



BIOSTATISTIK INFERENSIAL



Taruli Rohana Sinaga, Ribka Sabarina Panjaitan, Putri Pamungkas
A. Adriana Amal, Catherine Hermawan Salim
Rifki Sakinah Nompoo, Bunga Mardhotillah, Nurul Fadhillah Gani
Dika Betaditya, Hudriani Jamal

BIOSTATISTIK INFERENSIAL



UU 28 tahun 2014 tentang Hak Cipta

Fungsi dan sifat hak cipta Pasal 4

Hak Cipta sebagaimana dimaksud dalam Pasal 3 huruf a merupakan hak eksklusif yang terdiri atas hak moral dan hak ekonomi.

Pembatasan Perlindungan Pasal 26

Ketentuan sebagaimana dimaksud dalam Pasal 23, Pasal 24, dan Pasal 25 tidak berlaku terhadap:

- a. penggunaan kutipan singkat Ciptaan dan/atau produk Hak Terkait untuk pelaporan peristiwa aktual yang ditujukan hanya untuk keperluan penyediaan informasi aktual;
- b. Penggandaan Ciptaan dan/atau produk Hak Terkait hanya untuk kepentingan penelitian ilmu pengetahuan;
- c. Penggandaan Ciptaan dan/atau produk Hak Terkait hanya untuk keperluan pengajaran, kecuali pertunjukan dan Fonogram yang telah dilakukan Pengumuman sebagai bahan ajar; dan
- d. penggunaan untuk kepentingan pendidikan dan pengembangan ilmu pengetahuan yang memungkinkan suatu Ciptaan dan/atau produk Hak Terkait dapat digunakan tanpa izin Pelaku Pertunjukan, Produser Fonogram, atau Lembaga Penyiaran.

Sanksi Pelanggaran Pasal 113

1. Setiap Orang yang dengan tanpa hak dan/atau tanpa izin Pencipta atau pemegang Hak Cipta melakukan pelanggaran hak ekonomi Pencipta sebagaimana dimaksud dalam Pasal 9 ayat (1) huruf c, huruf d, huruf f, dan/atau huruf h untuk Penggunaan Secara Komersial dipidana dengan pidana penjara paling lama 3 (tiga) tahun dan/atau pidana denda paling banyak Rp500.000.000,00 (lima ratus juta rupiah).
2. Setiap Orang yang dengan tanpa hak dan/atau tanpa izin Pencipta atau pemegang Hak Cipta melakukan pelanggaran hak ekonomi Pencipta sebagaimana dimaksud dalam Pasal 9 ayat (1) huruf a, huruf b, huruf e, dan/atau huruf g untuk Penggunaan Secara Komersial dipidana dengan pidana penjara paling lama 4 (empat) tahun dan/atau pidana denda paling banyak Rp1.000.000.000,00 (satu miliar rupiah).

Biostatistik Inferensial

Taruli Rohana Sinaga, Ribka Sabarina Panjaitan, Putri Pamungkas
A. Adriana Amal, Catherine Hermawan Salim
Rifki Sakinah Nompo, Bunga Mardhotillah, Nurul Fadhilah Gani
Dika Betaditya, Hudriani Jamal



Penerbit Yayasan Kita Menulis

Biostatistik Inferensial

Copyright © Yayasan Kita Menulis, 2025

Penulis:

Taruli Rohana Sinaga, Ribka Sabarina Panjaitan, Putri Pamungkas
A. Adriana Amal, Catherine Hermawan Salim
Rifki Sakinah Nompoo, Bunga Mardhotillah, Nurul Fadhilah Gani
Dika Betaditya, Hudriani Jamal

Editor: Iko Mart Nadeak

Desain Sampul: Devy Dian Pratama, S.Kom.

Penerbit

Yayasan Kita Menulis

Web: kitamenulis.id

e-mail: press@kitamenulis.id

WA: 0813-9680-7167

IKAPI: 044/SUT/2021

Taruli Rohana Sinaga., dkk.

Biostatistik Inferensial

Yayasan Kita Menulis, 2025

xiv; 160 hlm; 16 x 23 cm

ISBN: 978-623-113-788-3

Cetakan 1, Maret 2025

I. Biostatistik Inferensial

II. Yayasan Kita Menulis

Katalog Dalam Terbitan

Hak cipta dilindungi undang-undang

Dilarang memperbanyak maupun mengedarkan buku tanpa
izin tertulis dari penerbit maupun penulis

Kata Pengantar

Puji syukur kami panjatkan kepada Tuhan Yang Maha Esa atas rahmat dan karunia-Nya, sehingga buku ini dapat diselesaikan. Buku ini disusun sebagai panduan praktis dalam memahami Biostatistik Inferensial, sebuah topik penting dalam dunia keperawatan dan bidang kesehatan lainnya. Biostatistik Inferensial sangat diperlukan untuk menganalisis data penelitian, yang pada gilirannya dapat digunakan untuk pengambilan keputusan yang berbasis bukti dalam praktik keperawatan dan kesehatan.

Biostatistik Inferensial tidak hanya mencakup analisis data yang berfokus pada pengujian hipotesis, estimasi, dan pengolahan data, tetapi juga berhubungan erat dengan penerapan hasil analisis untuk meningkatkan kualitas pelayanan kesehatan. Buku ini dirancang untuk memberikan pemahaman yang komprehensif mengenai konsep, teori, dan metode yang digunakan dalam analisis biostatistik, serta aplikasinya dalam berbagai disiplin ilmu, terutama keperawatan.

Buku ini membahas:

Bab 1 Konsep, Teori, dan Ruang Lingkup Biostatistik Inferensial

Bab 2 Konsep Dasar Analisis Bivariat

Bab 3 Estimasi

Bab 4 Analisis Hipotesis

Bab 5 Uji T-Independent dan Dependent

Bab 6 Uji ANOVA

Bab 7 Uji Korelasi

Bab 8 Analisis uji Chi Square dengan SPSS

Bab 9 Uji Regresi Linier Sederhana

Bab 10 Uji Regresi Logistik Sederhana

Kami berharap buku ini dapat membantu mahasiswa, peneliti, dan praktisi kesehatan dalam memahami dasar-dasar biostatistik inferensial serta aplikasinya. Dengan pengetahuan ini, diharapkan mereka dapat

mengolah data secara efektif untuk mendukung peningkatan kualitas layanan kesehatan yang lebih baik.

Akhir kata, kami mengucapkan terima kasih kepada semua pihak yang telah memberikan dukungan dalam penyusunan buku ini. Semoga buku ini bermanfaat dan menjadi referensi yang berguna bagi pembaca.

Maret 2025

Penulis

Daftar Isi

Kata Pengantar	v
Daftar Isi	vii
Daftar Gambar	xi
Daftar Tabel	xiii

Bab 1 Konsep, Teori, dan Ruang Lingkup Biostatistik Inferensial

1.1 Pendahuluan.....	1
1.2 Penggolongan dan Jenis Biostatistik.....	2
1.3 Istilah dan Simbol dalam Biostatistik Inferensial.....	6
1.3.1 Skala Pengukuran	6
1.3.2 Sumber Data.....	7
1.3.3 Tehnik Pengumpulan Data	8
1.3.4 Pengolahan Data	8
1.3.5 Analisis Data	9
1.4 Konsep, Teori, dan Ruang Lingkup.....	11

Bab 2 Konsep Dasar Analisis Bivariat

2.1 Pengertian Analisis Bivariat	13
2.2 Jenis Data dalam Analisis Bivariat	14
2.2.1 Data Kategorikal	14
2.2.2 Data Numerik	16
2.3 Teknik Analisis Bivariat	17
2.4 Statistik Analitis Bivariat	19
2.4.1 Tujuan Analisis Bivariat.....	20
2.4.2 Jenis Uji dalam Analisis Bivariat	20

Bab 3 Estimasi

3.1 Pengertian Estimasi	33
3.2 Jenis – Jenis Estimasi	37
3.3 Metode Estimasi	39
3.4 Cara Meningkatkan Akurasi Estimasi	41
3.5 Estimasi Probabilitas	44

Bab 4 Analisis Hipotesis

4.1	Pendahuluan.....	47
4.1.1	Peran Biostatistik dalam Keperawatan	47
4.1.2	Pengertian dan Fungsi Hipotesis	49
4.2	Jenis Hipotesis	51
4.2.1	Hipotesis Deskriptif.....	51
4.2.2	Hipotesis Kausal (Asosiasi).....	51
4.2.3	Hipotesis Komparatif	51
4.2.4	Hipotesis Korelasional	52
4.3	Langkah-langkah Uji Hipotesis.....	53
4.3.1	Menentukan Variabel	53
4.3.2	Menentukan Jenis Hipotesis	53
4.3.3	Menentukan Skala Pengukuran.....	54
4.3.4	Menentukan Berpasangan atau Tidak Berpasangan.....	55
4.3.5	Menentukan Jumlah Kelompok	56
4.3.6	Menentukan Uji Statistik	56
4.4	Aplikasi dalam Penelitian Keperawatan.....	57

Bab 5 Uji T-Independent dan Dependent

5.1	Pendahuluan.....	59
5.2	Uji t Independen	60
5.2.1	Pengertian.....	60
5.2.2	Asumsi Uji t Independen.....	61
5.2.3	Rumus Uji t Independen	61
5.2.4	Interpretasi Hasil.....	62
5.2.5	Contoh Kasus.....	62
5.3	Uji t Dependen.....	62
5.3.1	Pengertian.....	62
5.3.2	Asumsi Uji t Dependen.....	62
5.3.3	Rumus Uji t Dependen.....	63
5.3.4	Interpretasi Hasil.....	63
5.3.5	Alternatif Jika Asumsi Tidak Terpenuhi.....	64
5.3.6	Contoh Kasus.....	64
5.4	Keunggulan dan Kelemahan Uji t.....	64
5.5	Aplikasi Uji t dalam Berbagai Bidang.....	66
5.5.1	Uji t Independen dalam Kesehatan.....	67
5.5.2	Uji t Dependen dalam Kesehatan	68
5.5.3	Aplikasi Uji t dalam Bidang Psikologi.....	68
5.5.4	Aplikasi Uji t dalam Bidang Pendidikan	69

5.5.5 Aplikasi Uji t dalam Bidang Ekonomi dan Bisnis	69
Bab 6 Uji ANOVA	
6.1 Pengertian dan Sejarah Singkat ANOVA.....	71
6.2 Jenis-Jenis ANOVA.....	72
6.3 Asumsi ANOVA.....	74
6.4 Langkah-langkah Pengujian ANOVA.....	75
6.5 Penyelesaian ANOVA	78
6.5.1 One-Way ANOVA.....	78
6.5.2 Two-Way ANOVA.....	80
6.5.3 Repeated Measures ANOVA	81
6.5.4 Multivariate ANOVA (MANOVA)	81
6.5.5 Mixed-Design ANOVA.....	82
6.6 Langkah Melakukan Uji ANOVA dengan Penyelesaian SPSS	82
Bab 7 Uji Korelasi	
7.1 Korelasi	87
7.2 Korelasi Pearson untuk Analisis Parametrik	90
7.3 Korelasi Rank Spearman	92
Bab 8 Analisis uji Chi Square dengan SPSS	
8.1 Definisi	97
8.2 Asumsi dasar Uji Chi Square.....	100
8.3 Langkah-langkah uji Chi Square dengan SPSS	102
Bab 9 Uji Regresi Linier Sederhana	
9.1 Uji Regresi	109
9.2 Persamaan Regresi Linier	111
9.3 Koefisien Korelasi	115
9.4 Penaksiran Nilai Variabel Terikat	120
Bab 10 Uji Regresi Logistik Sederhana	
10.1 Konsep Uji Regresi Logistik	123
10.2 Tujuan Regresi Logistik	125
10.3 Asumsi Uji Regresi Logistik	126
10.4 Model Regresi Logistik.....	127
10.5 Langkah Uji Regresi Logistik	128
10.6 Contoh Uji Regresi Logistik Sederhana	130
10.7 Interpretasi Hasil Uji Regresi Logistik Sederhana.....	132
Daftar Pustaka	141
Biodata Penulis	153

Daftar Gambar

Gambar 1.1: Skema Penggolongan dan Jenis Statistik.....	4
Gambar 1.2: Tahapan Analisis Statistik.....	10
Gambar 1.3: Penggolongan Statistik Inferensial.....	10
Gambar 6.1: Daerah penerimaan H_0 dan H_a	75
Gambar 7.1: Tipikal Korelasi.....	88
Gambar 7.2: Beberapa Fungsi Monotonik.....	93
Gambar 8.1: Peta Uji Hipotesis Chi Square.....	101
Gambar 8.2: Pemilihan uji Chi-Square.....	102
Gambar 8.3: Aplikasi SPSS.....	103
Gambar 8.4: Nama variable.....	104
Gambar 8.5: Analyze-Descriptive Statistic-Crosstabs.....	104
Gambar 8.6: Variabel independen.....	105
Gambar 8.7: Chi-Square dan Risk.....	105
Gambar 8.8: Observed.....	106
Gambar 9.1: Uji Regresi.....	110
Gambar 9.2: Analisis Regresi Hubungan Usia dengan Kadar HbA1C.....	114
Gambar 10.1: Grafik Regresi Logistik.....	128
Gambar 10.2: Aplikasi SPSS.....	130
Gambar 10.3: Klik menu Analyze, pilih Regression, kemudian pilih Binary Logistic.....	131
Gambar 10.4: Kotak dialog Logistic Regression.....	131
Gambar 10.5: Categorical Covariates.....	132
Gambar 10.6: Kotak dialog Logistic Regression.....	137
Gambar 10.7: Statistic and Plots centang Hosmer-Lemeshow goodness-of-fit dan CI for exp (B): 95%.....	138

Daftar Tabel

Tabel 1.1: Skala Pengukuran	6
Tabel 2.1: Panduan interpretasi hasil uji hipotesi berdasarkan kekuatan korelasi, nilai p, dan arah korelasi	18
Tabel 4.1: Uji Statistik	56
Tabel 4.2: Aplikasi dalam Penelitian Keperawatan	57
Tabel 4.3: Perbedaan kadar gula darah	57
Tabel 4.4: Apakah ada hubungan beban kerja dengan stress kerja perawat?	58
Tabel 7.1: Korelasi X dan Y beserta Proses Penghitungannya	91
Tabel 8.1: Self Care Crosstabulation	106
Tabel 8.2: Chi-Square Test	107
Tabel 8.3: Rist Estimate	107
Tabel 9.1: Tipe Regresi	110
Tabel 9.2: Data besarnya jumlah produksi (X) dengan biaya produksi (Y)	112
Tabel 9.3: Tabel Bantu Data	113
Tabel 9.4: Menurut Guilford, kriteria untuk menentukan keeratan hubungan antara dua variabel	115
Tabel 9.5: Berikut data besarnya jumlah produksi (X) dengan biaya produksi (Y)	116
Tabel 9.6: Tabel Bantu Data	117
Tabel 9.7: Sampel Hari Kerja	118
Tabel 9.8: Tabel Bantu Data	119
Tabel 10.1: Pengaruh Paparan Iklan Rokok terhadap Perilaku Merokok	129
Tabel 10.2: Omnibus Test dan Model Summary	133
Tabel 10.3: Variables In the Equation	133
Tabel 10.4: Hasil Uji Regresi Logistik Tindakan Merokok Pelajar Indonesia Tahun 2014	138

Bab 1

Konsep, Teori, dan Ruang Lingkup Biostatistik Inferensial

1.1 Pendahuluan

Statistik adalah sekumpulan konsep dan metode yang digunakan untuk mengumpulkan dan menginterpretasikan data tentang bidang kegiatan tertentu serta mengambil kesimpulan dalam suatu kondisi dimana ada ketidakpastian dan adanya variasi. Secara lebih detail lagi pengertian statistik merupakan kumpulan data kuantitatif dan kualitatif yang disusun dalam tabel atau grafik yang menggambarkan suatu masalah baik masalah kesehatan, ekonomi, hukum dan lain-lain. Menyangkut defenisi ini maka statistik terkait dengan beberapa istilah seperti data, populasi, dan sampel (SP Hastono & L.Sabri, 2010).

Istilah “statistika” dan “statistik” mempunyai makna yang berbeda. Istilah statistika merupakan sebuah ilmu yang berkaitan dengan data sedangkan statistik itu sendiri menyatakan sebuah data, informasi, atau merupakan

perwujudan dari statistika yang berasal dari suatu data. Statistika adalah sekumpulan cara maupun aturan-aturan yang berkaitan dengan pengumpulan, pengolahan (analisis), penarikan kesimpulan atas data-data yang berbentuk angka dengan menggunakan suatu asumsi- asumsi tertentu.

Statistik merupakan salah satu cabang ilmu matematika yang mempelajari suatu pengukuran, observasi dan analisis. Statistik memiliki pengertian dasar yaitu suatu ringkasan dari suatu data yang berbentuk angka. Secara rinci statistik merupakan lmu yang mempelajari tentang cara pengumpulan data, pengolahan data, dan cara analisa data berdasarkan konsep probabilitas Sangila dan Luthfiah, 2018 & Husnul, dkk, 2020 dalam (Ristya Widi E.Y. dkk, 2023).

Biostatistik atau Statistik Kesehatan adalah suatu informasi dari sekumpulan data yang berkaitan dengan masalah kesehatan. Biostatistik digunakan dalam keperluan administratif misalnya untuk merencanakan program pelayanan kesehatan, sebagai indikator alternatif penyelesaian masalah kesehatan dan sebagai metode dalam melakukan analisis suatu penyakit dalam periode waktu tertentu.

Biostatistik terdiri dari dua kata dasar yaitu: bio dan statistik. Bio artinya hidup, sedangkan statistik kumpulan dari berbagai angka yang mempunyai variasi nilai. Dengan demikian secara harfiahnya biostatistik merupakan kumpulan dari berbagai angka yang berkaitan dengan kehidupan (Muhyidin, 2021).

1.2 Penggolongan dan Jenis Biostatistik

Penggolongan biostatistik menurut metode pengolahan data dibagi menjadi dua bagian yaitu: biostatistik deskriptif dan biostatistik inferensial. Biostatistik deskriptif berkaitan dengan bagaimana cara mendeskripsikan, menggambarkan atau menjabarkan serta menjelaskan suatu data sehingga akan lebih muda untuk dipahami.

Informasi yang dapat diperoleh dari biostatistik deskriptif antara lain:

1. Penentuan ukuran dari suatu data meliputi rata-rata atau mean, median atau nilai tengah dan modus;
2. Penentuan ukuran variabilitas data seperti: variasi (varian), tingkat penyimpangan (standar deviasi), dan jarak (range);
3. Penentuan ukuran bentuk suatu data misalnya: skewness, kurtosis, box plots.

Menurut Widodo, 2018 dalam (Ristya Widi E.Y. dkk, 2023) statistik deskriptif adalah metode statistik yang dapat digunakan untuk memaparkan hasil-hasil penelitian yang dilakukan dalam bentuk statistik populer sederhana. Rangkaian tahapan statistik deskriptif ini berguna untuk memberikan informasi secara detail sehingga lebih mudah untuk dipahami. Selanjutnya Budiarto 2012 dalam (Ristya Widi E.Y. dkk, 2023) menyatakan bahwa statistik deskriptif merupakan suatu metode atau cara pengumpulan data, pengolahan, penyajian data dan analisis secara sederhana berupa hasil perhitungan nilai tengah, variasi, rata-rata, rasio, populasi dan persentase.

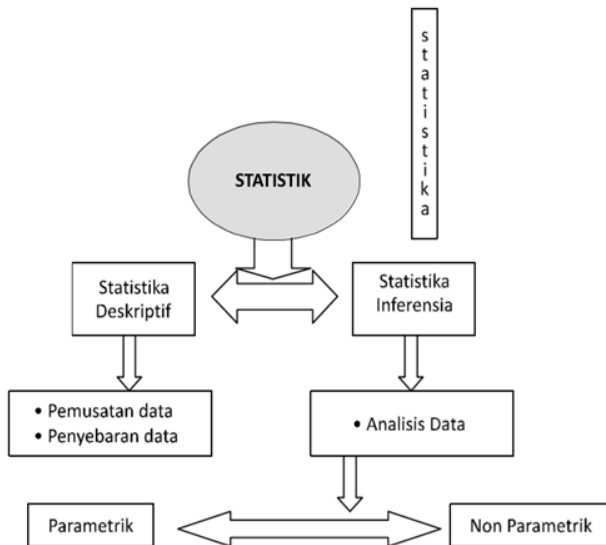
Statistika inferensial merupakan serangkaian cara atau tehnik yang digunakan untuk menaksir, mengkaji, menganalisis data serta mengambil kesimpulan berdasarkan data yang diperoleh dari sampel untuk menggambarkan atau mendeskripsikan karakteristik atau ciri dari suatu populasi. Metode ini sering disebut statistika induktif karena kesimpulan yang ditarik didasarkan pada informasi dari sebagian data. Statistika inferensia dibagi dalam dua kelompok yaitu statistika parametrik dan non parametrik (Syofian Siregar, 2014; VT Hulu & TR Sinaga, 2019a; Imani and Muslim, 2022).

Sementara itu menurut (Dharmawan Atik Mawarni, 2019) statistika inferensial bertujuan untuk menganalisis data serta mengambil kesimpulan (berkaitan dengan estimasi parameter dan pengujian hipotesis). Metode ini disebut statistika induktif karena kesimpulan yang diperoleh didasarkan pada informasi dari sebagian data sampel. Statistika inferensial terbagi dalam dua kelompok yaitu: statistika parametrik dan non parametrik.

Statistik parametrik adalah statistik dengan parameter populasi berdistribusi normal dan varians data bersifat homogen. Statistik non parametrik adalah statistik dengan parameter populasi tidak berdistribusi normal atau bebas dari persyaratan tertentu dan varians data tidak harus homogen.

Berbicara tentang penelitian khususnya penelitian kuantitatif, analisa data merupakan kegiatan setelah pengumpulan data dari lapangan. Umumnya tehnik analisa data dalam penelitian kuantitatif menggunakan alat-alat statistik. Ada dua jenis analisa dalam statistik yaitu analisa deskriptif dan analisa data inferensial. Analisa data inferensial merupakan penyajian statistik dari data yang diperoleh menggunakan sampel.

Analisis inferensial diterapkan untuk menganalisis data sampel yang dihasilkan dalam rangka mengestimasi data populasi. Penciri dari analisis statistik inferensial adalah penyajian data dalam bentuk tabel, grafik dan ukuran-ukuran statistik. Analisa data dalam statistik inferensial digunakan untuk generalisasi melalui dua pendekatan yaitu statistik parametrik dan non parametrik seperti yang sudah diuas sekilas di alinea sebelumnya.



Gambar 1.1: Skema Penggolongan dan Jenis Statistik (Yuantari and Handayani, 2017)

Statistik parametrik digunakan untuk menguji parameter populasi dengan menguji ukuran populasi melalui data sampel yang distribusi nilai pengamatannya diketahui. Pengujian populasi melalui sampel disebut uji hipotesis statistik. Dengan demikian penelitian yang mempunyai hipotesis adalah penelitian yang menggunakan data sampel. Sementara itu analisis statistik dengan pendekatan non parametrik diterapkan untuk menguji parameter populasi melalui data sampel yang distribusi nilai pengamatannya tidak diketahui (Purwoto A, 2007).

Hipotesis merupakan pernyataan yang akan dibuktikan secara empiris melalui penelitian. Jika didukung oleh data maka hipotesis terbukti benar dan jika tidak didukung oleh data maka hipotesis tidak terbukti benar. Pernyataan hipotesis menunjukkan adanya keterkaitan antara variabel penelitian. Pembuktian hipotesis selanjutnya dapat dilakukan dengan metode statistik inferensial. Hipotesis adalah pernyataan yg harus dibuktikan secara empiris.

Oleh karena itu pembahasan hipotesis terkait dengan variabel penelitian serta analisis statistik yang akan diterapkan untuk uji hipotesis. Biostatistik inferensial adalah statistik bidang kesehatan yang digunakan untuk menganalisis data sampel dan hasilnya akan digeneralisasikan/diinferensialkan kepada populasi dimana sampel diambil. Dengan demikian biostatistik inferensial merupakan lanjutan analisis uji perbedaan maupun uji hubungan antar variabel setelah didahului oleh analisis deskriptif (Alifiya V., 2023; Yuantari et al., 2020; VT Hulu & TR Sinaga, 2019; Muhyidin, 2021).

1.3 Istilah dan Simbol dalam Biostatistik Inferensial

1.3.1 Skala Pengukuran

Penelitian yang dilakukan dengan teknik pengumpulan data melalui alat ukur kuesioner, biasanya responden selalu memiliki jawaban yang berbeda-beda. Penggunaan skala ukur berkaitan dengan jenis data yang digunakan. Secara umum skala ukur yang digunakan dalam penelitian-penelitian kesehatan terdiri dari skala nominal, skala ordinal, skala interval dan skala rasio. Berikut penjelasan masing-masing skala ukur tersebut (Dahlan Sopiudin, 2021).

Tabel 1.1: Skala Pengukuran

Skala Ukur	Definisi	Contoh
Nominal	Merupakan skala ukur yang tidak memiliki tingkatan, nilainya sama atau sederajat. Pada skala nominal memiliki beberapa kategorik akan tetapi antara kategorik tidak dapat diketahui tingkat perbedaannya.	<ol style="list-style-type: none"> 1. Jenis kelamin (Laki-laki, perempuan) 2. Agama (Islam, Kristen, Katolik, Budha, Hindu)
Ordinal	Merupakan skala ukur yang memiliki urutan/tingkatan, nilainya bertingkat. Pada skala ordinal memiliki beberapa kategorik akan tetapi antara kategorik dapat diketahui tingkat perbedaannya.	<ol style="list-style-type: none"> 1. Pendidikan (SD, SMP, SMA, PT) 2. Pendapatan (Rendah, Menengah, Tinggi) 3. Tekanan darah tinggi (Rendah, Normal, Tinggi)
Interval	Merupakan skala ukur yang tidak memiliki nilai nol alami. Pada skala interval	<ol style="list-style-type: none"> 1. Suhu

Skala Ukur	Definisi	Contoh
	merupakan variabel numerik karena tidak memiliki kategori variabel.	
Rasio	Merupakan skala ukur yang memiliki nilai nol alami. Pada skala rasio juga merupakan variabel numerik karena tidak memiliki kategori variabel.	<ol style="list-style-type: none">1. Berat badan2. Tinggi badan3. Jarak

1.3.2 Sumber Data

Sumber data sangat penting dalam menentukan masalah penelitian, ketersediaan sumber data merupakan salah satu pertimbangan para peneliti dalam menentukan masalah penelitian. Sumber data penelitian berarti dari mana data penelitian dapat diperoleh, apakah berasal dari masyarakat atau responden maupun dari fasilitas pelayanan kesehatan seperti rumah sakit, puskesmas dan klinik. Untuk memperoleh data penelitian dapat dilakukan dengan cara menggunakan kuesioner atau wawancara kepada responden serta melalui kegiatan observasi.

Jika dalam penelitian tertentu menggunakan alat ukur seperti instrumen atau kuesioner dan wawancara dalam pengumpulan data maka sumber data adalah responden atau masyarakat, dalam hal ini responden yang menjawab pertanyaan peneliti baik secara tertulis maupun lisan. Secara umum sumber data dapat berasal dari masyarakat dan fasilitas atau institusi pelayanan. Data yang diperoleh dari masyarakat dapat berupa informasi yang terdapat dan diperoleh secara langsung dari masyarakat.

Sedangkan data yang diperoleh dari institusi pelayanan dapat berupa keseluruhan hasil dari berbagai kegiatan pelayanan kesehatan baik kegiatan pelayanan kesehatan yang dikumpulkan secara rutin maupun tidak rutin (VT Hulu & TR Sinaga, 2019).

1.3.3 Teknik Pengumpulan Data

Teknik pengumpulan data kesehatan sangat penting dalam melakukan penelitian-penelitian kesehatan. Jika pengumpulan data dilakukan dengan benar maka akan memberikan informasi menjadi sah serta dapat dipercaya. Oleh karena itu pengumpulan data penelitian kesehatan dapat dilakukan dengan berbagai cara seperti pengumpulan data dengan teknik wawancara, observasi, kuesioner dan seterusnya.

Pengumpulan data penelitian terdiri dari dua bagian yaitu teknik pengamatan dan teknik komunikasi. Pengumpulan data dengan menggunakan teknik pengamatan dapat dilakukan dengan pengamatan langsung kepada responden dan tidak langsung karena menggunakan alat bantu seperti laboratorium. Pengumpulan data dengan menggunakan teknik komunikasi dapat dilakukan dengan komunikasi langsung melalui wawancara kepada responden dan tidak langsung dengan menggunakan kuesioner tertulis (Dahlan Sopiyyudin, 2021) & (Purwoto Agus, 2007;).

1.3.4 Pengolahan Data

Penggunaan pengolahan data sangat penting dalam melakukan analisis data penelitian kuantitatif. Pengolahan data merupakan langkah-langkah yang digunakan untuk menganalisis data yang sudah diperoleh setelah melakukan penelitian. Misalkan seorang mahasiswa yang sedang penelitian dengan skripsi, maka hasil penelitian atau data yang sudah diperoleh dari lapangan, pasti diolah untuk mendapatkan informasi terkait dengan data yang sudah diperoleh selama penelitian.

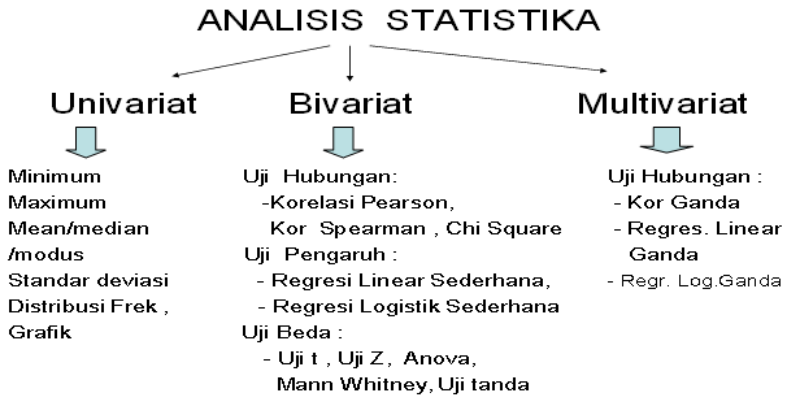
Perlu diketahui bahwa dalam melakukan pengolahan data harus sesuai dengan kode yang diberikan terhadap masing-masing variabel misalnya jenis kelamin yaitu laki-laki diberi kode 1(satu) dan perempuan diberi kode 2 (dua).

1.3.5 Analisis Data

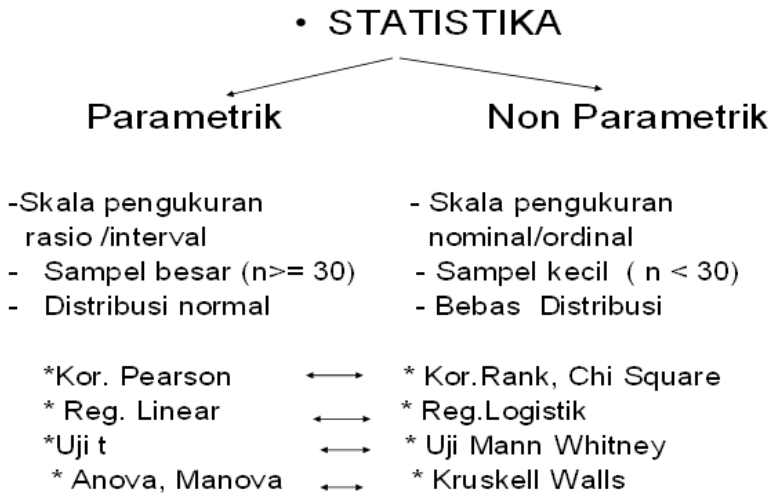
Setelah data dikumpulkan dari lapangan melalui kegiatan penelitian, maka data yang kumpulkan tersebut diproses dengan teknik pengolahan dan analisis data.

Secara umum analisis data penelitian terdiri dari analisis univariat, bivariat dan multivariate:

1. Analisis univariat merupakan analisis yang bertujuan untuk mengetahui distribusi frekuensi pada setiap variabel penelitian. Analisis univariat hanya mendeskripsikan masing-masing variabel penelitian. Analisis univariat dilakukan jika yang dianalisis hanya satu variabel. Ukuran nilai-nilai statistik deskriptif yang digunakan pada analisis ini adalah ukuran pemusatan data (misalnya rerata, median dan modus), ukuran penyebaran data (misalnya range, simpangan baku dan varians), serta melalui tabel distribusi frekuensi, grafik atau histogram.
2. Analisis bivariate. Pada analisis ini digunakan untuk menguji hubungan antara dua variabel, yaitu hubungan antara masing-masing variabel independen dengan variabel dependen. Teknik analisis bivariat digunakan jika terdapat dua variabel yang akan dianalisis.
3. Analisis multivariat. Analisis ini digunakan untuk melihat hubungan atau pengaruh beberapa variabel independen secara bersama-sama terhadap variabel dependen. Jika lebih dari dua variabel yang dianalisis dalam suatu penelitian disebut dengan teknik multivariat. Secara skematis tahapan analisis statistik dan penggolongan statistik inferensial diuraikan sebagai berikut (Dahlan Sopiudin, 2021).



Gambar 1.2: Tahapan Analisis Statistik



Gambar 1.3: Penggolongan Statistik Inferensial

1.4 Konsep, Teori, dan Ruang Lingkup

Biostatistik inferensial merupakan bagian dari biostatistik yang berfokus pada pengambilan kesimpulan tentang populasi berdasarkan data sampel.

Berikut adalah rangkuman tentang konsep, teori, dan ruang lingkup biostatistik inferensial:

Konsep:

1. Populasi dan Sampel: Populasi adalah keseluruhan unit analisis, sedangkan sampel adalah subset dari populasi yang digunakan untuk menarik kesimpulan.
2. Inferensi Statistik: Proses menggunakan data sampel untuk membuat estimasi atau keputusan tentang populasi.
3. Estimasi: Menghitung parameter populasi (seperti rata-rata atau proporsi) berdasarkan statistik dari sampel.
4. Uji Hipotesis: Metode untuk menguji asumsi tentang parameter populasi dengan menggunakan data sampel.

Teori:

1. Teori Sampling: Mempelajari bagaimana sampel diambil dan sifat-sifat distribusi sampel.
2. Distribusi Probabilitas: Menjelaskan bagaimana nilai variabel acak terdistribusi. Distribusi normal dan distribusi t adalah yang sering digunakan dalam biostatistik.
3. Interval Kepercayaan: Suatu rentang nilai di mana parameter populasi kemungkinan besar terletak berdasarkan data sampel.
4. Uji Statistik: Metode atau cara untuk menguji hipotesis melalui perhitungan statistik, seperti uji t, uji chi-square, dan ANOVA.

Ruang Lingkup:

1. Penelitian Kesehatan: Digunakan untuk analisis data dalam epidemiologi, studi klinis, dan survei kesehatan.
2. Analisis Data Genetik: Menerapkan metode inferensial untuk menarik kesimpulan tentang hubungan genetik dan fenotip.
3. Kesehatan Masyarakat: Menggunakan biostatistik untuk merancang strategi pencegahan dan intervensi berdasarkan data populasi.
4. Statistik Klinis: Mengkaji efektivitas intervensi kesehatan, seperti uji coba obat atau teknik medis.

Biostatistik inferensial memainkan peran penting dalam pengambilan keputusan di bidang kesehatan dan penelitian, memberikan kerangka kerja untuk menginterpretasi data dan hasil penelitian. Secara lebih detail Biostatistik Inferensial akan diuraikan pada bab selanjutnya.

Bab 2

Konsep Dasar Analisis Bivariat

2.1 Pengertian Analisis Bivariat

Data bivariat berisi dua variabel yang berbeda, dan tujuan dari analisis data semacam ini adalah untuk memahami hubungan atau koneksi antara kedua variabel tersebut. Analisis bivariat adalah jenis analisis data yang mengacu pada penyelidikan dua variabel untuk menemukan hubungan di antara keduanya. (Mulyana, et al., 2019). Analisis ini sering digunakan untuk menentukan pengaruh dan hubungan antara variabel x dan y , tetapi juga dapat digunakan untuk menentukan perbedaan variabel x dan z .

Analisis bivariat terdiri dari tiga macam yaitu (Widiyono, et al., 2023):

1. Menggunakan analisis korelasi chi-square untuk mengidentifikasi hubungan
2. Menggunakan regresi dasar untuk mengidentifikasi pengaruh
3. Uji t sampel independen, uji t satu sampel, uji t sampel berpasangan, uji mann-Whitney, Kruskal Wallis, uji Friedman,

dan uji tanda semuanya dapat digunakan untuk mencari perbedaan.

Analisis bivariat bertujuan untuk mengetahui hubungan atau perbedaan antara variabel-variabel dalam membuat tabulasi silang antar variabel yang berhubungan, dan dilihat pola kecenderungan hubngannya secara deskriptif dengan membaca secara silang. Karena itu tabel bivariat juga dikenal sebagai tabel silang (cross tabulation). Tabel bivariat menyajikan data dua variabel dalam kolom dan baris bersamaan (Mutalazimah, 2025).

2.2 Jenis Data dalam Analisis Bivariat

Analisis bivariat memeriksa hubungan antara dua variabel, yang dapat berupa kategorial atau numerik. Jenis dara ini menentukan metode statistik yang digunakan dalam analisis.

2.2.1 Data Kategorikal

Data kategorikal mewakili karakteristik kualitatif yang menggambarkan suatu fitur tetapi tidak memiliki makna numerik.

Data ini dibagi menjadi nominal dan ordinal:

1. Data Nominal

Data nominal terdiri dari kategori yang tidak memiliki peringkat atau urutan intrinsik.

Contoh:

- a. Jenis Kelamin (Laki-laki, Perempuan)
- b. Golongan Darah (A, B, AB, O)
- c. Status Pernikahan (Lajang, Menikah, Bercerai)

Metode statistik untuk analisis bivariat:

- a. Uji Chi-square: Menentukan apakah ada hubungan antara dua variable kategorikal.
- b. Tabel Kontigensi: Digunakan untuk merangkum hubungan antara dua variable kategorikal

Contoh Penerapan: Sebuah penelitian menyelidiki apakah status merokok (perokok vs bukan perokok) berhubungan dengan jenis kelamin (laki-laki vs perempuan). Uji chi-square dapat digunakan untuk mengetahui apakah ada perbedaan signifikan dalam kebiasaan merokok berdasarkan jenis kelamin.

2. Data Ordinal

Data ordinal memiliki peringkat yang bermakna, tetapi selisih antar kategori tidak selalu sama.

Contoh:

- a. Tingkat Nyeri (Ringan, Sedang, Berat)
- b. Tingkat Pendidikan (SMA, Sarjana, Pascasarjana)
- c. Skala Likert (Sangat Setuju, Setuju, Netral, Tidak Setuju, Sangat Tidak Setuju).

Metode statistik untuk analisis bivariat:

- a. Korelasi Rank Spearman: Mengukur kekuatan dan arah hubungan antara dua variable ordinal
- b. Uji Mann-Whitney U: Membandingkan dua kelompok independen ketika data tidak berdistribusi normal.

Contoh Penerapan: Sebuah rumah sakit meneliti apakah pasien dengan tingkat pendidikan tinggi lebih cenderung mengikuti pola makan sehat untuk diabetes. Data (tingkat pendidikan dan kepatuhan diet) dapat dianalisis menggunakan korelasi Spearman.

2.2.2 Data Numerik

Data numerik mewakili besaran yang dapat diukur dan memungkinkan perhitungan matematis.

Data ini dibagi menjadi interval dan rasio:

1. Data Interval

Data interval memiliki selisih yang sama antara nilai-nilai tetapi tidak memiliki nol absolut. Contoh: Suhu dalam Celsius atau Fahrenheit (0°C tidak berarti “tidak ada suhu”).

Metode statistik untuk analisis bivariat:

- a. Korelasi Pearson (r): Mengukur kekuatan dan arah hubungan linear antara dua variabel interval
- b. Uji T: Membandingkan rata-rata dua kelompok ketika variabel dependen berupa data interval.

Contoh Penerapan: Seorang peneliti menganalisis hubungan antara skor IQ siswa dan nilai ujian mereka. Korelasi Pearson dapat digunakan untuk mengukur hubungan antara kedua variabel numerik ini.

2. Data Rasio

Data rasio memiliki selisih yang sama antara nilai-nilai dan titik nol absolut, yang berarti nol menunjukkan ketiadaan variabel yang diukur.

Contoh:

- a. Berat Badan (0 Kg berarti tidak ada berat)
- b. Tinggi Badan (0 cm berarti tidak ada tinggi)
- c. Tekanan Darah (0 mmHg berarti tidak ada tekanan darah)

Metode statistik untuk analisis bivariat:

- a. Korelasi Pearson: Menentukan kekuatan dan arah hubungan antara dua variabel rasio.

- b. Regresi Linear: Memprediksi nilai satu variabel berdasarkan variabel lain.

Contoh Penerapan: Sebuah penelitian menyelidiki apakah indeks massa tubuh (IMT) berhubungan dengan tekanan darah. Korelasi Pearson atau regresi linear dapat digunakan untuk menganalisis hubungan antara kedua variabel rasio ini.

2.3 Teknik Analisis Bivariat

Analisis bivariat adalah metode analisis yang digunakan untuk memahami hubungan antara dua variabel dalam sebuah penelitian. Dalam analisis bivariat, fokus diberikan pada hubungan atau perbandingan antara dua variabel yang terkait. Tujuan utama analisis bivariat adalah untuk mengidentifikasi hubungan, asosiasi, atau perbedaan antara variabel-variabel tersebut.

Berikut adalah beberapa teknik analisis bivariat yang umum digunakan (Sarie, et al., 2023):

1. Korelasi

Analisis korelasi digunakan untuk mengukur kekuatan dan arah hubungan antara dua variabel numerik. Koefisien korelasi seperti koefisien Pearson digunakan untuk menggambarkan hubungan linier antara dua variabel. Nilai korelasi berkisar antara -1 hingga 1, dengan nilai positif menunjukkan hubungan positif dan nilai negatif menunjukkan hubungan negatif.

Interpretasi hasil uji korelasi didasarkan pada nilai p , kekuatan korelasi, serta arah korelasinya. Berikut Panduan lengkap dapat dilihat dari tabel (Dahlan, 2011):

Tabel 2.1: Panduan interpretasi hasil uji hipotesis berdasarkan kekuatan korelasi, nilai p, dan arah korelasi

No	Parameter	Nilai	Interpretasi
1.	Kekuatan Korelasi (r)	0,00-0,199	Sangat Lemah
		0,20-0,399	Lemah
		0,40-0,599	Sedang
		0,60-0,799	Kuat
		0,80-1,000	Sangat Kuat
2.	Nilai p	$P < 0,05$	Terdapat korelasi yang bermakna antara dua variabel yang diuji.
		$P > 0,05$	Tidak terdapat korelasi yang bermakna antara dua variabel yang diuji
3.	Arah Korelasi	+ (positif)	Searah, semakin besar nilai satu variabel semakin besar pula nilai variabel lainnya.
		- (negatif)	Berlawanan arah. Semakin besar nilai satu variabel, semakin kecil nilai variabel lainnya

2. Regresi

Analisis regresi digunakan untuk memodelkan hubungan antara variabel dependen (variabel respons) dan satu atau lebih variabel independent (variabel prediktor). Analisis regresi dapat

memberikan informasi tentang pengaruh variabel prediktor terhadap variabel respons, serta memprediksi nilai respons berdasarkan nilai prediktor.

3. Kontingensi

Analisis kontingensi (contingency analysis) digunakan untuk memeriksa hubungan antara dua variabel kategorikal. Dalam analisis ini, table kontingensi digunakan untuk menggambarkan frekuensi atau proporsi masing-masing kombinasi nilai dari kedua variabel.

4. Chi-Square

Uji chi-square digunakan dalam analisis kontingensi untuk menguji keberadaan hubungan yang signifikan antara dua variabel kategorikal. Uji chi-square antara dua variabel kategorikal. Uji chi-square menghitung nilai chi-square yang membandingkan frekuensi pengamatan dengan frekuensi yang diharapkan tidak ada hubungan antara dua variabel.

5. Uji T-Test/Uji Anova

Uji t-test atau uji analisis varians (ANOVA) digunakan untuk membandingkan rata-rata antara dua atau lebih kelompok pada satu variabel dependen. Uji ini membantu mengidentifikasi perbedaan signifikan antara kelompok-kelompok tersebut.

2.4 Statistik Analitis Bivariat

Statistik analitis bivariat adalah metode analisis statistik yang digunakan untuk menguji hubungan antara dua variabel. Analisis ini bertujuan untuk melihat ada atau tidaknya hubungan, arah hubungan, serta kekuatan hubungan antara dua variabel yang diamati.

2.4.1 Tujuan Analisis Bivariat

1. Menguji hubungan antara dua variabel, misalnya antara pola makan dan kadar gula darah.
2. Menentukan arah hubungan, apakah positif (searah), negatif (berlawanan arah), atau tidak ada hubungan.
3. Mengukur kekuatan hubungan menggunakan nilai koefisien korelasi atau uji statistik lainnya.

2.4.2 Jenis Uji dalam Analisis Bivariat

Pemilihan uji statistik dalam analisis bivariat bergantung pada jenis data yang digunakan:

1. Jika Kedua Variabel Berskala Nominal

Uji Chi-Square (χ^2) → Digunakan untuk melihat hubungan antar kategori, misalnya hubungan antara jenis kelamin dan kejadian hipertensi. Uji Chi-Square (χ^2) digunakan untuk menguji hubungan antara dua variabel kategori (nominal atau ordinal) dalam suatu tabel kontingensi.

Berikut adalah langkah-langkahnya:

- a. Tentukan Hipotesis
 - 1) H_0 (Hipotesis Nol): Tidak ada hubungan antara kedua variabel (independen).
 - 2) H_1 (Hipotesis Alternatif): Ada hubungan antara kedua variabel (dependen).

Contoh:

- 1) H_0 : Tidak ada hubungan antara kebiasaan merokok dan kejadian hipertensi.
- 2) H_1 : Ada hubungan antara kebiasaan merokok dan kejadian hipertensi.

b. Interpretasi Hasil

Jika $p\text{-value} < 0,05$, maka dapat disimpulkan bahwa ada hubungan yang signifikan antara kebiasaan merokok dan kejadian hipertensi.

Jika $p\text{-value} > 0,05$, maka tidak ada hubungan yang signifikan.

c. Contoh Penggunaan Software

Jika menggunakan SPSS, langkah-langkahnya adalah:

- 1) Input data dalam format kategori.
- 2) Pilih menu Analyze → Descriptive Statistics → Crosstabs.
- 3) Masukkan variabel ke dalam baris dan kolom.
- 4) Klik Statistics → Chi-Square dan klik OK.
- 5) Interpretasi hasil berdasarkan $p\text{-value}$.

Interpretasi Hasil:

- 1) $P\text{-value} (< 0,05)$ → Menunjukkan ada hubungan yang signifikan antara dua variabel.
- 2) Koefisien korelasi (r) → Menunjukkan arah dan kekuatan hubungan, dengan nilai berkisar antara -1 hingga 1:
 - $r > 0$ → Hubungan positif (searah).
 - $r < 0$ → Hubungan negatif (berlawanan arah).
 - $r = 0$ → Tidak ada hubungan.

2. Jika Satu Variabel Nominal Dan Satu Variabel Numerik

- a. Uji t (Independent t-test) → Untuk membandingkan rata-rata dua kelompok, misalnya perbedaan tekanan darah antara perokok dan non-perokok.
- b. Uji Mann-Whitney → Alternatif dari uji t jika data tidak berdistribusi normal.

Berikut adalah langkah-langkah melakukan Uji t (Independent t-test):

a. Tentukan Hipotesis

- 1) H_0 (Hipotesis Nol): Tidak ada perbedaan rata-rata antara kedua kelompok.
- 2) H_1 (Hipotesis Alternatif): Ada perbedaan rata-rata antara kedua kelompok.

Contoh:

- 1) H_0 : Tidak ada perbedaan rata-rata tekanan darah antara kelompok perokok dan non-perokok.
- 2) H_1 : Ada perbedaan rata-rata tekanan darah antara kelompok perokok dan non-perokok.

b. Mengumpulkan Data

Pastikan data memenuhi beberapa asumsi:

- 1) Data berskala interval atau rasio (misalnya berat badan, tekanan darah, kadar gula darah).
- 2) Sampel diambil dari dua kelompok yang independen (tidak saling berhubungan).
- 3) Data berasal dari populasi yang terdistribusi normal.
- 4) Varians kedua kelompok sebaiknya homogen (dapat diuji dengan uji Levene).

c. Melakukan Uji Asumsi

- 1) Uji Normalitas
 - Digunakan untuk mengecek apakah data berdistribusi normal.
 - Bisa menggunakan Uji Shapiro-Wilk atau Kolmogorov-Smirnov.

- 2) Uji Homogenitas Varians (Levene's Test)
 - Jika Sig. > 0,05 → Varians kedua kelompok dianggap homogen (gunakan Equal Variances Assumed).
 - Jika Sig. ≤ 0,05 → Varians tidak homogen (gunakan Equal Variances Not Assumed).
- d. Interpretasi Hasil
 - 1) Jika p-value < 0,05, maka terdapat perbedaan yang signifikan antara dua kelompok (tolak H_0).
 - 2) Jika p-value > 0,05, maka tidak terdapat perbedaan yang signifikan antara dua kelompok (gagal menolak H_0).

Contoh Interpretasi: Jika hasil uji menunjukkan $p = 0,03$ (< 0,05), maka dapat disimpulkan bahwa terdapat perbedaan rata-rata tekanan darah yang signifikan antara perokok dan non-perokok.

- e. Contoh Penggunaan Software (SPSS)

Jika menggunakan SPSS, langkah-langkahnya adalah:

 - 1) Input data dalam format angka (misalnya, tekanan darah) dan kategori kelompok (1 = Perokok, 2 = Non-Perokok).
 - 2) Pilih Analyze → Compare Means → Independent-Samples T Test.
 - 3) Masukkan variabel dependent (misalnya, tekanan darah) ke dalam kolom Test Variable.
 - 4) Masukkan variabel independen (misalnya, kelompok perokok dan non-perokok) ke dalam Grouping Variable.
 - 5) Klik Define Groups, lalu masukkan kode kelompok (misalnya, 1 dan 2), kemudian klik Continue.
 - 6) Klik OK untuk menjalankan analisis.

f. Interpretasi Hasil di SPSS

SPSS akan menghasilkan tabel berikut:

- 1) Group Statistics → Menampilkan rata-rata dan standar deviasi masing-masing kelompok.
- 2) Independent Samples Test → Menampilkan hasil uji t, termasuk:
 - Levene's Test for Equality of Variances:
 - Jika Sig. > 0,05 → Gunakan baris pertama (Equal variances assumed).
 - Jika Sig. ≤ 0,05 → Gunakan baris kedua (Equal variances not assumed).
 - t-test for Equality of Means:
 - Jika Sig. (2-tailed) ≤ 0,05 → Ada perbedaan yang signifikan.
 - Jika Sig. (2-tailed) > 0,05 → Tidak ada perbedaan yang signifikan

Berikut langkah – langkah melakukan Uji Mann-Whitney:

a. Tentukan Hipotesis

- 1) H_0 (Hipotesis Nol): Tidak ada perbedaan median antara kedua kelompok.
- 2) H_1 (Hipotesis Alternatif): Ada perbedaan median antara kedua kelompok.

Contoh:

- 1) H_0 : Tidak ada perbedaan tingkat nyeri antara pasien yang diberikan terapi A dan terapi B.
- 2) H_1 : Ada perbedaan tingkat nyeri antara pasien yang diberikan terapi A dan terapi B.

b. Mengumpulkan Data

Pastikan data memenuhi beberapa asumsi:

- 1) Data berskala ordinal atau interval/rasio yang tidak berdistribusi normal.
- 2) Sampel diambil dari dua kelompok yang independen.

c. Interpretasi Hasil

- 1) Jika $p\text{-value} < 0,05$, maka terdapat perbedaan yang signifikan antara dua kelompok (tolak H_0).
- 2) Jika $p\text{-value} > 0,05$, maka tidak terdapat perbedaan yang signifikan (gagal menolak H_0).

Contoh Interpretasi: Jika hasil uji menunjukkan $p = 0,02 (< 0,05)$, maka dapat disimpulkan bahwa terdapat perbedaan yang signifikan dalam tingkat nyeri antara pasien yang menerima terapi A dan terapi B.

d. Contoh Penggunaan Software (SPSS)

Jika menggunakan SPSS, langkah-langkahnya adalah:

- 1) Input data dalam format angka (misalnya, tingkat nyeri) dan kategori kelompok (1 = Terapi A, 2 = Terapi B).
- 2) Pilih Analyze → Nonparametric Tests → Legacy Dialogs → 2 Independent Samples.
- 3) Masukkan variabel dependent (misalnya, tingkat nyeri) ke dalam Test Variable List.
- 4) Masukkan variabel independen (misalnya, kelompok terapi) ke dalam Grouping Variable.
- 5) Klik Define Groups, lalu masukkan kode kelompok (misalnya, 1 dan 2), kemudian klik Continue.
- 6) Pilih Mann-Whitney U dan klik OK untuk menjalankan analisis.

e. Interpretasi Hasil di SPSS

SPSS akan menghasilkan tabel berikut:

- 1) Ranks → Menampilkan jumlah peringkat rata-rata dari masing-masing kelompok.
- 2) Test Statistics → Menampilkan nilai U dan p-value.
Jika $p\text{-value} < 0,05$, berarti terdapat perbedaan signifikan antara kedua kelompok.

3. Jika Kedua Variabel Berskala Numerik

- a. Korelasi Pearson → Jika data berdistribusi normal, misalnya hubungan antara tekanan darah dan kadar gula darah.
- b. Korelasi Spearman → Jika data tidak berdistribusi normal atau berbentuk ordinal.

Berikut langkah-langkah melakukan Korelasi Pearson: Uji Korelasi Pearson digunakan untuk mengetahui hubungan antara dua variabel kuantitatif (interval/rasio) yang berdistribusi normal.

a. Tentukan Hipotesis

Contoh:

- 1) H_0 : Tidak ada hubungan antara tekanan darah dan kadar kolesterol.
- 2) H_1 : Ada hubungan antara tekanan darah dan kadar kolesterol.

b. Mengumpulkan Data

Pastikan data memenuhi asumsi berikut:

- 1) Skala data harus interval atau rasio.
- 2) Hubungan linier antara kedua variabel (dapat diperiksa dengan scatter plot).
- 3) Distribusi normal (dapat diuji dengan uji normalitas seperti Kolmogorov-Smirnov atau Shapiro-Wilk).

c. Interpretasi Hasil

- 1) Jika $p\text{-value} < 0,05$, maka terdapat hubungan yang signifikan antara kedua variabel (tolak H_0).
- 2) Jika $p\text{-value} > 0,05$, maka tidak terdapat hubungan yang signifikan (gagal menolak H_0).

Contoh Interpretasi: Jika hasil uji menunjukkan $r = 0,75$ dan $p = 0,001 (< 0,05)$, maka dapat disimpulkan bahwa terdapat hubungan positif yang kuat dan signifikan antara tekanan darah dan kadar kolesterol.

d. Contoh Penggunaan Software (SPSS)

Jika menggunakan SPSS, langkah-langkahnya adalah:

- 1) Input data dalam format angka (misalnya, tekanan darah dan kadar kolesterol).
 - 2) Pilih Analyze → Correlate → Bivariate.
 - 3) Masukkan kedua variabel yang ingin diuji ke dalam kolom Variables.
 - 4) Centang opsi Pearson pada bagian Correlation Coefficients.
 - 5) Pilih Two-tailed (untuk uji dua arah) atau One-tailed (jika hanya ingin menguji satu arah).
 - 6) Klik OK untuk menjalankan analisis.
- e. Interpretasi Hasil di SPSS

SPSS akan menghasilkan tabel dengan:

- 1) Nilai r (Pearson Correlation) → Menunjukkan kekuatan dan arah hubungan.
- 2) $p\text{-value}$ (Sig. 2-tailed) → Menentukan apakah hubungan signifikan atau tidak.

Jika $p\text{-value} < 0,05$, maka ada hubungan yang signifikan antara kedua variabel.

Berikut langkah-langkah melakukan uji Korelasi Spearman: Uji Korelasi Spearman digunakan untuk mengetahui hubungan antara dua variabel yang berskala ordinal atau kuantitatif (interval/rasio) tetapi tidak berdistribusi normal.

a. Tentukan Hipotesis

- 1) H_0 (Hipotesis Nol): Tidak ada hubungan antara kedua variabel.
- 2) H_1 (Hipotesis Alternatif): Ada hubungan antara kedua variabel.

Contoh:

- 1) H_0 : Tidak ada hubungan antara tingkat stres dan kualitas tidur.
- 2) H_1 : Ada hubungan antara tingkat stres dan kualitas tidur.

b. Mengumpulkan Data

Pastikan data memenuhi asumsi berikut:

- 1) Variabel minimal berskala ordinal atau lebih tinggi (interval/rasio).
- 2) Hubungan antara variabel bersifat monotonic (bisa linier atau tidak).
- 3) Tidak memerlukan distribusi normal, sehingga cocok untuk data yang tidak berdistribusi normal.

c. Interpretasi Hasil

- 1) Jika $p\text{-value} < 0,05$, maka terdapat hubungan yang signifikan antara kedua variabel (tolak H_0).
- 2) Jika $p\text{-value} > 0,05$, maka tidak terdapat hubungan yang signifikan (gagal menolak H_0).

Contoh Interpretasi: Jika hasil uji menunjukkan $\rho = -0,65$ dan $p = 0,002 (< 0,05)$, maka dapat disimpulkan bahwa

terdapat hubungan negatif yang kuat dan signifikan antara tingkat stres dan kualitas tidur. Artinya, semakin tinggi tingkat stres, semakin buruk kualitas tidurnya.

d. Contoh Penggunaan Software (SPSS)

Jika menggunakan SPSS, langkah-langkahnya adalah:

- 1) Input data dalam format angka (misalnya, tingkat stres dan kualitas tidur).
- 2) Pilih Analyze → Correlate → Bivariate.
- 3) Masukkan kedua variabel yang ingin diuji ke dalam kolom Variables.
- 4) Centang opsi Spearman pada bagian Correlation Coefficients.
- 5) Pilih Two-tailed (untuk uji dua arah) atau One-tailed (jika hanya ingin menguji satu arah).
- 6) Klik OK untuk menjalankan analisis.

e. Interpretasi Hasil di SPSS

SPSS akan menghasilkan tabel dengan:

- 1) Nilai ρ (Spearman Correlation) → Menunjukkan kekuatan dan arah hubungan.
- 2) p-value (Sig. 2-tailed) → Menentukan apakah hubungan signifikan atau tidak.

Jika p-value < 0,05, maka ada hubungan yang signifikan antara kedua variabel.

4. Jika ingin melihat pengaruh satu variabel terhadap variabel lain

Regresi Linear Sederhana → Digunakan untuk melihat pengaruh variabel independen terhadap variabel dependen, misalnya pengaruh durasi olahraga terhadap indeks massa tubuh (IMT).

Uji Regresi Linear Sederhana digunakan untuk melihat pengaruh variabel independen terhadap variabel dependen. Contohnya, pengaruh durasi olahraga terhadap Indeks Massa Tubuh (IMT).

Berikut langkah-langkah melakukan uji Regresi Linear Sederhana:

a. Tentukan Hipotesis

- 1) H_0 (Hipotesis Nol): Tidak ada pengaruh variabel independen terhadap variabel dependen.
- 2) H_1 (Hipotesis Alternatif): Ada pengaruh variabel independen terhadap variabel dependen.

Contoh:

- 1) H_0 : Durasi olahraga tidak berpengaruh terhadap Indeks Massa Tubuh (IMT).
- 2) H_1 : Durasi olahraga berpengaruh terhadap Indeks Massa Tubuh (IMT).

b. Mengumpulkan Data

Pastikan data memenuhi asumsi berikut:

- 1) Variabel dependen harus berskala interval atau rasio (misalnya, IMT).
- 2) Variabel independen juga berskala interval atau rasio (misalnya, durasi olahraga dalam menit).
- 3) Hubungan antara variabel bersifat linear.
- 4) Tidak ada autokorelasi, heteroskedastisitas, atau multikolinearitas (karena ini regresi sederhana, hanya satu variabel independen).

c. Interpretasi Hasil

- 1) Jika $p\text{-value} < 0,05$, maka variabel independen memiliki pengaruh yang signifikan terhadap variabel dependen (tolak H_0).

- 2) Jika $p\text{-value} > 0,05$, maka variabel independen tidak memiliki pengaruh yang signifikan (gagal menolak H_0).

Selain itu, nilai R^2 (koefisien determinasi) menunjukkan seberapa besar kontribusi variabel independen dalam menjelaskan variabel dependen.

Kategori R^2 :

- 1) $R^2 < 0,20 \rightarrow$ Pengaruh sangat lemah.
- 2) $0,20 \leq R^2 < 0,40 \rightarrow$ Pengaruh lemah.
- 3) $0,40 \leq R^2 < 0,60 \rightarrow$ Pengaruh sedang.
- 4) $0,60 \leq R^2 < 0,80 \rightarrow$ Pengaruh kuat.
- 5) $R^2 \geq 0,80 \rightarrow$ Pengaruh sangat kuat.

Contoh Interpretasi: Jika hasil uji menunjukkan $b = -0,25$, $p = 0,001$, dan $R^2 = 0,55$, maka dapat disimpulkan bahwa durasi olahraga memiliki pengaruh negatif dan signifikan terhadap IMT, serta 55% variasi IMT dijelaskan oleh durasi olahraga.

d. Contoh Penggunaan Software (SPSS)

Jika menggunakan SPSS, langkah-langkahnya adalah:

- 1) Input data (misalnya, kolom durasi olahraga (X) dan IMT (Y)).
- 2) Pilih Analyze \rightarrow Regression \rightarrow Linear.
- 3) Masukkan:
 - Variabel Dependen (Y): IMT
 - Variabel Independen (X): Durasi olahraga
- 4) Centang Statistics \rightarrow Pilih Confidence Intervals dan R^2 Change.
- 5) Klik OK untuk menjalankan analisis.

e. Interpretasi Hasil di SPSS

SPSS akan menghasilkan tabel dengan:

- 1) Koefisien regresi (b) \rightarrow Menunjukkan pengaruh variabel independen terhadap variabel dependen.

- 2) p-value (Sig.) → Menentukan apakah pengaruh tersebut signifikan.
- 3) R^2 (Koefisien Determinasi) → Mengukur seberapa besar variabel independen menjelaskan variabel dependen.

Jika p-value $< 0,05$, maka variabel independen berpengaruh signifikan terhadap variabel dependen.

Bab 3

Estimasi

3.1 Pengertian Estimasi

Estimasi adalah suatu proses atau metode untuk menentukan perkiraan suatu nilai berdasarkan data atau informasi yang tersedia. Dalam berbagai bidang, estimasi digunakan untuk memperkirakan suatu nilai yang belum diketahui secara pasti dengan menggunakan pendekatan matematis, statistik, atau pengalaman sebelumnya (Basuki, 2019).

Estimasi adalah proses memperkirakan atau menaksir nilai dari suatu parameter berdasarkan data yang tersedia. Dalam Kamus Besar Bahasa Indonesia (KBBI), estimasi didefinisikan sebagai perkiraan, penilaian, atau pendapat. Secara lebih spesifik, estimasi merupakan suatu pengukuran yang didasarkan pada hasil kuantitatif atau hasil yang akurasinya dapat diukur dengan angka.

Beberapa ahli memberikan definisi estimasi sebagai berikut:

1. Menurut Harinaldi, estimasi adalah keseluruhan proses yang menggunakan sebuah estimator untuk menghasilkan sebuah estimate dari suatu parameter.
2. Tockey mendefinisikan estimasi sebagai pengukuran yang dilakukan pada suatu kegiatan, di mana hasilnya merupakan sesuatu yang kuantitatif dan memiliki akurasi yang bisa diukur menggunakan angka.
3. Budihartono (2021) menyatakan bahwa estimasi sering digunakan dalam memperkirakan populasi dengan proporsi variabel tertentu. Nilai estimasi diperoleh melalui perhitungan statistik sampel, bukan menghitung langsung dalam populasi.
4. Cahyono (2018) mendefinisikan estimasi sebagai suatu keseluruhan proses yang menggunakan estimator (nilai statistik: mean, median, varian, dan standar deviasi) untuk menghasilkan suatu estimasi yang diharapkan mendekati parameter populasi.

Ciri-ciri Estimasi yang Baik

Untuk dapat menghasilkan estimasi yang berkualitas dan dapat diandalkan, terdapat beberapa ciri atau karakteristik yang perlu diperhatikan.

Estimasi yang baik memiliki ciri-ciri sebagai berikut (Nasution, 2021):

1. Tidak Bias (Unbiased)

Estimasi yang tidak bias berarti nilai estimasi yang dihasilkan cenderung mendekati atau sesuai dengan nilai parameter populasi yang sebenarnya. Dalam hal ini, jika mean dari distribusi sampel ternyata sama dengan parameter populasi, maka dapat disebut sebagai estimator tak bias. Sebaliknya, jika mean dari distribusi sampel tidak sama dengan parameter populasi, maka disebut estimator bias.

Estimasi yang tidak bias penting untuk memastikan bahwa hasil perkiraan tidak secara sistematis overestimate (terlalu tinggi) atau underestimate (terlalu rendah) dari nilai sebenarnya. Hal ini membantu dalam pengambilan keputusan yang lebih akurat dan objektif.

2. Efisien (Efficient)

Efisiensi dalam estimasi berkaitan dengan tingkat presisi atau ketepatan dari estimasi tersebut. Estimasi dapat dikatakan efisien jika sampel dari dua statistik memiliki mean yang sama, maka statistik dengan varian lebih kecil dapat disebut estimator efisien dari mean. Sementara itu, statistik yang satunya lagi disebut estimator tidak efisien.

Estimator yang efisien menghasilkan estimasi dengan tingkat kesalahan atau penyimpangan yang lebih kecil. Ini berarti estimasi tersebut lebih dapat diandalkan dan memberikan gambaran yang lebih akurat tentang parameter populasi yang diestimasi.

3. Konsisten (Consistent)

Konsistensi dalam estimasi mengacu pada kemampuan estimator untuk menghasilkan estimasi yang semakin mendekati nilai parameter sebenarnya ketika ukuran sampel bertambah. Jika besaran sampel bertambah maka nilai estimasi akan mendekati skor populasi. Estimator inilah yang disebut konsisten.

Estimator konsisten cenderung sama dengan nilai sebenarnya walaupun ukuran sampel makin besar. Ini berarti semakin banyak data yang digunakan dalam estimasi, semakin akurat hasil yang diperoleh.

4. Akurat (Accurate)

Estimasi yang akurat adalah proses estimasi yang memastikan bahwa informasi yang disampaikan selama estimasi benar dan dapat dipertanggungjawabkan. Sebuah proses estimasi yang akurat akan membantu manajer proyek dalam membuat keputusan yang tepat dan menghindari risiko ketidakpastian.

Akurasi estimasi sangat penting terutama dalam proyek-proyek besar atau keputusan bisnis yang kritis, di mana kesalahan estimasi dapat berdampak signifikan pada hasil akhir.

5. Objektif (Objective)

Estimasi yang objektif berarti proses estimasi dilakukan tanpa bias personal atau kepentingan tertentu. Estimasi harus didasarkan pada data dan metode yang dapat diverifikasi, bukan pada asumsi subjektif atau preferensi pribadi.

Objektivitas dalam estimasi membantu memastikan bahwa hasil estimasi dapat dipercaya dan diterima oleh berbagai pihak yang terlibat dalam pengambilan keputusan.

6. Dapat Diverifikasi (Verifiable)

Estimasi yang baik harus dapat diverifikasi, artinya proses dan hasil estimasi dapat diperiksa dan divalidasi oleh pihak lain. Ini melibatkan dokumentasi yang jelas tentang metode, asumsi, dan data yang digunakan dalam proses estimasi.

Kemampuan untuk memverifikasi estimasi meningkatkan kredibilitas hasil dan memungkinkan perbaikan atau penyesuaian jika diperlukan.

7. Relevan (Relevant)

Estimasi harus relevan dengan tujuan dan konteks di mana estimasi tersebut akan digunakan. Ini berarti mempertimbangkan faktor-faktor yang paling penting dan berpengaruh dalam situasi tertentu, serta mengabaikan informasi yang tidak relevan atau kurang signifikan.

Relevansi estimasi memastikan bahwa hasil yang diperoleh benar-benar berguna dan dapat diterapkan dalam pengambilan keputusan atau perencanaan.

Dengan memperhatikan ciri-ciri estimasi yang baik ini, kita dapat meningkatkan kualitas dan keandalan estimasi yang kita lakukan. Hal ini

pada gilirannya akan mendukung pengambilan keputusan yang lebih baik dan perencanaan yang lebih akurat dalam berbagai bidang.

3.2 Jenis – Jenis Estimasi

Estimasi banyak digunakan dalam berbagai bidang, seperti teknik, ekonomi, statistik, dan manajemen proyek.

Berikut adalah beberapa jenis estimasi yang umum digunakan (Ananta, 2022):

1. Estimasi Berdasarkan Metode Perhitungan

a. Estimasi Point (Titik)

Estimasi titik adalah metode yang menghasilkan satu nilai tunggal sebagai perkiraan terbaik dari parameter populasi yang tidak diketahui. Estimasi ini sering digunakan dalam statistik untuk menentukan nilai suatu parameter berdasarkan sampel.

Contoh: Jika rata-rata sampel berat badan dari 100 orang adalah 65 kg, maka estimasi titik untuk rata-rata berat badan populasi juga 65 kg.

Kelebihan: Mudah dihitung dan dipahami, Cepat dalam memberikan hasil.

Kekurangan: Tidak memberikan informasi mengenai ketidakpastian atau keakuratan estimasi.

b. Estimasi Interval

Estimasi interval adalah metode yang memberikan rentang nilai sebagai perkiraan dari parameter populasi, disertai dengan tingkat kepercayaan tertentu (misalnya 95%).

Contoh: Jika rata-rata berat badan dalam sampel adalah 65 kg dengan margin of error ± 2 kg, maka estimasi intervalnya adalah (63 kg – 67 kg) dengan tingkat kepercayaan 95%.

Kelebihan: Memberikan gambaran yang lebih akurat dibanding estimasi titik, Memungkinkan analisis ketidakpastian dalam estimasi.

Kekurangan: Lebih kompleks dalam perhitungan dibanding estimasi titik.

2. Estimasi Berdasarkan Sumber Data

a. Estimasi Parametrik

Estimasi parametrik digunakan ketika asumsi distribusi data tertentu diketahui atau diasumsikan, seperti distribusi normal, binomial, atau poisson.

Contoh: Menggunakan metode Maximum Likelihood Estimation (MLE) untuk memperkirakan parameter distribusi normal.

Kelebihan: Memberikan hasil yang lebih akurat jika asumsi distribusi benar, Digunakan dalam berbagai bidang seperti ekonomi dan statistika.

Kekurangan: Kurang akurat jika asumsi distribusi tidak sesuai dengan data yang sebenarnya.

b. Estimasi Non-Parametrik

Estimasi non-parametrik tidak memerlukan asumsi distribusi tertentu sehingga lebih fleksibel dalam menangani berbagai jenis data.

Contoh: Metode Kernel Density Estimation (KDE) dalam analisis data.

Kelebihan: Lebih fleksibel dalam analisis data yang tidak diketahui distribusinya. Tidak membutuhkan asumsi yang ketat tentang bentuk distribusi data.

Kekurangan: Memerlukan sampel yang lebih besar agar estimasi lebih akurat.

3.3 Metode Estimasi

Estimasi merupakan proses memperkirakan nilai suatu variabel berdasarkan informasi yang tersedia. Dalam berbagai bidang seperti manajemen proyek, statistik, dan ilmu data, metode estimasi digunakan untuk memperoleh hasil yang mendekati kenyataan dengan tingkat ketelitian tertentu. Amaliah, E.N., Darnah, & Sifriyani. (2020).

Berikut adalah beberapa metode estimasi yang umum digunakan:

1. Metode Delphi

Metode Delphi adalah teknik estimasi yang bergantung pada konsensus dari sekelompok ahli. Para ahli memberikan estimasi mereka secara anonim dalam beberapa putaran, dan hasilnya dikompilasi serta dianalisis. Proses ini berulang hingga konsensus tercapai, mengurangi bias individu dan meningkatkan akurasi prediksi.

2. Metode Monte Carlo

Metode Monte Carlo menggunakan simulasi berbasis probabilitas untuk memperkirakan hasil dari suatu sistem atau proses. Dengan menjalankan ribuan hingga jutaan simulasi menggunakan variabel acak, metode ini memberikan distribusi probabilitas dari kemungkinan hasil yang dapat terjadi. Metode ini sering digunakan dalam analisis risiko dan keuangan.

3. Metode Analog

Metode Analog membandingkan proyek atau kasus yang sedang dianalisis dengan proyek atau kasus serupa di masa lalu. Jika suatu proyek baru memiliki karakteristik yang mirip dengan proyek sebelumnya, maka data dari proyek tersebut dapat digunakan untuk memperkirakan biaya, waktu, atau hasil proyek baru.

4. Metode Bottom-Up

Metode Bottom-Up melakukan estimasi dengan cara menganalisis bagian-bagian kecil terlebih dahulu, kemudian menjumlahkannya untuk mendapatkan gambaran keseluruhan. Metode ini memberikan tingkat akurasi yang lebih tinggi dibandingkan metode lain karena mempertimbangkan detail-detail kecil sebelum memberikan estimasi akhir.

5. Estimasi Tiga Titik (PERT Analysis)

Metode ini menggunakan tiga estimasi (optimis, realistis, dan pesimis) untuk mendapatkan hasil yang lebih akurat.

Rumus yang digunakan adalah:

$$E = \frac{O + 4M + P}{6}$$

di mana:

- O = Estimasi optimis
- M = Estimasi realistis
- P = Estimasi pesimis

3.4 Cara Meningkatkan Akurasi Estimasi

1. Memahami Faktor yang Memengaruhi Estimasi

Sebelum meningkatkan akurasi estimasi, penting untuk memahami faktor-faktor yang memengaruhi keakuratan estimasi, seperti:

- a. Variabilitas data: Data yang tidak stabil atau memiliki variabel yang sulit diprediksi akan mempersulit estimasi.
- b. Keterbatasan informasi: Estimasi yang dilakukan dengan informasi terbatas cenderung lebih rentan terhadap kesalahan.
- c. Keterampilan estimator: Pengalaman dan keterampilan individu dalam melakukan estimasi sangat memengaruhi akurasi.
- d. Ketidaktentuan eksternal: Faktor luar seperti perubahan pasar, kondisi ekonomi, atau bencana alam bisa memengaruhi akurasi estimasi.

Dengan memahami faktor-faktor ini, kita dapat mengembangkan strategi untuk meningkatkan akurasi estimasi secara lebih sistematis.

2. Menggunakan Data Historis

Salah satu cara paling efektif untuk meningkatkan akurasi estimasi adalah dengan menggunakan data historis.

Langkah-langkah yang dapat dilakukan antara lain:

- a. Mengumpulkan data historis yang relevan: Pastikan data yang digunakan sesuai dengan konteks estimasi yang dilakukan.
- b. Menganalisis tren dan pola: Identifikasi pola historis untuk mendapatkan gambaran mengenai kemungkinan hasil di masa depan.

- c. Menggunakan teknik statistik: Regresi linier, analisis varians, dan metode statistik lainnya dapat membantu meningkatkan akurasi estimasi.
 - d. Memperhitungkan outlier: Nilai ekstrem dalam data historis bisa memengaruhi estimasi. Gunakan teknik seperti trimming atau winsorizing untuk mengatasi outlier.
3. Menerapkan Teknik Estimasi yang Tepat
Beberapa teknik estimasi dapat digunakan untuk meningkatkan akurasi, tergantung pada jenis estimasi yang dibutuhkan. Beberapa teknik yang umum digunakan meliputi.
 4. Menerapkan Teknik Pembobotan
Teknik pembobotan dapat meningkatkan akurasi dengan memberikan bobot lebih tinggi pada faktor-faktor yang lebih signifikan dalam estimasi.
Beberapa metode yang dapat digunakan adalah:
 - a. Metode AHP (Analytical Hierarchy Process) untuk menilai faktor-faktor yang berpengaruh dalam estimasi.
 - b. Teknik Delphi, di mana sekelompok ahli memberikan estimasi independen yang kemudian dibandingkan dan disesuaikan untuk meningkatkan akurasi.
 5. Menggunakan Teknologi dan Kecerdasan Buatan
Teknologi dan kecerdasan buatan (AI) dapat membantu meningkatkan akurasi estimasi dengan memproses data dalam jumlah besar dan menemukan pola yang mungkin tidak terlihat oleh manusia.
Beberapa aplikasi AI dalam estimasi meliputi:
 - a. Machine Learning untuk mengidentifikasi pola dalam data historis.

- b. Neural Networks untuk melakukan estimasi non-linear yang kompleks.
 - c. Algoritma Prediktif untuk mengolah data besar dan menghasilkan estimasi yang lebih akurat.
6. Menganalisis dan Mengelola Risiko

Salah satu penyebab utama estimasi yang tidak akurat adalah risiko yang tidak terduga.

Untuk meningkatkan akurasi, perlu dilakukan analisis risiko dengan langkah-langkah berikut:

 - a. Mengidentifikasi risiko potensial yang dapat memengaruhi hasil estimasi.
 - b. Menganalisis dampak risiko terhadap estimasi yang dilakukan.
 - c. Menggunakan teknik mitigasi risiko, seperti perencanaan kontinjensi dan skenario alternatif.
7. Melakukan Kalibrasi dan Evaluasi Estimasi Secara Berkala

Akurasi estimasi dapat ditingkatkan dengan melakukan evaluasi berkala terhadap estimasi yang telah dibuat.

Langkah-langkah yang dapat diambil meliputi:

 - a. Membandingkan hasil estimasi dengan realisasi untuk mengukur tingkat akurasi.
 - b. Melakukan kalibrasi metode estimasi berdasarkan perbedaan antara estimasi dan realisasi.
 - c. Menggunakan teknik backtesting untuk menguji apakah metode estimasi yang digunakan masih relevan dengan kondisi terkini.
8. Menggunakan Pendekatan Kolaboratif

Kolaborasi dengan berbagai pihak dapat meningkatkan akurasi estimasi dengan mengurangi bias individu.

Beberapa strategi kolaboratif yang bisa diterapkan adalah:

- a. Diskusi lintas tim untuk mendapatkan berbagai perspektif dalam estimasi.
- b. Konsultasi dengan ahli yang memiliki pengalaman dalam bidang terkait.
- c. Menggunakan metode crowdsourcing, di mana estimasi dilakukan berdasarkan kontribusi banyak individu untuk mendapatkan hasil yang lebih akurat.

3.5 Estimasi Probabilitas

Estimasi probabilitas adalah proses untuk menentukan kemungkinan suatu peristiwa terjadi berdasarkan data yang tersedia. Dalam statistik dan pembelajaran mesin, estimasi probabilitas digunakan untuk membuat prediksi atau mengambil keputusan dalam situasi yang melibatkan ketidakpastian.

1. Metode Estimasi Probabilitas

a. Estimasi Probabilitas Empiris (Frekuentis)

Menghitung probabilitas suatu peristiwa berdasarkan frekuensi relatif dalam data historis.

b. Estimasi Berdasarkan Model Probabilitas (Bayesian)

1) Menggunakan Teorema Bayes untuk memperbarui probabilitas berdasarkan informasi tambahan.

2) Contoh: Jika kita tahu bahwa suatu penyakit jarang terjadi, tetapi ada tes medis dengan tingkat akurasi tertentu, kita dapat memperkirakan probabilitas seseorang benar-benar sakit setelah mendapatkan hasil tes positif.

-
- c. Estimasi Dengan Model Distribusi Statistik
 - 1) Mengasumsikan data mengikuti distribusi tertentu, seperti distribusi normal, binomial, atau Poisson.
 - 2) Parameter distribusi (misalnya mean dan varians) diestimasi dari data menggunakan teknik seperti Maximum Likelihood Estimation (MLE) atau Bayesian Estimation.
 - 2. Estimasi Probabilitas Dalam Pembelajaran Mesin
 - a. Digunakan dalam algoritma seperti Naïve Bayes, Logistic Regression, dan Neural Networks untuk memprediksi probabilitas suatu kelas atau hasil.
 - b. Contoh: Model klasifikasi spam dapat memperkirakan probabilitas sebuah email termasuk kategori spam berdasarkan kata-kata yang terkandung dalam email tersebut.
 - 3. Aplikasi Estimasi Probabilitas
 - a. Prediksi Cuaca: Menentukan kemungkinan hujan berdasarkan data historis.
 - b. Diagnosa Medis: Memperkirakan probabilitas pasien mengidap penyakit tertentu berdasarkan gejala.
 - c. Keuangan: Menghitung kemungkinan keuntungan atau risiko dalam investasi.
 - d. Sistem Rekomendasi: Memperkirakan kemungkinan pengguna akan menyukai suatu produk berdasarkan riwayat pembelian.

Bab 4

Analisis Hipotesis

4.1 Pendahuluan

4.1.1 Peran Biostatistik dalam Keperawatan

Biostatistik bukan hanya alat untuk analisis data, tetapi merupakan unsur penting dalam seluruh proses penelitian yang mendukung pengembangan praktik keperawatan yang berbasis bukti. Melalui pemahaman dan penerapan statistik yang tepat, para profesional keperawatan dapat memastikan bahwa penelitian mereka secara efektif berkontribusi pada peningkatan praktik klinis dan hasil kesehatan pasien.

Peran biostatistik dalam keperawatan yaitu:

1. Membantu dalam pengambilan keputusan berdasarkan bukti, statistik menyediakan alat yang diperlukan untuk menganalisis data secara objektif sehingga memungkinkan peneliti dan praktisi keperawatan untuk membuat keputusan berdasarkan bukti

2. Desain penelitian yang valid, statistik membantu dalam merancang studi penelitian yang valid dan dapat diandalkan. Melalui pemahaman konsep statistik, peneliti dapat memilih desain penelitian yang tepat dan metode analisis yang sesuai, serta menentukan ukuran sampel yang cukup untuk mendapatkan hasil yang akurat
3. Analisis data yang akurat, dengan menggunakan teknik statistik peneliti dapat menganalisis data yang dikumpulkan untuk mengidentifikasi pola, hubungan, atau perbedaan antara variabel. Dengan menggunakan teknik statistik, peneliti dapat menganalisis data yang dikumpulkan untuk mengidentifikasi pola, hubungan, atau perbedaan antara variabel. Hal ini penting dalam penelitian keperawatan, di mana hasil analisis dapat memengaruhi pemahaman tentang faktor-faktor yang memengaruhi kesehatan pasien
4. Menginterpretasikan hasil penelitian, statistik memberi peneliti cara untuk mengevaluasi kekuatan dan keakuratan hasil penelitian melalui pengujian hipotesis dan interval kepercayaan. Ini membantu dalam menentukan seberapa signifikan hasil studi, baik secara statistik maupun klinis
5. Meminimalkan bias dalam penelitian, statistik dapat digunakan untuk mengidentifikasi dan meminimalkan bias dalam penelitian, yang dapat membawa dampak besar pada hasil. Ini termasuk pemilihan metode sampling yang tepat dan penggunaan teknik analisis yang akurat
6. Membantu dalam pengembangan skala dan instrumen, statistik digunakan untuk mengembangkan dan mengevaluasi alat pengukuran, seperti kuesioner atau skala. Melalui analisis statistik,

peneliti dapat menentukan validitas dan reliabilitas instrumen yang digunakan dalam penelitian

7. Peningkatan kualitas layanan kesehatan, dengan menerapkan metode statistik dalam penelitian, keperawatan dapat terus meningkatkan kualitas layanan kesehatan melalui evaluasi program perawatan, analisis hasil klinis, dan pengembangan intervensi berbasis bukti (polit, 2014).

4.1.2 Pengertian dan Fungsi Hipotesis

Hipotesis adalah pernyataan atau dugaan yang diuji untuk mendapatkan bukti melalui data (Polit, 2014). Dalam penelitian keperawatan, hipotesis sering dibagi menjadi dua jenis: hipotesis nol (H_0) dan hipotesis alternatif (H_1). Hipotesis nol menyatakan tidak ada hubungan atau perbedaan antara variabel yang diteliti, sedangkan hipotesis alternatif menunjukkan adanya hubungan atau perbedaan.

Hipotesis Nol (H_0) adalah pernyataan yang menyatakan bahwa tidak ada pengaruh atau hubungan yang signifikan antara variabel yang diteliti. Dalam penelitian keperawatan, misalnya, sebuah studi mungkin berhipotesis bahwa kolaborasi antara perawat dan dokter tidak memiliki dampak signifikan terhadap hasil pasien.

Namun, penelitian yang dilakukan oleh Hattab et al., (2023) menunjukkan bahwa kolaborasi perawat-dokter ternyata memiliki dampak yang signifikan terhadap otonomi perawat. Dengan kata lain, meskipun hipotesis nol menyatakan tidak ada hubungan, bukti penelitian ini justru menunjukkan sebaliknya.

Selain itu, penelitian oleh Peršolja, (2022) menemukan adanya korelasi positif antara jumlah staf perawat dan tingkat kepuasan pasien. Hal ini menunjukkan bahwa dalam banyak kasus, hipotesis nol yang menyatakan tidak ada hubungan antara jumlah staf perawat dan kepuasan pasien sering kali tidak berlaku.

Hipotesis Alternatif (H_1) adalah pernyataan yang menyatakan bahwa terdapat pengaruh atau hubungan yang signifikan antara variabel yang diteliti. Sebagai contoh, penelitian yang dilakukan oleh (Haugan, 2021) membahas manfaat interaksi antara pasien dan perawat. Studi ini menemukan bahwa komunikasi positif antara perawat dan pasien dapat meningkatkan hasil kesehatan, yang mendukung hipotesis alternatif bahwa interaksi semacam itu memiliki dampak yang bermanfaat.

Selanjutnya, penelitian oleh (Parashakti et al., 2023) menyelidiki pengaruh pembelajaran organisasi terhadap kinerja perawat. Hasil penelitian ini menunjukkan adanya korelasi yang signifikan antara pembelajaran organisasi dan peningkatan kinerja perawat, yang dipengaruhi oleh faktor perilaku manajemen.

Fungsi hipotesis dalam penelitian sangat penting sebagai dasar dalam penyelidikan ilmiah, membantu merumuskan pertanyaan penelitian serta mengarahkan jalannya analisis. Dalam bidang keperawatan dan layanan kesehatan, hipotesis berperan sebagai prediksi awal atau dugaan mengenai hasil pasien, efektivitas intervensi, atau berbagai masalah kesehatan yang umum terjadi. Dengan adanya hipotesis, para praktisi dan peneliti dapat secara sistematis mengeksplorasi hubungan antar variabel.

Sebagai contoh, dalam metode Pembelajaran Berbasis Masalah (PBL), mahasiswa keperawatan sering kali menyusun diagnosis diferensial berdasarkan hipotesis yang dikembangkan dari data pasien. Mereka kemudian menyusun hipotesis tersebut berdasarkan tingkat kemungkinan terjadinya. Proses ini menunjukkan bagaimana pengujian hipotesis berperan dalam meningkatkan penalaran diagnostik di kalangan mahasiswa keperawatan (Brennan, 2020).

4.2 Jenis Hipotesis

4.2.1 Hipotesis Deskriptif

Hipotesis deskriptif bertujuan untuk menggambarkan atau merangkum pengamatan terhadap suatu sampel tanpa menyimpulkan adanya hubungan antarvariabel. Hipotesis ini memberikan gambaran dasar mengenai karakteristik dalam kumpulan data dan sering kali menjadi langkah awal untuk analisis lebih lanjut (Vlaški, 2020).

Sebagai contoh, hipotesis deskriptif dapat berbunyi: “Tekanan darah rata-rata pasien dalam kelompok demografi tertentu lebih tinggi dibandingkan dengan rata-rata nasional.” Jenis hipotesis ini sangat penting dalam menetapkan data awal serta membantu peneliti mengidentifikasi pola yang berpotensi untuk diteliti lebih lanjut (Hoskins, 2020).

4.2.2 Hipotesis Kausal (Asosiasi)

Hipotesis kausal menyatakan bahwa perubahan pada satu variabel (variabel independen) secara langsung menyebabkan perubahan pada variabel lainnya (variabel dependen). Jenis hipotesis ini sangat penting dalam penelitian yang bertujuan membangun hubungan sebab-akibat. Sebagai contoh, hipotesis kausal dapat dinyatakan sebagai: “Peningkatan aktivitas fisik di kalangan remaja akan menyebabkan penurunan tingkat obesitas.” Untuk membuktikan hubungan ini, diperlukan pengujian ketat melalui eksperimen atau studi longitudinal (Anh, 2024).

Penyelidikan hipotesis kausal sering kali menggunakan teknik analisis statistik, seperti analisis regresi atau pemodelan persamaan struktural, guna menentukan keberadaan dan kekuatan hubungan sebab-akibat antara variabel yang diteliti (Guerard et al., 2020).

4.2.3 Hipotesis Komparatif

Hipotesis komparatif digunakan ketika peneliti ingin menentukan apakah terdapat perbedaan antara dua atau lebih kelompok dalam suatu penelitian.

Sebagai contoh, hipotesis komparatif dapat dinyatakan sebagai: “Terdapat perbedaan yang signifikan dalam tingkat kecemasan antara siswa yang belajar di lingkungan dengan tingkat stres tinggi dan mereka yang belajar di lingkungan dengan tingkat stres rendah.”

Hipotesis ini memerlukan analisis statistik untuk mengevaluasi perbedaan yang ada, sehingga dapat memberikan wawasan penting mengenai efektivitas suatu kondisi atau perlakuan tertentu (Freiholtz et al., 2024). Jenis hipotesis ini sangat relevan dalam penelitian klinis, terutama dalam membandingkan kelompok perlakuan dan kelompok kontrol untuk menilai efektivitas suatu intervensi (Mertens & Recker, 2023).

4.2.4 Hipotesis Korelasional

Hipotesis korelasional menyatakan bahwa terdapat hubungan antara dua atau lebih variabel, yang menunjukkan sejauh mana variabel-variabel tersebut saling berkaitan. Hipotesis ini penting untuk memahami bagaimana perubahan pada satu variabel dapat berhubungan dengan perubahan pada variabel lainnya, tanpa menyiratkan adanya hubungan sebab-akibat. Terdapat tiga jenis hubungan dalam hipotesis korelasional, yaitu korelasi positif (kedua variabel bergerak searah), korelasi negatif (satu variabel meningkat, sementara yang lain menurun), dan korelasi nol (tidak ada hubungan antara variabel-variabel tersebut).

Hipotesis korelasi positif menyatakan bahwa ketika satu variabel meningkat, variabel lainnya juga mengalami peningkatan. Sebagai contoh, hasil penelitian menunjukkan adanya hubungan positif antara faktor-faktor keuangan mikro dan keberhasilan usaha kecil dan menengah (UKM) di Zimbabwe (Dzingirai & Baporikar, 2022).

Sebaliknya, hipotesis korelasi negatif menyatakan bahwa peningkatan pada satu variabel akan menyebabkan penurunan pada variabel lainnya. Sebagai contoh, penelitian yang bertujuan mengeksplorasi kelelahan dalam penggunaan aplikasi mHealth di kalangan orang dewasa yang lebih tua. Studi ini menemukan bahwa semakin banyak faktor yang berkontribusi

terhadap kelelahan pengguna, semakin tinggi perilaku resistensi terhadap teknologi (Engelsma et al., 2024).

Hipotesis korelasi nol menyatakan tidak ada hubungan antara variabel-variabel yang dimaksud. Sebagai contoh, penelitian Kappers menyoroti penerapan analisis korelasi dalam mengevaluasi berbagai jenis simulasi berbasis pemodelan, mengidentifikasi contoh-contoh di mana kondisi input tidak sama dengan hasil yang diharapkan, yang secara efektif mengindikasikan skenario korelasi nol (Kappers, 2020).

4.3 Langkah-langkah Uji Hipotesis

4.3.1 Menentukan Variabel

Langkah pertama dalam melakukan uji hipotesis adalah mengidentifikasi variabel yang akan diukur, yang umumnya dikategorikan sebagai variabel independen dan variabel dependen. Penting untuk memahami hubungan antar variabel dalam penelitian keperawatan, terutama dalam memilih metode analisis yang tepat (Bennett et al., 2022).

Setiap penelitian harus secara spesifik menjelaskan faktor-faktor yang memengaruhi variabel dependen (misalnya, kinerja perawat) serta menetapkan dasar teori yang kuat untuk hubungan tersebut. Hal ini sesuai dengan prinsip utama dalam menyusun kerangka kerja hipotesis yang valid dan sistematis. Contoh hubungan kompetensi spiritual dengan implementasi pelayanan keperawatan spiritual di rumah sakit, yang menjadi variabel dependen implementasi pelayanan keperawatan spiritual dan variabel independen kompetensi spiritual.

4.3.2 Menentukan Jenis Hipotesis

Hipotesis Deskriptif, menyatakan karakteristik dari satu variabel atau lebih tanpa membandingkan atau mencari hubungan. Contohnya: "Tingkat

kepuasan pasien di unit rawat inap lebih tinggi dibandingkan di unit gawat darurat."

Hipotesis Korelasional, menjelaskan hubungan antara dua atau lebih variabel tanpa menyiratkan hubungan sebab-akibat. Contohnya: "Ada hubungan positif antara tingkat stres perawat dan jumlah jam kerja per minggu."

Hipotesis Komparatif, digunakan untuk membandingkan dua atau lebih kelompok dalam penelitian. Contohnya: "Perawat yang bekerja di shift malam memiliki tingkat kelelahan yang lebih tinggi dibandingkan perawat yang bekerja di shift pagi."

Hipotesis Kausal (Asosiasi), menyatakan bahwa satu variabel menyebabkan perubahan pada variabel lain. Contohnya: "Peningkatan pelatihan komunikasi bagi perawat akan meningkatkan kepuasan pasien."

4.3.3 Menentukan Skala Pengukuran

Skala pengukuran merupakan salah satu aspek penting dalam penelitian. Skala yang umum digunakan meliputi nominal, ordinal, interval, dan rasio, di mana masing-masing menentukan cara variabel dianalisis dan diinterpretasikan. Pemilihan metode pengukuran yang tepat sangat penting untuk memastikan keakuratan hasil penelitian (Khuneswari & Hsieh, 2022).

1. Skala Pengukuran pada Hipotesis Komparatif
 - a. Skala Kategorik : Kategorik $\times\times$ Kategorik
 - b. Skala Numerik : Kategorik $\times\times$ Numerik atau Numerik $\times\times$ Numerik
2. Skala Pengukuran pada Hipotesis Korelatif
 - a. Skala Kategorik : Kategorik $\times\times$ kategorik atau Kategorik $\times\times$ Numerik
 - b. Skala Numerik : Numerik $\times\times$ numerik

4.3.4 Menentukan Berpasangan atau Tidak Berpasangan

Dalam penelitian keperawatan, membedakan antara data berpasangan dan tidak berpasangan sangat penting untuk memilih metode analisis yang tepat dan menginterpretasikan hasil secara akurat (Lee et al., 2024).

Data berpasangan mengacu pada pengukuran variabel yang sama pada subjek yang sama dalam beberapa waktu atau kondisi berbeda. Desain ini memperhitungkan variasi dalam kelompok yang sama. Sebagai contoh, penelitian yang mengevaluasi hasil sebelum dan sesudah intervensi pada individu yang sama merupakan contoh desain berpasangan yang khas. Metodologi ini sering diperkuat dengan teknik statistik yang mempertimbangkan korelasi antar pengukuran, sehingga meningkatkan kekuatan analisis dan memberikan wawasan lebih mendalam tentang efek suatu intervensi dari waktu ke waktu (Mbombi et al., 2023).

Sebaliknya, data tidak berpasangan merujuk pada pengamatan yang dikumpulkan dari subjek yang berbeda atau variabel yang berbeda, sehingga memungkinkan perbandingan antar kelompok yang terpisah. Format ini sangat penting ketika randomisasi memungkinkan, dan perbandingan dilakukan pada kelompok yang tidak memiliki keterkaitan langsung.

Sebagai contoh, penelitian yang menganalisis perbedaan kinerja atau tingkat kompetensi antara kelompok perawat yang menjalani program pelatihan yang berbeda memerlukan pengumpulan data tidak berpasangan (Hovenga, 2022). Teknik analisis seperti uji t independen atau ANOVA sering digunakan untuk mengevaluasi data semacam ini, yang membantu memahami perbedaan antar kelompok yang dibandingkan (Soilemezi et al., 2021).

Pendekatan data tidak berpasangan sangat berguna dalam penelitian yang berfokus pada perbandingan respons atau hasil antara populasi atau kondisi yang berbeda, sehingga memungkinkan generalisasi temuan ke dalam praktik klinis yang lebih luas (Menezes & Brasileiro, 2024).

4.3.5 Menentukan Jumlah Kelompok

Menentukan jumlah kelompok untuk analisis data, terutama dalam bidang keperawatan, memerlukan pemahaman mendalam tentang berbagai metode statistik. Pemilihan antara menganalisis dua atau lebih dari dua kelompok sangat bergantung pada desain penelitian, jenis data, dan hipotesis yang diuji. Kelompok data dapat dibagi menjadi 2 kelompok atau >2 kelompok.

Kelompok Data 2 Kelompok, contoh: sebuah penelitian membandingkan dua kelompok pasien yang menerima perawatan yang berbeda, misalnya kelompok intervensi (yang mendapatkan terapi khusus) dan kelompok kontrol (yang menerima perawatan standar) (Grove & Ciper, 2024).

Kelompok Data Lebih dari 2 Kelompok, contoh: penelitian yang membandingkan tiga kelompok pasien yang masing-masing menerima dosis tinggi, dosis rendah, atau plasebo untuk melihat perbedaan dalam tingkat kesembuhan (Grove & Ciper, 2024).

4.3.6 Menentukan Uji Statistik

Uji statistik ditentukan dengan melihat langkah-langkah sebelumnya dapat dilihat pada tabel uji statistik di bawah ini:

Tabel 4.1: Uji Statistik

SKALA PENGUKURAN VARIABEL	KOMPARATIF/ ASOSIASI				KORELASI
	TIDAK BERPASANGAN		BERPASANGAN		
	2 KELOMPOK	>2 KELOMPOK	2 KELOMPOK	>2 KELOMPOK	
INTERVAL/ RASIO	Uji t tidak berpasangan	One way anova	Uji t berpasangan	Repeated anova	Pearson
ORDINAL	Mann-Whitney	Kruskall Wallis	Wilcoxon	Friedman	Spearman
NOMINAL	Ragam Uji Chi-Square		Mc Nemar	Chochran	Koefisien kontingensial

4.4 Aplikasi dalam Penelitian Keperawatan

Contoh aplikasi langkah-langkah uji hipotesis dalam penelitian keperawatan sebagai berikut.

Apakah ada perbedaan pengetahuan perawat yang diberikan pelatihan modul intervensi keperawatan dengan intervensi biasa dalam mencegah preeklampsia?.

Tabel 4.2: Aplikasi dalam Penelitian Keperawatan

No	Langkah-langkah	Jawaban
1.	Variabel	Pengetahuan x Kelompok
2.	Menentukan jenis hipotesis	Komparatif
3.	Menentukan masalah skala pengukuran	Numerik x Numerik =Numerik
4.	Menentukan berpasangan/tidak berpasangan	Tidak berpasangan
5	Menentukan jumlah kelompok	2 kelompok
Uji statistik		Jika data berdistribusi normal maka uji statistik yang digunakan uji independent t test, jika data tidak berdistribusi normal maka menggunakan uji alternatif Mann-Whitney.

Apakah ada perbedaan kadar gula darah pada kelompok ekonomi tinggi, sedang, dan rendah?

Tabel 4.3: Perbedaan kadar gula darah

No	Langkah-langkah	Jawaban
1.	Variabel	Kadar gula darah x Status ekonomi
2.	Menentukan jenis hipotesis	Komparatif
3.	Menentukan masalah skala pengukuran	Numerik x Kategorik =Numerik
4.	Menentukan berpasangan/tidak berpasangan	Tidak berpasangan

5	Menentukan jumlah kelompok	>2 kelompok (3 kelompok)
Uji statistik		Jika data berdistribusi normal maka uji statistik yang digunakan uji one way anova, jika data tidak berdistribusi normal maka menggunakan uji alternatif Kruskal wallis.

Tabel 4.4: Apakah ada hubungan beban kerja dengan stress kerja perawat?

No	Langkah-langkah	Jawaban
1.	Variabel	Beban kerja x Stress kerja
2.	Menentukan jenis hipotesis	Komparatif
3.	Menentukan masalah skala pengukuran	Kategorik x Kategorik =Kategorik
4.	Menentukan berpasangan/tidak berpasangan	Tidak berpasangan
5	Menentukan jumlah kelompok	2 kelompok ($B \times K = 2 \times 2$)
Uji statistik		Jika tidak expected count < 5 maka menggunakan continuity correction, jika ada expected count < 5 maka menggunakan uji fisher

Bab 5

Uji T-Independent dan Dependent

5.1 Pendahuluan

Uji t merupakan salah satu uji statistik parametrik yang digunakan untuk membandingkan dua kelompok data. Uji ini sering digunakan dalam penelitian untuk menguji hipotesis terkait perbedaan rata-rata antara dua kelompok. Dalam analisis statistik, uji t merupakan salah satu metode yang digunakan untuk membandingkan rata-rata dua kelompok data. Uji ini sering diterapkan dalam berbagai bidang penelitian, seperti ilmu kesehatan, sosial, dan ekonomi, untuk menguji hipotesis mengenai perbedaan antara dua kelompok. Uji t termasuk dalam kategori uji parametrik, yang berarti penggunaannya mengharuskan data memenuhi beberapa asumsi tertentu (Gravetter, F. J., & Wallnau, 2017).

Uji t digunakan ketika kita ingin mengetahui apakah terdapat perbedaan yang signifikan antara dua kelompok berdasarkan variabel kuantitatif

tertentu. Misalnya, seorang peneliti ingin mengetahui apakah ada perbedaan signifikan dalam hasil ujian antara siswa yang belajar menggunakan metode konvensional dan siswa yang belajar dengan metode interaktif (Pallant, 2020). Untuk menguji perbedaan tersebut, peneliti dapat menggunakan uji t.

Terdapat dua jenis utama uji t yang umum digunakan:

1. Uji t independen (Independent t-test): digunakan untuk membandingkan rata-rata dua kelompok yang tidak berhubungan satu sama lain, seperti perbedaan berat badan antara pria dan wanita.
2. Uji t dependen (Dependent t-test atau Paired t-test): digunakan untuk membandingkan rata-rata dua kelompok yang berhubungan atau berasal dari subjek yang sama dalam dua kondisi yang berbeda, seperti sebelum dan sesudah intervensi pada individu yang sama.

Pemahaman mengenai kapan dan bagaimana menggunakan uji t sangat penting bagi para peneliti untuk memperoleh hasil yang valid dan dapat diandalkan dalam analisis data mereka (Rosner, 2015). Selain itu, memastikan bahwa asumsi uji t terpenuhi merupakan langkah penting dalam proses analisis untuk menghindari kesalahan interpretasi hasil penelitian.

5.2 Uji t Independen

5.2.1 Pengertian

Uji t independen digunakan untuk menguji perbedaan rata-rata antara dua kelompok yang berbeda secara independen (Trochim, W. M. K., & Donnelly, 2021). Contohnya, jika kita ingin mengetahui apakah ada

perbedaan signifikan antara rata-rata skor ujian siswa laki-laki dan perempuan.

5.2.2 Asumsi Uji t Independen

Sebelum melakukan uji t independen, terdapat beberapa asumsi yang harus dipenuhi:

1. Data berdistribusi normal: Setiap kelompok data harus berasal dari populasi yang terdistribusi normal.
2. Homogenitas varians: Kedua kelompok harus memiliki varians yang sama (dapat diuji dengan uji Levene).
3. Data berskala interval atau rasio: Data yang digunakan harus berupa data kuantitatif dengan skala interval atau rasio.
4. Observasi independen: Data dari kedua kelompok tidak saling berhubungan atau berasal dari individu yang berbeda.

5.2.3 Rumus Uji t Independen

Rumus dasar untuk uji t independen adalah:

$$t = \frac{(\bar{X}_1 - \bar{X}_2)}{\sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}}}$$

Di mana:

- \bar{x}_1 = rata-rata kelompok pertama
- \bar{x}_2 = rata-rata kelompok kedua
- s_1^2 = varians kelompok pertama
- s_2^2 = varians kelompok kedua
- Π_1 = jumlah sampel kelompok pertama
- Π_2 = jumlah sampel kelompok kedua

Jika asumsi homogenitas varians tidak terpenuhi, maka digunakan rumus Welch's t-test yang lebih robust terhadap ketidaksamaan varians.

5.2.4 Interpretasi Hasil

Setelah nilai t diperoleh, langkah berikutnya adalah membandingkan nilai p dengan tingkat signifikansi:

1. Jika, maka ada perbedaan signifikan antara kedua kelompok.
2. Jika, maka tidak ada perbedaan signifikan antara kedua kelompok.

5.2.5 Contoh Kasus

Seorang peneliti ingin mengetahui apakah ada perbedaan signifikan dalam skor ujian antara siswa laki-laki dan perempuan. Data yang dikumpulkan dianalisis dengan uji t independen, dan diperoleh nilai p sebesar 0,03. Karena $p < 0,05$, maka dapat disimpulkan bahwa terdapat perbedaan signifikan antara kedua kelompok.

5.3 Uji t Dependen

5.3.1 Pengertian

Uji t dependen digunakan untuk membandingkan dua rata-rata dari sampel yang sama dalam dua kondisi yang berbeda. Contohnya, jika seorang peneliti ingin mengetahui efektivitas terapi rehabilitasi dengan membandingkan skor tes sebelum dan sesudah terapi pada pasien yang sama (Howell, 2016).

5.3.2 Asumsi Uji t Dependen

Pada asumsi Uji t Dependen terdapat 3 asumsi diantaranya adalah:

1. Data berdistribusi normal: Perbedaan antara dua pengukuran harus terdistribusi normal.

2. Sampel berasal dari populasi yang sama: Data berasal dari individu yang sama yang diukur pada dua waktu atau dua kondisi berbeda.
3. Data berskala interval atau rasio: Data yang digunakan harus berupa data kuantitatif.

5.3.3 Rumus Uji t Dependen

Rumus dasar untuk uji t dependen adalah:

$$t = \frac{D}{sD / \sqrt{n}}$$

Di mana:

- D= rata-rata selisih antara pasangan data
- sD= standar deviasi dari selisih pasangan data
- n= jumlah pasangan data

5.3.4 Interpretasi Hasil

Sama seperti uji t independen, hasil uji t dependen dianalisis berdasarkan nilai p:

1. Jika $p < \alpha$, maka ada perbedaan signifikan antara dua kondisi.
2. Jika $p \geq \alpha$, maka tidak ada perbedaan signifikan antara dua kondisi.
3. Nilai p (p-value): Jika $p < 0.05$, maka hasilnya signifikan dan hipotesis nol ditolak.
4. Confidence Interval (CI): Interval kepercayaan membantu memahami kisaran kemungkinan perbedaan rata-rata.
5. Effect Size: Dapat dihitung dengan Cohen's d untuk mengetahui seberapa besar efek dari perbedaan yang ditemukan.

5.3.5 Alternatif Jika Asumsi Tidak Terpenuhi

Jika data tidak memenuhi asumsi uji t, alternatif yang bisa digunakan:

1. Uji Mann-Whitney U (untuk independent t-test dengan data non-normal).
2. Uji Wilcoxon signed-rank (untuk dependent t-test dengan data non-normal).
3. Transformasi data untuk mendekati distribusi normal.

5.3.6 Contoh Kasus

Seorang peneliti ingin mengetahui apakah program latihan fisik meningkatkan kebugaran. Sebelum dan sesudah program, kebugaran 30 peserta diukur. Hasil uji t dependen menunjukkan nilai $p = 0,01$, yang berarti ada peningkatan signifikan dalam kebugaran setelah program latihan.

5.4 Keunggulan dan Kelemahan Uji t

Terdapat beberapa keunggulan dan kelemahan dalam menggunakan Uji t, diantaranya adalah:

1. Mudah digunakan dan dipahami
Uji t adalah salah satu uji statistik yang paling sederhana untuk membandingkan dua kelompok. Dengan asumsi yang jelas dan prosedur yang sistematis, uji ini mudah dipahami oleh peneliti dari berbagai disiplin ilmu.
2. Cocok untuk ukuran sampel kecil hingga sedang
Uji t dapat digunakan dengan ukuran sampel yang relatif kecil (biasanya di bawah 30 sampel per kelompok) selama asumsi distribusi normal terpenuhi. Hal ini menjadikannya alat analisis yang berguna ketika jumlah data terbatas.

3. Banyak diterapkan dalam berbagai bidang penelitian
Uji t digunakan secara luas dalam bidang kesehatan, pendidikan, psikologi, bisnis, dan berbagai bidang lain untuk menguji perbedaan antara dua kelompok.
4. Dapat digunakan untuk data berpasangan atau independen
Dengan adanya dua jenis utama uji t (independen dan dependen), metode ini fleksibel untuk digunakan dalam berbagai desain penelitian, baik untuk membandingkan dua kelompok berbeda maupun dua kondisi dari kelompok yang sama.
5. Membantu dalam pengambilan keputusan statistic
Dengan membandingkan nilai p terhadap tingkat signifikansi (biasanya 0,05), uji t membantu peneliti dalam menentukan apakah ada perbedaan yang signifikan secara statistik antara dua kelompok (McDonald, 2014).

Kelemahan:

1. Hanya bisa digunakan untuk membandingkan dua kelompok
Salah satu keterbatasan utama uji t adalah hanya dapat digunakan untuk membandingkan dua kelompok. Jika ada lebih dari dua kelompok yang ingin dibandingkan, peneliti perlu menggunakan metode lain, seperti ANOVA (Analisis Varians).
2. Sensitif terhadap asumsi distribusi normal
Uji t mengasumsikan bahwa data dalam setiap kelompok berasal dari populasi yang terdistribusi normal. Jika data tidak normal dan jumlah sampel kecil, hasil uji t bisa menjadi tidak valid (Tabachnick, B. G., & Fidell, 2019).
3. Mengharuskan homogenitas varians untuk uji t independen
Dalam uji t independen, varians kedua kelompok harus sama (homogen). Jika asumsi ini tidak terpenuhi, hasil uji bisa bias. Jika

varians berbeda, maka harus menggunakan alternatif seperti Welch's t-test.

4. Kurang cocok untuk data ordinal atau non-parametrik

Uji t hanya dapat digunakan untuk data berskala interval atau rasio. Jika data yang digunakan adalah skala ordinal (seperti tingkat kepuasan pelanggan), maka lebih baik menggunakan uji non-parametrik seperti Mann-Whitney U test atau Wilcoxon signed-rank test (Keppel, G., & Wickens, 2004).

5. Kurang informatif tanpa perhitungan effect size

Walaupun uji t dapat menunjukkan apakah ada perbedaan signifikan, uji ini tidak menunjukkan seberapa besar perbedaan tersebut secara praktis. Oleh karena itu, perlu dihitung effect size (seperti Cohen's d) untuk memahami signifikansi praktis dari hasil penelitian.

6. Tidak mempertimbangkan faktor lain yang dapat memengaruhi hasil

Uji t hanya membandingkan dua kelompok berdasarkan satu variabel. Jika ada faktor lain yang juga memengaruhi hasil (misalnya usia, jenis kelamin, atau tingkat pendidikan), analisis yang lebih kompleks seperti regresi atau ANOVA mungkin lebih sesuai (McDonald, 2014).

5.5 Aplikasi Uji t dalam Berbagai Bidang

Dalam dunia medis dan kesehatan, uji t digunakan untuk mengevaluasi efektivitas pengobatan, membandingkan hasil intervensi medis, serta meneliti perbedaan antar kelompok pasien.

5.5.1 Uji t Independen dalam Kesehatan

Digunakan ketika membandingkan dua kelompok pasien yang tidak memiliki hubungan satu sama lain.

Contoh Aplikasi:

1. Efektivitas obat baru dibandingkan dengan plasebo.
2. Sebuah penelitian membandingkan apakah obat baru dapat menurunkan tekanan darah lebih baik dibandingkan dengan plasebo pada dua kelompok pasien yang berbeda.
3. Hipotesis nol (H_0): Tidak ada perbedaan antara kelompok yang menerima obat dan yang menerima plasebo.
4. Jika nilai $p < 0,05$, maka obat dianggap memiliki efek signifikan dibandingkan dengan plasebo.
5. Perbedaan kadar gula darah antara pasien diabetes tipe 1 dan tipe 2
6. Penelitian ingin mengetahui apakah terdapat perbedaan rata-rata kadar gula darah antara pasien dengan diabetes tipe 1 dan tipe 2.
7. Uji t independen digunakan untuk membandingkan dua kelompok yang berbeda.
8. Membandingkan tingkat kelelahan antara perawat di rumah sakit dan puskesmas
9. Apakah perawat yang bekerja di rumah sakit mengalami kelelahan kerja yang lebih tinggi dibandingkan dengan perawat di puskesmas?
10. Dengan uji t independen, dapat diketahui apakah ada perbedaan signifikan dalam tingkat kelelahan berdasarkan skor kuesioner kelelahan kerja.

5.5.2 Uji t Dependen dalam Kesehatan

Digunakan ketika membandingkan data dari kelompok yang sama, tetapi pada dua kondisi yang berbeda (sebelum dan sesudah intervensi) (Field, 2018).

Contoh Aplikasi:

1. Pengaruh program latihan rehabilitasi pada pasien stroke
2. Seorang peneliti ingin mengetahui apakah program rehabilitasi selama 3 bulan meningkatkan kemampuan berjalan pada pasien stroke.
3. Pasien diuji sebelum dan setelah mengikuti rehabilitasi menggunakan uji t dependen untuk melihat apakah ada perbedaan signifikan (Pagano, 2019).
4. Efektivitas terapi fisioterapi terhadap nyeri punggung
5. Sebuah studi mengukur tingkat nyeri pasien sebelum dan sesudah menjalani fisioterapi selama 4 minggu.
6. Jika hasil uji t dependen menunjukkan perbedaan signifikan, maka terapi dianggap efektif dalam mengurangi nyeri.
7. Pengaruh diet rendah kalori terhadap berat badan pasien obesitas
8. Berat badan pasien dicatat sebelum dan setelah menjalani diet rendah kalori selama 6 bulan.
9. Uji t dependen digunakan untuk mengetahui apakah ada perbedaan signifikan dalam penurunan berat badan setelah diet.

5.5.3 Aplikasi Uji t dalam Bidang Psikologi

Dalam psikologi, uji t digunakan untuk mengukur perubahan dalam perilaku atau kondisi mental sebelum dan sesudah perlakuan tertentu (Cohen, 1988).

Contoh Aplikasi:

1. Efektivitas terapi kognitif dalam mengurangi kecemasan

2. Peneliti membandingkan tingkat kecemasan pasien sebelum dan setelah terapi kognitif menggunakan uji t dependen.
3. Perbedaan tingkat stres antara mahasiswa kedokteran dan mahasiswa teknik
4. Dengan uji t independen, dapat diketahui apakah terdapat perbedaan tingkat stres yang signifikan antara dua kelompok mahasiswa dari jurusan yang berbeda.

5.5.4 Aplikasi Uji t dalam Bidang Pendidikan

Dalam pendidikan, uji t digunakan untuk mengevaluasi efektivitas metode pembelajaran dan pengaruh berbagai faktor terhadap prestasi akademik.

Contoh Aplikasi:

1. Efektivitas metode pembelajaran daring vs. luring
2. Uji t independen digunakan untuk membandingkan nilai ujian siswa yang belajar secara daring dan luring.
3. Perbedaan skor ujian sebelum dan sesudah pelatihan keterampilan belajar
4. Uji t dependen dapat digunakan untuk menguji apakah pelatihan keterampilan belajar meningkatkan nilai ujian siswa.

5.5.5 Aplikasi Uji t dalam Bidang Ekonomi dan Bisnis

Dalam ekonomi dan bisnis, uji t sering digunakan untuk membandingkan kinerja perusahaan, efektivitas strategi pemasaran, dan kepuasan pelanggan (Moore, D. S., McCabe, G. P., & Craig, 2021).

Contoh Aplikasi:

1. Perbandingan kepuasan pelanggan antara dua cabang restoran
2. Uji t independen digunakan untuk mengetahui apakah terdapat perbedaan tingkat kepuasan pelanggan di dua lokasi restoran yang berbeda.

3. Pengaruh program pelatihan terhadap produktivitas karyawan
4. Uji t dependen digunakan untuk mengetahui apakah ada peningkatan produktivitas karyawan setelah mengikuti pelatihan (Bland, 2015).

Uji t merupakan metode statistik yang sangat penting dalam analisis data untuk membandingkan dua kelompok. Uji t independen digunakan saat dua kelompok data tidak saling berhubungan, sedangkan uji t dependen digunakan untuk membandingkan dua kondisi dalam kelompok yang sama. Kedua uji ini memiliki asumsi yang harus dipenuhi agar hasilnya valid, seperti distribusi normal dan homogenitas varians.

Dalam bidang kesehatan, psikologi, pendidikan, dan bisnis, uji t banyak digunakan untuk mengevaluasi efektivitas suatu intervensi atau membandingkan metode yang berbeda. Uji t memberikan hasil yang jelas dan mudah diinterpretasikan, tetapi juga memiliki keterbatasan, seperti ketergantungan pada distribusi normal dan ketidakmampuan membandingkan lebih dari dua kelompok tanpa metode tambahan seperti ANOVA.

Pemilihan jenis uji t yang tepat sangat bergantung pada desain penelitian dan hubungan antara data yang dianalisis. Dengan memahami konsep, asumsi, dan penerapan uji t secara benar, peneliti dapat menghasilkan analisis data yang akurat dan dapat dipertanggungjawabkan.

Bab 6

Uji ANOVA

6.1 Pengertian dan Sejarah Singkat ANOVA

Analysis of Variance (ANOVA) adalah teknik statistik yang digunakan untuk membandingkan rata-rata dari dua atau lebih kelompok untuk menentukan apakah terdapat perbedaan yang signifikan secara statistik diantara kelompok-kelompok tersebut. ANOVA digunakan ketika terdapat satu atau lebih variabel independen kategorikal dan satu variabel dependen kuantitatif yang berskala interval atau rasio. ANOVA pertama kali diperkenalkan oleh Sir Ronald A.

Fisher pada tahun 1920-an dalam konteks eksperimen agrikultur. Fisher mengembangkan ANOVA sebagai metode untuk menganalisis variasi dalam hasil pertanian yang dipengaruhi oleh berbagai faktor, seperti jenis pupuk dan kondisi tanah. Metode ini dipublikasikan secara lebih formal dalam bukunya *Statistical Methods for Research Workers* (1925) dan *The Design of Experiments* (1935). Seiring waktu, ANOVA berkembang dan

diterapkan dalam berbagai bidang ilmu, termasuk psikologi, kedokteran, ekonomi, dan ilmu sosial.

ANOVA bekerja dengan membandingkan variabilitas antara kelompok (antara grup) dengan variabilitas dalam kelompok (dalam grup). Jika variabilitas antara kelompok lebih besar dibandingkan variabilitas dalam kelompok, maka kemungkinan terdapat perbedaan yang signifikan antara kelompok-kelompok yang diuji (Montgomery, 2017; Field, 2018; Tabachnick & Fidell, 2019).

6.2 Jenis-Jenis ANOVA

Menurut Howell (2012); Field (2013); Montgomery (2017); Tabachnick & Fidell (2019), terdapat beberapa jenis ANOVA yang digunakan sesuai dengan desain penelitian dan jumlah faktor yang dibandingkan diantaranya:

1. One-Way ANOVA atau ANOVA satu faktor: adalah jenis yang paling dasar, digunakan untuk membandingkan rata-rata lebih dari dua kelompok berdasarkan satu variabel independen (faktor). Misalnya, dalam penelitian tentang metode pembelajaran, One-Way ANOVA dapat digunakan untuk menguji apakah terdapat perbedaan rata-rata skor ujian antara tiga metode belajar yang berbeda.
2. Two-Way ANOVA atau ANOVA dua faktor: digunakan untuk mengevaluasi efek dua faktor independen secara bersamaan. Two-Way ANOVA tidak hanya mengevaluasi efek utama dari masing-masing faktor, tetapi juga interaksi antara faktor-faktor tersebut, misalnya pengaruh jenis metode belajar dan gender terhadap skor ujian.

3. Repeated Measures ANOVA: digunakan saat data diambil dari subjek yang sama dalam kondisi yang berbeda pada waktu yang berbeda. Teknik ini berguna ketika data dikumpulkan dari peserta yang sama dalam beberapa kondisi atau waktu yang berbeda, seperti dalam penelitian psikologi yang mengukur tingkat stres seseorang sebelum, selama, dan setelah menjalani terapi.
4. Multivariate ANOVA (MANOVA): digunakan untuk menganalisis lebih dari satu variabel dependen dalam satu analisis. MANOVA memperhitungkan korelasi antara variabel dependen dan menganalisis perbedaan antara kelompok berdasarkan beberapa variabel hasil sekaligus, misalnya dalam penelitian tentang efek latihan fisik terhadap tekanan darah dan detak jantung.
5. Mixed-Design ANOVA: merupakan gabungan antara ANOVA dua faktor dan Repeated Measures ANOVA. Teknik ini digunakan ketika penelitian memiliki faktor independen yang melibatkan kelompok yang berbeda serta pengukuran berulang dalam setiap kelompok. Contohnya adalah penelitian tentang efektivitas metode belajar di mana metode belajar berbeda antar kelompok, tetapi performa peserta diuji beberapa kali dalam jangka waktu tertentu. Setiap jenis ANOVA memiliki kegunaan spesifik tergantung pada desain penelitian dan struktur data, sehingga pemilihan metode yang tepat sangat penting untuk mendapatkan hasil analisis yang valid dan akurat.

6.3 Asumsi ANOVA

Menurut Howell (2012); Field (2013); Montgomery (2017), ANOVA memiliki beberapa asumsi dasar yang harus dipenuhi agar hasil analisis valid dan dapat diinterpretasikan dengan benar diantaranya:

1. Independensi, yang berarti setiap observasi dalam sampel harus independen satu sama lain. Hal ini memastikan bahwa tidak ada hubungan antara data dalam satu kelompok dengan data dalam kelompok lain, sehingga tidak terjadi bias dalam perhitungan varians. Pelanggaran terhadap asumsi ini dapat menyebabkan hasil ANOVA menjadi tidak valid, sehingga perancangan penelitian yang baik, seperti penggunaan randomisasi, sangat diperlukan.
2. Normalitas, yang mengharuskan data dalam setiap kelompok berdistribusi normal, terutama ketika ukuran sampel kecil ($n < 30$). Normalitas dapat diuji dengan Shapiro-Wilk Test atau Kolmogorov-Smirnov Test, serta divisualisasikan menggunakan histogram atau Q-Q plot. Jika data tidak berdistribusi normal, transformasi data seperti log atau square root dapat digunakan, atau alternatif lain seperti Kruskal-Wallis Test bisa dipertimbangkan.
3. Homogenitas varians atau kesamaan varians antar kelompok, merupakan asumsi penting dalam ANOVA. Homogenitas varians dapat diuji dengan Levene's Test, Bartlett's Test, atau Brown-Forsythe Test. Jika varians antar kelompok berbeda secara signifikan, Welch's ANOVA dapat digunakan sebagai alternatif karena tidak bergantung pada asumsi homogenitas varians.
4. Variabel dependen harus berskala interval atau rasio, yang berarti data memiliki nilai numerik dengan selisih yang bermakna. Jika

data berskala ordinal, ANOVA tidak dapat digunakan sehingga Kruskal-Wallis Test lebih sesuai sebagai metode alternatif.

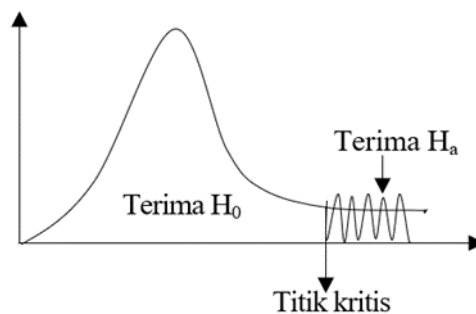
Peneliti harus memastikan semua asumsi ini terpenuhi sangat penting agar hasil ANOVA tidak bias dan dapat memberikan kesimpulan yang akurat. Jika satu atau lebih asumsi tidak terpenuhi, maka perlu dipertimbangkan metode analisis alternatif atau strategi penyesuaian seperti transformasi data dan penggunaan uji statistik yang lebih sesuai.

6.4 Langkah-langkah Pengujian ANOVA

Menurut Sujarweni (2015); Field (2018), pengujian ANOVA dilakukan untuk menentukan apakah terdapat perbedaan yang signifikan antara rata-rata lebih dari dua kelompok.

Berikut adalah langkah-langkah umum dalam melakukan uji ANOVA:

1. Langkah 1: Menentukan Hipotesis Statistik
 - a. Hipotesis nol (H_0): Tidak ada perbedaan rata-rata antar kelompok.
 - b. Hipotesis alternatif (H_a): Setidaknya ada satu kelompok yang memiliki rata-rata berbeda secara signifikan.



Gambar 6.1: Daerah penerimaan H_0 dan H_a (Sujarweni, 2015)

2. Langkah 2: Menentukan Tingkat Signifikansi (α)

Tingkat signifikansi menunjukkan besarnya kemungkinan penerimaan H_a . Besarnya signifikansi ditentukan oleh peneliti berkisar antara 25% sampai 0,1%, namun pada umumnya digunakan tingkat signifikansi $\alpha = 0.05$, yang berarti ada kemungkinan 5% untuk melakukan kesalahan tipe I.

3. Langkah 3: Mengumpulkan dan Memeriksa Asumsi Data

Pada langkah ini, asumsi-asumsi pada uji ANOVA menjadi tolak ukur uji dilakukan. Uji normalitas, homogenitas varians, dan independensi menjadi pertimbangan utamanya.

4. Langkah 4: Menghitung Statistik ANOVA

Pada langkah ini peneliti menentukan varians antar kelompok (MSB) dan dalam kelompok (MSW). Selanjutnya menghitung nilai F statistik. Lalu nilai F dibandingkan dengan nilai kritis pada tabel distribusi F.

5. Langkah 5: Menentukan p-Value dan Keputusan Uji

- a. Jika $p\text{-value} < \alpha$, maka tolak hipotesis nol (H_0) dan simpulkan bahwa terdapat perbedaan yang signifikan antar kelompok.
- b. Jika $p\text{-value} \geq \alpha$, maka gagal menolak hipotesis nol, artinya tidak ada cukup bukti untuk menyatakan adanya perbedaan signifikan.

6. Langkah 6: Analisis Lanjutan (Post Hoc Test, jika diperlukan)

Jika hasil ANOVA signifikan, dilakukan uji post hoc (misalnya Tukey's HSD, Bonferroni, Scheffé Test, Duncan's Multiple Range Test, Games-Howell Test) untuk mengetahui kelompok mana yang berbeda secara signifikan.

- a. Tukey's HSD (Honestly Significant Difference): digunakan untuk membandingkan semua pasangan kelompok dalam ANOVA, cocok jika jumlah sampel antar kelompok seimbang,

dan mengontrol tingkat kesalahan tipe I dengan baik. Contoh: Menentukan metode pembelajaran mana yang berbeda secara signifikan dari metode lainnya.

- b. Bonferroni Correction: metode konservatif yang menyesuaikan tingkat signifikansi dengan membagi α (misalnya, 0.05) dengan jumlah perbandingan yang dilakukan, cocok untuk penelitian dengan jumlah kelompok kecil karena mengurangi kemungkinan kesalahan tipe I. Contoh: Digunakan saat membandingkan efektivitas tiga jenis terapi dalam mengurangi kecemasan.
- c. Scheffé Test: lebih fleksibel dibandingkan Tukey karena dapat digunakan untuk perbandingan kompleks antar kelompok. Namun, lebih konservatif, sehingga memerlukan perbedaan yang lebih besar agar signifikan. Contoh: Menentukan apakah kombinasi metode belajar tertentu memiliki dampak lebih besar dibanding metode lainnya.
- d. Duncan's Multiple Range Test: mirip dengan Tukey tetapi lebih liberal, sehingga lebih mungkin menemukan perbedaan signifikan, cocok untuk penelitian eksploratif yang tidak terlalu ketat dalam mengontrol kesalahan tipe I. Contoh: Meneliti pengaruh variasi dosis obat terhadap efektivitas pengobatan.
- e. Games-Howell Test: digunakan jika asumsi homogenitas varians tidak terpenuhi, cocok untuk kelompok dengan varians yang tidak sama. Contoh: Menguji apakah perbedaan gaji karyawan di berbagai industri signifikan meskipun variabilitas gaji sangat berbeda.

- Langkah 7: Menarik Kesimpulan dan Interpretasi data
Kesimpulan dibuat berdasarkan hasil uji F dan p-value. Jika ada perbedaan signifikan, interpretasi harus mempertimbangkan makna praktis dan relevansi hasil penelitian.

6.5 Penyelesaian ANOVA

Menurut Field (2018); Tabachnick & Fidell (2019), berikut penyelesaian berdasarkan jenis ANOVA yang digunakan:

6.5.1 One-Way ANOVA

Statistik uji F dalam One-Way ANOVA dihitung dengan:

$$F = \frac{MSB}{MSW}$$

Namun, sebelumnya peneliti harus mengetahui varians ANOVA dalam suatu kelompok berbeda dari rata-rata kelompok tersebut. Varians dihitung sebagai rata-rata dari kuadrat deviasi setiap nilai terhadap rata-rata kelompok.

Terdapat dua jenis varians utama:

- Varians Antar Kelompok (Between-Group Variance): mengukur perbedaan rata-rata antara kelompok yang dibandingkan. Jika varians antar kelompok besar, maka ada indikasi bahwa kelompok-kelompok tersebut memiliki perbedaan yang signifikan.

Rumusnya:

$$MSB = \frac{SSB}{df_B} = \frac{\sum n_i(\bar{X}_i - \bar{X}_T)^2}{k - 1}$$

Keterangan:

- a. SSB = Sum of squares between (jumlah kuadrat antar kelompok)
 - b. n_i = jumlah sampel dalam kelompok ke-i
 - c. \bar{X}_i = rata-rata kelompok ke-i
 - d. \bar{X}_T = rata-rata keseluruhan (grand mean)
 - e. k = jumlah kelompok
 - f. $dfB = k-1$ = derajat kebebasan antara kelompok
2. Varians Dalam Kelompok (Within-Group Variance): mengukur variasi data di dalam masing-masing kelompok yang dibandingkan. Jika varians dalam kelompok besar, maka ada kemungkinan bahwa perbedaan antar kelompok hanya disebabkan oleh variasi acak dalam kelompok.

Rumusnya:

$$MSW = \frac{SSW}{df_W} = \frac{\sum \sum (X_{ij} - \bar{X}_i)^2}{N - k}$$

Keterangan:

- a. SSW = Sum of squares within (jumlah kuadrat dalam kelompok)
- b. X_{ij} = nilai individu dalam kelompok ke-i
- c. \bar{X}_i = rata-rata kelompok ke-i
- d. N = jumlah total sampel
- e. k = jumlah kelompok
- f. $dfW = N-k$ = derajat kebebasan dalam kelompok

Jika varians antar kelompok lebih besar dibandingkan varians dalam kelompok, maka nilai F akan tinggi, yang menunjukkan kemungkinan adanya perbedaan signifikan antar kelompok. Nilai F yang dihitung dibandingkan dengan nilai F kritis dari tabel distribusi F untuk

menentukan apakah perbedaan antar kelompok signifikan secara statistik. Jika p -value lebih kecil dari tingkat signifikansi (misalnya 0.05), maka hipotesis nol ditolak, yang berarti ada perbedaan nyata antara kelompok yang diuji.

Misalnya, dalam sebuah penelitian tentang metode pembelajaran (tradisional, online, dan hybrid), diperoleh nilai $F = 5.23$ dengan p -value = 0.008 pada tingkat signifikansi $\alpha = 0.05$. Karena p -value lebih kecil dari 0.05, maka hipotesis nol ditolak, yang berarti ada perbedaan signifikan antara metode pembelajaran terhadap skor ujian.

Distribusi F memiliki beberapa karakteristik utama:

1. Asimetris positif, yang berarti nilainya selalu positif karena varians tidak pernah bernilai negatif.
2. Derajat kebebasan (df), yang terdiri dari df_1 (numerator) yang dihitung dari jumlah kelompok dikurangi satu ($k - 1$) dan df_2 (denominator) yang berasal dari total jumlah observasi dikurangi jumlah kelompok ($N - k$). Bentuk distribusi F akan berubah tergantung pada kombinasi kedua parameter tersebut.
3. Perbandingan varians, di mana nilai F dihitung sebagai rasio antara varians antar kelompok (MSB) terhadap varians dalam kelompok (MSW).

6.5.2 Two-Way ANOVA

Total variasi dibagi menjadi tiga komponen:

$$SS_{Total} = SS_A + SS_B + SS_{AB} + SS_W$$

Keterangan:

- SS_A = variasi yang disebabkan oleh faktor A
- SS_B = variasi yang disebabkan oleh faktor B

- SSAB = interaksi antara faktor A dan B
- SSW = variasi dalam kelompok

Contoh: peneliti ingin mengetahui pengaruh Metode Pembelajaran (A, B, C) dan Jenis Kelamin (Laki-laki, Perempuan) terhadap skor ujian. Maka, hitung rata-rata setiap kombinasi kelompok, hitung SS_A , SS_B , SS_{AB} , dan SS_W . lalu hitung MSA , MSB , $MSAB$, dan MSW , dan hitung F untuk setiap faktor dan interaksi.

6.5.3 Repeated Measures ANOVA

Rumus ini digunakan ketika data dikumpulkan dari subjek yang sama pada beberapa titik waktu atau dalam kondisi yang berbeda.

$$F = \frac{MS_{Treatment}}{MS_{Error}}$$

Dimana variasi dalam subjek diperhitungkan untuk mengurangi error. Contoh: skor ujian siswa diukur setelah 3 sesi pelatihan berbeda, maka hitung rata-rata tiap sesi, hitung varians dalam individu dan antar sesi, dan gunakan formula F untuk menentukan signifikansi perbedaan.

6.5.4 Multivariate ANOVA (MANOVA)

Dapat digunakan matriks kovarians dan statistik Wilks' Lambda:

$$\Lambda = \frac{Det(W)}{Det(B + W)}$$

Dimana W adalah matriks varians dalam kelompok dan B adalah matriks varians antar kelompok. Contoh: peneliti ingin menguji pengaruh metode pembelajaran terhadap skor matematika dan skor bahasa Inggris, maka lakukan hitungan matriks varians-kovarians antar variable, lalu hitung statistik Wilks' Lambda, dan bandingkan dengan nilai kritis dari tabel distribusi F .

6.5.5 Mixed-Design ANOVA

Rumus yang digunakan adalah gabungan antara rumus Two-Way ANOVA dan Repeated Measures ANOVA. Contoh Manual:

Misalkan kita ingin menguji pengaruh Metode Pembelajaran (A, B, C) pada skor ujian sebelum dan sesudah pelatihan:

1. Hitung efek metode pembelajaran,
2. Hitung efek waktu (before vs. after), dan
3. Hitung interaksi metode \times waktu.

6.6 Langkah Melakukan Uji ANOVA dengan Penyelesaian SPSS

Berikut ialah langkah-langkah untuk melakukan uji ANOVA berdasarkan jenisnya menggunakan SPSS dan interpretasinya:

1. One-Way ANOVA
 - a. Masukkan data: buat dua variabel, variabel kategori (faktor) dan variabel dependen (nilai yang diukur).
 - b. Buka menu analisis: klik analyze \rightarrow compare means \rightarrow one-way ANOVA.
 - c. Atur variabel: masukkan variabel dependen ke dalam kolom dependent list dan masukkan variabel kategori (independen) ke dalam kolom faktor.
 - d. Pilih uji post hoc (jika diperlukan): klik tombol post hoc, pilih Tukey atau metode lainnya.
 - e. Klik OK dan lihat hasil di output SPSS.

- f. Interpretasi:
 - 1) Lihat nilai Sig. (p-value) pada tabel ANOVA. Jika $p < 0.05$, berarti ada perbedaan signifikan antara kelompok.
 - 2) Jika hasil signifikan, lakukan post hoc test untuk melihat kelompok mana yang berbeda.
2. Two-Way ANOVA
 - a. Masukkan data: buat tiga variabel: dua variabel independen (kategori/faktor) dan satu variabel dependen.
 - b. Buka menu analisis: klik analyze → general linear model → univariate.
 - c. Atur variabel: masukkan variabel dependen ke kolom dependent variable, dan masukkan dua variabel independen ke kolom fixed factors.
 - d. Pilih uji interaksi: klik tombol model → custom, centang main effects dan interaction effects.
 - e. Klik OK dan lihat hasil di output SPSS.
 - f. Interpretasi:
 - 1) Perhatikan kolom Sig. (p-value) untuk melihat efek dari masing-masing faktor serta interaksi antara keduanya.
 - 2) Jika ada interaksi signifikan ($p < 0.05$), berarti kedua faktor memengaruhi variabel dependen secara bersama-sama.
3. Repeated Measures ANOVA
 - a. Masukkan data: setiap kolom mewakili pengukuran pada waktu yang berbeda (misalnya, pre-test, post-test, follow-up).
 - b. Buka menu analisis: klik analyze → general linear model → repeated measures.
 - c. Tentukan faktor waktu: pada bagian within-subject factor name, beri nama faktor (misalnya, waktu) dan tentukan jumlah level (misalnya, 3 untuk pre-test, post-test, follow-up).

- d. Atur variabel dependen: klik define, lalu masukkan variabel pengukuran ke dalam kolom within-subjects variables.
 - e. Klik OK dan lihat hasil di output SPSS.
 - f. Interpretasi: Lihat p-value pada Wilks' Lambda di tabel Multivariate Tests untuk menentukan apakah ada perbedaan antar waktu.
4. Multivariate ANOVA (MANOVA)
- a. Masukkan data: buat dua atau lebih variabel dependen dan satu atau lebih variabel independen (faktor/kategori).
 - b. Buka menu analisis: klik analyze → general linear model → multivariate.
 - c. Atur variabel: masukkan variabel dependen ke kolom dependent variables, dan masukkan variabel independen ke kolom fixed factors.
 - d. Klik OK dan lihat hasil di output SPSS.
 - e. Interpretasi: Perhatikan nilai Wilks' Lambda, Pillai's Trace, atau Hotelling's Trace untuk melihat apakah faktor independen berpengaruh pada kombinasi variabel dependen.
5. Mixed-Design ANOVA
- a. Masukkan data: buat satu variabel faktor antara subjek (kelompok) dan beberapa variabel pengukuran berulang (waktu/kondisi).
 - b. Buka menu analisis: klik analyze → general linear model → repeated measures.
 - c. Tentukan faktor waktu: pada within-subjects factor name, beri nama faktor (misalnya, waktu) dan jumlah levelnya (misalnya, 3).

- d. Tambahkan faktor between-subjects: masukkan faktor antara subjek (misalnya, metode pembelajaran) ke dalam kolom between-subjects factor.
- e. Klik OK dan lihat hasil di output SPSS.
- f. Interpretasi:
 - 1) Perhatikan hasil within-subjects effects untuk melihat efek faktor berulang.
 - 2) Perhatikan between-subjects effects untuk melihat pengaruh kelompok perlakuan.
 - 3) Jika ada interaksi signifikan, berarti pengaruh faktor waktu berbeda tergantung pada kelompok.

Bab 7

Uji Korelasi

7.1 Korelasi

Korelasi mengukur derajat dan arah keterkaitan dua variabel. Korelasi tidak membuat garis melalui titik data. Namun, hanya menghitung koefisien korelasi yang menunjukkan seberapa besar kecenderungan satu variabel berubah saat variabel lain berubah. Jika r adalah 0,0, tidak ada hubungan. ketika r positif, ada tren bahwa satu variabel naik saat variabel lain naik. Ketika r negatif, ada tren bahwa satu variabel naik saat variabel lain turun. Dengan korelasi, tidak perlu memikirkan sebab dan akibat.

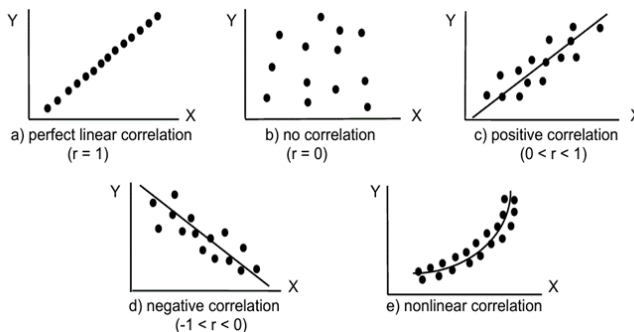
Tidak masalah variabel mana dari kedua variabel yang bergantung pada panggilan dan yang tidak bergantung pada panggilan, jika kedua variabel ditukar, tingkat koefisien korelasi akan sama. Tanda (+, -) dari koefisien korelasi menunjukkan arah asosiasi. Besarnya koefisien korelasi menunjukkan kekuatan asosiasi, misalnya. Korelasi $r = -0,8$ menunjukkan asosiasi negatif yang kuat (tren terbalik) antara dua variabel, sedangkan korelasi $r = 0,4$ menunjukkan asosiasi positif yang lemah. Korelasi yang mendekati nol menunjukkan tidak ada asosiasi linier antara dua variabel

kontinu sebagaimana ditunjukkan pada Gambar 7.1. (Mardhotillah, 2012; Zaid, 2015).

Statistik parametrik adalah jenis statistik inferensial yang paling umum, yang dihitung dengan tujuan menggeneralisasi temuan sampel ke populasi yang diwakilinya. Uji parametrik membuat asumsi tentang parameter populasi, sedangkan uji nonparametrik tidak menyertakan asumsi tersebut atau menyertakan lebih sedikit.

Misalnya, uji parametrik mengasumsikan bahwa sampel telah dipilih secara acak dari populasi yang diwakilinya dan bahwa distribusi data dalam populasi memiliki distribusi dasar yang diketahui. Asumsi distribusi yang paling umum adalah bahwa distribusinya normal. Distribusi lainnya termasuk distribusi binomial (regresi logistik) dan distribusi Poisson (regresi Poisson), dan uji nonparametrik terkadang disebut uji "bebas distribusi". Selain itu, statistik parametrik mengharuskan data diukur menggunakan skala interval atau rasio, sedangkan statistik nonparametrik menggunakan data yang diukur dengan skala nominal atau ordinal.

Terdapat tiga jenis koefisien korelasi nonparametrik yang umum digunakan (Spearman R, Kendall Tau, dan koefisien Gamma), dimana koefisien korelasi parametrik (Pearson). (Zaid, 2015; Mardhotillah, dkk., 2022; Chen & Popovich, 2022). Pada Bab ini akan dibahas Korelasi Pearson untuk Analisis Parametrik dan Korelasi Rank Spearman untuk Analisis Non Parametrik.



Gambar 7.1: Tipikal Korelasi

Perhitungan koefisien korelasi secara statistik sedikit rumit, namun akan lebih mudah jika dilakukan berbantuan software statistik. Korelasi menggunakan distribusi t untuk menguji hipotesis nol bahwa tidak ada hubungan antara kedua variabel (yaitu, $r = 0$). Seperti halnya uji-t lainnya, korelasi mengasumsikan bahwa kedua variabel terdistribusi normal. Jika salah satu atau kedua variabel condong ke satu arah atau yang lain, koefisien korelasi yang dihasilkan mungkin tidak mewakili data dan hasil uji t tidak valid.

Jika diagram sebar data tidak mengasumsikan beberapa bentuk pola elips, salah satu atau kedua variabel mungkin condong (seperti pada Gambar 7.1e). Masalah variabel yang tidak berdistribusi normal dapat diatasi dengan mengubah data menjadi distribusi normal atau menggunakan metode nonparametrik untuk menghitung korelasi pada peringkat data (Cohen, dkk, 2003; Hidayanti & Mandalika, 2023; Jabnabillah & Margina, 2022; Rojas, 2021; Al-Hameed, 2022).

Pada keilmuan Statistika, "signifikan" pada Uji Korelasi berarti penting, sedangkan dalam Statistik "signifikan" berarti mungkin benar (bukan karena kebetulan). Temuan penelitian mungkin benar tanpa menjadi penting. Ketika ahli statistik mengatakan suatu hasil "sangat signifikan", maksud mereka kemungkinan besar benar. Tingkat signifikansi menunjukkan seberapa besar kemungkinan pola dalam data yang diteliti terjadi karena kebetulan. Tingkat yang paling umum, yang digunakan untuk menunjukkan sesuatu cukup baik untuk dipercaya, adalah "0,95".

Ini berarti bahwa temuan tersebut memiliki peluang 95% untuk menjadi benar yang juga berarti bahwa temuan tersebut memiliki tingkat keyakinan 95% untuk menjadi benar. Tidak ada paket statistik yang akan menunjukkan "95%" atau "0,95" untuk menunjukkan tingkat ini. Sebaliknya, paket tersebut akan menunjukkan "0,05," yang berarti bahwa temuan tersebut memiliki peluang lima persen (0,05) untuk tidak menjadi "kesalahan" yang merupakan kebalikan dari peluang 95% untuk menjadi benar.

Untuk menemukan tingkat signifikansi, kurangi angka yang ditunjukkan dari angka satu. Misalnya, nilai "0,01" berarti ada tingkat keyakinan 99% ($1 - 0.01 = .99$) kemungkinan bahwa hal itu benar. Dan secara populer kekeliruan dalam pengujian hipotesis yang digunakan adalah alpha. (Zaid, 2015; Mardhotillah dkk, 2022).

Adapun manfaat menganalisis korelasi di antaranya untuk membantu dalam menentukan tingkat dan arah dua atau lebih variabel secara tepat, dengan mengetahui koefisien korelasi, peneliti dapat mengurangi rentang ketidakpastian prediksi, sehingga Prediksi yang didasarkan pada analisis korelasi akan lebih dapat diandalkan dan mendekati kenyataan. Selain itu Analisis korelasi berkontribusi pada pemahaman perilaku ekonomi, membantu menemukan variabel yang sangat penting yang menjadi sandaran variabel lain, dapat mengungkapkan kepada para peneliti hubungan yang menyebabkan beberapa permasalahan berhubungan dengan permasalahan lainnya serta memberikan beberapa rekomendasi pada para peneliti di berbagai disiplin ilmu, termasuk biostatistics.

Teori ekonomi dan studi bisnis menunjukkan hubungan antara variabel seperti harga dan jumlah yang diminta, pengeluaran iklan dan promosi penjualan, dll. Ukuran koefisien korelasi adalah ukuran relatif perubahan. (Bujang & Baharum, 2016; Schober, dkk., 2018; Chen, 2003; Hazra & Gogtay, 2016).

7.2 Korelasi Pearson untuk Analisis Parametrik

Koefisien korelasi dilambangkan dengan r . Karl Pearson (1857 – 1936), seorang ahli biometrika Inggris, mengembangkan rumus untuk Koefisien Korelasi. Koefisien korelasi antara dua variabel X dan Y dilambangkan dengan $r(X,Y)$ atau $r_{x,y}$. Misalkan x_1, x_2, \dots, x_n merupakan himpunan observasi dari variabel X dan misalkan y_1, y_2, \dots, y_n merupakan nilai-nilai Y yang bersesuaian, secara matematis dirumuskan:

$$r_{xy} = \frac{\frac{1}{n} \sum (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\frac{1}{n} \sum (x_i - \bar{x})^2} \sqrt{\frac{1}{n} \sum (y_i - \bar{y})^2}} = \frac{\text{Cov}(X, Y)}{\sigma_x \sigma_y}$$

Dimana, $\text{Cov}(X, Y)$ disebut sebagai kovariansi antara X dan Y dengan σ_x dan σ_y masing-masing adalah simpangan baku dari X dan Y.

Contoh penghitungan koefisien korelasi Pearson bernilai nol yang menunjukkan x dan y tidak berhubungan, dapat dilihat pada Tabel 7.1 dengan ilustrasi dan proses penghitungan sebagai berikut:

Tabel 7.1: Korelasi X dan Y beserta Proses Penghitungannya

x	y = x ²	xy	x ²	y ²
-3	9	-27	9	81
-2	4	-8	4	16
-1	1	-1	1	1
0	0	0	0	0
1	1	1	1	1
2	4	8	4	16
3	9	27	9	81
$\Sigma x = 0$	$\Sigma y = 28$	$\Sigma xy = 0$	$\Sigma x^2 = 28$	$\Sigma y^2 = 196$

$$\bar{x} = \frac{\sum x}{n} = \frac{0}{n} = 0$$

$$r_{xy} = \frac{\sum xy - n\bar{x}\bar{y}}{\sqrt{\sum x^2 - n\bar{x}^2} \sqrt{\sum y^2 - n\bar{y}^2}} = \frac{0 - n \times 0 \times \bar{y}}{\sqrt{\sum x^2 - n\bar{x}^2} \sqrt{\sum y^2 - n\bar{y}^2}} = 0$$

Korelasi Pearson kerap digunakan dalam Biostatistika untuk mengukur hubungan linear antara dua variabel kontinu.

Berikut beberapa contoh penerapannya:

1. Hubungan antara kadar gula darah dan tekanan darah: Peneliti ingin mengetahui apakah ada hubungan antara kadar gula darah (mg/dL) dengan tekanan darah sistolik (mmHg) pada pasien diabetes
2. Hubungan antara tingkat aktivitas fisik dan indeks massa tubuh (IMT): Studi dapat mengevaluasi apakah ada hubungan linear

antara jumlah waktu yang dihabiskan untuk berolahraga (jam/minggu) dengan IMT

3. Hubungan antara asupan kalori harian dan kadar kolesterol: Analisis dilakukan untuk mengetahui apakah jumlah kalori yang dikonsumsi sehari-hari berkaitan dengan kadar kolesterol total dalam darah
4. Hubungan antara usia dan kapasitas paru-paru: Dalam studi kesehatan pernapasan, peneliti dapat memeriksa apakah usia seseorang memiliki korelasi dengan volume kapasitas paru-parunya
5. Hubungan antara tingkat paparan polutan udara dan fungsi paru-paru: Penelitian biostatistik ini dapat menggunakan korelasi Pearson untuk mengevaluasi hubungan antara konsentrasi polutan di udara dan kapasitas vital paru-paru
6. Hubungan antara lama tidur dan kesehatan mental: Dalam studi epidemiologi, korelasi Pearson bisa digunakan untuk mengevaluasi hubungan antara durasi tidur (jam/malam) dengan skor kesehatan mental berdasarkan skala tertentu. Penerapan-penerapan tersebut sangat berguna untuk membantu memahami hubungan antar variabel dalam konteks Biostatistika Inferensial.

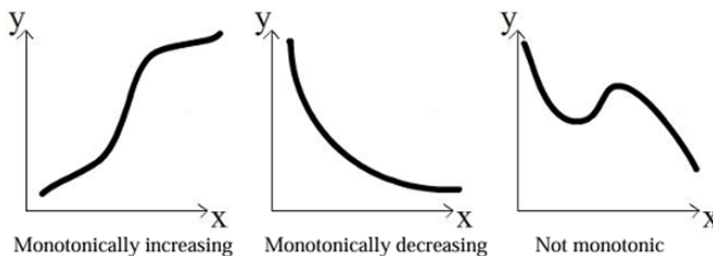
7.3 Korelasi Rank Spearman

Sebelum mempelajari korelasi Spearman, penting untuk memahami korelasi Pearson yang merupakan ukuran statistik kekuatan hubungan linear antara data berpasangan.

Perhitungan dan pengujian signifikansi selanjutnya memerlukan asumsi data berikut agar berlaku data:

1. Skala interval atau rasio
2. Memiliki keterkaitan secara linear
3. Bivariat berdistribusi normal. Jika data biostatistika yang diteliti tidak memenuhi asumsi tersebut, dapat digunakan korelasi rank spearman.

Untuk memahami korelasi Spearman, perlu diketahui apa itu fungsi monotonik. Fungsi monotonik adalah fungsi yang tidak pernah meningkat atau tidak pernah menurun seiring dengan peningkatan variabel independennya. Grafik berikut menggambarkan fungsi monotonik:



Gambar 7.2: Beberapa Fungsi Monotonik

Gambar 7.2 menjelaskan tiga kondisi, antara lain:

1. Meningkat secara monoton, ketika variabel x meningkat, variabel y tidak pernah menurun
2. Menurun secara monoton, ketika variabel x meningkat, variabel y tidak pernah meningkat
3. Tidak monoton, ketika variabel x meningkat, variabel y terkadang menurun dan terkadang meningkat.

Koefisien korelasi Spearman adalah ukuran statistik kekuatan hubungan monoton antara data berpasangan. Dalam sampel, koefisien ini dilambangkan dengan dan dirancang dengan batasan sebagai berikut:

$$-1 \leq r_s \leq 1$$

Dan interpretasinya mirip dengan interpretasi Pearson, misalnya, semakin dekat hubungan monotonik, \pm hubungan semakin kuat. Korelasi adalah ukuran efek sehingga kita dapat menggambarkan kekuatan korelasi secara verbal menggunakan panduan berikut untuk nilai absolut:

- 0,00 – 0,19 “Korelasi sangat lemah”
- 0,20 – 0,39 “Korelasi lemah
- 0,40 – 0,59 “Korelasi moderat atau sedang”
- 0,60 – 0,79 “Korelasi kuat”
- 0,80 – 1 “Korelasi sangat kuat”

Perhitungan koefisien korelasi Spearman dan pengujian signifikansi selanjutnya memerlukan asumsi berikut agar berlaku data berskala interval atau tingkat rasio atau ordinal (minimal data berskala ordinal) dan kedua variabel terkait secara monoton. Tidak seperti korelasi Pearson, pada Korelasi Rank Spearman tidak terdapat persyaratan kenormalan dan karenanya merupakan statistik nonparametrik (Chen & Popovich, 2022; Al-Hameed, 2022).

Korelasi Rank Spearman bekerja dengan menghitung nilai-nilai peringkat data yang diteliti. Peringkat (dari rendah ke tinggi) diperoleh dengan menetapkan peringkat 1 untuk nilai terendah, 2 untuk nilai terendah berikutnya, dan seterusnya.

Formulasi Korelasi Rank Spearman adalah sebagai berikut:

$$R = 1 - \frac{6\sum D^2}{N^3 - N}$$

di mana, D merupakan Selisih peringkat antara item berpasangan dalam dua seri dan N menunjukkan Jumlah pasangan peringkat. Dalam beberapa kasus, mungkin perlu untuk menetapkan peringkat yang sama pada dua atau lebih elemen atau individu atau entri. Dalam situasi seperti itu, biasanya untuk memberikan peringkat rata-rata pada setiap individu atau entri.

Dapat disimpulkan bahwa Korelasi Rank Spearman digunakan untuk mengukur hubungan antara dua variabel ordinal atau variabel kontinu yang tidak memenuhi asumsi distribusi normal.

Berikut beberapa contoh penerapannya dalam biostatistika:

1. Hubungan antara tingkat pendidikan dan skor kesadaran kesehatan: Peneliti dapat menggunakan Korelasi Spearman untuk melihat apakah ada hubungan antara tingkat pendidikan (misalnya, SD, SMP, SMA, Perguruan Tinggi) dengan skor kesadaran kesehatan berdasarkan survei.
2. Hubungan antara tingkat nyeri dan kepatuhan minum obat: Studi dapat mengevaluasi apakah tingkat nyeri (diberi peringkat, misalnya, dari 1 sampai 10) berkorelasi dengan seberapa patuh pasien mengambil obat mereka.
3. Hubungan antara tingkat kebisingan dan kualitas tidur: Peneliti dapat memeriksa hubungan antara tingkat kebisingan (rendah, sedang, tinggi) di lingkungan tempat tinggal seseorang dengan kualitas tidur yang dinilai menggunakan skala ordinal.
4. Hubungan antara urutan kelahiran dan status gizi anak: Dalam studi kesehatan anak, analisis Spearman dapat digunakan untuk melihat apakah posisi anak dalam urutan kelahiran (anak pertama, kedua, ketiga, dan seterusnya) berkorelasi dengan status gizi mereka.

5. Hubungan antara skor aktivitas fisik dan tingkat stres: Korelasi Spearman dapat digunakan untuk mengetahui hubungan antara skor aktivitas fisik (berdasarkan skala ordinal) dan tingkat stres yang diukur dengan skala ordinal juga.
6. Hubungan antara kualitas air minum dan kepuasan masyarakat: Peneliti dapat mengevaluasi apakah peringkat kualitas air minum (buruk, cukup, baik, sangat baik) berkorelasi dengan tingkat kepuasan masyarakat.

Penerapan metode ini sangat berguna saat data yang kita miliki tidak linear atau melanggar asumsi parametrik.

Bab 8

Analisis uji Chi Square dengan SPSS

8.1 Definisi

Dalam dunia keperawatan dan kesehatan, analisis statistik berperan penting dalam menilai hubungan antarvariabel dan mendukung pengambilan keputusan berbasis bukti (evidence-based practice). Salah satu uji statistik yang banyak digunakan adalah uji Chi-Square, yang memungkinkan peneliti mengevaluasi hubungan antara dua variabel kategori, seperti hubungan antara tingkat pengetahuan pasien dan kepatuhan dalam menjalani terapi, atau hubungan antara status gizi dengan kejadian anemia pada ibu hamil.

Dengan bantuan perangkat lunak SPSS, analisis ini dapat dilakukan secara lebih sistematis dan akurat, membantu tenaga kesehatan dalam mengidentifikasi pola dan tren yang dapat berdampak pada intervensi keperawatan. Bab ini akan membahas konsep dasar uji Chi-Square, jenis-

jenisnya, serta cara mengaplikasikannya dalam penelitian keperawatan dan kesehatan menggunakan SPSS, sehingga dapat menjadi panduan praktis bagi akademisi maupun praktisi di bidang kesehatan.

Uji Chi-Square digunakan untuk menentukan apakah terdapat perbedaan yang signifikan antara frekuensi yang diamati dan frekuensi yang diharapkan dalam satu atau lebih kategori:

1. Uji Chi Square merupakan analisis yang digunakan untuk untuk menguji perbedaan proporsi/persentase antara beberapa kelompok data. Uji ini dapat pula menilai seberapa besar proporsi tersebut (Dahlan, 2014; Yuandari & Rahman, 2017)
2. Dalam penelitian kesehatan, seringkali ditemukan dua variabel dan setiap variabel terdiri atas dua atau lebih kategori, misalnya saja beratnya penyakit atau tingkat kesembuhan. Untuk mengetahui hubungan antara keduanya, dapat digunakan uji Chi-Square (Budiarto, 2001; Tiro, 2007). Lebih dari itu, Chi-Square juga dapat digunakan untuk menguji kesesuaian ataupun homogenitas. Terdapat beberapa jenis uji Chi-Square yang relevan dalam penelitian keperawatan dan kesehatan:

- a. Chi-Square Goodness of Fit (Uji Kesesuaian)

Uji Goodness of Fit digunakan untuk menentukan apakah distribusi frekuensi dalam sampel sesuai dengan distribusi yang diharapkan dalam populasi. Dengan kata lain, uji ini membandingkan frekuensi yang diamati dengan frekuensi yang diharapkan untuk melihat apakah ada perbedaan yang signifikan.

Contoh: Seorang peneliti ingin mengetahui apakah distribusi pasien berdasarkan tingkat kepatuhan dalam menjalani perawatan diabetes (tinggi, sedang, rendah) sesuai dengan yang diharapkan berdasarkan studi sebelumnya.

- 1) Hipotesis Nol (H_0): Tidak ada perbedaan antara distribusi yang diamati dengan yang diharapkan.
- 2) Hipotesis Alternatif (H_a): Ada perbedaan yang signifikan antara distribusi yang diamati dan yang diharapkan.

Jika hasil uji menunjukkan $p\text{-value} < 0,05$, maka H_0 ditolak, yang berarti distribusi data tidak sesuai dengan distribusi yang diharapkan.

b. Chi-Square Test of Independence (Uji Independensi)

Uji Chi-Square Test of Independence digunakan untuk menentukan apakah ada hubungan antara dua variabel kategorik dalam satu populasi. Dengan kata lain, uji ini menguji apakah dua variabel saling independen atau memiliki keterkaitan.

Contoh: Seorang peneliti ingin menganalisis hubungan antara tingkat literasi kesehatan dan kepatuhan ibu hamil dalam pemeriksaan antenatal. Maka, dengan analisis Chi-Square akan didapatkan table kontingensi terkait proporsi masing-masing variabel dan nilai p value hasil analisis.

- 1) Hipotesis Nol (H_0): Tidak ada hubungan antara tingkat literasi kesehatan dengan kepatuhan ibu hamil dalam pemeriksaan antenatal.
- 2) Hipotesis Alternatif (H_a): ada hubungan antara tingkat literasi kesehatan dengan kepatuhan ibu hamil dalam pemeriksaan antenatal.

Jika nilai p value didapatkan $< 0,05$, maka ada hubungan antara Tingkat literasi dengan kepatuhan ibu hamil dalam pemeriksaan antenatal. Uji Chi square jenis ini yang paling sering digunakan.

c. Chi-Square Test for Homogeneity (Uji Homogenitas)

Uji Chi-Square for Homogeneity digunakan untuk membandingkan distribusi satu variabel kategorik dalam dua atau lebih kelompok yang berbeda. Uji ini bertujuan untuk menentukan apakah distribusi proporsi antara beberapa kelompok populasi adalah sama atau tidak.

Contoh:

Seorang peneliti ingin mengetahui apakah tingkat stres perawat berbeda secara signifikan antara tiga unit kerja: IGD, ICU, dan Bangsal Rawat Inap.

- 1) Jika $p\text{-value} < 0,05$, maka tolak $H_0 \rightarrow$ Artinya distribusi tingkat stres berbeda secara signifikan di antara tiga unit kerja.
- 2) Jika $p\text{-value} > 0,05$, maka terima $H_0 \rightarrow$ Artinya distribusi tingkat stres serupa di ketiga unit kerja, sehingga dianggap homogen.

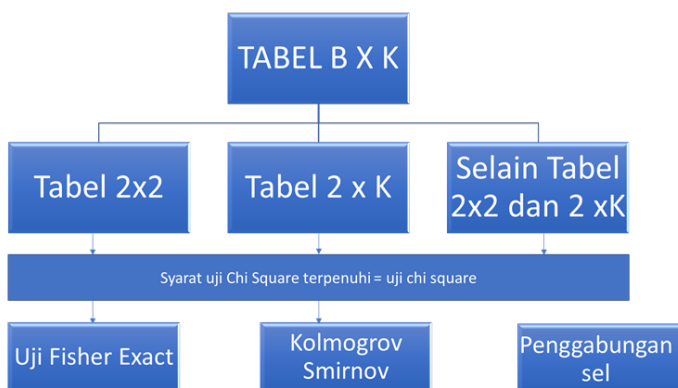
8.2 Asumsi dasar Uji Chi Square

Agar hasil uji Chi-Square valid dan dapat diinterpretasikan dengan benar, terdapat beberapa asumsi yang harus dipenuhi (Dahlan, 2014; Budiarto, 2001; Rustika, 2013):

1. Data harus dalam bentuk kategorik.
2. Variabel yang dianalisis harus dalam bentuk kategorik (nominal atau ordinal), bukan data numerik kontinu.
3. Frekuensi observasi tidak boleh terlalu kecil

4. Idealnya, tidak ada sel dalam tabel kontingensi yang memiliki frekuensi harapan < 5 . Jika ada, uji alternatif seperti Fisher's Exact Test bisa digunakan.
5. Setiap observasi bersifat independen
Setiap individu hanya boleh masuk dalam satu kategori dan tidak boleh dihitung lebih dari satu kali.
6. Ukuran sampel cukup besar
Semakin besar ukuran sampel, semakin valid hasil uji Chi-Square.

Peneliti harus memberikan perhatian khusus pada kriteria uji Chi-Square agar dapat melakukan analisis dan interpretasi hasil secara akurat. Memahami serta memastikan terpenuhinya asumsi dasar merupakan aspek krusial dalam menentukan ketepatan penggunaan uji Chi-Square. Asumsi dasar ini mencakup beberapa aspek penting, seperti independensi antar kategori, distribusi ekspektasi frekuensi minimal yang memadai, serta ukuran sampel yang cukup.



Gambar 8.1: Peta Uji Hipotesis Chi Square (Sumber: Yuandari & Rahman, 2014)

Jika asumsi tersebut tidak terpenuhi, pemilihan uji alternatif yang sesuai. Langkah ini menjadi tahapan penting untuk menjaga validitas dan reliabilitas hasil penelitian. Oleh karena itu, sebelum menerapkan uji Chi-

Square, peneliti harus melakukan pemeriksaan awal terhadap data dan mempertimbangkan metode analisis yang paling sesuai agar kesimpulan yang dihasilkan dapat dipertanggungjawabkan secara ilmiah.

Berdasarkan peta, dapat dijelaskan sebagai berikut:



Gambar 8.2: Pemilihan uji Chi-Square.

8.3 Langkah-langkah uji Chi Square dengan SPSS

Contoh Kasus: Peneliti ingin mengetahui hubungan antara Health literacy dengan Self Care pada Ibu Hamil.

Langkah pertama yang dilakukan adalah menentukan langkah-langkah pemilihan uji.

1. Variabel
 - a. Variabel Independen : Health literacy
 - b. Variabel dependen: Self care Ibu Hamil
2. Jenis Data: Kategorik

Data Health literacy dikelompokkan menjadi dua kategori yaitu:

 - a. Health Literacy Kurang

b. Health Literacy Baik

Data Self Care juga dikelompokkan menjadi dua kategori, yaitu:

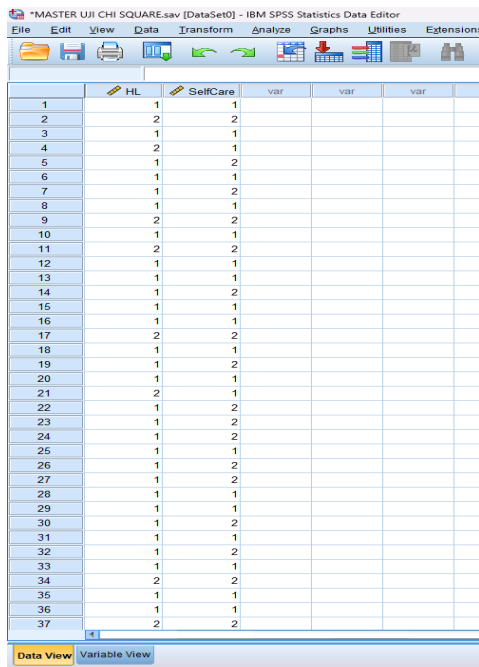
a. 1: Self Care Kurang

b. 2: Self Care Baik

3. Jenis Hipotesis: Komparatif
4. Berpasangan/Tidak: Tidak Berpasangan
5. Pengukuran data: 1 kali pengukuran
6. Jumlah kelompok: 2 kelompok

Langkah-langkah uji Chi-Square:

1. Buka Aplikasi SPSS
2. Input data ke dalam data View



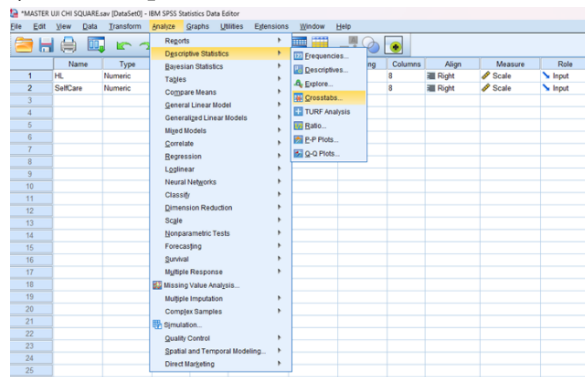
Gambar 8.3: Aplikasi SPSS

3. Sesuaikan nama variabel, value, dan skala data pada variabel view. Isi bagian value sesuai dengan kategori yang sudah dibuat sebelumnya.

	Name	Type	Width	Decimals	Label	Values	Missing	Columns	Align	Measure	Role
1	HL	Numeric	8	0	Health Literacy	{1, Kurang}...	None	8	Right	Scale	Input
2	SelfCare	Numeric	8	0	Self Care	{1, Kurang}...	None	8	Right	Scale	Input
3											
4											
5											
6											
7											
8											
9											
10											
11											

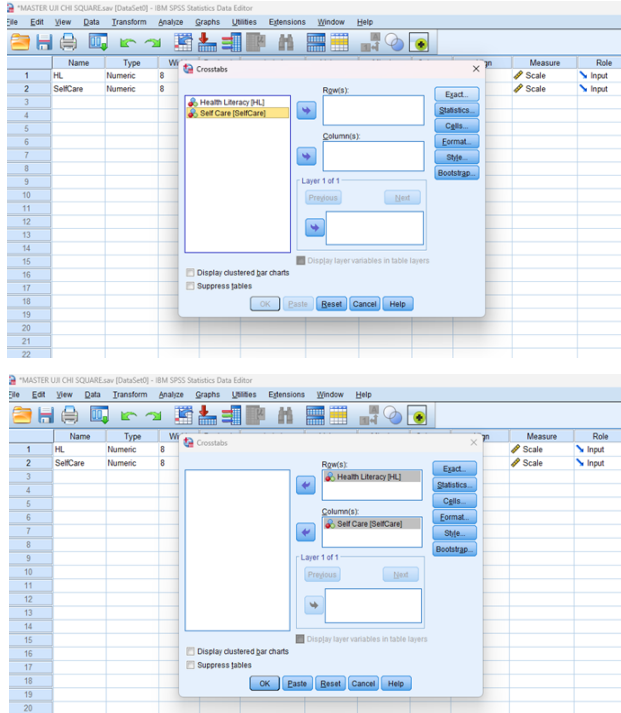
Gambar 8.4: Nama variabel

4. Pilih Analyze-Descriptive Statistic-Crosstabs



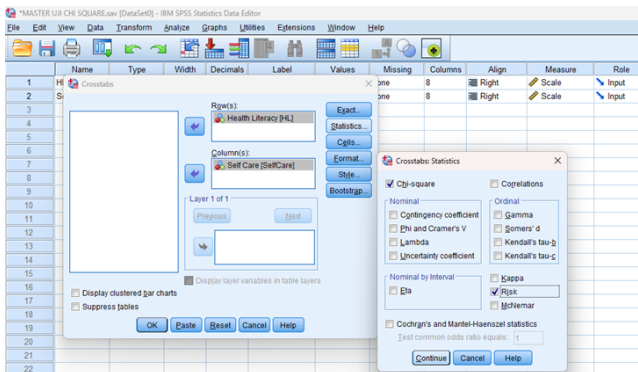
Gambar 8.5: Analyze-Descriptive Statistic-Crosstabs

5. Masukkan variabel independen pada bagian Row (s) dan variabel dependen pada bagian column (s).



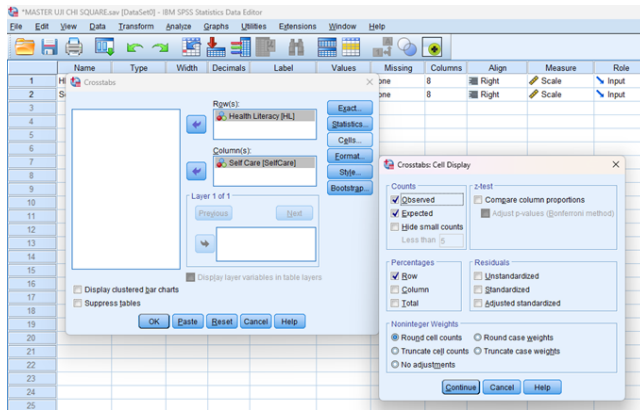
Gambar 8.6: Variabel independen

6. Pilih statistic, centang Chi-Square dan Risk, klik Continue



Gambar 8.7: Chi-Square dan Risk

- Pilih Cells, centang observed, centang coloumn, klik Continue, lalu OK.



Gambar 8.8: Observed

- Output

Tabel 8.1: Self Care Crosstabulation

Health Literacy * Self Care Crosstabulation					
			Self Care		Total
			Kurang	Baik	
Health Literacy	Kurang	Count	45	18	63
		Expected Count	40.2	22.8	63.0
		% within Health Literacy	71.4%	28.6%	100.0%
	Baik	Count	8	12	20
		Expected Count	12.8	7.2	20.0
		% within Health Literacy	40.0%	60.0%	100.0%
Total		Count	53	30	83
		Expected Count	53.0	30.0	83.0
		% within Health Literacy	63.9%	36.1%	100.0%

Tabel 8.2: Chi-Square Test

Chi-Square Tests					
	Value	df	Asymptotic Significance (2- sided)	Exact Sig. (2- sided)	Exact Sig. (1-sided)
Pearson Chi-Square	6.497 ^a	1	.011		
Continuity Correction ^b	5.206	1	.023		
Likelihood Ratio	6.302	1	.012		
Fisher's Exact Test				.016	.012
Linear-by-Linear Association	6.419	1	.011		
N of Valid Cases	83				

a. 0 cells (0.0%) have expected count less than 5. The minimum expected count is 7.23.

Tabel 8.3: Rist Estimate

Risk Estimate			
	Value	95% Confidence Interval	
		Lower	Upper
Odds Ratio for Health Literacy (Kurang / Baik)	3.750	1.314	10.699
For cohort Self Care = Kurang	1.786	1.021	3.123
For cohort Self Care = Baik	.476	.280	.809
N of Valid Cases	83		

9. Interpretasi

- Tabel pertama merupakan tabel silang antara variabel independen dan dependen, menunjukkan proporsi masing-masing variabel independen terhadap dependen
- Tabel kedua merupakan tabel hasil chi square, tabel 2 x 2 tidak ada $ec < 5$, sehingga nilai yang dibaca adalah continuity correction, didapatkan p value $0.023 < \alpha 0.05$, berarti ada

hubungan antara health literacy dengan self care pada ibu hamil.

- c. Pada tabel Risk Estimate diperoleh nilai $OR = 3.750$. Berarti, health literacy yang baik akan berdampak pada self care yang baik sebesar 3.75 kali dibandingkan health literacy yang kurang.

Bab 9

Uji Regresi Linier Sederhana

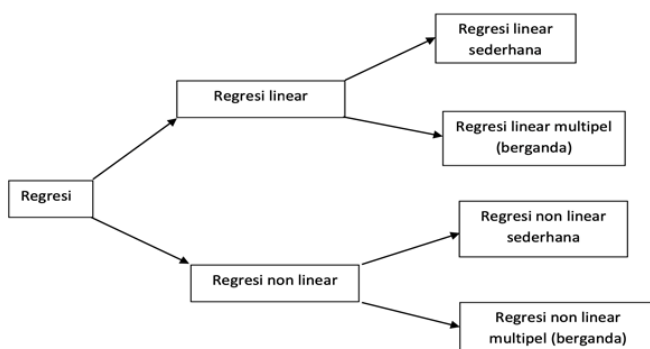
9.1 Uji Regresi

Uji regresi adalah alat uji statistik untuk mengetahui hubungan antar variabel (Montgomery et al., 2021). Peneliti dapat memastikan efek kausal dari satu variabel pada variabel lainnya, misalnya pengaruh tinggi badan orang tua terhadap tinggi badan anak atau efek pemberian asupan makan terhadap status gizi anak.

Untuk mengeksplorasi masalah-masalah tersebut, peneliti mengumpulkan data pada variabel yang mendasari yang diminati dan menggunakan regresi untuk memperkirakan efek kuantitatif dari variabel kausal pada variabel yang memengaruhi (Almuntazah et al., 2021). Peneliti juga biasanya menilai "signifikansi statistik" dari perkiraan hubungan, yaitu, tingkat kepercayaan bahwa hubungan yang sebenarnya dekat dengan perkiraan hubungan.

Pada dasarnya dalam penelitian kesehatan, ada tiga jenis analisis regresi umum yang digunakan yaitu regresi linear, regresi logistik dan regresi Cox

(Marcus et al., 2012). Penggunaan jenis regresi tergantung pada jenis variabel yang kita hadapi (Gogtay et al., 2017; Trianggana, 2020). Regresi linier dapat berupa regresi linier sederhana atau berganda sementara regresi logistik dapat berupa sederhana dan juga berganda (Gogtay et al., 2017; Nuryadi et al., 2017).



Gambar 9.1: Uji Regresi

Perbedaan setiap uji regresi dapat dilihat pada Tabel 9.1. Regresi linear sederhana mempelajari bentuk hubungan dan pengaruh yang diduga bersifat konstan antara satu variabel bebas (X) terhadap variabel terikat (Y).

Tabel 9.1: Tipe Regresi (Gogtay et al., 2017)

Tipe Regresi	Variabel Dependen dan Sifatnya	Variabel Independen dan Sifatnya	Hubungan Antar Variabel
Linier Sederhana	1, Data kontinyu, distribusi normal	1, Data kontinue, distribusi normal	Linier
Linier Multipel	1, Data kontinyu	2 atau lebih, Data dapat kontinyu atau kategorikal	Linier
Logistik	1, Data biner	2 atau lebih, Data dapat kontinyu atau kategorikal	Tidak Perlu Linier

Tipe Regresi	Variabel Dependen dan Sifatnya	Variabel Independen dan Sifatnya	Hubungan Antar Variabel
Logistik Multipel	Non biner	2 atau lebih, Data dapat kontinyu atau kategorikal	Tidak Perlu Linier

Analisis regresi linier sederhana melibatkan dua buah variabel yaitu independent yang memengaruhi dan variabel lain yang dipengaruhi (dependen). Kedua variabel merupakan data kontinyu (kuantitatif) dan garis yang menggambarkan hubungan adalah garis lurus (linear). Linier yang dimaksudkan adalah asumsi yang digunakan bahwa hubungan antara dua variabel yang dianalisis menunjukkan hubungan linier (Bengnga & Ishak, 2018; Nuryadi et al., 2017). Analisis korelasi dengan scatter plot dan garis regresi adalah prasyarat untuk regresi dan kedua analisis sering dilakukan bersama-sama.

Tahapan analisis regresi dilakukan dalam 3 langkah:

1. Menganalisis korelasi (kekuatan dan arah hubungan)
2. Menyesuaikan regresi atau garis kuadrat terkecil, dan
3. Mengevaluasi validitas dan kegunaan model.

Langkah 1: Ini telah dijelaskan dalam artikel tentang analisis korelasi, langkah 2: Menyesuaikan garis regresi (Poni Egistin et al., 2025).

9.2 Persamaan Regresi Linier

Analisis regresi yang baik dapat digunakan untuk mentaksir nilai variabel dependen.

Persamaan regresi linier adalah:

$$y = a + bx$$

- a = intercept (perkiraan besarnya rata-rata nilai y pada saat $x=0$)

- b = slope (kemiringan/perubahan nilai y pada saat nilai x berubah 1 unit)
- y = variabel yang nilainya dipengaruhi variabel lain (dependent variabel)
- x = variabel yang nilainya memengaruhi variabel lain (independent variabel)

dimana:

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

$$a \text{ (koefisien)} = \bar{Y} - b\bar{X}$$

$$\bar{Y} = \frac{\sum Y}{n}$$

$$\bar{X} = \frac{\sum X}{n}$$

Dimana :

$$\bar{Y} : \text{Nilai Y rata-rata} / \bar{Y} = \frac{\sum Y}{n}$$

$$\bar{X} : \text{Nilai X rata-rata} / \bar{X} = \frac{\sum X}{n}$$

n : Jumlah data observasi

Contoh 1: Diketahui perusahaan Sepatu Onitsuka ingin mengetahui hubungan fungsional antara biaya produksi dengan jumlah produksi. Tentukan persamaan regresi dari hubungan tersebut. Dengan $\alpha = 5\%$.

Tabel 9.2: Data besarnya jumlah produksi (X) dengan biaya produksi (Y)

Jumlah Produksi (X)	Biaya Produksi (Y)	XY	X ²	Y ²
20	64	1280	400	4096
16	61	976	256	3721
34	84	2856	1156	7056
23	70	1610	529	4900

27	88	2376	729	7744
32	92	2944	1024	8464
18	72	1296	324	5184
22	77	1694	484	5929
192	608	15032	4902	47094

Penyelesaian:

1. Tetapkan persamaan regresi

$$Y = a + bX$$

- Y = taksiran biaya produksi pada jumlah tertentu
- a = biaya produksi apabila tidak berproduksi ($X=0$)
- b = perubahan biaya produksi apabila terjadi perubahan satu unit output
- X = jumlah produksi

2. Buatlah Tabel Bantu

Tabel 9.3: Tabel Bantu Data

Jumlah Produksi (X)	Biaya Produksi (Y)	XY	X ²	Y ²
20	64	1280	400	4096
16	61	976	256	3721
34	84	2856	1156	7056
23	70	1610	529	4900
27	88	2376	729	7744
32	92	2944	1024	8464
18	72	1296	324	5184
22	77	1694	484	5929
192	608	15032	4902	47094
$\Sigma X = 192$	$\Sigma Y = 608$	$\Sigma XY = 15032$	$\Sigma X^2 = 4902$	$\Sigma Y^2 = 47094$

3. Menghitung Persamaan Regresi Linier

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

$$b = \frac{8(15032) - (192)(608)}{8(4902) - (192)^2}$$

$$b = \frac{3520}{2325} = 1,49$$

$$a = \bar{Y} - b\bar{X}$$

$$\bar{Y} = \frac{\sum Y}{n} = \frac{608}{8} = 76$$

$$\bar{X} = \frac{\sum X}{n} = \frac{192}{8} = 24$$

$$a = \bar{Y} - b\bar{X}$$

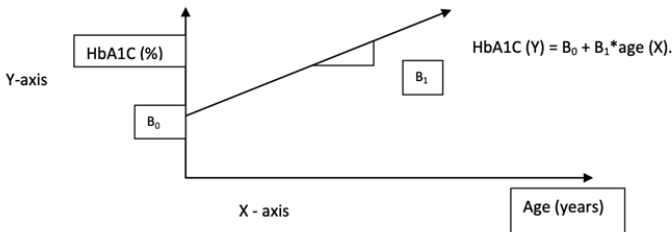
$$= 76 - 1,5(24)$$

$$= 40$$

$$Y = a + bX$$

$$= 40 + 1,5X$$

Sehingga diketahui persamaan regresi linier yang terbentuk adalah $Y = 40 + 1,5x$. Pada gambar 9.2 menunjukkan hasil regresi pada penelitian lainnya, yaitu seorang peneliti yang ingin mengetahui hubungan usia dengan kadar HbA1C.



Gambar 9.2: Analisis Regresi Hubungan Usia dengan Kadar HbA1C

Berdasarkan gambar 9.2, menunjukkan hubungan antara usia dan tingkat HbA1C, persamaan akan diberikan oleh tingkat HbA1C (y) = $B_0 + B_1$ *usia (x). Intercept (B_0) adalah nilai Y atau variabel dependen (HbA1C, dalam

hal ini), ketika nilai variabel prediktor adalah nol (usia). Pada kenyataannya usia tidak pernah bisa nol, jadi ini adalah nilai HbA1C yang ada terlepas dari usia. Persamaan $y = B_0 + B_1x$ menggambarkan regresi linier sederhana (Gogtay et al., 2017).

9.3 Koefisien Korelasi

Persamaan regresi yang terbentuk tersebut belum dapat menyimpulkan hasil seberapa jauh dapat mentaksir hubungan variabel independent terhadap variabel dependent, perlu dilakukan beberapa pengujian yaitu pengujian terhadap koefisien regresi, pengujian terhadap varians dan penentuan keeratan hubungan. Koefisien korelasi (r) menunjukkan tinggi rendahnya hubungan antara dua variabel. Besarnya koefisien korelasi berkisar dari $-1 \leq r \leq +1$.

Koefisien korelasi tersebut menunjukkan semakin mendekati 1 menunjukkan hubungan yang kuat atau sempurna. Jika $r=0$ maka menunjukkan tidak ada hubungan. Koefisien korelasi dapat memiliki nilai positif atau negatif yang menunjukkan arah hubungan setiap variabelnya. Arah hubungan positif terjadi ketika kenaikan nilai x akan meningkatkan nilai y , sedangkan arah negatif ketika kenaikan nilai x akan menurunkan nilai y (Dahlan, 2011; Nuryadi et al., 2017).

Tabel 9.4: Menurut Guilford, kriteria untuk menentukan keeratan hubungan antara dua variabel

Nilai Koefisien Korelasi	Kriteria
0,00–0,20	Sangat Lemah
0,20–0,40	Lemah
0,40–0,70	Moderate/ Sedang
0,70–0,90	Kuat
0,90–0,99	Sangat Kuat
1	Sempurna

Menetapkan nilai koefisien korelasi dengan formula sebagai berikut:

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

Contoh 2: Berdasarkan Contoh 1 dimana diketahui Perusahaan Sepatu Onitsuka ingin mengetahui hubungan fungsional antara biaya produksi dengan jumlah produksi.

Tabel 9.5: Berikut data besarnya jumlah produksi (X) dengan biaya produksi (Y)

Jumlah Produksi (X)	Biaya Produksi (Y)
20	64
16	61
34	84
23	70
27	88
32	92
18	72
22	77
192	608

Penyelesaian:

1. Menetapkan Persamaan Regresi:

$$Y = a + bX$$

- Y = taksiran biaya produksi pada jumlah tertentu
- a = biaya produksi apabila tidak berproduksi (X=0)
- b = perubahan biaya produksi apabila terjadi perubahan satu unit output
- X = jumlah produksi

2. Buat Tabel bantu

Tabel 9.6: Tabel Bantu Data

Jumlah Produksi (X)	Biaya Produksi (Y)	XY	X ²	Y ²
20	64	1280	400	4096
16	61	976	256	3721
34	84	2856	1156	7056
23	70	1610	529	4900
27	88	2376	729	7744
32	92	2944	1024	8464
18	72	1296	324	5184
22	77	1694	484	5929
192	608	15032	4902	47094
$\Sigma X=192$	$\Sigma Y=608$	$\Sigma XY=15032$	$\Sigma X^2=4902$	$\Sigma Y^2=47094$

3. Menghitung Persamaan Regresi Linier

$$\begin{aligned}
 b &= \frac{n \Sigma XY - \Sigma X \Sigma Y}{n \Sigma X^2 - (\Sigma X)^2} \\
 b &= \frac{8(15032) - (192)(608)}{8(4902) - (192)^2} \\
 b &= \frac{3520}{2325} = 1,49 \\
 a &= \bar{Y} - b\bar{X} \\
 \bar{Y} &= \frac{\Sigma Y}{n} = \frac{608}{8} = 76 \\
 \bar{X} &= \frac{\Sigma X}{n} = \frac{192}{8} = 24 \\
 a &= \bar{Y} - b\bar{X} \\
 &= 76 - 1,5(24) \\
 &= 40 \\
 Y &= a + bX \\
 &= 40 + 1,5X
 \end{aligned}$$

4. Menetapkan nilai koefisien korelasi

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

$$r = \frac{8(15032) - (192)(608)}{\sqrt{8(4902) - (192)^2} \sqrt{8(47094) - (608)^2}}$$

$$r = \frac{3520}{4083} = 0,86$$

$$r_{\text{tabel}} = r(\alpha; 9 - 2)$$

$$r_{\text{tabel}} = r(5\%, 6-2) = 0,6664$$

5. Kesimpulan:

- $r_{\text{hit}} > r_{\text{tabel}}$ artinya H_0 ditolak atau terdapat hubungan biaya produksi dengan jumlah output yang dihasilkan.
- Keeratan hubungan antara biaya produksi dengan jumlah output yang dihasilkan adalah 0,86 atau 86% atau memiliki keeratan hubungan yang kuat.

Contoh 3: Seorang dokter Puskesmas ingin mengetahui hubungan antara jumlah pengunjung dengan obat tetrasiklin yang digunakan. Diketahui $\alpha=5\%$. Untuk itu diambil 5 sampel hari kerja dengan hasil sebagai berikut:

Tabel 9.7: Sampel Hari Kerja

Hari	Jumlah Kunjungan	Jumlah tetrasiklin
1	60	150
2	50	140
3	70	205
4	40	130
5	60	165
6	70	210

Penyelesaian:

1. Buat Persamaan

$$Y = a + bX$$

- Y = taksiran obat tetrasiklin yang digunakan
- a = obat tetrasiklin yang digunakan apabila tidak berproduksi (X=0)
- b = perubahan penggunaan obat tetrasiklin apabila terjadi perubahan satu unit output
- X = jumlah pengunjung

2. Buat Tabel bantu

Tabel 9.8: Tabel Bantu Data

Hari	Jumlah Kunjungan (X)	Jumlah tetrasiklin (Y)	XY	x ²	y ²
1	60	150	9000	3600	22500
2	50	140	7000	2500	19600
3	70	205	14350	4900	42025
4	40	130	5200	1600	16900
5	60	165	9900	3600	27225
6	70	210	14700	4900	44100
Σ	350	1000	60150	21100	172350

3. Menetapkan nilai konstanta dan koefisien

$$\begin{aligned}
 b &= \frac{n \sum XY - \sum X \sum Y}{n \sum X^2 - (\sum X)^2} \\
 b &= \frac{6(60150) - (350)(1000)}{6(21100) - (350)^2} \\
 b &= \frac{360900 - 350000}{126600 - 122500} = \frac{10900}{4100} = 2,66 \\
 a &= \bar{Y} - b\bar{X} \\
 \bar{Y} &= \frac{\sum Y}{n} = \frac{1000}{6} = 166,67 \\
 \bar{X} &= \frac{\sum X}{n} = \frac{350}{6} = 58,33
 \end{aligned}$$

$$\begin{aligned}
 a &= \bar{Y} - b\bar{X} \\
 &= 166,67 - 2,66(58,33) \\
 &= 166,67 - 155,16 \\
 &= 11,51
 \end{aligned}$$

$$Y = a + bX$$

$$Y = 11,51 + 2,66X \text{ (Persamaan yang terbentuk)}$$

4. Menetapkan nilai koefisien korelasi

$$\begin{aligned}
 r &= \frac{n\sum xy - (\sum x)(\sum y)}{\sqrt{\{n\sum y^2 - (\sum y)^2\}}\sqrt{\{n\sum x^2 - (\sum x)^2\}}} \\
 r &= \frac{6(60150) - (350)(1000)}{\sqrt{6(172350) - (1000)^2}\sqrt{6(21100) - (350)^2}} \\
 r &= \frac{360900 - 350000}{\sqrt{6(172350) - (1000)^2}\sqrt{6(350)^2 - (21100)}} \\
 &= \frac{10900}{144520} = 0,07
 \end{aligned}$$

5. Mencari r tabel

$$r \text{ tabel} = r(\alpha; n - 2)$$

$$r \text{ tab} = r_{5\%, 6-2} = 0,8114$$

6. Kesimpulan:

- $r \text{ hit} < r \text{ tabel}$ maka H_0 diterima atau tidak terdapat perbedaan signifikan.
- Keeratan hubungan antara jumlah kunjungan dengan penggunaan obat tetrasiklin sebesar 0,07 atau korelasi yang sangat lemah.

9.4 Penaksiran Nilai Variabel Terikat

Analisis regresi dapat membantu peneliti dalam mengestimasi nilai rata-rata variabel dependen berdasarkan variabel independennya (Aini & Irawati, 2019). Penaksiran nilai variabel dependen menggunakan persamaan regresi yang diperoleh dengan memasukkan nilai variabel independennya dalam persamaan regresi yang diperoleh.

Contoh 4: Buatlah taksiran biaya total pada produksi 100 unit Sepatu Onitsuka dengan menggunakan persamaan regresi $Y = 40 + 1,5X$.

Dengan memasukkan jumlah output ($x = 100$) ke dalam persamaan tersebut, maka taksiran biaya produksi (Y).

$$\begin{aligned} Y &= 40 + 1,5 (100) \\ &= 40 + 150 \end{aligned}$$

Jadi biaya produksi yang harus dikeluarkan untuk memproduksi sebanyak 100 unit ditaksir sebesar 190.

Bab 10

Uji Regresi Logistik Sederhana

10.1 Konsep Uji Regresi Logistik

Regresi merupakan salah satu analisis data yang berkaitan dengan penggambaran hubungan antara variabel dependen dan satu atau lebih variabel independen (prediktor). Regresi adalah salah satu cara paling umum untuk memperkirakan parameter sebenarnya dengan cara yang sebisa mungkin tidak bias. Artinya, regresi biasanya digunakan untuk membuat model data populasi yang akurat.

Proses pemodelan regresi adalah metode yang digunakan untuk memahami dan mengendalikan ketidakpastian yang melekat dalam memperkirakan parameter sebenarnya dari distribusi yang menggambarkan data populasi. Hal ini penting karena prediksi yang dibuat dari suatu model diasumsikan berasal dari populasi tersebut (Hosmer & Lemeshow, 2000).

Contoh umum dari pemodelan, misalnya model regresi linier biasa dimana variabel dependen diasumsikan bersifat kontinu (skala interval atau rasio), seperti pendapatan, tekanan darah atau skor ujian. Namun, sering kali

variabel dependen bersifat diskrit, yang hanya memiliki dua atau lebih nilai yang mungkin, seperti variabel status sosial-ekonomi sering kali bersifat kategori (nominal dan ordinal), bukan skala interval dan rasio.

Sedangkan dalam banyak kasus, penelitian difokuskan pada model yang variabel dependennya bersifat kategori. Misalnya, variabel dependennya berkategori 'sakit' atau 'tidak' yang kemudian menarik untuk diselidiki bagaimana variabel ini berkaitan dengan jenis kelamin, usia, kelompok etnis, dan lain-lain. Dalam kasus ini, tidak dapat dilakukan regresi linier karena banyak asumsi di teknik ini yang tidak bisa terpenuhi. Sebaliknya, akan dianalisis dengan menggunakan regresi logistik.

Regresi logistik telah berkembang pesat menjadi metode analisis standar dalam berbagai bidang selama beberapa dekade terakhir. Sejak diterima pertama kali dalam penelitian epidemiologi, metode ini sekarang umum digunakan di berbagai bidang tidak hanya terbatas pada penelitian biomedis, bisnis dan keuangan, kriminologi, ekologi, teknik, kebijakan kesehatan, linguistik dan biologi (Hilbe, 2015).

Selain konsep dan pengertian regresi logistik, pada bab ini akan dibahas tentang kegunaan atau tujuan regresi logistik, asumsi dan model regresi logistik, langkah-langkah uji regresi logistik sederhana dan contohnya, serta interpretasi hasil uji regresi logistik. Penyelesaian contoh studi kasus dianalisis dengan menggunakan program komputer SPSS.

Sebelum membahas lebih dalam mengenai regresi logistik, penting untuk memahami bahwa tujuan analisis menggunakan metode ini adalah sama dengan setiap teknik pembuatan model yang digunakan dalam statistik, yaitu untuk menemukan model terbaik, model yang paling tepat, dan paling hemat, tetapi secara biologis juga masuk akal untuk menggambarkan hubungan antara suatu variabel dependen (terikat, akibat, atau respon) dan sekumpulan variabel independen (bebas, penyebab, atau prediktor). Variabel independen ini sering disebut kovariat (Hosmer & Lemeshow, 2000).

Regresi logistik terutama digunakan untuk memodelkan variabel biner atau dikotom (kode 0,1) berdasarkan satu atau lebih variabel lain, yang disebut prediktor. Variabel biner yang dimodelkan umumnya disebut sebagai variabel respon, variabel terikat, atau variabel dependen. Secara umum, regresi logistik diartikan sebagai salah satu pendekatan atau model matematis yang digunakan untuk mempelajari hubungan / pengaruh antara satu atau beberapa variabel independen dengan satu atau beberapa variabel dependen kategorik yang bersifat dikotom atau binari.

Variabel kategorik yang dikotom adalah variabel yang hanya mempunyai dua nilai (0 dan 1), misalnya sakit diberi kode 1 dan sehat diberi kode 0, atau kategori antara BBLR (Berat Badan Lahir Rendah)-BBL normal, merokok-tidak merokok, dan lain-lain (Fauziyah, 2019). Regresi logistik terdiri atas regresi logistik sederhana dan regresi logistik berganda.

Analisis regresi logistik sederhana digunakan jika hanya memiliki satu variabel prediktor, sedangkan regresi logistik berganda digunakan jika memiliki lebih dari satu variabel prediktor. Regresi logistik adalah sebuah pendekatan untuk membuat model prediksi seperti halnya regresi linier atau yang biasa disebut dengan istilah Ordinary Least Squares (OLS) regression. Perbedaannya adalah pada regresi logistik, peneliti memprediksi variabel terikat yang berskala dikotomi, sedangkan regresi linier memprediksi variabel terikat berskala kontinu (Husnul, 2018).

10.2 Tujuan Regresi Logistik

Regresi logistik sebagai salah jenis uji statistik bertujuan untuk menguji apakah probabilitas terjadinya variabel terikat (dependen) dapat diprediksi dengan variabel bebas (independen).

Sebagai contoh, studi yang analisisnya pada umumnya menggunakan regresi logistik, antara lain:

1. Seorang peneliti akan menentukan seberapa besar probabilitas seorang pelajar merokok berdasarkan data keterpaparan terhadap iklan rokok, serta kebiasaan merokok orang tua, teman, dan guru.
2. Seorang dokter akan menganalisis probabilitas atau kecenderungannya salah seorang pasien terjangkit penyakit jantung dengan memprediksinya menggunakan tekanan darah, kadar kolesterol, pola konsumsi, serta aktivitas fisik pasien.

Adapun kegunaan dari uji regresi logistik (Stang, 2014), antara lain:

1. Meramalkan terjadinya variabel dependen pada individu berdasarkan nilai-nilai sejumlah variabel prediktor yang ada pada individu tersebut.
2. Mengukur pengaruh variabel independen terhadap variabel dependen.
3. Dapat mengkonversi koefisien regresi (B_i) menjadi Odds ratio (OR) dengan rumus $OR = \exp(B_i)$
4. Dapat menaksir probabilitas individu untuk sakit atau mengalami event berdasarkan nilai-nilai sejumlah variabel prediktor.

10.3 Asumsi Uji Regresi Logistik

Berikut beberapa syarat atau asumsi yang harus dipenuhi dalam regresi logistik (Hilbe, 2015; Basuki, 2020), antara lain:

1. Regresi logistik tidak membutuhkan hubungan linier antara variabel independen dengan variabel dependen, namun prediktor memiliki hubungan yang signifikan dengan respon.

2. Variabel independen tidak memerlukan asumsi multivariate normality seperti pada regresi linier.
3. Asumsi homokedastisitas juga tidak diperlukan
4. Variabel independen tidak perlu diubah ke dalam bentuk metrik (skala interval atau skala ratio).
5. Variabel dependen harus bersifat dikotomi (2 kategori, misal: sehat dan sakit atau merokok dan tidak merokok)
6. Variabel independen tidak harus memiliki keragaman yang sama antar kelompok variabel.
7. Kategori dalam variabel independen harus terpisah satu sama lain atau bersifat eksklusif. Artinya, prediktor tidak berkorelasi satu sama lain.
8. Regresi logistik dapat menyeleksi hubungan karena menggunakan pendekatan non linier log transformasi untuk memprediksi odds ratio. Odd dalam regresi logistik sering dinyatakan sebagai probabilitas.

10.4 Model Regresi Logistik

Berikut ini model regresi logistik sederhana (Stang, 2014; Field, 2017):

$$f(z) = \frac{1}{1 + e^{-(\beta_0 + \beta_1 X_1)}}$$

Keterangan:

$f(z)$ = Probabilitas terjadinya suatu masalah kesehatan

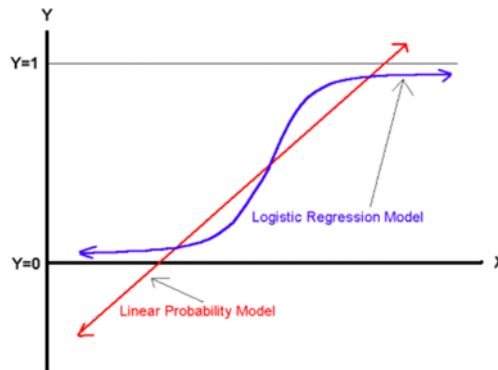
β_0 = Intercept

β_1 = Koefisien regresi logistik

X_1 = Variabel Independen

e = Bilangan log natural (2,72)

Berbeda dengan regresi linier, regresi logistik tidak mengasumsikan hubungan antara variabel independen dan dependen secara linier. Regresi logistik merupakan regresi non linier di mana model yang ditentukan akan mengikuti pola kurva seperti gambar di bawah ini.



Gambar 10.1: Grafik Regresi Logistik

Grafik fungsi logistik berupa kurve sigmoid yang menyerupai huruf 'S' mencerminkan efek satu atau lebih faktor risiko disebut prediktor dalam menyebabkan suatu penyakit/event disebut respon.

10.5 Langkah Uji Regresi Logistik

Langkah-langkah uji regresi logistik sederhana akan diuraikan dengan menggunakan contoh penelitian sebagai berikut.

Contoh 1: Seorang peneliti ingin mengetahui pengaruh keterpaparan iklan rokok (terpapar atau tidak terpapar) terhadap terjadinya perilaku merokok pada remaja (merokok dan tidak merokok). Peneliti mengambil sampel secara acak sebesar 50 orang. Data dapat dilihat pada tabel 10.1 berikut.

Tabel 10.1: Pengaruh Paparan Iklan Rokok terhadap Perilaku Merokok

No	Iklan	Merokok	No	Iklan	Merokok	No	Iklan	Merokok
1	0	1	18	0	1	35	1	0
2	1	1	19	0	1	36	1	0
3	0	1	20	0	1	37	1	0
4	1	1	21	0	1	38	0	0
5	0	1	22	0	1	39	1	0
6	0	1	23	1	0	40	1	0
7	0	1	24	1	0	41	1	0
8	1	1	25	1	0	42	1	0
9	1	1	26	1	0	43	0	0
10	0	1	27	0	0	44	0	0
11	0	1	28	0	0	45	1	0
12	0	1	29	0	0	46	1	0
13	0	1	30	1	0	47	1	0
14	1	1	31	1	0	48	0	0
15	1	1	32	1	0	49	1	0
16	1	1	33	1	0	50	1	0
17	0	1	34	1	0			

Keterangan:

Iklan : 1 = Terpapar; 0 = Tidak Terpapar

Merokok : 1 = Merokok; 0 = Tidak Merokok

Berikut ini adalah langkah menentukan uji statistik yang akan digunakan:

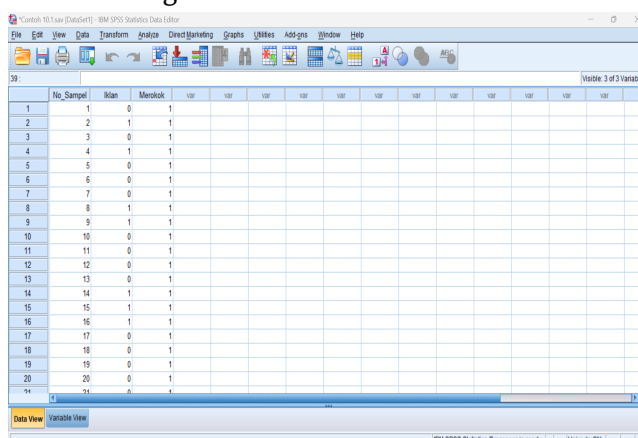
1. Mengidentifikasi tujuan penelitian, yaitu untuk melihat pengaruh antara variabel independen (paparan iklan rokok) dengan variabel dependen (perilaku merokok remaja)
2. Mengidentifikasi skala variabel independen paparan iklan rokok yaitu nominal, karena paparan iklan rokok diukur dalam bentuk kategori terpapar dan tidak terpapar.

3. Mengidentifikasi skala variabel dependen, yaitu nominal dikotom yang diukur dalam dua kategori yaitu merokok dan tidak merokok.
4. Mengidentifikasi jumlah variabel independen, yaitu berjumlah satu variabel paparan iklan rokok.
5. Sehingga dapat disimpulkan, karena mempunyai tujuan untuk melihat pengaruh, skala variabel dependen nominal dikotom dan jumlah variabel independen hanya satu, maka uji yang sesuai adalah regresi logistik sederhana.

10.6 Contoh Uji Regresi Logistik Sederhana

Langkah-langkah analisis regresi logistik sederhana dengan menggunakan program komputer SPSS berdasarkan Contoh 1 adalah sebagai berikut:

1. Buka program SPSS dengan klik SPSS
2. Masukkan data sebagai berikut:

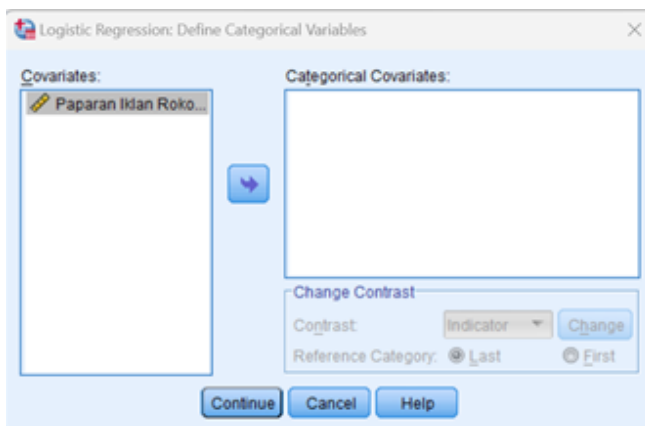


The screenshot shows the IBM SPSS Statistics Data Editor window. The data is as follows:

No_Sampel	Iklan	Merokok
1	1	0
2	2	1
3	3	0
4	4	1
5	5	0
6	6	0
7	7	0
8	8	1
9	9	1
10	10	0
11	11	0
12	12	0
13	13	0
14	14	1
15	15	1
16	16	1
17	17	0
18	18	0
19	19	0
20	20	0
21	21	0

Gambar 10.2: Aplikasi SPSS

5. Karena variabel paparan iklan rokok adalah data kategori, maka klik Categorical, pindahkan variabel paparan iklan rokok ke kolom Categorical Covariates dengan mengklik tanda panah kanan, pilih First.



Gambar 10.5: Categorical Covariates

6. Klik Continue, kemudian klik OK

10.7 Interpretasi Hasil Uji Regresi Logistik Sederhana

Setelah dilakukan langkah-langkah uji regresi logistik seperti yang telah diuraikan pada bagian sebelumnya, maka akan muncul hasil atau output SPSS sebagai berikut:

Block 1: Method = Enter

Block 1 adalah tahap memasukkan variabel independen ke dalam model penelitian. Cara memasukkan variabel independen ini menggunakan

metode Enter yaitu seluruh variabel independen secara bersama-sama dimasukkan ke dalam model.

Tabel 10.2: Omnibus Test dan Model Summary

Omnibus Tests of Model Coefficients				
		Chi-square	df	Sig.
Step 1	Step	9.581	1	.002
	Block	9.581	1	.002
	Model	9.581	1	.002

Model Summary			
Step	-2 Log likelihood	Cox & Snell R Square	Nagelkerke R Square
1	59.012 ^a	.174	.234

a. Estimation terminated at iteration number 4 because parameter estimates changed by less than .001.

Output tabel Model Summary di atas memperlihatkan bahwa model dengan memasukkan variabel independen ternyata terjadi perbedaan dalam penaksiran parameternya (-2 Log likelihood) sebesar 59,012 poin. Jika dilihat nilai R Square sebesar 17,4% (Cox & Snell) dan 23,4% (Nagelkerke) nilai-nilai ini hampir mirip interpretasinya dengan nilai determinasi dalam regresi linier biasa.

Tabel 10.3: Variables In the Equation

Variables in the Equation							
		B	S.E.	Wald	df	Sig.	Exp(B)
Step 1 ^a	Iklan(1)	1.861	.632	8.656	1	.003	6.429
	Constant	-1.099	.436	6.336	1	.012	.333

a. Variable(s) entered on step 1: Iklan.

Output tabel Variabel in The Equation ini merupakan tabel utama dari hasil analisis data dengan menggunakan regresi logistik. Nilai koefisien regresi logistik diperoleh "B" 0 = -1,099 dan "B" 1 = 1,861 dengan nilai p untuk variabel independen paparan iklan rokok = 0,003 ($p < 0,05$). H_a ini berarti

ada pengaruh paparan iklan rokok terhadap terjadinya perilaku merokok pada remaja. Berdasarkan tabel tersebut model persamaan regresi logistik yang terbentuk adalah sebagai berikut.

$$f(z) = \frac{1}{1 + e^{-(\beta_0 + \beta_1 X_1)}}$$

$$f(z) = \frac{1}{1 + e^{-(-1,099 + 1,861(1))}}$$

Hasil persamaan regresi logistik di atas tidak bisa langsung diinterpretasikan dari nilai koefisiennya seperti dalam regresi linier biasa. Interpretasi bisa dilakukan dengan melihat nilai $\exp(B) = OR$ atau nilai eksponen dari koefisien persamaan regresi yang terbentuk. Diperoleh nilai $OR = 6,429$ dapat diartikan bahwa remaja yang terpapar iklan rokok mempunyai kecenderungan atau berisiko 6,429 kali untuk berperilaku merokok dibandingkan dengan remaja yang tidak terpapar iklan rokok.

Dengan model persamaan regresi logistik ini juga bisa diprediksi probabilitas seseorang berperilaku merokok jika orang tersebut terpapar iklan rokok, yaitu:

$$f(z) = \frac{1}{1 + e^{-(-1,099 + 1,861(1))}}$$

$$f(z) = \frac{1}{1 + e^{-(0,762)}}$$

$$f(z) = 0,682 = 68,2\%$$

Artinya, remaja yang terpapar iklan rokok memiliki probabilitas untuk berperilaku merokok adalah sebesar 67,2%.

Sebaliknya, probabilitas merokok seorang remaja jika tidak terpapar iklan rokok dapat diprediksi dengan persamaan sebagai berikut:

$$f(z) = \frac{1}{1 + e^{-(-1,099 + 1,861(0))}}$$

$$f(z) = 0,249 = 24,9\%$$

Dengan demikian, probabilitas seorang remaja untuk berperilaku merokok jika dia tidak terpapar iklan rokok hanya sebesar 24,9%.

Bab sebelumnya telah membahas mengenai berbagai analisis bivariat, maka teknik analisis berikutnya yang perlu dipahami adalah analisis multivariat.

Analisis multivariat merupakan teknik analisis perluasan atau pengembangan dari analisis bivariat. Jika analisis bivariat melihat hubungan atau keterkaitan antar dua variabel, maka teknik analisis multivariat bertujuan melihat atau mempelajari hubungan beberapa variabel (lebih dari satu variabel) independen dengan satu atau beberapa variabel dependen (umumnya satu variabel dependen).

Selain regresi linier, salah satu jenis uji statistik untuk analisis multivariat yang digunakan untuk melihat hubungan beberapa variabel (dua atau lebih variabel) independen dengan variabel dependen adalah uji regresi logistik berganda.

Model Regresi Logistik Berganda:

$$f(z) = \frac{1}{1 + e^{-(\beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \dots + \beta_n X_n)}}$$

Keterangan:

- $f(z)$ = Probabilitas terjadinya suatu masalah kesehatan
- β_0 = Intercept
- β_n = Koefisien regresi logistik
- X_n = Variabel Independen
- e = Log natural (2,72)

Contoh 2:

Sebuah studi cross sectional yang melakukan analisis data sekunder dari hasil Global Youth Tobacco Survey (GYTS) di Indonesia tahun 2014. Adapun populasi studi ini adalah populasi survei GYTS, yaitu seluruh pelajar dari semua sekolah di Indonesia pada kelas 7, 8 dan 9 dengan jumlah sampel sebesar 5093 orang. Tindakan merokok adalah variabel dependen dalam penelitian ini, sedangkan variabel independennya terdiri atas paparan iklan rokok, pengetahuan, dan sikap.

Penelitian ini bertujuan untuk mengetahui faktor yang paling dominan berhubungan antara iklan rokok, pengetahuan, dan sikap dengan perilaku merokok pelajar di Indonesia. Tindakan merokok terdiri atas 2 kategori, yaitu 1 = merokok dan 0 = tidak merokok. Paparan iklan rokok dengan 2

kategori, yaitu 1 = terpapar dan 0 = tidak terpapar. Pengetahuan juga terdiri atas 2 kategori, yaitu 1 = kurang dan 0 = cukup, serta variabel sikap dengan 2 kategori, yaitu 1 = negatif dan 0 = positif (Jamal et al., 2024).

Berikut ini cara untuk menentukan uji statistik yang digunakan adalah uji regresi logistik:

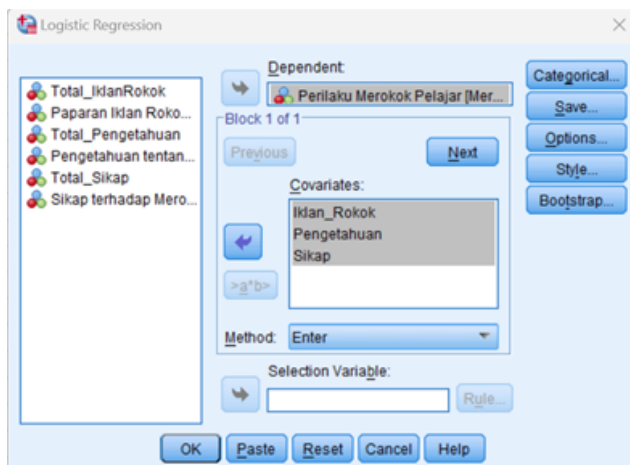
1. Identifikasi tujuan penelitian, yaitu untuk melihat faktor yang paling dominan berhubungan antara variabel independen (paparan iklan rokok, pengetahuan, dan sikap) dengan variabel dependen (tindakan merokok).
2. Identifikasi skala variabel independen, seluruhnya berskala nominal.
3. Identifikasi skala variabel dependen, yaitu nominal dikotom diukur dalam dua kategori yaitu merokok dan tidak merokok.
4. Identifikasi jumlah variabel independen, yaitu tiga variabel independen.
5. Sehingga dapat disimpulkan, karena mempunyai tujuan untuk melihat faktor yang paling dominan berhubungan antara variabel independen terhadap variabel dependen, skala variabel dependen nominal dikotom dan jumlah variabel independen lebih dari satu, yaitu tiga variabel, maka uji yang sesuai adalah regresi logistik berganda.

Langkah-langkah analisis regresi logistik dengan menggunakan program komputer SPSS adalah sebagai berikut:

1. Buka program SPSS atau klik file data SPSS GYTS yang telah diolah
2. Klik menu Analyze, pilih Regression, kemudian pilih Binary Logistic

3. Kemudian muncul kotak dialog Logistic Regression, masukkan variabel tindakan merokok sebagai variabel dependen pada kotak Dependent dan variabel paparan iklan rokok, pengetahuan, dan sikap sebagai variabel independen di kotak Covariates. Kemudian pilih metode yang Enter.

Selain metode Enter, terdapat pula metode alternatif pemilihan variabel secara bertahap (stepwise), yang terdiri atas metode forward dan backward. Metode forward dilakukan dengan cara memasukkan prediktor secara bertahap berdasarkan korelasi parsial terbesar ke dalam model, dengan kriteria nilai p variabel $< 0,05$. Sebaliknya, metode backward atau metode eliminasi langkah mundur dimulai dengan memasukkan semua variabel ke dalam model, kemudian secara bertahap mengeliminasi atau mengeluarkan satu per satu prediktor dari model yang memiliki kriteria nilai p variabel $> 0,1$.

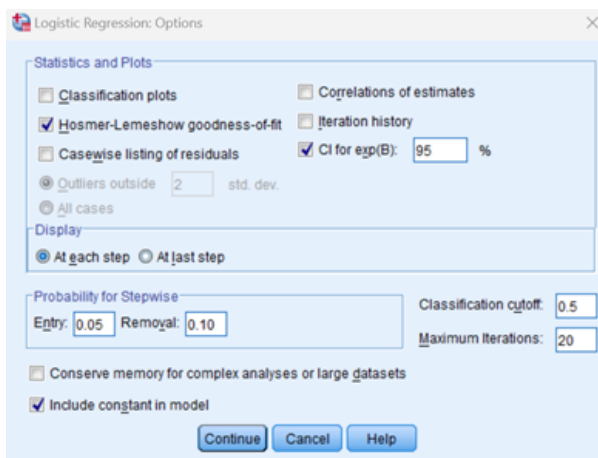


Gambar 10.6: Kotak dialog Logistic Regression

4. Karena variabel paparan iklan rokok, pengetahuan, dan sikap adalah data kategori, maka klik Categorical, pindahkan variabel

paparan iklan rokok, pengetahuan, dan sikap ke kolom Categorical Covariates, pilih First.

- Klik Continue, kemudian klik Options akan muncul sebagai berikut. Pada Statistic and Plots centang Hosmer-Lemeshow goodness-of-fit dan CI for exp (B): 95%.



Gambar 10.7: Statistic and Plots centang Hosmer-Lemeshow goodness-of-fit dan CI for exp (B): 95%.

- Klik Continue, kemudian klik OK

Maka diperoleh hasil analisis data regresi logistik berganda sebagai berikut.

Tabel 10.4: Hasil Uji Regresi Logistik Tindakan Merokok Pelajar Indonesia Tahun 2014 (Jamal et al., 2024)

Nama Variabel	B	SE	Wald	Sig.	Exp(B)	95% CI	
						LL	UL
Paparan Iklan Rokok	0,851	0,114	56,033	0,000	2,341	1,874	2,925
Pengetahuan	0,047	0,104	0,204	0,651	1,048	0,855	1,284
Sikap	1,654	0,093	319,421	0,000	5,229	4,361	6,269
Constant	-3,400	0,125	742,889	0,000	0,033		

Tabel 10.4 merupakan hasil analisis uji regresi logistik untuk melihat variabel independen yang paling dominan berhubungan dengan tindakan merokok. Semua variabel independen kecuali pengetahuan $\text{Exp}(B) = 1,048$ (95% CI: 0,855-1,284) berhubungan signifikan dengan perilaku merokok. Hasil analisis juga menemukan bahwa sikap sebagai faktor yang memiliki hubungan paling dominan dengan tindakan merokok pelajar yakni $\text{Exp}(B) = 5,229$ (95% CI: 4,361-6,269), sedangkan iklan rokok memiliki nilai $\text{Exp}(B) = 2,341$ (95% CI: 1,874-2,925).

Jadi, disimpulkan bahwa ada hubungan yang signifikan antara sikap dan paparan iklan rokok dengan tindakan merokok pada pelajar, yaitu pelajar dengan sikap negatif mempunyai risiko sebesar 5,229 kali untuk merokok dibanding pelajar dengan sikap positif, dan pelajar yang terpapar iklan rokok mempunyai risiko sebesar 2,341 kali untuk merokok dibanding yang tidak terpapar iklan rokok.

Output Tabel 10.4 juga menunjukkan nilai koefisien regresi logistik diperoleh "B" 0 = -3,400, "B" 1 = 0,851, "B" 2 = 0,047, serta "B" 3 = 1,654. Karena variabel pengetahuan tidak signifikan berhubungan ($p = 0,651$) dengan tindakan merokok pelajar, maka model persamaan regresi logistik yang terbentuk adalah sebagai berikut.

$$f(z) = \frac{1}{1 + e^{-(B_0 + B_1 X_1 + B_2 X_2 + B_3 X_3)}}$$

$$f(z) = \frac{1}{1 + e^{-(-3,400 + 0,851(1) + 1,654(1))}}$$

$$f(z) = \frac{1}{1 + e^{-(-0,895)}}$$

$$f(z) = 0,29 = 29,0\%$$

Artinya, pelajar yang memiliki sikap negatif dan terpapar iklan rokok memiliki probabilitas untuk bertindak merokok adalah sebesar 29,0%. Sebaliknya, probabilitas merokok seorang pelajar jika memiliki sikap positif dan tidak terpapar iklan rokok dapat diprediksi dengan persamaan sebagai berikut:

$$f(z) = \frac{1}{1 + e^{-(-3,400 + 0,851(0) + 1,654(0))}}$$

$$f(z) = \frac{1}{1 + e^{-(-3,400)}}$$

$$f(z) = 0,032 = 3,2\%$$

Dengan demikian, probabilitas seorang pelajar untuk merokok jika dia memiliki sikap positif dan tidak terpapar iklan rokok hanya sebesar 3,2%.

Daftar Pustaka

- 'Dahlan, S. (2021) Statistik Untuk Kedokteran dan Kesehatan. 6th edn. Edited by Agus Kurniawan. Jakarta: PT Epidemiologi Indonesia Jakarta.
- 'Purwoto, A. (2007) Panduan Laboratorium Statistik Inferensial. Edited by Eneste Pramusuk. Jakarta: PT Grasindo.
- A. Purwoto (2007) Panduan Laboratorium Statistik Inferensial. Edited by O. Heriyani. Jakarta: PT Gramedia Pustaka Utama.
- Adli, D.N. dan Sjojfan, O. (2020). "Estimasi dan Validasi Kandungan Energi Bekatul Sebagai Pakan Unggas dari Komposisi Kimia Pakan." *Jurnal Nutrisi Ternak Tropis*, 3(2), 90–96.
- Aini, S. D., & Irawati, S. (2019). Strategi Pembelajaran Quick on the Draw untuk Meningkatkan Aktivitas dan Hasil Belajar Mahasiswa Pada Materi Interpolasi. Numerical: *Jurnal Matematika Dan Pendidikan Matematika*, 19–30.
- Al-Hameed, K. A. A. (2022). Spearman's Correlation Coefficient in Statistical Analysis. *International Journal of Nonlinear Analysis and Applications*, 13 (1), 3249 – 3255.
- Alifiya, V. (2023) STATISTIK INFERENSIAL, https://www.academia.edu/39155657/STATISTIK_INFERENSIAL.

- Almumtazah, N., Azizah, N., Putri, Y. L., & Novitasari, D. C. R. (2021). Prediksi Jumlah Mahasiswa Baru Menggunakan Metode Regresi Linier Sederhana. *JURNAL ILMIAH MATEMATIKA DAN TERAPAN*, 18(1), 31–40. <https://doi.org/10.22487/2540766x.2021.v18.i1.15465>
- Amaliah, E.N., Darnah, & Sifriyani. (2020). "Regresi Data Panel dengan Pendekatan Common Effect Model (CEM), Fixed Effect Model (FEM) dan Random Effect Model (REM) (Studi Kasus: Persentase Penduduk Miskin Menurut Kabupaten/Kota di Kalimantan Timur Tahun 2015–2018)." *ESTIMASI: Journal of Statistics and Its Application*, 1(2), 106–115.
- Ananta, R.D. (2022). "Estimasi Model Kunjungan Wisatawan di Pulau Sumatera dengan Spatial Autoregressive." Universitas
- Anh, V. (2024). Causal Relationship Among Bank Capitalization, Efficiency, and Risk-Taking in ASEAN Commercial Banks. <https://doi.org/10.5772/intechopen.109120>
- Arikunto, S. (2010). *Prosedur Penelitian Suatu Pendekatan Praktik*. Rineka Cipta.
- Basuki, Agus Tri. (2020). *Regresi Logistik*. Universitas Muhammadiyah Yogyakarta.
- Basuki, H. (2019). *Statistika Deskriptif dan Inferensial*. Jakarta: Penerbit Erlangga.
- Bengnga, A., & Ishak, R. (2018). Prediksi Jumlah Mahasiswa Registrasi Per Semester Menggunakan Linier Regresi Pada Universitas Ichsan Gorontalo. *ILKOM Jurnal Ilmiah*, 10(2), 136–143.
- Bennett, M., Kleczyk, E. J., Hayes, K., & Mehta, R. (2022). Evaluating Similarities and Differences Between Machine Learning and Traditional Statistical Modeling in Healthcare Analytics. <https://doi.org/10.5772/intechopen.105116>

- Bland, M. (2015). *An Introduction to Medical Statistics*. Oxford University Press, 4th ed.
- Boone, H. N., & Boone, D. A. (2012). "Analyzing Likert Data." *Journal of Extension*, 50(2), 1-5.
- Brennan, M. M. (2020). *Cultivating Diagnostic Decision-Making With Problem-Based Learning: From Most Likely to Least Likely*. <https://doi.org/10.1891/9780826161215.0001>
- Bujang, M. A. & Baharum, N. (2016). Sample Size Guideline for Correlation Analysis. *World Journal of Social Science Research*, 3 (1), 37 – 49.
- Chan, Y. H. (2003). *Biostatistics 104: Correlational Analysis*. Singapore Medical Journal, 44 (12), 614 – 619.
- Chen, P. Y. & Popovich, P. M. (2022). *Correlation Parametric and Nonparametric Measures*. Thousand Oaks, London, New Delhi: Sage Publications.
- Cohen, J. (1988). *Statistical Power Analysis for the Behavioral Sciences*. Lawrence Erlbaum Associates., 2nd ed.
- Cohen, J., Cohen, P., West, S. G. & Aiken, L. S. (2003). *Applied Multiple Regression/Correlation Analysis for The Behavioral Sciences*, Third Edition, Mahwah, New Jersey, London: Lawrence Erlbaum Associates Publishers.
- Dahlan, M. S. (2011). *Statistik untuk Kedokteran dan Kesehatan* . Jakarta: Salemba Medika.
- Dahlan, M. S. (2014). *Statistik Untuk Kedokteran dan kesehatan (Edisi 6)*. Epidemiologi Indonesia.
- Dahlan, M.S. (2015) *Statistik Untuk Kedokteran dan Kesehatan*. Jakarta: Epidemiologi Indonesia.
- Dharmawan Atik Mawarni, Y. (2019) *BUKU AJAR BIOSTATISTIK INFERENSIAL*.

- Dzingirai, M., & Baporikar, N. (2022). Role of Microfinance for Entrepreneurial Success. 193–210. <https://doi.org/10.4018/978-1-6684-7552-2.ch011>
- Eko Budiarto. (2001). *Biostatistika Untuk Kedokteran dan kesehatan masyarakat*. EGC.
- Engelsma, T., Jaspers, M. W. M., & Peute, L. (2024). From Theory to Practice: Practical Applications of the MOLD-US Framework. <https://doi.org/10.3233/shti240337>
- Fauziyah, Rr Nur. (2019). Analisis Data Menggunakan Multiple Logistic Regression Test di Bidang Kesehatan Masyarakat dan Klinik (Gurid Pramintarto Eko Mulyo (ed.)). Politeknik Kesehatan Kemenkes Bandung.
- Field, A. (2013). *Discovering Statistics Using IBM SPSS Statistics (4th ed.)*. SAGE Publications.
- Field, A. (2018). *Discovering Statistics Using IBM SPSS Statistics (5th ed.)*. SAGE Publications.
- Field, Andy. (2017). *Discovering Statistics Using IBM SPSS Statistics (5th ed.)*. Sage edge. <http://repo.darmajaya.ac.id/5678/1/>
- Fisher, R. A. (1925). *Statistical Methods for Research Workers*. Oliver & Boyd.
- Fisher, R. A. (1935). *The Design of Experiments*. Oliver & Boyd.
- Freiholtz, D., Eriksson, P., & Björck, H. M. (2024). Ascending Aortic Aneurysm in Relation to Aortic Valve Phenotype. <https://doi.org/10.5772/intechopen.112883>
- Gani, N.F, dkk.(2017). Health Literacy and Self-Care Management of Pregnant Women at Level 1 Health Service in Makassar. *Indonesian Contemporary Nursing Journal*. 1(2):94-100.
- Ghozali, I. (2020). *Aplikasi Analisis Multivariate dengan Program IBM SPSS 26*. Semarang: Universitas Diponegoro.

- Gogtay, N. J., Deshpande, S. P., & Thatte, U. M. (2017). Principles of Regression Analysis. In *Journal of The Association of Physicians of India* ■ (Vol. 65).
- Gravetter, F. J., & Wallnau, L. B. (2017). *Statistics for the Behavioral Sciences*. Cengage Learning, 10th ed.
- Grove, S. K., & CIPHER, D. J. (2024). *Statistics for nursing research: A workbook for evidence-based practice* (3rd Editio). Elsevier.
- Guerard, J. B., Saxena, A., & Gültekin, M. (2020). Time Series Modeling and the Forecasting Effectiveness of the US Leading Economic Indicators. 301–381. https://doi.org/10.1007/978-3-030-43547-9_13
- Harlan, J. (2018). *Analisi Regresi Logistik* (1st ed.). Gunadarma. <https://www.academia.edu/39277804/>
- Hattab, S. A., Alsadi, M., Musleh, G., Eshah, N. F., Hani, M. B., & ALBashtawy, M. (2023). The Impact of Nurse-Physician Collaboration on Critical Care Nurses' Autonomy. 265–277. https://doi.org/10.2991/978-94-6463-248-4_22
- Haugan, G. (2021). Nurse-Patient Interaction: A Vital Salutogenic Resource in Nursing Home Care. 117–136. https://doi.org/10.1007/978-3-030-63135-2_10
- Hazra, A. & Gogtay, N. (2016). Biostatistics Series Module 6: Correlation and Linear Regression, *Indian Journal of Dermatology* 61, 593 – 601.
- Heryana, A. (2020). Uji Chi-Square. Universitas Esa Unggul. https://www.researchgate.net/publication/341539841_UJI_CHI_SQUARE
- Hidayanti, A. A. & Mandalika, E. N. D. (2023). Analisis Korelasi Pearson Biaya Produksi terhadap Luas Lahan Petani Garam di Kecamatan Bolo Kabupaten Bima. *Journal Inovasi Pendidikan dan Sains*, 4 (1), 5 – 10.

- Hilbe, Joseph M. (2015). *Practical Guide to Logistic Regression*. CRS Press Taylor & Francis Group, LLC. <https://ftp.idu.ac.id/wp-content/uploads/ebook/>
- Hoskins, K. (2020). *Research Design and Methodological Approaches*. 137–162. <https://doi.org/10.1108/978-1-78973-563-520201008>
- Hosmer, D. W., Lemeshow, S.. (2000). *Applied Logistic Regression* (2nd ed.). John Wiley & Sons, Inc.
- Hovenga, E. (2022). *Nursing Informatics Integration Into Mainstream Health Informatics*. <https://doi.org/10.3233/shti220947>
- Howell, D. C. (2012). *Statistical Methods for Psychology* (8th ed.). Cengage Learning.
- Howell, D. C. (2016). *Fundamental Statistics for the Behavioral Sciences*. Cengage Learning, 9th ed.
- Husnul, N. R. I. (2018). *Pertemuan 18 Regresi Logistik*. In *Statistik Inferensial. Sistem Pembelajaran Daring (SPADA) Indonesia - Program Studi S1 Akuntansi Universitas Pamulang*. <https://lmsspada.kemdiktisaintek.go.id/course/>
- IBM. (2013). *IBM SPSS Statistics 22 Algorithms*. IBM Corporation 1989. www.sussex.ac.uk/its/
- IBM. (2021). *IBM SPSSRegression 28* (28; p. 42). International Business Machines Corporation. [https://www.ibm.com/docs/en/SSLVMB_28.0.0/](https://www.ibm.com/docs/en/SSLVMB_28.0.0)
- Imani, A. and Muslim, I. (2022) 'Statistika deskriptif', (December).
- Irwan, S.N. (2024). *Hubungan Health Literacy dengan Self Care pada Ibu Hamil di Wilayah Puskesmas Samata*. *Jurnal media Keperawatan* 15(1):.83-90.
- Jabnabillah, F. & Margina, N. (2022). *Analisis Korelasi Pearson dalam Menentukan Hubungan antara Motivasi Belajar dengan Kemandirian Belajar pada Pembelajaran Daring*. *Jurnal Sintak*, 1 (1), 14 – 18.

- Jamal, H., Abdullah, A. Z., & Alfira, N. (2024). Faktor yang Berhubungan dengan Tindakan Merokok Pelajar di Indonesia. 04(02), 139–149. <https://doi.org/10.31603/bnur.11602>
- Junaidi, J. (2015). Prosedur Uji Chi-Square. Universitas Jambi. https://www.researchgate.net/publication/277953672_Prosedur_Uji_Chi-Square
- Keppel, G., & Wickens, T. D. (2004). *Design and Analysis: A Researcher's Handbook*. Pearson, 4th ed.
- Khuneswari, G. P., & Hsieh, C. C. (2022). Survey on Factors Influencing the Unemployment Among UTHM Fresh Graduates. 255–266. https://doi.org/10.2991/978-94-6463-014-5_24
- Lee, M. A., Vyas, P., D'Agostino, F., Wieben, A., Coviak, C. P., Mullen-Fortino, M., Park, S., Sileo, M., Brown, S., Souza, E. N. d., Pruinelli, L., & Role, J. (2024). Data Literacy and Data Science Literacy for Nurses: State of the Art Literature Review. <https://doi.org/10.3233/shti240133>
- Malay, M. N. (2022). *Belajar Mudah & Praktis Analisa Data dengan SPSS dan JASP*. Bandar Lampung: CV. Madani Jaya.
- Marcus, G. L., Wattimanela, H. J., & Lesnussa, Y. A. (2012). Analisis regresi komponen utama untuk mengatasi masalah multikolinieritas dalam analisis regresi linier berganda. *BAREKENG: Jurnal Ilmu Matematika Dan Terapan*, 6(1), 31–40.
- Mardhotillah, B. (2012). Analisis Korelasional Perilaku Berinternet terhadap Prestasi Belajar dan Prestasi Ekstrakurikuler Siswa SMA Kabupaten Batang Hari, *Jurnal Gema Litbang Khazanah Intelektual*, 1 (14), 441 – 450.
- Mardhotillah, B., Asyhar, R. & Elisa, E. (2022). Filosofi Keilmuan Statistika Terapan pada Era Smart Society 5.0. *Multi Proximity: Jurnal Statistika*, 1 (2), 57 – 70.
- Mark Tranmer, M. E. (2008). Binary Logistic Regression. In *Binary Logistic Regression* (pp. 1–43). Cathie Marsh Centre for and Survey Research.

- Mbombi, M. O., Ophilia, M. D., Bopape, M. A., & Muthelo, L. (2023). Obstacles in the Nursing Training Programs. <https://doi.org/10.5772/intechopen.109191>
- McDonald, J. H. (2014). Handbook of Biological Statistics. Sparky House Publishing, 3rd ed.
- Menezes, L. de C. B. B., & Brasileiro, T. S. A. (2024). Online Data Collection Strategies in Qualitative Nursing Research. <https://doi.org/10.56238/sevened2023.007-018>
- Mertens, W., & Recker, J. (2023). New Guidelines for Null Hypothesis Significance Testing in Hypothetico-Deductive IS Research. 385–437. https://doi.org/10.1007/978-3-031-38719-7_13
- Montgomery, D. C. (2017). Design and Analysis of Experiments (9th ed.). John Wiley & Sons.
- Montgomery, D. C., Peck, E. A., & Vining, G. G. (2021). Introduction to linear regression analysis. John Wiley & Sons.
- Moore, D. S., McCabe, G. P., & Craig, B. A. (2021). Introduction to the Practice of Statistics. W. H. Freeman., 10th ed.
- Muhyidin (2021) Pendekatan Biostatistik Dalam Kesehatan Masyarakat.
- Mulyana , A., Susilawati, E., Fransiska , Y., Arismawati, M., Madrapriya, F., Phety, D., . . . Sumiati, I. (2019). Metode Penelitian Kuantitatif . Makassar: CV. Tohar Media.
- Mutalazimah. (2025). Metode Penelitian Kesehatan. Surakarta: Muhammadiyah University Press.
- Nasution, M. (2021). Metode Kuantitatif dalam Manajemen dan Bisnis. Jakarta: Salemba Empat.
- Nurhamdi, M.Q., & Ikhsan. (2022). "Analisa Perbandingan Estimasi Anggaran Biaya Antara Metode AHSP SNI 2016 dengan Metode Perhitungan Kontraktor." Journal of Applied Civil and Environmental Engineering, 2(1), 62–70.

- Nursalam. (2011). *Konsep dan Penerapan Metodologi Penelitian Ilmu Keperawatan Pedoman Skripsi, Tesis, dan Instrumen Penelitian Keperawatan*. Jakarta Selatan : Salemba Medika .
- Nuryadi, Tutut Dewi Astuti, Endang Sri Utami, & M.Budiantara. (2017). *Dasar- dasar Statistik Penelitian*. SIBUKU MEDIA. www.sibuku.com
- Pagano, R. R. (2019). *Understanding Statistics in the Behavioral Sciences*. Cengage Learning., 10th ed.
- Pallant, J. (2020). *SPSS Survival Manual: A Step by Step Guide to Data Analysis Using IBM SPSS*. Open University Press., 7th ed.
- Parashakti, R. D., Sudiro, A., Sumiati, S., & Rahayu, M. (2023). Impact of Organizational Learning on Nurse Performance With Organizational Commitment and Citizenship Behavior (OCB) as a Mediation Variable and in Leadership Moderation. 173–182. https://doi.org/10.2991/978-94-6463-214-9_19
- Peršolja, M. (2022). Congruence of Nurse Staffing and Activities With Patient Needs. <https://doi.org/10.5772/intechopen.96589>
- Polit, D. F. (2014). *Statistics and Data Analysis for Nursing Research*. In British Library Cataloguing-in-Publication Data.
- Polit, D. F., & Beck, C. T. (2017). *Nursing Research: Generating and Assessing Evidence for Nursing Practice*. Wolters Kluwer Health.
- Poni Egistin, D., Yahdi Rauza, M., Has Ramadhan, R., & Ramadani, S. (2025). Analisis regresi linier sederhana dan penerapannya (Vol. 1, Issue 2). <https://creativecommons.org/licenses/by/4.0/>
- Purnawinadi, I. G., Wardani, Y. S., Koro, S., Utami, R. A., Rahmadani, P., Dewi, R. K., . . . Elizawarda. (2023). *Manajemen & Analisis Data Penelitian Kuantitatif Kesehatan*. Medan: Yayasan Kita Menulis.
- Ristya Widi E.Y. dkk (2023) *Buku Ajar Biostatistika*. Jember: UPT Penerbitan Universitas Jember.

- Rojas, J. C. (2021). Intraclass Correlation Coefficient: Applications to Estimate the Temporal Stability of a Measuring Instrument. *Ciencias Psicológicas*, 15 (2), 1 – 13.
- Rosner, B. (2015). *Fundamentals of Biostatistics*. Cengage Learning, 8th ed.
- Rozak, A., & Hidayati, W. (2019). *Pengolahan Data dengan SPSS*. Yogyakarta: Erhaka Utama.
- Sarie, F., Sutaguna, N. T., Suiroka, I., Damanik, D., Efrina, G., Sari, R., . . . Massenga, T. W. (2023). *Metodologi Penelitian*. Batam: Yayasan Cndikia Mulia Mandiri .
- Schober, P., Boer, C. & Schwarte, L. A. (2018). Correlation Coefficients: Appropriate Use and Interpretation. *International Anesthesia Research Society, Special Article*, 126 (5), 1763 – 1768.
- Siburian, D.P., Kritiana, W., & P., V.H. (2022). "Analisis Perbandingan Estimasi Biaya Menggunakan Metode SNI 2017 dan ASHP 2016 (Studi Kasus: Proyek Pembangunan Gedung Kuliah Terpadu Universitas Palangka Raya)." *Jurnal TRANSUKMA*, 4(2), 138–143.
- Soilemezi, C., Liaskos, J., & Mantas, J. (2021). A Smartphone App for Bedside Recording of Nursing Handovers in Haemodialysis Units. <https://doi.org/10.3233/shti210256>
- SP Hastono & L.Sabri (2010) *Statistik Kesehatan*. 5th edn. Jakarta: PT RAJAGRAFINDO PERSADA.
- Stang. (2014). *Cara Praktis Penentuan Uji Statistik dalam Penelitian Kesehatan dan Kedokteran* (1st ed.). Mitra Wacana Media.
- Sugiyono. (2013). *Statistik untuk Penelitian*. Jakarta: Alfabeta.
- Sugiyono. (2022). *Metode Penelitian Kuantitatif, Kualitatif, dan R&D*. Bandung: Alfabeta.
- Sujarweni, W. V. (2015). *Statistik untuk Kesehatan*. Gava Media.
- Supardi, S & Rustika. (2013). *Metodologi Riset Keperawatan*. Trans Info Media.

- Syofian Siregar (2014) *Statistik Deskriptif Untuk Penelitian*. 1st edn. Jakarta: PT RAJAGRAFINDO PERSADA.
- Tabachnick, B. G., & Fidell, L. S. (2019). *Using Multivariate Statistics*. Pearson, 7th ed.
- Tim Dosen. (2018). *Uji Regresi dengan SPSS*. Universitas Esa Unggul.
- Tiro, M.A. (2007). *Dasar-Dasar Statistika*. State University makassar Press.
- Trianggana, D. A. (2020). Peramalan jumlah siswa-siswi melalui pendekatan metode regresi linear. *Jurnal Media Infotama*, 16(2).
- Trochim, W. M. K., & Donnelly, J. P. (2021). *Research Methods Knowledge Base*. Atomic Dog Publishing, 4th ed.
- Vlaški, A. (2020). Research Objectives, Hypotheses and Methodology. 31–37. <https://doi.org/10.1201/9781003073154-2>
- VT Hulu & TR Sinaga (2019a) *Analisis Data Statistik Parametrik Aplikasi SPSS dan STATCAL*. 1st edn. Edited by J Simarmata. Yayasan Kita Menulis.
- VT Hulu & TR Sinaga (2019b) *Analisis Data Statistik Parametrik Aplikasi SPSS dan STATCAL*. 1st edn. Edited by J Simarmata. Yayasan Kita Menulis.
- Wardhani, D.S. dan Utami, D.P. (2020). "Perhitungan Estimasi Ketidakpastian."
- Widiyono, Aryani, Putra, F. A., Herawati, V. D., Indiyati, Suwarni, A., . . . Azmi, L. D. (2023). *Buku Mata Ajar Konsep Dasar Metodologi Penelitian Keperawatan*. Kediri: Chakra Brahmanda Lentera .
- Wuensch, K. L. (2021). *Binary Logistic Regression with SPSS*. <https://www.researchgate.net/profile/David-Booth-7/>
- Yuantari, C. and Handayani, S. (2017) *Buku Ajar Statistik Deskriptif & Inferensial*, Badan Penerbit Universitas Dian Nuswantoro.

- Yusuf, Hendra Andrianto. (2017). Modul Praktikum Analisis Regresi. Universitas Negeri Gorontalo.
- Zaid, M. A. (2015). Correlation and Regression Analysis. Ankara: SESRIC.

Biodata Penulis



Taruli Rohana Sinaga lahir di Kampung Baru, pada tanggal 16 Oktober 1971. Penulis menyelesaikan studi S1 pada Program Studi Gizi Masyarakat dan Sumberdaya Keluarga di Institut Pertanian Bogor. Selanjutnya penulis menyelesaikan studi S2 pada Program Studi Ilmu Kesehatan Masyarakat FKM UI pada tahun 2009. Penulis menyelesaikan studi S3 di Lincoln University College Kuala Lumpur Malaysia. Sejak tahun 2000 sampai sekarang penulis memulai karir sebagai dosen di Universitas Sari Mutiara

Indonesia Medan. Di samping mengajar penulis juga aktif di Asosiasi Instituti Pendidikan Tinggi Kesehatan Masyarakat sebagai Wakil Ketua Divisi Pengembangan Organisasi Periode Tahun 2022-2025 dan Asosiasi Profesi Kesehatan Masyarakat IAKMI Provinsi Sumatera Utara sebagai Wakil Ketua IV dari tahun 2017 sampai tahun 2020. Mendapat penghargaan sebagai reviewer dan panel expert dari Komite Nasional UKSKMI berturut-turut tahun 2017, 2018 dan 2019. Pernah membawa oral presentasi tingkat nasional pada Forum Ilmiah Tahunan IAKMI tahun 2020, 2021, 2022 dan 2023 dan tingkat internasional di Thailand pada tahun 20019, 2020, 2021, 2022, 2023 dan 2024. Juga sebagai narasumber tingkat nasional dan internasional tahun 2017, 2018, 2019, 2020 dan 2021. Sebagai guest lecturer di Thailand dan Laos pada tahun 2020, 2021, 2022, 2023 dan 2024.



Ribka Sabarina Panjaitan lahir di Manado, pada 26 Januari 1994. Ia tercatat sebagai lulusan Magister Keperawatan di Universitas Padjadjaran pada tahun 2022. Sebelumnya mengikuti pendidikan Program S1 Keperawatan UNSRIT di Tomohon dan mengikuti Program Ners di STIKes Immanuel Bandung dan sempat bekerja selama 2 tahun di Santosa Hospital Bandung. Wanita yang kerap disapa Ribka Sabrina ini adalah anak dari pasangan Robert Panjaitan (ayah)

dan Selfie Sumangando (ibu). Selama ini telah menjadi Dosen Keperawatan Medikal Bedah (KMB) di STIKes RS Husada Jakarta.

Email : sabrinapanjaitan26@gmail.com



Putri Pamungkas S.Kep.,Ns.,M.K.M. Dosen Program Studi Ilmu Keperawatan dan Profesi Ners Institut Kesehatan dan Bisnis Surabaya. Penulis lahir di Mataram, tanggal 3 November 1992. Penulis adalah dosen tetap pada Program Studi Pendidikan Ilmu Keperawatan dan Profesi Ners, Institut Kesehatan dan Bisnis Surabaya. Menyelesaikan pendidikan S1 pada Jurusan Ilmu Keperawatan Universitas Kadiri dan melanjutkan S2 pada Jurusan Kesehatan Masyarakat Universitas Sebelas Maret. Penulis menekuni bidang Epidemiologi dan Biostatistik Penelitian fokus bidang penyakit menular.

Penulis juga aktif menjabat sebagai Sekertaris Program Studi Pendidikan Ilmu Keperawatan dan Profesi Ners, Institut Kesehatan dan Bisnis Surabaya. Penulis mulai aktif menulis buku pada tahun 2020.

E-mail : pamungkasputri95@gmail.com



A. Adriana Amal, S.Kep.Ns.,M.Kep, anak pertama dari A.Akmal, S.Pd dan Hj. Darmawati, S.Pd (Alm) yang lahir di Bone, 16 November 1988. Menamatkan studi sarjana pada Program Studi Keperawatan pada tahun 2010 di UIN Alauddin Makassar dan melanjutkan Pendidikan Profesi Ners pada tahun 2011 dan selesai pada tahun 2012 di STIK GIA Makassar. Pada tahun 2016, penulis menyelesaikan pendidikan Magister Keperawatan Di Universitas Hasanuddin. Pengalaman sebagai dosen non PNS UIN Alauddin Makassar pada tahun 2017-2019. Kemudian menjadi

dosen PNS UIN Alauddin Makassar pada tahun 2019 hingga saat ini. Selain itu, penulis juga sebagai Anggota DPK PPNI UIN Alauddin Makassar, Anggota Himpunan Perawat Manajer Indonesia (HPMI) Sulawesi selatan, Wakil

Sekretaris Himpunan Perawat Kesehatan Kerja (HIPERKES) Sulawesi Selatan dan Anggota TWG SATU SEHAT Sulawesi Selatan.



Ftr. Catherine Hermawan Salim, S.Ft, M.M. adalah seorang dosen dan praktisi fisioterapi dengan pengalaman di bidang fisioterapi muskuloskeletal. Beliau memperoleh gelar Sarjana Fisioterapi dari Universitas Esa Unggul dan melanjutkan studi magister di bidang Ilmu Manajemen di PPM School of Management. Sebagai fisioterapis, beliau telah menangani berbagai kasus gangguan muskuloskeletal sejak tahun 2017 di klinik-klinik fisioterapi dan rumah sakit Tzu Chi Hospital Jakarta. Keahlian utama beliau terletak pada penanganan gangguan tulang belakang

lumbopelvic, osteoarthritis, dry needling dan fisioterapi cedera olahraga. Beliau juga telah menulis beberapa artikel ilmiah serta sering menjadi pembicara dalam seminar fisioterapi. Dengan dedikasi tinggi pada peningkatan kualitas hidup pasien, beliau terus berkomitmen dalam pengembangan fisioterapi berbasis bukti di Indonesia.

Email Penulis: cathysalim8@gmail.com



Rifki Sakinah Nampo, M.Kep. Telah menyelesaikan Studi Program Magister Keperawatan di Universitas Padjadjaran, dengan bidang peminatan Mental Health-Psychiatric Nursing. Saat ini, bekerja sebagai Dosen di Sekolah Tinggi Ilmu Kesehatan (STIKes) Papua – Sorong. Aktif dalam bidang penelitian kesehatan, publikasi penelitian kesehatan dan mengikuti berbagai pelatihan pengembangan diri terkait Metodologi Penelitian dan Statistika Penelitian.

Saat ini, penulis aktif sebagai salah satu dosen pengajar Keperawatan Jiwa, Metodologi Penelitian, dan Biostatistika di Program Studi Keperawatan dan Program Studi K3 (Kesehatan dan Keselamatan Kerja) STIKes Papua.

Penulis juga mengembangkan kemampuan diri dengan mengikuti berbagai pelatihan seperti: NLP (Neuro-Linguistic Programming), Hipnotherapy, akupuntur, akupresure, berbagai terapi komplementer, dan terapi modalitas. Aktif dalam organisasi keperawatan PPNI DPD Kabupaten Jayapura sebagai Ketua Bidang Informasi & Penelitian, dan Organisasi IPKJI DPW Papua sebagai sekretaris.

Adapun karya buku yang penulis publikasi secara online diantaranya: “Buku Riset Keperawatan”, “Populasi Khusus dalam Keperawatan Jiwa”, dan “Keperawatan Paliatif Care”. Selain itu, berbagai karya penelitian penulis juga telah dipublikasi diantaranya: “Effect of Neuro-Linguistic Programming (NLP) on Anxiety: A Systematic Literature Review”, dan “Pengaruh Aplikasi Akupuntur Pada Pasien Hipertensi”.



Bunga Mardhotillah. Saat ini tengah menempuh Pendidikan Strata-3 Program Doktor Pendidikan MIPA di Universitas Jambi dengan topik disertasi tentang Pengembangan Model Pembelajaran Case Based Learning yang Berbasis Statistica preneurship. Ia adalah dosen tetap Program Studi Matematika, Fakultas Sains dan Teknologi Univeristas Jambi. Penulis telah menyelesaikan Pendidikan Strata 1 dan 2 Bidang Keilmuan Statistika Terapan di Universitas Padjadjaran, Bandung.

Mengampu Mata Kuliah Metode Statistika, Statistika Matematika, Kapita Seleкта Statistika, Statistika Non Parametrik, Ekonometrika, Rancangan Percobaan, Proses Stokastik, dan Analisis Multivariat. Selama ini terlibat aktif sebagai Dosen Pembimbing mahasiswa pada berbagai LKTI Tingkat Daerah dan Nasional yang mengedepankan Analisis Statistik. Memiliki pengalaman sebagai narasumber dalam berbagai webinar nasional, serta pernah menjadi Ketua Panitia Workshop Kewirausahaan MIPA UNJA Berbasis Teknopreneurship dan Ketua Pelaksana Seleksi Satria Data Tingkat Universitas Jambi pada Tahun 2024.

Telah menulis beberapa artikel media cetak dan online serta telah menulis sekitar 50 Artikel Ilmiah yang berkolaborasi bersama beberapa penulis baik dari Universitas Jambi maupun dengan penulis dari PTN/PTS lainnya di Indonesia. Penulis telah memiliki 22 Hak Cipta (HKI) dan 2 Paten Sederhana. Penulis juga

berperan sebagai Managing Editor pada Multi Proximity: Jurnal Statistika sejak Tahun 2022 hingga saat ini.

E-mail: bunga.mstat08@unja.ac.id



Nurul Fadhilah Gani., S.Kep., Ns., M.Kep, Lahir di Polewali pada Tanggal 12 Juni 1989. Menamatkan Sekolah Dasar di SDN Mangga Tiga Makassar. Sekolah Menengah Pertama Negeri 12 Makassar pada Tahun 2003. Sekolah Menengah Atas Negeri 5 Makassar pada tahun 2006. Melanjutkan Program Sarjana Keperawatan pada UIN Alauddin Makassar Tahun 2006 dan selesai pada tahun 2010, Program Profesi Ners di STIK GIA Makassar Pada Tahun 2012. Setelah itu penulis melanjutkan Pendidikan Program Magister Keperawatan di Universitas

Hasanuddin Makassar jurusan Manajemen Keperawatan, dan menyelesaikan studi pada Tahun 2016. Pengalaman mengajar dan pengelola di berbagai institusi pendidikan keperawatan sejak tahun 2010-2019. Saat ini tercatat sebagai Dosen PNS di Jurusan Ilmu Keperawatan UIN Alauddin Makassar sejak Tahun 2019 dan berperan serta dalam berbagai penelitian, pengabdian masyarakat. Penulis juga aktif dalam organisasi profesi dan kemasyarakatan, diantaranya menjadi anggota PPNI dan sebagai pengurus Ikatan Perawat Maternitas Indonesia Provinsi Sulawesi Selatan tahun 2023 sampai sekarang. Menjadi editor serta reviewer di beberapa jurnal nasional terakreditasi juga menjadi aktivitas penulis. Penulis juga aktif berkolaborasi dalam penulisan buku, artikel, serta publikasi ilmiah. Email: nurul.fadhilah@uin-alauddin.ac.id



Dika Betaditya lahir di Ciamis, pada 2 September 1992 menyelesaikan program Strata 1 di Program Studi Gizi dan Kesehatan UGM pada Tahun 2014. Pada Tahun 2017 menyelesaikan program S-2 di Program Studi Ilmu Kesehatan Masyarakat Minat Gizi Kesehatan UGM, dan pada tahun 2021 menyelesaikan Program Studi Pendidikan Profesi Dietisien di UGM. Penulis merupakan staff pengajar di Jurusan Ilmu Gizi Fakultas Ilmu-ilmu Kesehatan Universitas Jenderal Soedirman.

Penulis mengampu mata kuliah Dietetik Penyakit Defisiensi dan Infeksi, Dietetik Penyakit Tidak Menular, Gizi Dalam Daur Kehidupan, Patofisiologis Penyakit Tidak Menular, Patofisiologis Penyakit Menular, Gizi Olahraga dan Statistik Kesehatan.

Penulis aktif menulis karya ilmiah terkait Gizi dan melakukan beberapa kegiatan tridharma perguruan tinggi di bidang Gizi dan Kesehatan. Telah menulis Buku referensi dan yaitu Buku Ilmu Gizi, Buku Gizi dalam Keperawatan, Buku Penyakit Sistem Pencernaan, dan Buku Kebidanan dan Edukasi Gizi Ibu Hamil yang diterbitkan dari Penerbit Yayasan Kita Menulis.

E-mail: dika.betaditya@unsoed.ac.id



Hudriani Jamal. Lahir di Bulukumba pada 32 tahun yang lalu dan telah menyelesaikan Pendidikan Program S1 dan S2 di Fakultas Kesehatan Masyarakat Universitas Hasanuddin Makassar. Saat ini, bekerja sebagai dosen tetap Program Studi S-1 Kesehatan Masyarakat, Fakultas Kesehatan, Universitas Patria Artha. Sejak kuliah, tertarik dengan studi epidemiologi dan biostatistik serta pemanfaatannya dalam penyelesaian masalah kesehatan masyarakat.

Ditugaskan mengampu mata kuliah Biostatistik Deskriptif, Biostatistik Inferensial, Manajemen Data, dan Dasar Epidemiologi. Biostatistik dan Epidemiologi sangat menarik, terkhusus dalam pengumpulan dan pengolahan data untuk mengukur besar masalah kesehatan dan hubungan kausalitas agar dapat dilakukan upaya pencegahan dan intervensi masalah kesehatan masyarakat yang tepat serta efektif dan efisien.

Selama menjadi seorang akademisi atau dosen, telah menulis banyak artikel di bidang kesehatan masyarakat. Beberapa artikel telah dipublikasikan di jurnal terkemuka seperti Jurnal Promkes: The Indonesian Journal of Health Promotion and Health Education, Jurnal Kesehatan Vokasional, dan juga telah mengikuti acara 9th Indonesian Conference on Tobacco or Health (ICTOH) 2024. Selain mengajar dan meneliti, juga aktif dalam berbagai kegiatan sosial dan kemanusiaan atau pengabdian di universitas dan masyarakat.

E-mail: hudriani.jamal@patria-artha.ac.id, hudrianijamal@gmail.com

BIOSTATISTIK INFERENSIAL

Biostatistik Inferensial tidak hanya mencakup analisis data yang berfokus pada pengujian hipotesis, estimasi, dan pengolahan data, tetapi juga berhubungan erat dengan penerapan hasil analisis untuk meningkatkan kualitas pelayanan kesehatan. Buku ini dirancang untuk memberikan pemahaman yang komprehensif mengenai konsep, teori, dan metode yang digunakan dalam analisis biostatistik, serta aplikasinya dalam berbagai disiplin ilmu, terutama keperawatan.

Buku ini membahas:

Bab 1 Konsep, Teori, dan Ruang Lingkup Biostatistik Inferensial

Bab 2 Konsep Dasar Analisis Bivariat

Bab 3 Estimasi

Bab 4 Analisis Hipotesis

Bab 5 Uji T-Independent dan Dependent

Bab 6 Uji ANOVA

Bab 7 Uji Korelasi

Bab 8 Analisis uji Chi Square dengan SPSS

Bab 9 Uji Regresi Linier Sederhana

Bab 10 Uji Regresi Logistik Sederhana



YAYASAN KITA MENULIS
press@kitamenulis.id
www.kitamenulis.id

