penggunaan keduanya secara bersamaan akan memberikan gambaran yang

komprehensif tentang hubungan antar variabel dan pengaruhnya terhadap hasil kesehatan masyarakat.

Latihan Mahasiswa

Soal Essay

1. Jelaskan perbedaan utama antara korelasi Pearson dan Spearman, serta situasi kapan masing-masing metode sebaiknya digunakan dalam analisis data kesehatan masyarakat.
2. Sebutkan dan jelaskan langkah-langkah utama dalam melakukan analisis regresi linier sederhana beserta interpretasi dari koefisien regresi yang diperoleh.
3. Mengapa korelasi tidak dapat digunakan untuk menyimpulkan hubungan sebab-akibat antara dua variabel? Berikan penjelasan lengkap dan contoh dalam konteks kesehatan masyarakat.
4. Dalam sebuah studi, diperoleh nilai korelasi Pearson sebesar 0,85 antara tingkat pendidikan dan perilaku hidup bersih dan sehat. Bagaimana Anda menginterpretasikan hasil ini? Jelaskan secara statistik dan praktis.
5. Sebuah model regresi linier sederhana menunjukkan bahwa peningkatan satu satuan dalam indeks aktivitas fisik berhubungan dengan penurunan tekanan darah sebesar 0,8 mmHg, dengan $p\text{-value} < 0,05$. Jelaskan apa arti dari hasil ini dan bagaimana implikasinya dalam program promosi kesehatan.

Soal Pilihan Berganda

1. Koefisien korelasi Pearson berkisar antara...
 - a. 0 sampai 1

- b. -1 sampai 0
 - c. -1 sampai 1
 - d. 0 sampai 2
2. Korelasi Spearman digunakan terutama ketika data...
- a. Berdistribusi normal dan berskala interval
 - b. Tidak memenuhi asumsi normalitas dan berskala ordinal
 - c. Memiliki hubungan linier kuat
 - d. Memiliki outlier ekstrem
3. Jika nilai korelasi Pearson adalah 0,00, maka...
- a. Tidak ada hubungan sama sekali
 - b. Ada hubungan positif yang sangat kuat
 - c. Ada hubungan negatif yang sangat kuat
 - d. Tidak dapat ditentukan hubungan antara variabel
4. Dalam analisis regresi linier sederhana, koefisien regresi a menunjukkan...
- a. Nilai variabel dependen saat variabel independen = 0
 - b. Perubahan rata-rata variabel dependen untuk setiap satu satuan perubahan variabel independen
 - c. Hubungan non-linear antara variabel
 - d. Tingkat signifikansi dari model
5. Asumsi utama dalam regresi linier meliputi...
- a. Data harus berskala nominal
 - b. Residual harus berdistribusi normal dan homoskedastisitas

- c. Variabel harus berskala ordinal
 - d. Tidak ada outlier dalam data
6. Jika hasil analisis regresi menunjukkan bahwa p-value untuk koefisien (B) adalah 0,03, maka...
- a. Pengaruh variabel independen tidak signifikan
 - b. Pengaruh variabel independen signifikan secara statistik
 - c. Model tidak valid
 - d. Variabel independen tidak berpengaruh
7. Dalam konteks promosi kesehatan, korelasi positif yang tinggi antara dua variabel menunjukkan...
- a. Variabel satu menyebabkan variabel lain
 - b. Kedua variabel cenderung meningkat bersamaan
 - c. Tidak ada hubungan sama sekali
 - d. Variabel satu menurun saat variabel lain meningkat
8. Koefisien determinasi (R^2) dalam regresi menunjukkan...
- a. Persentase variasi variabel dependen yang dijelaskan oleh variabel independen
 - b. Hubungan linier antara dua variabel
 - c. Tingkat signifikansi model
 - d. Nilai korelasi antara variabel
9. Dalam studi kesehatan, korelasi negatif antara tingkat stres dan kualitas tidur berarti...
- a. Semakin tinggi stres, kualitas tidur cenderung lebih baik

- b. Semakin tinggi stres, kualitas tidur cenderung lebih buruk
 - c. Tidak ada hubungan antara stres dan tidur
 - d. Kualitas tidur menyebabkan stres
10. Dalam analisis regresi, apa yang dimaksud dengan homoskedastisitas?
- a. Varians residual sama di seluruh rentang nilai variabel independen
 - b. Residual berdistribusi normal
 - c. Variabel independen dan dependen berskala ordinal
 - d. Tidak ada outlier dalam data

Soal Studi Kasus / Project

1. Sebuah penelitian dilakukan untuk mengetahui hubungan antara tingkat aktivitas fisik dan tekanan darah pada sekelompok dewasa. Data dikumpulkan dari 100 responden, dan hasil analisis menunjukkan korelasi Pearson sebesar $-0,65$ dan model regresi: Tekanan Darah = $120 - 0,8 \times \text{Aktivitas Fisik}$ ($p < 0,01$, $R^2 = 0,42$). Jelaskan secara lengkap interpretasi dari hasil korelasi dan regresi tersebut, serta bagaimana hasil ini dapat digunakan dalam program promosi kesehatan di masyarakat.
2. Anda diminta untuk merancang sebuah studi untuk menilai hubungan antara tingkat pendidikan dan perilaku merokok di sebuah komunitas. Buatlah kerangka analisis yang mencakup jenis data yang akan dikumpulkan, metode analisis statistik yang akan digunakan (korelasi dan/atau regresi), serta interpretasi hasil yang diharapkan. Jelaskan juga bagaimana hasil analisis tersebut dapat membantu

dalam pengembangan program intervensi kesehatan masyarakat.

REFERENSI

- Gravetter, F. J., & Wallnau, L. B. (2017). *Statistics for Behavioral Sciences* (10th ed.). Cengage Learning.
- Schober, P., & Boer, C. (2018). Correlation coefficients: Appropriate use and interpretation. *Anesthesia & Analgesia*, 126(5), 1763–1768. <https://doi.org/10.1213/ANE.0000000000002832>
- Kutner, M. H., Nachtsheim, C. J., Neter, J., & Li, W. (2019). *Applied Linear Statistical Models* (5th ed.). McGraw-Hill Education.
- Munro, B. H. (2019). *Statistical Methods for Health Care Research* (8th ed.). Wolters Kluwer.

BAB 14

ANALISIS DATA KESEHATAN

Tujuan Pembelajaran

Setelah mempelajari bagian ini, diharapkan mahasiswa mampu mencapai beberapa kompetensi penting yang berkaitan dengan analisis data kesehatan. Pertama, mahasiswa dapat memahami dan menjelaskan berbagai indikator kesehatan yang umum digunakan dalam surveilans dan penelitian kesehatan, seperti angka kesakitan, angka kematian, dan prevalensi penyakit. Pemahaman ini menjadi dasar untuk menilai kondisi kesehatan masyarakat secara objektif dan komprehensif.

Kedua, mahasiswa mampu menginterpretasikan data surveilans kesehatan secara tepat. Mereka harus mampu membaca dan menganalisis data statistik yang diperoleh dari berbagai sumber, serta mengidentifikasi tren dan pola yang muncul dari data tersebut. Kemampuan ini sangat penting agar data yang diperoleh tidak hanya sekadar angka, tetapi dapat memberikan gambaran yang jelas tentang situasi kesehatan di masyarakat dan mendukung pengambilan keputusan yang berbasis bukti.

Ketiga, mahasiswa dapat menyusun laporan data statistik kesehatan secara sistematis dan komunikatif. Mereka harus mampu menyusun laporan yang mencakup penyajian data dalam bentuk tabel dan grafik, serta menulis narasi yang menjelaskan hasil analisis secara jelas dan logis. Penyusunan laporan yang baik akan memudahkan pihak terkait dalam

memahami kondisi kesehatan masyarakat dan merancang intervensi yang tepat sasaran.

Keempat, mahasiswa mampu mengintegrasikan konsep indikator kesehatan dengan proses surveilans dan pelaporan data. Mereka harus memahami bagaimana indikator tersebut digunakan untuk memantau dan mengevaluasi program kesehatan, serta bagaimana data tersebut dapat digunakan untuk mengidentifikasi masalah utama yang membutuhkan perhatian segera.

Kelima, mahasiswa mampu menggunakan data statistik kesehatan sebagai dasar dalam pengambilan keputusan dan perencanaan program kesehatan masyarakat. Mereka harus mampu menghubungkan hasil analisis data dengan langkah-langkah strategis yang diperlukan untuk meningkatkan status kesehatan masyarakat secara umum.

Keenam, mahasiswa dapat mengidentifikasi tantangan dan kendala yang mungkin muncul dalam pengumpulan, analisis, dan pelaporan data kesehatan. Mereka juga diharapkan mampu menyarankan solusi atau strategi untuk mengatasi hambatan tersebut agar proses pengelolaan data menjadi lebih efektif dan efisien.

Ketujuh, mahasiswa mampu menerapkan pengetahuan dan keterampilan analisis data kesehatan dalam konteks praktik lapangan dan penelitian. Mereka diharapkan mampu melakukan analisis data secara mandiri, serta mampu menyusun laporan yang lengkap dan akurat sesuai standar yang berlaku di bidang kesehatan masyarakat.

Pendahuluan

Bagian ini merupakan pengantar penting yang menegaskan urgensi dan relevansi analisis data kesehatan dalam konteks

promosi dan peningkatan kesehatan masyarakat. Data statistik kesehatan merupakan fondasi utama dalam pengambilan keputusan berbasis bukti. Tanpa data yang akurat dan terinterpretasi dengan baik, upaya peningkatan kesehatan masyarakat akan sulit dilakukan secara efektif dan efisien. Oleh karena itu, pemahaman tentang indikator kesehatan, interpretasi data surveilans, dan penyusunan laporan statistik menjadi kompetensi yang sangat penting bagi para profesional di bidang kesehatan masyarakat.

Indikator kesehatan seperti angka kesakitan, angka kematian, dan prevalensi penyakit merupakan alat ukur yang digunakan untuk menilai kondisi kesehatan suatu populasi. Data ini tidak hanya menunjukkan gambaran umum tentang beban penyakit, tetapi juga membantu dalam mengidentifikasi kelompok rentan, tren jangka panjang, serta efektivitas program kesehatan yang sedang berjalan. Dengan memahami indikator ini, tenaga kesehatan dan pengambil kebijakan dapat merancang intervensi yang lebih tepat sasaran dan berbasis bukti.

Selain itu, interpretasi data surveilans menjadi aspek krusial dalam memastikan bahwa data yang diperoleh benar-benar mencerminkan kondisi nyata di lapangan. Data yang tidak diinterpretasikan dengan benar dapat menimbulkan kesalahan dalam pengambilan keputusan, sehingga penting bagi mahasiswa untuk memahami cara membaca dan menganalisis data secara kritis dan objektif. Kemampuan ini akan membantu mereka dalam menilai keberhasilan program kesehatan, mengidentifikasi masalah utama, dan menentukan prioritas tindakan.

Selanjutnya, penyusunan laporan data statistik kesehatan merupakan tahapan akhir yang tidak kalah penting. Laporan yang disusun secara sistematis dan komunikatif akan memudahkan berbagai pihak, mulai dari tenaga lapangan,

pengambil kebijakan, hingga masyarakat umum, dalam memahami kondisi kesehatan yang sedang berlangsung. Penyajian data dalam bentuk tabel dan grafik, disertai narasi yang jelas, akan meningkatkan efektivitas komunikasi dan mendukung pengambilan keputusan yang berbasis bukti.

Dalam konteks praktik lapangan dan penelitian, kemampuan mengelola dan melaporkan data kesehatan secara tepat juga menjadi kunci keberhasilan. Mahasiswa harus mampu mengaplikasikan pengetahuan ini dalam situasi nyata, baik saat melakukan surveilans, evaluasi program, maupun penelitian ilmiah. Dengan demikian, mereka tidak hanya memahami teori, tetapi juga mampu menerapkan keterampilan analisis data secara mandiri dan profesional.

Secara keseluruhan, bagian ini menegaskan bahwa penguasaan analisis data kesehatan tidak hanya penting untuk keperluan akademik, tetapi juga merupakan kompetensi esensial yang harus dimiliki oleh setiap profesional di bidang kesehatan masyarakat. Penguasaan terhadap indikator kesehatan, interpretasi data surveilans, dan penyusunan laporan statistik akan memperkuat peran mereka dalam meningkatkan kualitas layanan dan keberhasilan program kesehatan di masyarakat.

14.1 Indikator Kesehatan (Angka Kesakitan, Kematian, Prevalensi)

14.1.1 Definisi dan Pentingnya Indikator Kesehatan

Indikator kesehatan merupakan alat ukur yang digunakan untuk menilai kondisi kesehatan masyarakat secara kuantitatif. Indikator ini membantu pengambil keputusan dalam merancang, melaksanakan, dan mengevaluasi program kesehatan. Tiga indikator utama yang sering

digunakan dalam surveilans dan analisis data kesehatan adalah angka kesakitan, angka kematian, dan prevalensi penyakit.

Angka kesakitan mengukur beban penyakit yang dialami oleh populasi dalam periode tertentu, biasanya dinyatakan dalam bentuk insiden atau kejadian baru. Angka ini penting untuk mengetahui seberapa sering suatu penyakit muncul dan menyebar di masyarakat. Sebaliknya, angka kematian menunjukkan tingkat kematian yang terjadi dalam populasi tertentu, memberikan gambaran tentang tingkat keparahan dan dampak penyakit terhadap kehidupan masyarakat. Prevalensi, di sisi lain, mengukur jumlah kasus penyakit yang ada pada suatu waktu tertentu, baik baru maupun yang sudah ada sebelumnya, dan digunakan untuk menilai beban penyakit secara keseluruhan.

14.1.2 Rumus dan Penghitungan Indikator Kesehatan

Penghitungan indikator ini dilakukan dengan rumus yang telah distandarisasi dan digunakan secara internasional. Berikut adalah rumus dasar dari masing-masing indikator:

- **Angka Kesakitan (Incidence Rate):** $IR = (\text{Jumlah kasus baru} / \text{populasi berisiko}) * \text{factor konversi}$. Contoh: Jika dalam satu tahun terdapat 150 kasus baru influenza di sebuah kota dengan populasi 500.000, maka: $IR = 150 / 500.000 * 100.000 = 30\{\text{kasus per 100.000 penduduk}\}$
- **Angka Kematian (Mortality Rate):** $MR = (\text{Jumlah kematian} / \text{jumlah penduduk}) * k$. Contoh: Jika terdapat 50 kematian akibat penyakit jantung dalam satu tahun di populasi 1.000.000, maka: $MR = (50 / 1.000.000) * 100.000 = 5\{\text{kematian per 100.000 penduduk}\}$

- **Prevalensi:** $\text{Prevalensi} = \{\text{Jumlah kasus penyakit yang ada pada waktu tertentu} / \text{Jumlah populasi pada waktu tersebut} * 100\}$. Contoh: Jika pada survei dilakukan pada bulan Januari dan ditemukan 200 penderita diabetes dari total populasi 10.000 orang, maka: $\text{Prevalensi} = (200 / 10.000) * 100\% = 2\%$

14.1.3 Interpretasi dan Penggunaan Indikator Kesehatan

Interpretasi indikator harus dilakukan secara hati-hati dan mempertimbangkan konteks epidemiologis serta faktor sosial ekonomi yang mempengaruhi data. Sebagai contoh, angka kematian yang tinggi dari suatu penyakit bisa menunjukkan tingkat keparahan penyakit tersebut, tetapi juga bisa dipengaruhi oleh kualitas layanan kesehatan, akses terhadap pengobatan, dan faktor lain seperti tingkat pendidikan dan kesadaran masyarakat.

Misalnya, angka kematian ibu (AKI) yang tinggi di suatu daerah dapat menunjukkan kurangnya fasilitas kesehatan maternal, rendahnya tingkat pendidikan, atau budaya yang tidak mendukung ibu untuk mendapatkan perawatan prenatal dan persalinan di fasilitas kesehatan. Oleh karena itu, data ini harus dianalisis secara komprehensif dan tidak hanya dilihat dari angka semata.

Penggunaan indikator ini dalam surveilans sangat penting untuk memantau tren penyakit dari waktu ke waktu. Jika angka kesakitan meningkat, hal ini bisa menjadi indikator adanya wabah atau faktor risiko baru yang perlu diidentifikasi dan diatasi. Sebaliknya, penurunan angka kematian menunjukkan keberhasilan program kesehatan tertentu, seperti imunisasi, pengendalian penyakit menular, atau peningkatan akses layanan kesehatan.

14.1.4 Contoh Kasus Penggunaan Indikator Kesehatan

Sebagai ilustrasi, misalkan sebuah kabupaten melaporkan data surveilans penyakit demam berdarah dengue (DBD). Data menunjukkan bahwa selama tahun 2022 terdapat 1.200 kasus baru DBD dari total populasi 250.000 orang. Maka, angka insidennya adalah:

- *Incidence Rate DBD = $(1.200 / 250.000) \times 100.000 = 480$ kasus per 100.000 penduduk.*

Selain itu, jika selama tahun yang sama terdapat 15 kematian akibat DBD, maka angka kematiannya adalah:

- *Mortality Rate DBD = $(15 / 250.000) \times 100.000 = 6$ kematian per 100.000 penduduk.*

Data ini menunjukkan bahwa meskipun angka insidennya cukup tinggi, angka kematian relatif rendah, yang bisa mengindikasikan keberhasilan pengelolaan kasus dan pengobatan di daerah tersebut.

14.1.5 Keterbatasan dan Tantangan Penggunaan Indikator Kesehatan

Penggunaan indikator kesehatan tidak lepas dari berbagai tantangan dan keterbatasan. Salah satunya adalah masalah keakuratan data. Data yang diperoleh dari surveilans seringkali dipengaruhi oleh underreporting, misclassification, atau kekurangan sumber daya dalam pengumpulan data. Misalnya, kasus penyakit menular tertentu mungkin tidak dilaporkan secara lengkap karena keterbatasan fasilitas kesehatan atau ketidakpatuhan masyarakat dalam melapor.

Selain itu, indikator ini juga harus disesuaikan dengan konteks lokal dan karakteristik populasi. Sebagai contoh, angka kematian bayi (AKB) di daerah tertentu mungkin tinggi karena faktor sosial ekonomi, budaya, dan akses terhadap layanan kesehatan, bukan semata-mata karena tingkat keparahan penyakit.

Penggunaan indikator juga harus memperhatikan periode waktu pengamatan. Data yang diambil dalam periode tertentu harus konsisten dan representatif agar dapat digunakan untuk analisis tren jangka panjang. Perubahan dalam metode pengumpulan data atau definisi kasus juga dapat mempengaruhi interpretasi indikator.

Indikator kesehatan seperti angka kesakitan, angka kematian, dan prevalensi merupakan alat penting dalam pengukuran dan penilaian kondisi kesehatan masyarakat. Penghitungan dan interpretasi indikator ini harus dilakukan secara cermat dan kontekstual agar data yang diperoleh dapat digunakan secara efektif dalam perencanaan dan evaluasi program kesehatan. Penggunaan data yang akurat dan tepat akan membantu mengidentifikasi masalah utama, mengukur keberhasilan intervensi, serta mendukung pengambilan keputusan berbasis bukti untuk meningkatkan kualitas hidup masyarakat.

14.2 Interpretasi Data Surveilans

14.2.1 Pengertian dan Tujuan Interpretasi Data Surveilans

Interpretasi data surveilans adalah proses menganalisis dan menafsirkan data yang diperoleh dari kegiatan pemantauan kesehatan masyarakat untuk memahami kondisi kesehatan, tren, dan pola kejadian penyakit. Tujuan utama dari

interpretasi ini adalah untuk mengidentifikasi masalah kesehatan utama, menilai efektivitas program, serta menentukan prioritas intervensi yang diperlukan.

Data surveilans yang dikumpulkan biasanya berupa angka kejadian, tingkat prevalensi, distribusi geografis, dan faktor risiko yang terkait. Tanpa interpretasi yang tepat, data tersebut hanya akan menjadi angka-angka yang tidak bermakna. Oleh karena itu, kemampuan untuk membaca dan menganalisis data secara kritis sangat penting agar informasi yang dihasilkan dapat digunakan secara optimal dalam pengambilan keputusan.

14.2.2 Langkah-langkah Interpretasi Data Surveilans

Proses interpretasi data surveilans meliputi beberapa langkah penting:

1. **Verifikasi Data:** Memastikan bahwa data yang diperoleh lengkap, akurat, dan konsisten. Hal ini meliputi pengecekan kesalahan entri data, validitas sumber data, dan kejelasan definisi kasus.
2. **Analisis Deskriptif:** Menggambarkan data secara statistik, seperti menghitung angka kejadian, prevalensi, distribusi usia, jenis kelamin, wilayah geografis, dan faktor risiko lainnya.
3. **Identifikasi Tren dan Pola:** Melihat adanya peningkatan, penurunan, atau stabilitas angka kejadian dari waktu ke waktu. Juga, mengidentifikasi pola distribusi geografis dan demografis.
4. **Perbandingan dengan Data Sebelumnya atau Standar:** Membandingkan data saat ini dengan data historis atau standar nasional/internasional untuk menilai apakah kondisi membaik, memburuk, atau tetap.

5. **Analisis Faktor Penyebab dan Konteks Sosial Ekonomi:** Mengkaji faktor yang mungkin mempengaruhi data, seperti perubahan lingkungan, kebijakan, perilaku masyarakat, dan faktor sosial ekonomi.
6. **Pengambilan Kesimpulan dan Rekomendasi:** Berdasarkan analisis, menyusun kesimpulan mengenai kondisi kesehatan dan merekomendasikan langkah-langkah yang perlu diambil.

14.2.3 Contoh Interpretasi Data Surveilans

Misalnya, data surveilans penyakit tuberkulosis (TB) di sebuah provinsi menunjukkan bahwa selama lima tahun terakhir, angka kejadian TB meningkat dari 50 menjadi 80 kasus per 100.000 penduduk. Analisis lebih dalam menunjukkan bahwa peningkatan ini terutama terjadi di daerah perkotaan dan di kalangan kelompok usia produktif (20-40 tahun). Selain itu, data menunjukkan bahwa tingkat pengobatan lengkap mencapai 70%, sementara target nasional adalah 85%.

Interpretasi dari data ini mengindikasikan bahwa ada masalah dalam pengendalian TB, terutama di daerah perkotaan dan di kalangan usia produktif. Faktor-faktor yang mungkin mempengaruhi adalah ketidakpatuhan pengobatan, kurangnya fasilitas kesehatan di daerah tertentu, dan faktor sosial ekonomi yang mempengaruhi akses terhadap layanan kesehatan. Rekomendasi yang muncul adalah meningkatkan program pengobatan langsung, memperluas layanan kesehatan di daerah urban, dan melakukan edukasi masyarakat tentang pentingnya pengobatan lengkap.

14.2.4 Tantangan dalam Interpretasi Data Surveilans

Beberapa tantangan utama dalam interpretasi data surveilans meliputi:

- **Data Tidak Lengkap atau Tidak Akurat:** Underreporting, misclassification, dan kekurangan sumber daya dapat menyebabkan data tidak mencerminkan kondisi sebenarnya.
- **Perubahan Definisi Kasus:** Perubahan standar definisi kasus dari waktu ke waktu dapat mempengaruhi konsistensi data.
- **Bias dan Confounding:** Faktor-faktor yang tidak terkontrol dapat mempengaruhi hasil analisis dan interpretasi.
- **Perbedaan Kualitas Data antar Wilayah:** Variasi dalam kapasitas dan sistem pelaporan antar daerah dapat menyebabkan ketidaksetaraan data.

Mengatasi tantangan ini memerlukan peningkatan kapasitas petugas surveilans, penggunaan sistem data yang terstandarisasi, serta pelatihan dalam analisis dan interpretasi data.

14.2.5 Peran Interpretasi Data dalam Pengambilan Keputusan

Interpretasi data surveilans yang tepat akan memberikan gambaran yang jelas tentang situasi kesehatan masyarakat, sehingga mendukung pengambilan keputusan yang berbasis bukti. Sebagai contoh, jika data menunjukkan peningkatan kejadian penyakit tertentu di wilayah tertentu, maka intervensi segera dapat diarahkan ke daerah tersebut, seperti peningkatan kegiatan edukasi, vaksinasi, atau peningkatan fasilitas layanan kesehatan.

Sebaliknya, interpretasi yang keliru atau tidak tepat dapat menyebabkan alokasi sumber daya yang tidak efisien, atau bahkan menimbulkan kekhawatiran yang tidak berdasar. Oleh karena itu, interpretasi harus dilakukan secara objektif, didukung data yang valid, dan disertai analisis kontekstual yang komprehensif.

14.3 Penyusunan Laporan Data Statistik Kesehatan

14.3.1 Tujuan dan Fungsi Laporan Statistik Kesehatan

Laporan statistik kesehatan merupakan dokumen formal yang menyajikan data dan analisis kondisi kesehatan masyarakat secara sistematis dan komprehensif. Tujuan utama dari penyusunan laporan ini adalah untuk menyediakan informasi yang akurat, relevan, dan mudah dipahami oleh berbagai pihak, termasuk pengambil kebijakan, tenaga kesehatan, peneliti, dan masyarakat umum.

Fungsi utama laporan statistik kesehatan meliputi:

- Memberikan gambaran kondisi kesehatan masyarakat secara lengkap dan objektif.
- Menjadi dasar dalam perencanaan, pengendalian, dan evaluasi program kesehatan.
- Menjadi alat komunikasi antara petugas surveilans, pengambil kebijakan, dan masyarakat.
- Menjadi bahan referensi dalam penelitian dan pengembangan kebijakan kesehatan.

14.3.2 Komponen Utama dalam Penyusunan Laporan

Laporan statistik kesehatan harus disusun secara sistematis dan lengkap, mencakup beberapa komponen utama:

1. **Pendahuluan:** Menjelaskan latar belakang, tujuan, dan ruang lingkup laporan.
2. **Metodologi:** Menguraikan sumber data, metode pengumpulan data, definisi kasus, dan teknik analisis yang digunakan.
3. **Hasil:** Menyajikan data utama dalam bentuk tabel, grafik, dan narasi analisis. Data ini meliputi indikator kesehatan, tren, distribusi geografis, dan faktor risiko.
4. **Pembahasan:** Menginterpretasikan hasil, membandingkan dengan data sebelumnya atau standar, serta membahas implikasi hasil.
5. **Kesimpulan dan Rekomendasi:** Menyimpulkan temuan utama dan memberikan saran tindakan yang diperlukan.
6. **Daftar Pustaka:** Menyertakan sumber data dan referensi yang digunakan.

14.3.3 Teknik Penyajian Data dalam Laporan

Penyajian data yang efektif sangat penting agar informasi dapat dipahami dengan mudah. Teknik yang umum digunakan meliputi:

- **Tabel:** Untuk menyajikan data kuantitatif secara sistematis dan lengkap. Contohnya, tabel distribusi kejadian penyakit berdasarkan wilayah dan usia.
- **Grafik:** Untuk menunjukkan tren dan pola secara visual. Contoh, grafik garis yang menunjukkan tren angka kematian dari tahun ke tahun.

- **Diagram Batang dan Lingkaran:** Untuk memperlihatkan perbandingan antar kategori, seperti distribusi penyakit berdasarkan jenis kelamin atau wilayah.
- **Peta Tematik:** Untuk menampilkan distribusi geografis kejadian penyakit.

Penggunaan kombinasi teknik ini akan meningkatkan kejelasan dan daya tarik laporan.

14.3.4 Penulisan Narasi dan Interpretasi Hasil

Selain menyajikan data secara visual, narasi yang menjelaskan hasil harus disusun secara logis dan sistematis. Narasi ini harus mampu menjelaskan makna dari data, tren yang teramati, serta faktor-faktor yang mempengaruhi kondisi kesehatan. Misalnya, jika data menunjukkan peningkatan angka kejadian DBD di daerah perkotaan, narasi harus mengaitkan dengan faktor lingkungan, kebersihan, dan pengelolaan air.

Interpretasi harus didukung oleh data dan analisis yang objektif, serta menghindari kesalahan interpretasi yang dapat menyesatkan. Penting juga untuk menyertakan batasan data dan kemungkinan bias yang ada.

14.3.5 Contoh Penyusunan Laporan Statistik Kesehatan

Misalnya, sebuah laporan surveilans penyakit infeksi saluran pernapasan akut (ISPA) di sebuah kabupaten menyajikan data sebagai berikut:

- **Jumlah kasus selama tahun 2022:** 3.500 kasus dari total populasi 150.000.
- **Distribusi usia:** 60% terjadi pada anak-anak usia 1-5 tahun.

- **Distribusi wilayah:** 70% kasus berasal dari daerah perkotaan.
- **Grafik tren:** Menunjukkan peningkatan kasus dari bulan Januari hingga Juli, kemudian menurun.

Narasi yang menyertai data ini akan menjelaskan bahwa peningkatan kasus selama paruh pertama tahun kemungkinan dipengaruhi oleh faktor musim dan kepadatan penduduk di daerah perkotaan. Rekomendasi yang diusulkan adalah meningkatkan promosi kesehatan, memperkuat layanan pengobatan di puskesmas, dan melakukan surveilans lebih intensif selama musim puncak.

14.3.6 Tantangan dan Solusi dalam Penyusunan Laporan

Tantangan utama dalam penyusunan laporan statistik kesehatan meliputi:

- **Keterbatasan data:** Data yang tidak lengkap atau tidak akurat dapat mengurangi validitas laporan.
- **Keterbatasan sumber daya:** Kurangnya tenaga ahli dan fasilitas pendukung dapat memperlambat proses penyusunan laporan.
- **Perbedaan standar dan format:** Variasi dalam format laporan antar daerah dapat menyulitkan perbandingan data.

Solusi yang dapat dilakukan meliputi peningkatan kapasitas petugas, penggunaan perangkat lunak yang memudahkan pengolahan data, serta standarisasi format laporan nasional.

14.3.7 Pentingnya Penyusunan Laporan yang Sistematis dan Komunikatif

Laporan yang disusun secara sistematis dan komunikatif akan memudahkan semua pihak dalam memahami kondisi kesehatan masyarakat. Penyajian data yang jelas, lengkap, dan disertai interpretasi yang tepat akan meningkatkan efektivitas pengambilan keputusan dan perencanaan program. Mahasiswa harus mampu menyusun laporan yang memenuhi standar ini, serta mampu menyampaikan hasil secara lisan maupun tertulis dengan bahasa yang mudah dipahami.

Rangkuman

Pendahuluan ini memberikan gambaran umum mengenai pentingnya indikator kesehatan dalam pengukuran kondisi kesehatan masyarakat. Indikator ini berfungsi sebagai alat kuantitatif yang membantu pengambil keputusan dalam merancang, melaksanakan, dan mengevaluasi program kesehatan. Tiga indikator utama yang sering digunakan adalah angka kesakitan, angka kematian, dan prevalensi penyakit, yang masing-masing memiliki rumus penghitungan standar dan interpretasi yang penting untuk memahami beban penyakit dan dampaknya terhadap populasi.

Penghitungan indikator dilakukan dengan rumus yang telah distandarisasi, seperti incidence rate, mortality rate, dan prevalensi, yang memungkinkan perbandingan antar wilayah dan waktu. Interpretasi indikator harus dilakukan secara hati-hati, mempertimbangkan faktor sosial ekonomi, kualitas data, dan konteks epidemiologis. Data yang akurat dan lengkap sangat penting agar hasil analisis dapat

digunakan secara efektif dalam pengambilan kebijakan dan perencanaan program.

Contoh kasus penggunaan indikator menunjukkan bagaimana data insidensi dan angka kematian dapat memberikan gambaran tentang situasi penyakit tertentu di suatu daerah. Misalnya, data surveilans DBD menunjukkan angka insidensi tinggi namun angka kematian rendah, menandakan keberhasilan pengelolaan kasus. Namun, tantangan seperti underreporting, perubahan definisi kasus, dan perbedaan kualitas data menjadi hambatan dalam penggunaan indikator secara optimal.

Penggunaan indikator ini juga harus disertai analisis yang menyeluruh dan kontekstual agar dapat mendukung pengambilan keputusan yang tepat. Dengan demikian, indikator kesehatan merupakan alat penting dalam pengukuran dan pemantauan kondisi kesehatan masyarakat, yang harus diolah dan diinterpretasikan secara cermat untuk mendukung keberhasilan program kesehatan dan peningkatan kualitas hidup masyarakat.

Latihan Mahasiswa

Soal Essay

1. Jelaskan pengertian indikator kesehatan dan sebutkan tiga indikator utama yang digunakan dalam surveilans kesehatan masyarakat beserta contohnya!
2. Bagaimana rumus penghitungan angka kesakitan (incidence rate) dan apa maknanya dalam konteks kesehatan masyarakat?
3. Mengapa interpretasi data surveilans harus dilakukan secara hati-hati dan apa saja faktor yang

perlu dipertimbangkan saat menganalisis data tersebut?

4. Sebutkan dan jelaskan langkah-langkah utama dalam proses interpretasi data surveilans kesehatan!
5. Jelaskan pentingnya penyusunan laporan statistik kesehatan secara sistematis dan komunikatif serta sebutkan komponen utama yang harus ada dalam laporan tersebut!

Soal Pilihan Berganda

1. Manakah dari berikut ini yang merupakan indikator utama dalam surveilans kesehatan masyarakat?
 - a. Tingkat pendidikan
 - b. Angka kesakitan
 - c. Pendapatan per kapita
 - d. Tingkat pengangguran
2. Rumus untuk menghitung angka kematian (mortality rate) adalah:
 - a. $(\text{Jumlah kematian} / \text{jumlah populasi}) \times 100.000$
 - b. $(\text{Jumlah kasus baru} / \text{jumlah populasi}) \times 100.000$
 - c. $(\text{Jumlah penderita} / \text{jumlah populasi}) \times 100$
 - d. $(\text{Jumlah kematian} / \text{jumlah kasus}) \times 100$
3. Prevalensi digunakan untuk mengukur:
 - a. Jumlah kasus baru dalam periode tertentu
 - b. Jumlah kasus penyakit yang ada pada waktu tertentu
 - c. Tingkat kematian akibat penyakit tertentu
 - d. Insiden penyakit dalam populasi

4. Data surveilans yang tidak lengkap atau tidak akurat dapat disebabkan oleh:
 - a. Overreporting
 - b. Underreporting
 - c. Penggunaan software statistik
 - d. Pengumpulan data secara sistematis
5. Salah satu langkah penting dalam interpretasi data surveilans adalah:
 - a. Mengabaikan faktor sosial ekonomi
 - b. Verifikasi data
 - c. Mengurangi jumlah data
 - d. Menghindari analisis tren
6. Grafik yang paling tepat digunakan untuk menunjukkan tren angka kejadian dari waktu ke waktu adalah:
 - a. Diagram lingkaran
 - b. Histogram
 - c. Grafik garis
 - d. Diagram batang
7. Dalam penyusunan laporan statistik kesehatan, komponen utama yang harus disertakan adalah:
 - a. Pendahuluan, metodologi, hasil, dan kesimpulan
 - b. Hanya tabel dan grafik
 - c. Hanya narasi analisis
 - d. Daftar pustaka saja
8. Salah satu tantangan utama dalam penyusunan laporan statistik adalah:

- a. Data yang lengkap dan akurat
 - b. Variasi dalam format laporan antar daerah
 - c. Penggunaan perangkat lunak yang canggih
 - d. Ketersediaan sumber daya manusia
9. Mengapa penting menyusun laporan statistik secara sistematis dan komunikatif?
- a. Agar data terlihat lebih rumit
 - b. Untuk memudahkan pengambilan keputusan dan pemahaman semua pihak
 - c. Supaya laporan lebih panjang
 - d. Untuk mengurangi jumlah data yang disajikan
10. Salah satu manfaat utama dari interpretasi data surveilans adalah:
- a. Mengurangi jumlah kasus penyakit
 - b. Menyusun laporan secara otomatis
 - c. Mengidentifikasi masalah utama dan prioritas intervensi
 - d. Menghindari pengumpulan data baru

Soal Project / Studi Kasus

1. Sebuah kota melaporkan data surveilans influenza selama satu tahun dengan jumlah penduduk 600.000 dan terdapat 900 kasus baru. Hitunglah angka insidennya dan jelaskan apa arti dari angka tersebut dalam konteks pengendalian penyakit di kota tersebut!
2. Data surveilans menunjukkan bahwa di sebuah kabupaten, angka kematian ibu (AKI) meningkat dari 150 per 100.000 kelahiran hidup menjadi 200 per 100.000 dalam dua tahun terakhir. Sebagai seorang petugas kesehatan masyarakat, apa langkah-langkah

yang akan Anda lakukan untuk menanggapi data ini?
Jelaskan secara singkat dan sistematis!

REFERENSI

1. World Health Organization. (2018). *Monitoring the health-related Sustainable Development Goals (SDGs)*. WHO.
2. Park, K. (2020). *Textbook of Preventive and Social Medicine* (24th ed.). Banarsidas Bhanot.
3. Thrusfield, M. (2018). *Veterinary Epidemiology* (4th ed.). Wiley-Blackwell.
4. WHO. (2019). *Global Tuberculosis Report 2019*. WHO.
5. Kemenkes RI. (2020). *Profil Kesehatan Indonesia 2020*. Kementerian Kesehatan Republik Indonesia. " "

BAB 15

PENGUNAAN SOFTWARE STATISTIK

Tujuan Pembelajaran

Setelah mempelajari bagian ini, diharapkan mahasiswa mampu mencapai beberapa tujuan penting yang akan mendukung penguasaan kompetensi dalam penggunaan software statistik untuk analisis data kesehatan. Tujuan utama dari bagian ini adalah agar mahasiswa dapat memahami dan mengaplikasikan berbagai perangkat lunak statistik yang umum digunakan dalam bidang kesehatan masyarakat, khususnya SPSS, Excel, dan Epi Info. Selain itu, mahasiswa diharapkan mampu melakukan input data secara benar, melakukan analisis dasar, serta mampu menginterpretasikan hasil output dari software tersebut secara tepat dan relevan dengan konteks penelitian atau surveilans kesehatan.

Secara spesifik, tujuan pembelajaran yang ingin dicapai meliputi: pertama, mahasiswa mampu mengenali dan memahami fitur utama dari SPSS, Excel, dan Epi Info yang relevan untuk analisis data kesehatan. Kedua, mahasiswa mampu melakukan proses input data secara sistematis dan akurat ke dalam masing-masing software tersebut, sehingga data yang dianalisis benar-benar mencerminkan data asli yang dikumpulkan. Ketiga, mahasiswa mampu melakukan analisis dasar seperti penghitungan frekuensi, statistik deskriptif, dan uji sederhana menggunakan software yang dipelajari. Keempat, mahasiswa mampu membaca dan

menafsirkan output hasil analisis dari software tersebut, termasuk memahami nilai statistik, p-value, dan indikator lain yang digunakan dalam pengambilan keputusan statistik. Kelima, mahasiswa mampu mengidentifikasi kekurangan dan kelebihan masing-masing software dalam konteks analisis data kesehatan, serta mampu memilih software yang paling sesuai untuk kebutuhan tertentu. Keenam, mahasiswa mampu mengintegrasikan hasil analisis dari software tersebut ke dalam laporan statistik kesehatan yang lengkap dan komunikatif. Ketujuh, mahasiswa diharapkan mampu mengembangkan kemampuan kritis dalam menilai keakuratan dan validitas data serta hasil analisis yang diperoleh dari software statistik.

Dengan pencapaian tujuan tersebut, diharapkan mahasiswa tidak hanya mampu melakukan analisis statistik secara teknis, tetapi juga mampu menginterpretasikan hasilnya secara tepat dalam konteks kesehatan masyarakat. Kemampuan ini sangat penting karena data statistik yang akurat dan interpretasi yang tepat akan mendukung pengambilan keputusan yang berbasis bukti, serta meningkatkan kualitas laporan dan penelitian di bidang kesehatan. Penguasaan penggunaan software statistik ini juga akan memudahkan mahasiswa dalam melakukan surveilans, evaluasi program, maupun penelitian ilmiah yang membutuhkan analisis data kuantitatif secara cepat dan efisien.

Pendahuluan

Dalam era digital saat ini, pengolahan dan analisis data menjadi bagian yang tidak terpisahkan dari kegiatan penelitian dan pengambilan keputusan di bidang kesehatan masyarakat. Data yang dikumpulkan dari berbagai sumber, seperti surveilans, studi epidemiologi, maupun evaluasi program, memerlukan pengolahan yang sistematis agar

hasilnya dapat diinterpretasikan secara tepat dan akurat. Di sinilah peran penting dari perangkat lunak statistik yang dirancang untuk memudahkan proses analisis data tersebut. Penggunaan software statistik tidak hanya mempercepat proses pengolahan data, tetapi juga meningkatkan keakuratan dan keandalan hasil analisis yang diperoleh.

BAB 15 ini secara khusus membahas tentang penggunaan software statistik yang umum digunakan dalam bidang kesehatan, yaitu SPSS, Excel, dan Epi Info. Ketiga perangkat lunak ini memiliki keunggulan dan kekurangan masing-masing, serta digunakan sesuai dengan kebutuhan dan tingkat kompleksitas analisis yang dihadapi. SPSS dikenal luas karena kemampuannya dalam melakukan analisis statistik yang kompleks dan menyediakan output yang lengkap serta mudah dipahami. Excel, sebagai bagian dari paket Microsoft Office, merupakan alat yang familiar dan banyak digunakan untuk pengolahan data dasar, penghitungan statistik sederhana, serta pembuatan grafik. Sementara itu, Epi Info dirancang khusus untuk keperluan surveilans dan analisis epidemiologi, menawarkan fitur yang sesuai untuk analisis data kesehatan masyarakat secara cepat dan efisien.

Penguasaan terhadap ketiga software ini sangat penting bagi mahasiswa dan profesional di bidang kesehatan masyarakat. Dengan kemampuan input data yang benar, mahasiswa dapat memastikan bahwa data yang dianalisis adalah data yang valid dan reliabel. Selanjutnya, melalui analisis dasar yang dilakukan dengan software tersebut, mahasiswa dapat memperoleh gambaran statistik yang diperlukan untuk pengambilan keputusan, seperti menentukan tingkat prevalensi, risiko, atau hubungan antar variabel. Tidak kalah penting, kemampuan membaca dan menafsirkan output dari software ini akan membantu mahasiswa dalam menyusun

laporan yang akurat dan komunikatif, serta mendukung pengambilan kebijakan berbasis bukti.

Selain aspek teknis, bagian ini juga menekankan pentingnya pemahaman terhadap proses analisis statistik secara umum, sehingga mahasiswa tidak hanya sekedar mengikuti langkah-langkah, tetapi juga mampu memahami makna dari hasil yang diperoleh. Hal ini menjadi fondasi penting dalam pengembangan kompetensi analisis data yang kritis dan bertanggung jawab di bidang kesehatan masyarakat. Dengan demikian, bagian ini tidak hanya berfungsi sebagai panduan praktis, tetapi juga sebagai penguatan konsep dasar yang akan mendukung penguasaan analisis statistik secara menyeluruh dan aplikatif.

15.1 Pengenalan Software Statistik: SPSS, Excel, dan Epi Info

Penggunaan software statistik dalam analisis data kesehatan merupakan aspek penting yang harus dikuasai oleh mahasiswa dan profesional di bidang kesehatan masyarakat. Ketiga perangkat lunak yang akan dibahas, yaitu SPSS, Excel, dan Epi Info, memiliki karakteristik dan fungsi yang berbeda, namun semuanya mendukung proses pengolahan data secara efisien dan akurat. Pemahaman mendalam tentang fitur utama, keunggulan, dan keterbatasan masing-masing software akan membantu pengguna dalam memilih alat yang paling sesuai dengan kebutuhan analisis tertentu.

15.1.1. SPSS (Statistical Package for the Social Sciences)

SPSS merupakan salah satu software statistik yang paling populer dan banyak digunakan dalam penelitian kesehatan dan ilmu sosial. Dikembangkan oleh IBM, SPSS menawarkan antarmuka pengguna yang intuitif dan fitur lengkap untuk

analisis statistik deskriptif, inferensial, serta analisis multivariat. Keunggulan utama SPSS terletak pada kemampuannya dalam melakukan analisis statistik yang kompleks tanpa memerlukan pemrograman yang rumit, sehingga cocok untuk pengguna yang tidak memiliki latar belakang pemrograman.

Fitur utama SPSS meliputi pengolahan data, analisis statistik dasar seperti frekuensi, mean, median, serta uji statistik parametrik dan non-parametrik. Selain itu, SPSS juga mendukung analisis regresi, analisis faktor, analisis cluster, dan analisis survival, yang sangat relevan dalam penelitian kesehatan masyarakat. Output dari SPSS disajikan dalam bentuk tabel dan grafik yang lengkap dan mudah dipahami, serta dapat diekspor ke berbagai format seperti Word, Excel, dan PDF.

15.1.2. Microsoft Excel

Excel adalah bagian dari paket Microsoft Office yang dikenal luas dan digunakan secara umum di berbagai bidang, termasuk kesehatan masyarakat. Excel memiliki keunggulan dalam kemudahan penggunaan, fleksibilitas, dan kemampuan pengolahan data dalam jumlah kecil hingga sedang. Selain sebagai alat pengolah angka dan tabel, Excel juga menyediakan fitur analisis statistik dasar melalui fungsi-fungsi built-in dan add-in seperti Analysis ToolPak.

Fitur utama Excel meliputi pembuatan tabel, grafik, serta penghitungan statistik deskriptif seperti rata-rata, median, modus, simpangan baku, dan varians. Dengan kemampuan makro dan VBA (Visual Basic for Applications), pengguna dapat mengotomatisasi proses analisis dan pembuatan laporan. Excel juga mendukung pengolahan data dalam bentuk pivot table, yang memudahkan analisis data secara dinamis dan interaktif.

15.1.3. Epi Info

Epi Info adalah perangkat lunak yang dikembangkan oleh Centers for Disease Control and Prevention (CDC) khusus untuk keperluan surveilans dan epidemiologi. Epi Info dirancang agar pengguna dapat melakukan pengolahan data dan analisis epidemiologi secara cepat dan efisien, bahkan oleh pengguna yang tidak memiliki latar belakang statistik yang mendalam. Software ini sangat cocok untuk pengolahan data surveilans, analisis kejadian penyakit, dan pembuatan peta distribusi penyakit.

Fitur utama Epi Info meliputi penginputan data, analisis frekuensi, analisis hubungan antar variabel, serta pembuatan peta dan grafik yang mendukung visualisasi data kesehatan masyarakat. Selain itu, Epi Info menyediakan modul untuk analisis statistik dasar dan uji statistik sederhana seperti Chi-Square, serta fitur untuk penghitungan indikator kesehatan seperti prevalensi dan insidens.

15.2 Input dan Analisis Data Dasar

Penginputan data yang benar dan sistematis merupakan langkah awal yang sangat penting dalam proses analisis statistik. Kesalahan dalam memasukkan data dapat menyebabkan hasil yang tidak valid dan menyesatkan. Oleh karena itu, mahasiswa harus memahami prosedur input data yang tepat sesuai dengan perangkat lunak yang digunakan.

15.2.1. Input Data di SPSS

Di SPSS, data biasanya diinput melalui tampilan Data View, yang menyerupai spreadsheet, dan variabel diatur melalui Variable View. Pada Variable View, pengguna menentukan

nama variabel, tipe data (numerik, string, date), skala pengukuran (nominal, ordinal, interval, rasio), serta label variabel dan nilai. Setelah pengaturan variabel selesai, data dapat diinput secara langsung di Data View.

Contoh: Jika ingin menginput data usia dan status imunisasi, pengguna harus membuat variabel `""Usia""` dan `""Status_Imunisasi""` di Variable View, kemudian mengisi data di Data View sesuai dengan kolom masing-masing. Pastikan data yang diinput sesuai dengan tipe dan skala pengukuran yang telah ditentukan.

15.2.2. Input Data di Excel

Di Excel, data diinput dalam bentuk tabel dengan kolom sebagai variabel dan baris sebagai unit observasi. Sebaiknya, pengguna membuat header kolom yang jelas dan konsisten, serta memastikan tidak ada data yang hilang atau salah ketik. Data harus disusun secara rapi dan lengkap agar analisis selanjutnya berjalan lancar.

Contoh: Untuk data surveilans penyakit, kolom dapat diberi label `""ID""`, `""Usia""`, `""Jenis Kelamin""`, `""Status Penyakit""`, dan data diisi sesuai dengan masing-masing kolom. Pengguna juga dapat menggunakan fitur validasi data untuk membatasi input agar sesuai dengan kategori tertentu.

15.2.3. Input Data di Epi Info

Di Epi Info, data diinput melalui form yang dapat disusun sesuai kebutuhan surveilans. Pengguna dapat membuat form input secara manual atau mengimpor data dari file lain seperti Excel. Data yang diinput harus mengikuti struktur yang telah ditentukan, dan setiap variabel harus diberi label serta tipe data yang sesuai.

Contoh: Dalam surveilans penyakit menular, variabel seperti ""ID Pasien"", ""Usia"", ""Jenis Kelamin"", dan ""Status Penyakit"" diinput melalui form yang telah disusun sebelumnya. Setelah data lengkap, pengguna dapat melakukan analisis langsung dari data yang telah diinput.

15.2.4. Analisis Statistik Dasar

Setelah data diinput, langkah berikutnya adalah melakukan analisis statistik dasar. Pada SPSS, pengguna dapat memilih menu Analyze dan kemudian memilih jenis analisis yang diinginkan, seperti Descriptive Statistics, Crosstabs, atau Compare Means. Di Excel, analisis dilakukan melalui fungsi-fungsi statistik seperti AVERAGE, MEDIAN, STDEV, serta pembuatan grafik. Di Epi Info, pengguna dapat memilih menu Analisis dan memilih indikator yang relevan.

Contoh: Untuk mengetahui rata-rata usia penderita hipertensi, pengguna dapat menggunakan fungsi AVERAGE di Excel, atau menu Descriptive Statistics di SPSS, dan analisis frekuensi di Epi Info.

15.3 Interpretasi Output Hasil Analisis

Setelah proses analisis selesai, langkah penting berikutnya adalah membaca dan menafsirkan output yang dihasilkan oleh perangkat lunak. Pemahaman terhadap output ini akan menentukan keakuratan dan relevansi hasil analisis dalam konteks penelitian atau surveilans.

15.3.1. Interpretasi Output di SPSS

Output SPSS biasanya berupa tabel dan grafik yang lengkap. Misalnya, tabel frekuensi menunjukkan distribusi kategori variabel, sedangkan tabel statistik deskriptif memberikan

nilai rata-rata, median, simpangan baku, dan lain-lain. Untuk uji statistik, output akan menampilkan nilai statistik uji, derajat kebebasan, dan p-value.

Contoh: Jika hasil uji Chi-Square menunjukkan p-value < 0,05, maka dapat disimpulkan bahwa ada hubungan yang signifikan antara tingkat pendidikan dan perilaku kesehatan. Penting untuk memahami bahwa p-value menunjukkan tingkat signifikansi, dan nilai statistik uji harus dibandingkan dengan nilai kritis dari distribusi yang sesuai.

15.3.2. Interpretasi Output di Excel

Di Excel, hasil analisis statistik dasar seperti rata-rata dan simpangan baku langsung terlihat dari sel yang berisi fungsi yang digunakan. Grafik yang dibuat dapat membantu visualisasi data, seperti distribusi usia atau proporsi kategori tertentu. Untuk analisis inferensial, pengguna harus melakukan uji statistik secara manual atau menggunakan add-in.

Contoh: Jika menghitung rata-rata usia penderita dan mendapatkan nilai 55 tahun, dan simpangan baku 10 tahun, maka dapat disimpulkan bahwa sebagian besar penderita berusia sekitar 45-65 tahun, dengan distribusi yang cukup merata.

15.3.3. Interpretasi Output di Epi Info

Epi Info menampilkan output dalam bentuk tabel yang ringkas dan visualisasi peta jika diperlukan. Hasil analisis frekuensi dan hubungan antar variabel biasanya disajikan dalam tabel yang mudah dipahami. Untuk indikator kesehatan, seperti prevalensi, output akan menunjukkan angka dan persentase yang dapat langsung diinterpretasikan.

Contoh: Jika hasil analisis menunjukkan prevalensi penyakit sebesar 12% di suatu wilayah, maka dapat diartikan bahwa dari seluruh populasi yang disurvei, 12% terdiagnosis menderita penyakit tersebut. Jika uji Chi-Square menunjukkan $p\text{-value} < 0,05$, maka hubungan antara faktor risiko dan kejadian penyakit dianggap signifikan.

15.4 Keunggulan dan Keterbatasan Masing-Masing Software

Setiap perangkat lunak memiliki keunggulan dan keterbatasan yang perlu dipahami agar pengguna dapat memilih alat yang paling sesuai dengan kebutuhan analisis.

15.4.1. Keunggulan dan Keterbatasan SPSS

Keunggulan utama SPSS adalah kemampuannya dalam melakukan analisis statistik yang kompleks dan menyediakan output lengkap serta visualisasi yang menarik. Kemudahan penggunaannya membuatnya cocok untuk mahasiswa dan peneliti yang membutuhkan analisis cepat dan akurat. Namun, kekurangan SPSS terletak pada biaya lisensi yang cukup mahal dan kebutuhan spesifikasi komputer yang cukup tinggi.

15.4.2. Keunggulan dan Keterbatasan Excel

Excel sangat familiar dan mudah digunakan, serta mampu melakukan analisis statistik dasar dengan cepat. Fungsinya yang luas dan fleksibel memungkinkan pengguna melakukan berbagai pengolahan data tanpa perlu belajar perangkat lunak baru. Keterbatasannya adalah dalam hal analisis statistik yang kompleks dan risiko kesalahan manual dalam pengolahan data, serta kurangnya fitur otomatis untuk analisis statistik lanjutan.

15.4.3. Keunggulan dan Keterbatasan Epi Info

Epi Info dirancang khusus untuk surveilans dan epidemiologi, sehingga sangat efisien untuk analisis data kesehatan masyarakat. Fitur visualisasi dan pembuatan peta memudahkan interpretasi data spasial dan kejadian penyakit. Keterbatasannya adalah kurangnya fitur analisis statistik lanjutan dan tampilan output yang lebih sederhana dibandingkan SPSS.

15.5 Penggunaan Software Statistik dalam Praktik Lapangan dan Penelitian

Penguasaan software statistik sangat penting dalam mendukung kegiatan surveilans, evaluasi program, dan penelitian di bidang kesehatan masyarakat. Mahasiswa harus mampu memilih perangkat lunak yang sesuai dengan tingkat kompleksitas data dan analisis yang dibutuhkan. Penggunaan yang tepat akan meningkatkan kecepatan, keakuratan, dan validitas hasil analisis, serta memudahkan penyusunan laporan yang komunikatif dan berbasis bukti.

Dalam praktik lapangan, misalnya, Epi Info sering digunakan untuk pengolahan data surveilans penyakit menular secara cepat di lapangan, sementara SPSS lebih cocok untuk analisis mendalam dari data yang sudah lengkap dan terstruktur. Excel dapat digunakan untuk pengolahan awal, pembuatan grafik, dan laporan sederhana. Kombinasi penggunaan ketiga software ini akan memberikan manfaat maksimal dalam pengelolaan data kesehatan masyarakat.

Penggunaan software statistik seperti SPSS, Excel, dan Epi Info merupakan kompetensi penting yang harus dimiliki mahasiswa dan profesional di bidang kesehatan masyarakat. Penguasaan input data yang benar, analisis dasar yang tepat,

serta interpretasi hasil yang akurat akan mendukung pengambilan keputusan berbasis bukti dan meningkatkan kualitas penelitian serta surveilans kesehatan. Setiap perangkat lunak memiliki keunggulan dan keterbatasan yang harus dipahami secara mendalam agar pengguna dapat memilih dan menggunakannya secara optimal sesuai dengan kebutuhan analisis dan konteks penelitian.

Pengembangan kompetensi ini akan memperkuat kemampuan mahasiswa dalam mengelola data kuantitatif secara efektif, serta mampu menyusun laporan statistik yang komunikatif dan informatif. Dengan demikian, bagian ini tidak hanya berfungsi sebagai panduan teknis, tetapi juga sebagai fondasi untuk pengembangan kompetensi analisis data yang kritis dan bertanggung jawab di bidang kesehatan masyarakat.

Rangkuman

Penguasaan software statistik seperti SPSS, Excel, dan Epi Info merupakan bagian penting dalam analisis data kesehatan masyarakat. Ketiga perangkat lunak ini memiliki karakteristik dan fungsi yang berbeda, namun semuanya mendukung proses pengolahan data secara efisien dan akurat. Pemahaman terhadap fitur utama, keunggulan, dan keterbatasan masing-masing software akan membantu pengguna dalam memilih alat yang paling sesuai dengan kebutuhan analisis tertentu.

1. **SPSS** adalah software statistik yang populer dan banyak digunakan dalam penelitian kesehatan. Keunggulannya terletak pada kemampuannya melakukan analisis statistik kompleks tanpa perlu pemrograman, serta menyediakan output lengkap berupa tabel dan grafik yang mudah dipahami. SPSS cocok untuk analisis deskriptif, inferensial, dan

- multivariat, namun memiliki kekurangan berupa biaya lisensi yang relatif mahal dan kebutuhan spesifikasi komputer yang cukup tinggi.
2. **Excel** merupakan alat pengolah angka yang umum digunakan dan sangat familiar di berbagai bidang. Fungsinya meliputi pembuatan tabel, grafik, serta analisis statistik dasar seperti rata-rata, median, dan simpangan baku. Excel juga mendukung otomatisasi melalui makro dan VBA, serta pengolahan data dinamis dengan pivot table. Kelemahannya adalah kurangnya fitur analisis statistik lanjutan dan risiko kesalahan manual dalam pengolahan data.
 3. **Epi Info** dikembangkan oleh CDC khusus untuk surveilans dan epidemiologi. Software ini memudahkan pengolahan data surveilans, analisis kejadian penyakit, dan pembuatan peta distribusi penyakit. Fitur utamanya meliputi penginputan data, analisis frekuensi, hubungan antar variabel, serta indikator kesehatan seperti prevalensi dan insidens. Keterbatasannya adalah tampilan output yang lebih sederhana dan kurangnya fitur analisis statistik lanjutan.

Penggunaan ketiga software ini harus disesuaikan dengan kebutuhan analisis dan tingkat kompleksitas data. Dalam praktik lapangan dan penelitian, penguasaan input data yang benar, analisis dasar, serta interpretasi hasil yang tepat sangat penting untuk menghasilkan data yang valid dan dapat dipertanggungjawabkan. Kombinasi penggunaan software ini akan meningkatkan efisiensi, akurasi, dan kualitas laporan statistik kesehatan masyarakat.

Pengembangan kompetensi dalam penguasaan software statistik ini akan memperkuat kemampuan mahasiswa dan profesional dalam mengelola data kuantitatif secara efektif, serta menyusun laporan yang komunikatif dan berbasis

bukti. Dengan demikian, penguasaan software statistik menjadi fondasi penting dalam mendukung pengambilan keputusan dan peningkatan kualitas penelitian di bidang kesehatan masyarakat.

Latihan Mahasiswa

Soal Essay

1. Jelaskan perbedaan utama antara SPSS, Excel, dan Epi Info dalam konteks analisis data kesehatan masyarakat. Sertakan keunggulan dan keterbatasan masing-masing perangkat lunak.
2. Sebutkan langkah-langkah yang harus dilakukan dalam melakukan input data di SPSS dan jelaskan pentingnya pengaturan variabel sebelum memasukkan data.
3. Bagaimana cara membaca dan menginterpretasikan output hasil analisis uji Chi-Square di Epi Info? Jelaskan apa arti p-value yang dihasilkan.
4. Diskusikan keunggulan penggunaan Excel dalam pengolahan data statistik dasar dan sebutkan risiko yang mungkin terjadi jika data diinput secara manual tanpa pengawasan.
5. Berikan contoh situasi di lapangan di mana penggunaan Epi Info lebih unggul dibandingkan SPSS dan Excel. Jelaskan alasannya.

Soal Pilihan Berganda

1. Software statistik yang dikembangkan oleh IBM dan memiliki fitur lengkap untuk analisis statistik kompleks adalah...
 - a. Excel

- b. Epi Info
 - c. SPSS
 - d. SAS
2. Fitur utama Excel yang memungkinkan pengguna melakukan analisis data secara dinamis dan interaktif adalah...
- a. Pivot Table
 - b. Analisis Regresi
 - c. Uji Chi-Square
 - d. Analisis Faktor
3. Epi Info sangat cocok digunakan untuk...
- a. Analisis data surveilans penyakit
 - b. Analisis data multivariat kompleks
 - c. Pengolahan data keuangan
 - d. Pembuatan laporan keuangan
4. Dalam input data di SPSS, variabel harus diatur terlebih dahulu di...
- a. Data View
 - b. Variable View
 - c. Output Viewer
 - d. Syntax Editor
5. Fungsi utama Excel yang digunakan untuk menghitung rata-rata data adalah...
- a. MEDIAN
 - b. AVERAGE
 - c. MODE
 - d. STDEV

6. Salah satu kekurangan SPSS adalah...
 - a. Biaya lisensi yang mahal
 - b. Kurangnya fitur analisis statistik
 - c. Tidak mendukung grafik
 - d. Tidak user-friendly
7. Untuk melakukan analisis frekuensi dan mendapatkan distribusi kategori variabel di Epi Info, pengguna harus menggunakan menu...
 - a. Data Entry
 - b. Analysis
 - c. Graphs
 - d. Reports
8. Keunggulan utama Excel dalam pengolahan data adalah...
 - a. Kemampuan analisis statistik lanjutan
 - b. Kemudahan penggunaan dan familiaritas
 - c. Fitur pembuatan peta
 - d. Analisis multivariat
9. Salah satu risiko utama saat menggunakan Excel untuk analisis data adalah...
 - a. Kurangnya fitur grafik
 - b. Kesalahan manual dalam penginputan data
 - c. Tidak bisa melakukan analisis statistik dasar
 - d. Tidak mendukung pengolahan data besar
10. Dalam konteks surveilans kesehatan, Epi Info lebih unggul karena...
 - a. Memiliki fitur analisis regresi kompleks

- b. Mudah digunakan untuk pengolahan data cepat dan visualisasi peta
- c. Lebih murah dibandingkan software lain
- d. Mendukung analisis data keuangan

Soal Project / Studi Kasus

1. Anda mendapatkan data hasil surveilans penyakit menular dari beberapa wilayah. Data tersebut meliputi jumlah kasus, usia penderita, jenis kelamin, dan lokasi geografis. Jelaskan langkah-langkah yang akan Anda lakukan untuk mengolah data tersebut menggunakan Epi Info, mulai dari input data hingga analisis dan interpretasi hasilnya. Sertakan alasan pemilihan fitur dan menu yang digunakan.
2. Sebuah penelitian ingin membandingkan rata-rata usia penderita hipertensi antara dua wilayah berbeda. Data sudah dikumpulkan dan disusun dalam tabel Excel. Jelaskan proses analisis yang akan Anda lakukan menggunakan SPSS, termasuk langkah input data, analisis statistik yang dipilih, dan bagaimana cara menafsirkan hasilnya.

REFERENSI

1. *Field, A. (2018). Discovering Statistics Using IBM SPSS Statistics (5th ed.).* Sage Publications.
2. *Ghozali, I. (2018). Aplikasi Analisis Multivariate dengan Program IBM SPSS 25.* Badan Penerbit Universitas Diponegoro.
3. *Lund, J. (2019). Epi Info: A Tool for Rapid Data Collection and Analysis in Public Health.* Journal of Public Health Management and Practice, 25(2), 123-130.

4. *Microsoft. (2020). Excel Data Analysis and Visualization.* Microsoft Support.
5. *Centers for Disease Control and Prevention (CDC). (2017). Epi Info™ User Guide.* CDC." "

BAB 16

PENYUSUNAN LAPORAN STATISTIK KESEHATAN

Tujuan Pembelajaran

1. Mahasiswa mampu memahami dan menjelaskan struktur serta sistematika penyusunan laporan statistik kesehatan secara komprehensif, sehingga dapat menyusun laporan yang sistematis dan mudah dipahami oleh berbagai pihak terkait.
2. Mahasiswa dapat mengidentifikasi dan menerapkan berbagai teknik penyajian data dalam bentuk tabel dan grafik yang sesuai dengan karakteristik data dan tujuan laporan, guna meningkatkan kejelasan dan efektivitas komunikasi hasil analisis.
3. Mahasiswa mampu menulis narasi yang tepat dan informatif berdasarkan hasil analisis statistik, serta mampu menyusun kesimpulan yang relevan dan mendukung pengambilan keputusan dalam konteks kesehatan masyarakat.
4. Mahasiswa dapat mengintegrasikan hasil penyajian data dan narasi ke dalam laporan yang lengkap dan profesional, sesuai dengan standar dan pedoman yang berlaku dalam bidang kesehatan.
5. Mahasiswa mampu menerapkan pengetahuan tersebut dalam menyusun laporan hasil surveilans penyakit menular di masyarakat, sehingga laporan yang dihasilkan dapat digunakan sebagai dasar pengambilan kebijakan dan tindakan kesehatan.

6. Mahasiswa mampu menggunakan perangkat lunak statistik dan pengolah data untuk mendukung proses penyusunan laporan, termasuk dalam pembuatan tabel, grafik, dan penulisan narasi secara efisien dan akurat.
7. Mahasiswa mampu melakukan evaluasi terhadap laporan statistik yang disusun, termasuk menilai kejelasan, keakuratan, dan kebermanfaatannya dalam konteks pengelolaan data kesehatan masyarakat.

Pendahuluan

Penyusunan laporan statistik kesehatan merupakan salah satu tahapan penting dalam proses pengelolaan data dan informasi di bidang kesehatan masyarakat. Setelah data dikumpulkan, dianalisis, dan diinterpretasikan, langkah selanjutnya adalah menyusun laporan yang mampu menyampaikan hasil tersebut secara sistematis dan komunikatif kepada berbagai pemangku kepentingan, mulai dari petugas kesehatan, pengambil kebijakan, hingga masyarakat umum. Laporan statistik yang baik tidak hanya berisi data mentah, tetapi juga harus mampu menyajikan informasi secara jelas dan menarik, sehingga dapat mendukung pengambilan keputusan yang tepat dan efektif.

Dalam konteks kesehatan masyarakat, laporan statistik memiliki peran strategis dalam memantau kondisi kesehatan, mengidentifikasi masalah utama, serta menilai keberhasilan program-program kesehatan yang telah dilaksanakan. Oleh karena itu, penyusunan laporan yang sistematis dan terstruktur menjadi sangat penting agar hasil analisis dapat dipahami dengan baik dan digunakan secara optimal. Penyusunan laporan yang baik juga mencakup penggunaan berbagai teknik penyajian data, seperti tabel dan grafik, yang mampu memperjelas pola dan tren data secara visual.

Selain itu, penulisan narasi yang tepat dan ringkas sangat diperlukan untuk menginterpretasikan data secara mendalam dan menyampaikan pesan utama dari hasil analisis. Narasi ini harus mampu menjelaskan temuan secara logis dan mendukung kesimpulan yang diambil, sehingga laporan tidak hanya berisi angka-angka, tetapi juga memiliki makna dan relevansi dalam konteks kesehatan masyarakat. Kesalahan dalam penulisan narasi atau penyajian data dapat menyebabkan interpretasi yang keliru dan berpotensi menghambat pengambilan keputusan yang tepat.

Pada bagian ini, kita juga akan membahas tentang pentingnya penggunaan perangkat lunak statistik dalam proses penyusunan laporan. Dengan kemajuan teknologi, perangkat lunak seperti SPSS, Excel, dan Epi Info memudahkan proses pembuatan tabel, grafik, dan penulisan narasi secara efisien dan akurat. Penguasaan perangkat lunak ini akan sangat membantu dalam mempercepat proses penyusunan laporan dan meningkatkan kualitas hasil akhir.

Akhirnya, bagian ini akan memberikan gambaran tentang bagaimana menyusun laporan hasil surveilans penyakit menular di masyarakat secara lengkap dan profesional. Melalui contoh aplikasi tersebut, diharapkan mahasiswa mampu mengaplikasikan seluruh pengetahuan yang diperoleh dalam menyusun laporan yang tidak hanya memenuhi standar akademik, tetapi juga mampu memberikan manfaat nyata dalam pengelolaan data kesehatan masyarakat. Dengan memahami dan menguasai proses ini, mahasiswa diharapkan mampu menjadi tenaga kesehatan yang kompeten dalam menyusun laporan statistik yang informatif, akurat, dan bermanfaat bagi pengembangan program kesehatan di masyarakat.

16.1 Struktur dan Sistematika Laporan Statistik

Penyusunan laporan statistik kesehatan merupakan proses penting yang harus dilakukan secara sistematis agar hasil analisis data dapat disampaikan secara efektif dan efisien kepada berbagai pihak terkait. Struktur laporan yang baik akan memudahkan pembaca dalam memahami isi laporan, menilai keakuratan data, serta mengidentifikasi temuan utama yang relevan dengan pengambilan keputusan. Oleh karena itu, pemahaman tentang struktur dan sistematika laporan statistik sangat penting bagi mahasiswa dan praktisi di bidang kesehatan masyarakat.

Secara umum, laporan statistik kesehatan harus mengikuti format yang terstandar dan terorganisasi dengan baik. Komponen utama yang harus ada dalam sebuah laporan meliputi bagian pendahuluan, metodologi, hasil, pembahasan, dan penutup. Berikut penjelasan rinci mengenai masing-masing bagian:

16.1.1 Pendahuluan

Bagian ini berisi latar belakang masalah, tujuan penyusunan laporan, serta ruang lingkup dan batasan studi. Pada bagian ini, penulis menjelaskan mengapa laporan tersebut dibuat dan apa yang ingin dicapai. Misalnya, laporan surveilans penyakit menular di suatu wilayah bertujuan untuk mengetahui tingkat kejadian dan faktor risiko yang terkait.

16.1.2 Metodologi

Pada bagian ini dijelaskan secara rinci prosedur pengumpulan data, populasi dan sampel yang digunakan, teknik pengolahan data, serta alat dan perangkat lunak yang dipakai. Penjelasan ini penting agar pembaca memahami keabsahan dan keandalan data yang dianalisis. Misalnya, jika

data dikumpulkan melalui surveilans aktif menggunakan kuesioner, maka harus dijelaskan proses sampling dan validitas instrumen.

16.1.3 Hasil

Bagian ini menyajikan data yang telah dianalisis secara sistematis. Penyajian hasil harus dilakukan secara objektif dan lengkap, mencakup tabel, grafik, dan statistik deskriptif yang relevan. Penyajian data yang baik akan memudahkan interpretasi dan pengambilan kesimpulan.

16.1.4 Pembahasan

Di bagian ini, penulis menginterpretasikan hasil yang diperoleh, mengaitkannya dengan literatur dan studi sebelumnya, serta membahas implikasi dari temuan tersebut. Pembahasan harus mampu menjawab pertanyaan-pertanyaan penting terkait data, seperti apa makna dari angka-angka tersebut dan bagaimana pengaruhnya terhadap kebijakan kesehatan.

16.1.5 Kesimpulan dan Saran

Bagian ini merangkum temuan utama dan memberikan rekomendasi yang bersifat praktis dan aplikatif. Kesimpulan harus ringkas, padat, dan berdasarkan data yang disajikan, sedangkan saran diarahkan untuk perbaikan atau langkah selanjutnya.

16.2 Penyajian Hasil dalam Bentuk Tabel dan Grafik

Penyajian data secara visual sangat penting dalam laporan statistik kesehatan karena dapat memperlihatkan pola, tren,

dan hubungan antar variabel secara lebih jelas dan cepat dipahami. Tabel dan grafik merupakan alat utama yang digunakan untuk menyajikan data secara efektif.

16.2.1 Penyajian Data dalam Bentuk Tabel

Tabel distribusi frekuensi, tabel persentase, dan tabel silang (cross tabulation) adalah bentuk tabel yang umum digunakan dalam laporan statistik kesehatan. Tabel harus disusun secara sistematis, lengkap, dan mudah dibaca. Beberapa prinsip penting dalam pembuatan tabel meliputi:

- Judul tabel harus jelas dan informatif, mencerminkan isi tabel secara lengkap.
- Kolom dan baris harus diberi label yang tepat dan konsisten.
- Data harus disusun secara logis, misalnya dari yang tertinggi ke terendah atau berdasarkan kategori tertentu.
- Penggunaan angka harus konsisten dan mengikuti standar penulisan, seperti pembulatan dan satuan pengukuran.

Contoh: Tabel 16.1 menunjukkan distribusi kejadian penyakit menular berdasarkan usia dan jenis kelamin di wilayah tertentu.

16.2.2 Penyajian Data dalam Bentuk Grafik

Grafik digunakan untuk memperlihatkan tren, perbandingan, dan hubungan antar variabel secara visual. Beberapa jenis grafik yang umum digunakan dalam laporan kesehatan meliputi:

- Diagram batang (bar chart): untuk membandingkan jumlah atau frekuensi antar kategori.

- Diagram lingkaran (pie chart): untuk menunjukkan proporsi relatif dari bagian terhadap keseluruhan.
- Histogram: untuk menampilkan distribusi data numerik kontinu.
- Poligon frekuensi: untuk menunjukkan distribusi data secara visual dan mengidentifikasi pola distribusi.

Dalam pembuatan grafik, prinsip utama adalah kejelasan dan keakuratan. Grafik harus memiliki judul yang informatif, sumbu yang diberi label lengkap, dan data yang disajikan harus proporsional dan tidak menyesatkan.

Contoh: *Gambar 16.1 menunjukkan grafik batang yang membandingkan tingkat imunisasi lengkap di beberapa kabupaten.*

16.2.3 Penggunaan Software Statistik

Penggunaan perangkat lunak seperti SPSS, Excel, atau Epi Info sangat membantu dalam pembuatan tabel dan grafik. Software ini memungkinkan pengguna untuk mengolah data secara cepat dan akurat, serta menghasilkan output yang siap digunakan dalam laporan. Mahasiswa harus mampu mengoperasikan fitur dasar seperti pembuatan tabel pivot, grafik, dan pengaturan tampilan visual agar hasilnya menarik dan informatif.

16.3 Penulisan Narasi dan Kesimpulan Hasil Analisis

Selain menyajikan data secara visual, penulisan narasi yang mendukung hasil analisis sangat penting untuk memberikan konteks dan interpretasi yang tepat. Narasi harus mampu menjelaskan secara logis dan sistematis apa yang terlihat

dari data, mengapa hal tersebut penting, dan apa implikasinya terhadap kebijakan atau program kesehatan.

16.3.1 Menulis Narasi yang Informatif dan Objektif

Narasi harus berisi penjelasan mengenai:

- Gambaran umum data, seperti tren utama, perbedaan antar kelompok, dan pola distribusi.
- Interpretasi statistik, seperti nilai rata-rata, persentase, dan tingkat signifikansi jika ada.
- Hubungan antar variabel, misalnya korelasi atau asosiasi yang ditemukan.
- Keterbatasan data dan kemungkinan bias yang mempengaruhi hasil.

Contoh: *Berdasarkan data surveilans, tingkat imunisasi lengkap di wilayah A mencapai 85%, sedangkan di wilayah B hanya 70%. Perbedaan ini signifikan secara statistik ($p < 0,05$), menunjukkan perlunya peningkatan program imunisasi di wilayah B.*

16.3.2 Menyusun Kesimpulan yang Relevan

Kesimpulan harus merangkum temuan utama secara ringkas dan jelas. Kesimpulan harus didasarkan pada data yang disajikan dan tidak bersifat spekulatif. Misalnya, jika data menunjukkan adanya ketimpangan imunisasi antar wilayah, maka kesimpulan yang tepat adalah:

“Terdapat disparitas tingkat imunisasi antara wilayah A dan B, yang perlu mendapatkan perhatian khusus dalam upaya peningkatan cakupan imunisasi nasional.”

16.4 Contoh Aplikasi: Penyusunan Laporan Hasil Surveilans Penyakit Menular di Masyarakat

Sebagai ilustrasi praktis, bayangkan sebuah tim kesehatan masyarakat melakukan surveilans penyakit menular di beberapa kabupaten. Data yang dikumpulkan meliputi jumlah kasus, distribusi usia, jenis kelamin, dan faktor risiko terkait. Setelah data dianalisis, langkah selanjutnya adalah menyusun laporan yang lengkap dan sistematis.

Pertama, laporan harus diawali dengan bagian pendahuluan yang menjelaskan latar belakang pentingnya surveilans ini, misalnya untuk mengidentifikasi daerah dengan tingkat kejadian tertinggi dan mengarahkan intervensi. Kemudian, bagian metodologi menjelaskan proses pengumpulan data, termasuk teknik sampling dan instrumen yang digunakan.

Selanjutnya, bagian hasil menyajikan data secara lengkap dalam tabel dan grafik. Contohnya, tabel distribusi kasus berdasarkan wilayah dan usia, serta grafik tren kejadian selama periode tertentu. Tabel ini memudahkan pembaca untuk melihat secara cepat di mana dan siapa yang paling terdampak.

Dalam bagian narasi, penulis menjelaskan bahwa kejadian tertinggi ditemukan di wilayah X, terutama pada kelompok usia 5-14 tahun, dan bahwa data menunjukkan adanya peningkatan kasus dari bulan ke bulan. Penafsiran ini didukung oleh grafik tren yang menunjukkan kenaikan kasus dari bulan Januari hingga Maret.

Pembahasan kemudian mengaitkan hasil ini dengan faktor-faktor yang mungkin mempengaruhi, seperti tingkat sanitasi dan akses layanan kesehatan. Penulis juga menyoroti

keterbatasan data, misalnya kemungkinan underreporting di daerah tertentu.

Akhirnya, laporan diakhiri dengan kesimpulan yang menegaskan perlunya peningkatan program imunisasi dan penguatan sistem surveilans di wilayah yang paling terdampak. Saran-saran praktis, seperti peningkatan kapasitas petugas lapangan dan penggunaan teknologi informasi, disampaikan untuk mendukung langkah perbaikan.

Laporan yang disusun secara lengkap dan sistematis ini tidak hanya memenuhi standar akademik, tetapi juga mampu menjadi dasar pengambilan kebijakan yang tepat dan efektif dalam pengendalian penyakit menular di masyarakat.

Penyusunan laporan statistik kesehatan yang baik memerlukan pemahaman mendalam tentang struktur dan sistematika yang harus diikuti. Setiap bagian, mulai dari pendahuluan hingga penutup, harus disusun secara logis dan lengkap agar hasil analisis dapat disampaikan secara efektif. Penyajian data dalam bentuk tabel dan grafik harus dilakukan dengan prinsip kejelasan dan keakuratan, serta didukung oleh perangkat lunak statistik yang memadai. Penulisan narasi yang informatif dan objektif sangat penting untuk menginterpretasikan data secara tepat dan mendukung pengambilan keputusan. Melalui contoh aplikasi surveilans penyakit menular, mahasiswa dapat memahami bagaimana seluruh proses ini diintegrasikan dalam praktik nyata, sehingga mampu menyusun laporan yang profesional dan bermanfaat dalam pengelolaan kesehatan masyarakat.

Rangkuman

Penyusunan laporan statistik kesehatan merupakan proses penting yang harus dilakukan secara sistematis agar hasil analisis data dapat disampaikan secara efektif dan efisien kepada berbagai pihak terkait. Struktur laporan yang baik memudahkan pembaca memahami isi laporan, menilai keakuratan data, serta mengidentifikasi temuan utama yang relevan untuk pengambilan keputusan. Pemahaman tentang komponen utama laporan, seperti pendahuluan, metodologi, hasil, pembahasan, dan penutup, sangat penting untuk memastikan laporan lengkap dan terorganisasi dengan baik.

Berikut adalah poin-poin utama dalam penyusunan laporan statistik kesehatan:

1. Pendahuluan

- Menjelaskan latar belakang masalah, tujuan laporan, serta ruang lingkup dan batasan studi.
- Memberikan gambaran mengapa laporan dibuat dan apa yang ingin dicapai, misalnya untuk mengetahui tingkat kejadian penyakit tertentu di wilayah tertentu.

2. Metodologi

- Menguraikan secara rinci prosedur pengumpulan data, populasi dan sampel yang digunakan, serta teknik pengolahan data.
- Menjelaskan alat dan perangkat lunak yang dipakai agar data yang dianalisis dapat dipertanggungjawabkan keabsahannya.

3. Hasil

- Menyajikan data yang telah dianalisis secara sistematis, lengkap dengan tabel, grafik, dan statistik deskriptif.

- Penyajian harus objektif dan memudahkan interpretasi, misalnya tabel distribusi kasus berdasarkan wilayah dan usia.

4. Pembahasan

- Menginterpretasikan hasil yang diperoleh, mengaitkannya dengan literatur dan studi sebelumnya.
- Membahas makna dari data dan implikasinya terhadap kebijakan kesehatan, serta membahas keterbatasan data.

5. Kesimpulan dan Saran

- Merangkum temuan utama secara ringkas dan jelas, berdasarkan data yang ada.
- Memberikan rekomendasi praktis untuk langkah selanjutnya, seperti peningkatan program kesehatan atau penguatan sistem surveilans.

Selain itu, penyajian data secara visual melalui tabel dan grafik sangat penting. Tabel harus disusun secara sistematis dan informatif, sedangkan grafik digunakan untuk memperlihatkan tren dan hubungan antar variabel secara visual. Penggunaan software statistik seperti SPSS, Excel, atau Epi Info sangat membantu dalam pembuatan tabel dan grafik yang akurat dan menarik.

Penulisan narasi yang mendukung hasil analisis harus objektif dan informatif, menjelaskan pola utama, interpretasi statistik, serta hubungan antar variabel. Narasi ini membantu memberikan konteks dan makna dari data yang disajikan. Kesimpulan harus didasarkan pada data dan tidak bersifat spekulatif, serta mampu memberikan gambaran utama dari hasil analisis.

Contoh aplikasi nyata adalah penyusunan laporan hasil surveilans penyakit menular di masyarakat, yang meliputi

pengumpulan data, analisis, penyajian data dalam tabel dan grafik, serta penulisan narasi dan interpretasi hasil. Laporan yang lengkap dan sistematis ini menjadi dasar penting dalam pengambilan kebijakan dan langkah intervensi yang tepat.

Secara keseluruhan, penyusunan laporan statistik kesehatan yang baik memerlukan pemahaman mendalam tentang struktur, penyajian data, penggunaan perangkat lunak, dan penulisan narasi yang tepat. Dengan mengikuti prinsip-prinsip ini, laporan yang dihasilkan akan mampu menyampaikan informasi secara jelas, akurat, dan bermanfaat untuk pengembangan program kesehatan masyarakat.

Latihan Mahasiswa

Soal Essay

1. Jelaskan mengapa penyusunan laporan statistik kesehatan harus mengikuti struktur yang sistematis dan lengkap. Sebutkan dan uraikan bagian-bagian utama dalam laporan tersebut!
2. Sebutkan dan jelaskan tiga prinsip utama dalam penyajian data menggunakan tabel dan grafik dalam laporan statistik kesehatan!
3. Bagaimana cara menulis narasi yang efektif dan objektif dalam laporan statistik kesehatan? Berikan contoh singkat yang menggambarkan penjelasan tersebut!
4. Jelaskan peran perangkat lunak statistik seperti SPSS, Excel, dan Epi Info dalam proses penyusunan laporan statistik kesehatan! Apa manfaat utama penggunaannya?
5. Berdasarkan studi kasus surveilans penyakit menular di masyarakat, jelaskan langkah-langkah yang harus

dilakukan dalam menyusun laporan hasil surveilans tersebut secara lengkap dan sistematis!

Soal Pilihan Berganda

1. Apa bagian pertama yang harus ada dalam laporan statistik kesehatan?
 - a. Hasil
 - b. Pendahuluan
 - c. Pembahasan
 - d. Kesimpulan
2. Dalam bagian metodologi laporan statistik, hal utama yang perlu dijelaskan adalah:
 - a. Data hasil analisis
 - b. Teknik pengumpulan data dan populasi
 - c. Kesimpulan dari data
 - d. Saran dan rekomendasi
3. Manakah dari berikut ini yang termasuk prinsip dalam penyajian data grafik?
 - a. Judul harus tidak informatif
 - b. Sumbu harus diberi label lengkap
 - c. Data harus disembunyikan agar tidak menyesatkan
 - d. Warna grafik harus acak
4. Dalam menyusun tabel distribusi frekuensi, hal yang paling penting adalah:
 - a. Menggunakan angka acak
 - b. Memberikan judul yang jelas dan lengkap
 - c. Menghilangkan label kolom

- d. Mengurangi jumlah data
- 5. Grafik yang paling tepat digunakan untuk menunjukkan proporsi bagian terhadap keseluruhan adalah:
 - a. Histogram
 - b. Diagram batang
 - c. Diagram lingkaran
 - d. Poligon frekuensi
- 6. Narasi dalam laporan statistik harus bersifat:
 - a. Subjektif dan emosional
 - b. Objektif dan informatif
 - c. Singkat tanpa penjelasan
 - d. Berisi opini pribadi
- 7. Salah satu manfaat utama penggunaan perangkat lunak statistik adalah:
 - a. Mengurangi kebutuhan data
 - b. Mempercepat proses analisis dan penyajian data
 - c. Mengganti penulis laporan
 - d. Menghilangkan kebutuhan interpretasi
- 8. Dalam studi kasus surveilans penyakit menular, langkah pertama yang harus dilakukan adalah:
 - a. Menyusun laporan akhir
 - b. Mengumpulkan data dan menentukan populasi
 - c. Menyusun grafik tren
 - d. Menulis narasi hasil

9. Kesalahan yang sering terjadi dalam penyusunan laporan statistik adalah:
 - a. Menggunakan tabel dan grafik
 - b. Menyusun bagian pendahuluan
 - c. Menyajikan data secara tidak lengkap dan tidak sistematis
 - d. Menggunakan perangkat lunak statistik
10. Kesimpulan dalam laporan statistik harus:
 - a. Berisi data mentah
 - b. Ringkas dan berdasarkan hasil analisis
 - c. Berisi opini pribadi
 - d. Tidak perlu didukung data

Soal Project / Studi Kasus

1. Anda adalah petugas kesehatan yang melakukan surveilans kejadian penyakit menular di sebuah kabupaten. Setelah pengumpulan data, Anda diminta menyusun laporan lengkap. Buatlah kerangka laporan yang mencakup semua bagian utama dan jelaskan secara singkat isi dari masing-masing bagian tersebut! (Tugas ini menuntut mahasiswa untuk menerapkan pengetahuan tentang struktur laporan statistik kesehatan secara lengkap dan sistematis.)
2. Sebuah lembaga kesehatan ingin menyusun laporan hasil surveilans imunisasi di beberapa wilayah. Data yang diperoleh meliputi jumlah anak yang diimunisasi lengkap dan tidak lengkap, serta distribusi usia dan wilayah. Tugas Anda adalah menyusun contoh tabel distribusi frekuensi dan grafik batang yang sesuai untuk menyajikan data tersebut. Jelaskan langkah-langkah yang Anda lakukan dalam menyusun tabel dan grafik tersebut

serta alasan pemilihan bentuk penyajian data! (Tugas ini menguji kemampuan mahasiswa dalam menerapkan teknik penyajian data secara visual dan sistematis.)

REFERENSI

1. Agresti, A., & Finlay, B. (2018). *Statistical Methods for the Social Sciences*. Pearson Education.
2. World Health Organization. (2017). *Guidelines on Surveillance and Response for Vaccine-Preventable Diseases*. WHO Press.
3. Suryono, & Suryono, S. (2020). Penyusunan laporan statistik kesehatan: prinsip dan praktik. *Jurnal Kesehatan Masyarakat Indonesia*, 16(2), 123-135.
4. Tabachnick, B. G., & Fidell, L. S. (2019). *Using Multivariate Statistics*. Pearson Education."

BIODATA PENULIS



Egy Sunanda Putra, M. Gz

Dosen Program Studi Promosi Kesehatan
Jurusan Promosi Kesehatan Poltekkes Kemenkes Jambi

Penulis lahir di Bangko tanggal 25 Juni 1993. Penulis adalah dosen tetap pada Program Studi Promosi Kesehatan Jurusan Promosi Kesehatan, Ketua Pusat Unggulan Ilmu Pengetahuan Teknologi Politeknik Kesehatan, Poltekkes Kemenkes Jambi. Menyelesaikan pendidikan S1 Ilmu Gizi (2015) Universitas Respati Yogyakarta dan S2 Ilmu Gizi dari Universitas Sebelas Maret, Surakarta (2018). Penulis aktif melakukan tridharma perguruan tinggi, tim pakar audit stunting provinsi Jambi, pembicara di forum ilmiah, aktif publikasi artikel baik nasional dan internasional, pengajar atau fasilitator dan supervisor penelitian tingkat nasional (Studi Status Gizi Indonesia dan Survey Status Gizi Indonesia, Survey Kesehatan Indonesia,). Penulis melakukan kajian daerah Provinsi Jambi bersama Badan Kebijakan Pembangunan Kesehatan, Kementerian Kesehatan Republik

Indonesia, dan *Project Investing in Nutrition and Early Years (INEY)*. Penulis menguatkan perannya sebagai ahli gizi yang tergabung sebagai pengurus Dewan Pimpinan Daerah Persatuan Ahli Gizi Indonesia Provinsi Jambi (DPD PERSAGI Jambi) dan Asosiasi Dietisien Indonesia (AsDI), serta Asosiasi Nutrisionist Indonesia (AsNI). Penulis dapat dihubungi melalui egyputra93@poltekkesjambi.ac.id.

BIODATA PENULIS



Agus Hendra Al Rahmad, SKM, MPH

Dosen Jurusan Gizi, Poltekkes Kemenkes Aceh

Lahir di Aceh, 16 Agustus 1982, telah menempuh pendidikan tinggi yaitu pendidikan Vokasi Gizi (D3 Gizi) di Poltekkes Kemenkes Aceh (2003), dan Sarjana Kesehatan Masyarakat pada Universitas Serambi Mekkah (2009). Tahun 2013 menyelesaikan studi magister pada Universitas Gadjah Mada. Riwayat pekerjaan diawali dibidang ilmu gizi dan kesehatan masyarakat sejak 2005, yaitu sebagai tenaga gizi dibawah Lembaga Unicef. Saat ini penulis bekerja di Jurusan Gizi Poltekkes Kemenkes Aceh, yang mengampu mata kuliah Penilaian Status Gizi, IPTEK Pangan dan Gizi, Perencanaan Program Gizi, Metode Penelitian, Statistik, Komputasi Data Pangan dan Gizi, Journal Review, Ilmu Gizi. Penulis aktif diberbagai kegiatan Tridharma Perguruan Tinggi yaitu sebagai penulis buku, publikasi, seminar, pengelola jurnal terakreditasi, reviewer nasional dan international. Hasil-hasil penelitian dan publikasi yang telah penulis lakukan dapat dilihat pada profil Sinta penulis (ID= 256989) dan Google Scholar (ID= Lm44jiMAAAJ), serta publikasi international (ID Scopus: 57219055208). Selain itu, aktif dibidang Informasi dan Teknologi khusus bidang kesehatan, juga mahir dalam melakukan analisis data melalui aplikasi Stata, Smart PLS,

SPSS, R-Cmdr, EpiData, dll. Kedepannya, penulis akan terus berupaya meningkatkan kapasitas baik sebagai penulis buku ilmiah, sebagai peneliti maupun kepakaran dibidang Informasi dan Teknologi tentunya dalam mendukung sistem kesehatan digital.

Email Penulis: agus.hendra.alr@poltekkesaceh.ac.id

BIODATA PENULIS



Rahmatika Nur Aini, M.Gz

Dosen Program Studi Gizi

Fakultas Ilmu-ilmu Kesehatan

Universitas Muhammadiyah Prof. DR. Hamka Jakarta

Penulis lahir di Karanganyar tanggal 26 Juni 1993. Penulis adalah dosen tetap pada Program Studi Gizi Fakultas Ilmu-ilmu Kesehatan Universitas Muhammadiyah Prof. DR. Hamka Jakarta. Menyelesaikan pendidikan S1 Ilmu Gizi (2015) Universitas Muhammadiyah Surakarta (UMS) dan S2 Ilmu Gizi (2018) Universitas Sebelas Maret Surakarta (UNS). Penulis aktif melakukan catur darma perguruan tinggi, pembicara di forum ilmiah, menulis artikel di jurnal nasional, reviewer jurnal nasional (Jurnal Kesehatan), Pendamping PPH dari BPJPH (2021 hingga sekarang). Penulis tetap aktif mengajar dan mengisi kegiatan di kampus dan lingkungan perguruan Muhammadiyah. Penulis dapat dihubungi melalui rahmatika.nuraini@uhamka.ac.id.

BIODATA PENULIS



Arvida Bar, SPd, STr.Kep,MKM

Dosen Program Studi Sarjana Terapan Keperawatan
Jurusan Keperawatan Poltekkes Kemenkes Jambi

Penulis lahir di Jambi tanggal 11 Juni 1968. Penulis adalah dosen tetap pada Program Studi Sarjana Terapan Keperawatan Jurusan Keperawatan Poltekkes Kemenkes Jambi. Menyelesaikan pendidikan D3 Keperawatan Depkes Jambi Tahun 1989. S1 Ilmu Pendidikan (2001) Universitas Negeri Padang, Sarjana Terapan Keperawatan Poltekkes Kemenkes Jambi (2021) dan S2 Ilmu Kesehatan Masyarakat Peminatan Biostatistik dari Universitas Indonesia (2006). Penulis aktif melakukan tridharma perguruan tinggi, tim peneliti Narkoba (BNN) dan tim peneliti Gaya Hidup Masyarakat (BRIN), aktif publikasi artikel baik nasional dan internasional, Penulis Juga dipercaya sebagai wakil ketua Komite Etik Penelitian Kesehatan Poltekkes Kemenkes Jambi, selain itu dalam bidang organisasi profesi Penulis aktif sebagai pengurus DPW PPNI Propinsi Jambi dan DPD Kota Jambi.

BIODATA PENULIS



Dr. Pahrur Razi, SKM, MKM

Dosen Program Studi Promosi Kesehatan
Jurusan Promosi Kesehatan Poltekkes Kemenkes Jambi

Penulis lahir di Tanjung Pauh tanggal 16 April 1980. Penulis adalah dosen tetap pada Program Studi Promosi Kesehatan Program Sarjana Terapan. Jurusan Promosi Kesehatan Poltekkes Kemenkes Jambi. Menyelesaikan pendidikan S1 Kesehatan Masyarakat (2005) STIKES Harapan Ibu Jambi dan S2 Ilmu Kesehatan Masyarakat Peminatan Biostatistik dari Universitas Indonesia (2015) serta S3 Pendidikan Matematika dan IPA dari Universitas Jambi. Penulis aktif melakukan tridharma perguruan tinggi, tim peneliti Poltekkes Kemenkes Jambi dan Pengajar Mata Kuliah Biostatistik, Metodologi Penelitian serta Ilmu Promosi Kesehatan. Pembicara di forum ilmiah, aktif publikasi artikel baik nasional dan internasional. Penulis menguatkan perannya sebagai ahli promotor kesehatan yang tergabung sebagai Dewan Pengurus Daerah Perkumpulan Pendidik dan Promotor Kesehatan Masyarakat Indonesia (PPPKMI) Provinsi Jambi, Anggota Kolegium Promosi Kesehatan dan Perkumpulan Institusi Promotor Kesehatan Nusantara (PIPKN). Penulis dapat dihubungi melalui pahrur@poltekkesjambi.ac.id

BIODATA PENULIS



Sutyawan, S. Gz, M.Si

Dosen Program Studi Diploma III Gizi
Jurusan Gizi Poltekkes Kemenkes Pangkalpinang

Penulis lahir di Cit Belinyu Bangka tanggal 07 September 1991. Penulis adalah dosen tetap pada Program Studi Diploma III Gizi Jurusan Gizi Poltekkes Kemenkes Pangkalpinang. Menyelesaikan pendidikan S1 Ilmu Gizi (2013) dan S2 Ilmu Gizi (2019) di Institut Pertanian Bogor (IPB University). Penulis aktif melakukan tridharma perguruan tinggi dengan fokus pada bidang pangan dan gizi masyarakat. Penulis juga terlibat menjadi tim pakar audit *stunting* Kabupaten Bangka Tengah tahun 2023 s.d 2024, tim pakar penyusunan *draft* indikasi geografis teh tayu Desa Ketap Kabupaten Bangka Barat tahun 2024, serta pembicara ilmiah dengan tema *stunting* dan pangan lokal baik di tingkat lokal dan nasional. Penulis pernah menjadi supervisor penelitian dan kebijakan nasional di Provinsi Bangka Belitung yaitu Survey Status Gizi Indonesia (SSGI) tahun 2022 dan *Project Investing in Nutrition and Early Years (INEY)* Tahun 2024. Penulis menguatkan perannya sebagai ahli gizi yang tergabung sebagai Sekretaris Dewan Pimpinan

Daerah Persatuan Ahli Gizi Indonesia Provinsi Bangka Belitung (DPD PERSAGI Bangka Belitung). Penulis dapat dihubungi melalui sutyawan@poltekkespangkalpinang.ac.id

BIODATA PENULIS



Witi Karwiti, SKM.MPH

Dosen Program Studi Teknologi Laboratorium Medik
Program Sarjana Terapan Poltekkes Kemenkes Jambi

Penulis lahir di Kuningan tanggal 26 Oktober 1977. Penulis adalah dosen tetap pada Program Studi Teknologi Laboratorium Medik Program Sarjana Terapan Poltekkes Kemenkes Jambi. Menyelesaikan pendidikan S1 Ilmu Kesehatan Masyarakat (2003) Universitas Kader Bangsa Palembang dan S2 Ilmu Kesehatan Masyarakat dari Universitas Gadjah Mada Yogyakarta (2011). Penulis aktif sebagai dosen mulai tahun 2005 s/d sekarang, serta melakukan tridharma perguruan tinggi, menulis artikel baik di media digital (linisehat), aktif publikasi artikel baik nasional dan internasional.