



STATISTIK KESEHATAN TEORI & APLIKASI



I Made Sudarma **Adiputra** • Deborah **Siregar** • Dina Dewi **Anggraini**
Ahmad **Irfandi** • Ni Wayan **Trisnadewi** • Marlynda Happy Nurmalita **Sari**
Ni Putu Wiwik **Oktaviani** • Puji Laksmi • Agus **Supinganto**
Martina **Pakpahan** • Yana **Listyawardhani** • Fahrul **Islam** • Murti **Ani**

STATISTIK KESEHATAN

TEORI & APLIKASI



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).

Statistik Kesehatan: Teori dan Aplikasi

I Made Sudarma Adiputra, Deborah Siregar, Dina Dewi Anggraini
Ahmad Irfandi, Ni Wayan Trisnadewi, Marlynda Happy Nurmalita Sari
Ni Putu Wiwik Oktaviani, Puji Laksmi, Agus Supinganto
Martina Pakpahan, Yana Listyawardhani, Fahrul Islam, Murti Ani



Penerbit Yayasan Kita Menulis

Statistik Kesehatan: Teori dan Aplikasi

Copyright © Yayasan Kita Menulis, 2021

Penulis:

I Made Sudarma Adiputra, Deborah Siregar, Dina Dewi Anggraini
Ahmad Irfandi, Ni Wayan Trisnadewi, Marlynda Happy Nurmalita
Sari Ni Putu Wiwik Oktaviani, Puji Laksmi, Agus Supinganto
Martina Pakpahan, Yana Listyawardhani, Fahrul Islam, Murti Ani

Editor: Abdul Karim

Desain Sampul: Devy Dian Pratama, S.Kom.

Penerbit

Yayasan Kita Menulis

Web: kitamenulis.id

e-mail: press@kitamenulis.id

WA: 0821-6453-7176

Anggota IKAPI: 044/SUT/2021

I Made Sudarma Adiputra, dkk.

Statistik Kesehatan: Teori dan Aplikasi

Yayasan Kita Menulis, 2021

xiv; 190 hlm; 16 x 23 cm

ISBN: 978-623-342-042-6

Cetakan 1, April 2021

- I. Statistik Kesehatan: Teori dan Aplikasi
- 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 yang telah memberikan berkat, rahmat dan karunia-Nya sehingga kami dapat menyelesaikan buku "Statistik Kesehatan (Teori dan Aplikasi)" yang disusun secara kolaborasi oleh penulis-penulis yang berdedikasi dalam penyebaran dan pengembangan ilmu pengetahuan.

Dalam kehidupan sehari-hari, kita acapkali disuguhkan data statistik, statistik digunakan sebagai alat untuk mengubah data menjadi Informasi yang mudah dipahami. Statistik mempunyai peranan yang sangat penting dalam berbagai bidang, tak terkecuali dalam dunia Pendidikan dan penelitian. Penggunaan statistik dalam penelitian khususnya dalam penelitian kesehatan sangat berguna dalam melakukan finalisasi data sehingga dapat menguji hipotesis suatu penelitian.

Pada buku ini akan lebih detail membahas tentang:

1. Pengantar Statistik Kesehatan
2. Probabilitas
3. Populasi dan Sampling
4. Uji Hipotesis
5. Statistik Deskriptif
6. Uji Beda Berpasangan (Kategorik dan Numerik)
7. Uji Beda Dua Kelompok (Kategorik dan Numerik)
8. Uji Beda Lebih Dua Kelompok (Kategorik dan Numerik)
9. Analisis Data Proporsi
10. Uji Chi-Square
11. Korelasi dan Regresi Linier
12. Analisis Multivariat
13. Regresi Logistik

Para penulis mengucapkan terima kasih kepada berbagai pihak atas kerjasama, dukungan, bimbingan dan kritik yang diberikan sehingga buku ini dapat diwujudkan. Para penulis berharap semoga buku ini dapat bermanfaat bagi pembaca serta dapat memberikan sumbangsih bagi pengembangan ilmu pengetahuan terutama di bidang kesehatan.

Salam sejahtera teruntuk kita semua.

Denpasar, 1 April 2021

I Made Sudarma Adiputra, dkk

Daftar Isi

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

Bab 1 Pengantar Statistik Kesehatan

1.1 Pendahuluan.....	1
1.2 Penerapan Statistik dalam Penelitian Kesehatan.....	2
1.3 Fungsi Statistik.....	3
1.3.1 Untuk Menyajikan Informasi atau Data Dalam Bentuk Yang Lebih Mudah Dimengerti.....	3
1.3.2 Untuk Menyederhanakan Data Yang Komplek.....	3
1.3.3 Statistik Dapat Digunakan Untuk Melakukan Perbandingan.....	3
1.3.4 Untuk Memperluas Pengalaman	3
1.3.5 Memberikan Pertimbangan Dalam Perumusan Kebijakan	4
1.3.6 Untuk melakukan pengukuran pada fenomena yang besar dan luas	4
1.4 Jenis Statistik	4
1.4.1 Statistik Deskriptif	5
1.4.2 Statistik Inferensial	5
1.5 Tahapan analisis statistic	5
1.5.1 Pengumpulan data.....	5
1.5.2 Pengolahan data	6
1.5.3 Penyajian data	6
1.5.4 Analisis	6
1.5.5 Interpretasi data	6
1.6 Variabel	7
1.6.1 Variabel kualitatif.....	8
1.6.2 Variabel kuantitatif	8
1.7 Skala Data variabel	9
1.7.1 Skala Nominal	9
1.7.2 Skala Ordinal	10

1.7.3 Skala Interval	10
1.7.4 Skala Rasio.....	11
Bab 2 Probabilitas	
2.1 Pendahuluan.....	13
2.2 Pengertian Probabilitas.....	14
2.2 Konsep Probabilitas.....	15
2.3 Unsur-Unsur Probabilitas.....	16
2.4 Asas Perhitungan Probabilitas	17
Bab 3 Populasi dan Sampling	
3.1 Pendahuluan.....	23
3.2 Populasi	24
3.3 Sampel.....	28
3.4 Teknik Sampling	34
3.4.1 Random Sampling/ Probability Sampling.....	34
3.4.2 Non Random Sampling/ Non Probability Sampling	36
Bab 4 Uji Hipotesis	
4.1 Pendahuluan.....	41
4.2 Defenisi Hipotesis.....	43
4.3 Jenis Hipotesis	45
4.4 Pemilihan Jenis Uji Parametrik atau Non Parametrik.....	45
4.5 Langkah Uji Hipotesis.....	47
Bab 5 Statistik Deskriptif	
5.1 Pendahuluan.....	51
5.2 Pengertian Statistik Deskriptif	52
5.3 Manfaat Statistik Deskriptif.....	52
5.4 Metode Dasar Statistik Deskriptif.....	53
5.4.1 Distribusi Frekuensi/ Distribusi Data.....	54
5.4.2 Ukuran Pemusatan (Central Tendency)	54
5.4.3 Dispersion	61
5.5 Penyajian Data Statistika Deskriptif.....	63
5.5.1 Penyajian Data dalam Bentuk Grafik	63
5.5.2 Penyajian Data Numerik.....	66

Bab 6 Uji Beda Berpasangan (Kategorik Dan Numerik)

6.1 Pendahuluan.....	67
6.2 Uji beda Berpasangan/ Uji Beda Dua Mean Dependen (Uji t Dependen)...	67
6.3 Uji Beda Dua Mean Dependen (Uji t Wilcoxon).....	70
6.4 Uji T Dependen dengan menggunakan SPSS.....	75
6.5 Penyajian dan Interpretasi di Laporan Penelitian:.....	78

Bab 7 Uji Beda Dua Kelompok

7.1 Pendahuluan.....	79
7.2 Parametrik.....	80
7.2.1 Berpasangan atau Dependent (Uji t-paired).....	80
7.2.2 Tidak Berpasangan atau Independent (Uji t).....	82
7.3 Non Parametrik.....	85
7.3.1 Berpasangan (Dependent).....	85
7.3.2 Uji Sign.....	86
7.3.2 Wilcoxon.....	88
7.3.3 Mc.Nemar.....	90
7.3.4 Tidak Berpasangan (Independent).....	92
7.3.5 Mann-Whitney.....	92
7.3.6 Kolmogorov Smirnov.....	95
7.3.7 Run Wald-Wol Fowitz.....	98

Bab 8 Uji Beda Rata-rata (MEAN) Lebih dari Dua Kelompok (ANOVA)

8.1 Pendahuluan.....	101
8.2 Jenis ANOVA.....	102
8.3 Rumus ANOVA.....	104
8.4 Contoh Soal Uji ANOVA.....	105
8.5 Uji Banding Ganda / Multiple Comparison (Post Hoc Test).....	108
8.6 Uji ANOVA dengan SPSS.....	108

Bab 9 Analisis Data Proporsi

9.1 Pendahuluan.....	113
9.1.1 Pengertian.....	113
9.1.2 Pengumpulan Data.....	114
9.2 Proses Analisis Dan Penyajian Data.....	118

Bab 10 Uji Chi-Square

10.1 Karakteristik Uji Chi-Square.....	123
10.1.1 Tujuan Uji Chi-Square.....	124

10.1.2 Prinsip Dasar Uji Chi-Square	124
10.1.3 Sifat Uji Chi-Square	126
10.2 Analisis Uji Chi-Square	126
10.2.1 Pengujian Non Parametrik	126
10.2.2 Cross-Tabulation	128
10.3 Interpretasi Uji Chi-Square	131
Bab 11 Korelasi dan Regresi Linier	
11.1 Karakteristik Uji Chi-Square	135
11.2 Analisis Korelasi	136
11.3 Koefisien Korelasi	137
11.4 Uji Hipotesis Korelasi	139
11.5 Interval Kepercayaan Korelasi	140
11.6 Koefisien Korelasi Rank Spearman	141
11.7 Analisis Regresi Linier	141
11.8 Asumsi Regresi Linier	146
Bab 12 Analisis Multivariat	
12.1 Pendahuluan	149
12.2 Konfounding dan Interaksi	151
12.2.1 Konfounding	151
12.2.2 Interaksi	154
12.3 Teknik Analisis Multivariat	156
12.3.1 Teknik Dependen	156
12.3.2 Teknik Interdependen	159
12.3.3 Teknik Persamaan Struktural	160
12.4 Analisis Multivariat dalam Penelitian Kesehatan	161
Bab 13 Regresi Logistik	
13.1 Pendahuluan	165
13.2 Analisis Regresi Logistik	166
13.3 Regresi Logistik Sederhana	170
13.4 Regresi Logistik Ganda	171
13.5 Aplikasi Regresi Logistik	172
Daftar Pustaka	175
Biodata Penulis	183

Daftar Gambar

Gambar 1.1: pengujian penelitian.....	2
Gambar 1.2: Variabel	7
Gambar 2.1: Skala Probabilitas	14
Gambar 4.1: Kurva Normal Standar	49
Gambar 5.1: Histogram	64
Gambar 5.2: Pie Chart	64
Gambar 5.3: Poligon.....	65
Gambar 5.4: Ogive	65
Gambar 6.1: Tampilan SPSS	76
Gambar 6.2: Uji Coba SPSS	76
Gambar 7.1: Ilustrasi Sampel Berpasangan.....	80
Gambar 7.2: Ilustrasi Sampel Tidak Berpasangan.....	82
Gambar 13.1: Grafik hubungan umur dengan kejadian preeklamsi.....	167
Gambar 13.2: Peningkatan proporsi kejadian preeklamsi pada kelompok Umur	169

Daftar Tabel

Tabel 2.1: Jumlah perawat berdasarkan kelompok usia dan jenis kelamin..	21
Tabel 4.1: Kesalahan dalam Membuat Hipotesis	44
Tabel 4.2: Tabel uji hipotesis alur menuju pemilihan uji hipotesis komparatif variabel numerik	46
Tabel 4.3: Tabel uji hipotesis: alur menuju pemilihan hipotesis variabel kategorikal	46
Tabel 5.1 Data Nilai Mata Kuliah Biostatistik Mahasiswa	56
Tabel 6.1: Penyajian data	72
Tabel 6.2: Bantuan Rangking	72
Tabel 6.3: Distribusi Rata-Rata Kadar Hb Responden Menurut Pengukuran Pertama dan Kedua di... Tahun	78
Tabel 7.1: Model Uji Mc.Nemar	91
Tabel 8.1: Contoh Soal uji anova	106
Tabel 8.2: Distribusi Rata-rata Skor Kecemasan Mahasiswa Berdasarkan Jenjang Pendidikan	112
Tabel 10.1: Guide to Major Bivariate Statistical Tests	124
Tabel 13.1 Proporsi kejadian preeklamsi dengan umur.....	168

Bab 1

Pengantar Statistik Kesehatan

1.1 Pendahuluan

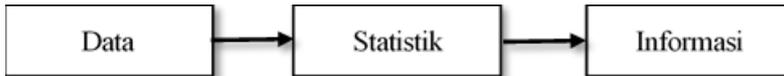
Pertama kali mendengar kata statistik, pasti akan timbul pertanyaan dalam diri kita, “Apa guna dari statistik ini dan kapan kita menggunakan statistik dalam kehidupan sehari-hari?”. Sering kali tanpa kita sadari dalam kehidupan sehari-hari kita selalu disajikan data statistik, sebagai contoh dalam acara berita di televisi menayangkan jumlah penduduk usia tua di Indonesia tahun 2020, atau angka hitung cepat hasil pilkada serentak dan masih banyak lagi. Dengan informasi ini, kita dapat membuat keputusan tentang kebenaran pernyataan, klaim, atau "fakta". Metode statistik dapat membantu kita membuat perkiraan terbaik terhadap suatu hal (Bridges, Weiss and Hassett, 1992).

Statistik mempunyai peranan yang sangat penting dalam berbagai bidang, tak terkecuali dalam dunia Pendidikan dan penelitian. Penggunaan statistik dalam penelitian khususnya dalam penelitian kesehatan sangat berguna dalam melakukan finalisasi data sehingga dapat menguji hipotesis suatu penelitian (Lotka and Odell, 1926).

Dalam dunia kesehatan statistik sering digunakan untuk menyajikan data tentang penyakit ataupun masalah kesehatan yang dewasa ini sering terjadi, sebagai contoh pandemi covid 19 yang melanda dunia mulai dari akhir tahun 2019, statistik digunakan untuk mengetahui bagaimana trend penyebaran

covid 19 di seluruh dunia, berapa negara yang terkena, berapa Case Fatality Rate (CFR) dari covid 19 dan lain sebagainya.

Statistik seringkali digunakan untuk menganalisis sepuluh besar penyakit menular dan tidak menular yang terjadi di suatu negara, menganalisis data tentang angka kematian ibu dan anak, menganalisis cakupan suatu imunisasi sampai. Di dalam suatu penelitian kesehatan, statistik kerap kali digunakan untuk menguji kebenaran suatu hipotesis, sebagai contoh untuk mencari hubungan suatu paparan sesuatu dengan kejadian penyakit tertentu. Statistik juga memiliki peran yang penting pada pengujian suatu terapi atau obat baru yang sebelum disebarluaskan ke masyarakat.



Gambar 1.1: pengujian penelitian

1.2 Penerapan Statistik dalam Penelitian Kesehatan

Peranan statistik sangat penting dalam penelitian kesehatan, saat kita akan melakukan suatu penelitian, statistik berperan mulai dari bagaimana penelitian direncanakan. Peran statistik sangat vital terutama dalam penelitian kuantitatif. Statistik akan dijadikan acuan dalam penentuan rancangan penelitian, populasi dan sampel, pemilihan sampel supaya representatif atau mewakili populasi harus sejalan dengan kaidah-kaidah statistik. Tahap selanjutnya statistik akan digunakan dalam proses pengolahan data atau yang sering kita sebut sebagai proses analisis, statistik akan membantu merubah data kita sehingga dapat menguji hipotesis yang sudah ditetapkan.

1.3 Fungsi Statistik

Statistik mempunyai beberapa fungsi penting dalam bidang kesehatan maupun penelitian kesehatan antara lain:

1.3.1 Untuk Menyajikan Informasi atau Data Dalam Bentuk Yang Lebih Mudah Dimengerti

Setiap hari kita menemukan ratusan bahkan ribuan informasi dalam bidang kesehatan. Ketika potongan informasi tersebut dianalisis dengan statistik tertentu dan disajikan dalam bentuk yang mudah dimengerti (tabel atau gambar). Misalnya, saat kita mengatakan bahwa ada beberapa kasus kematian Ibu melahirkan yang terjadi di Provinsi A, contoh ada 1000 ibu hamil yang melahirkan dan 10 orang meninggal dunia, dengan analisis statistik tertentu akan didapatkan dapat angka kematian ibu melahirkan di Provinsi A 1 per 100 orang, jadi pembaca akan lebih mudah memahami informasi yang akan disampaikan (Bajwa, 2017).

1.3.2 Untuk Menyederhanakan Data Yang Komplek

Dalam suatu penelitian ataupun kehidupan sehari-hari, seringkali kita mendapat Informasi yang komplek dan beragam, statistik akan menyederhanakan data untuk lebih mudah dimengerti, dengan cara mengambil beberapa data untuk dijadikan sampel yang representatif. Data yang kompleks dapat disederhanakan dengan disajikan dalam bentuk, grafik, diagram, atau merepresentasikannya melalui rata-rata.

1.3.3 Statistik Dapat Digunakan Untuk Melakukan Perbandingan

Seringkali data penelitian menjadi lebih jelas dan signifikan pada saat dibandingkan dengan data lain. Beberapa teknik statistik dapat digunakan untuk melakukan perbandingan data baik pada satu kelompok ataupun lebih, seperti perbandingan mean dari suatu data.

1.3.4 Untuk Memperluas Pengalaman

Sebagai individu, pengetahuan kita terbatas pada apa yang kita lihat dan amati di sekitar kita, ini merupakan bagian kecil dari luasnya ilmu pengetahuan.

Statistik memperluas pengetahuan dan pengalaman dengan menyajikan berbagai kesimpulan dan hasil, yang sebelumnya sudah dilakukan investigasi. Misalnya kita setiap hari melihat dan mendengarkan tentang kenaikan biaya hidup disuatu daerah, melihat perbedaan pendapatan perorangan masing-masing daerah dan berapa biaya hidup yang dibutuhkan masing-masing daerah, dalam bidang kesehatan kita bisa melihat kejadian penyakit yang berbeda-beda antara satu daerah yang ada di pada suatu negara, dan juga mengetahui kenapa terdapat perbedaan penyakit yang terjadi pada suatu daerah.

1.3.5 Memberikan Pertimbangan Dalam Perumusan Kebijakan

Statistik dapat menjadi pertimbangan dalam merumuskan kebijakan kesehatan yang akan diambil oleh pemangku kepentingan. Statistik akan membantu pemangku kepentingan dalam mempertimbangan kebijakan yang akan dirumuskan. Perumusan kebijakan kesehatan akan sangat memerlukan analisis statistik dalam menentukan program kesehatan mana yang akan menjadi prioritas. Tentunya perumusan kebijakan yang menjadi prioritas akan mempertimbangkan besarnya masalah kesehatan yang ada.

1.3.6 Untuk melakukan pengukuran pada fenomena yang besar dan luas

Statistik memungkinkan kita mengukur fenomena yang luas dan besar, Perkiraan populasi suatu negara atau jumlah pasangan usia produktif, cakupan imunisasi dan angka kematian ibu dan anak di suatu negara.

1.4 Jenis Statistik

Para peneliti menggunakan berbagai macam prosedur statistik dalam menginterpretasikan data, prosedur ini dapat digolongkan menjadi dua jenis yaitu: statistik deskriptif dan statistik inferensial.

1.4.1 Statistik Deskriptif

Statistik deskriptif digunakan untuk mendeskripsikan atau menggambarkan sebaran data penelitian. Jika merujuk pada kata deskriptif berarti statistik yang digunakan untuk menganalisis sebaran atau distribusi frekuensi suatu data tanpa melakukan analisis lebih lanjut, statistik deskriptif tidak digunakan untuk membuat kesimpulan yang lebih luas (generalisasi/inferensial).

1.4.2 Statistik Inferensial

Statistik inferensial adalah statistik yang digunakan untuk menganalisis data penelitian, hasil analisis data dengan menggunakan statistik inferensial dapat digeneralisasikan pada populasi di mana data atau penelitian dilakukan.

Terdapat dua jenis statistik inferensial yaitu statistik parametrik dan statistik nonparametrik. Statistik parametrik digunakan untuk menganalisis data yang berskala Numerik (interval dan rasio) dengan syarat distribusi atau sebaran data harus berdistribusi normal. Statistik nonparametrik digunakan untuk menganalisis data dengan skala kategorik (nominal dan ordinal), selain untuk menganalisis data kategorik, statistik nonparametrik juga digunakan untuk menganalisis data dengan skala interval dan rasio dengan distribusi data tidak normal, pada statistik parametrik tidak membutuhkan syarat data yang berdistribusi normal seperti statistik parametrik.

1.5 Tahapan analisis statistik

Dalam melakukan analisis statistik ada lima tahapan penting yang harus dilakukan, antara lain: pengumpulan data, pengolahan, analisis dan kesimpulan.

1.5.1 Pengumpulan data

Tahap awal dari suatu analisis adalah tahap pencarian Informasi atau pengumpulan data, pada tahap ini akan dilakukan pengumpulan Informasi dari masalah yang diteliti. Tahap pengumpulan data ini merupakan tahapan yang sangat penting, karena ketepatan dalam melakukan proses pengumpulan data akan sangat memengaruhi kualitas data yang akan dianalisis. Berbagai macam teknik biasa digunakan dalam pengumpulan data baik secara langsung maupun

tidak langsung, pengumpulan data dapat dilakukan dengan berbagai alat ukur seperti: kuesioner, pedoman wawancara mendalam ataupun menggunakan alat (termometer, sphygmomanometer dll)

1.5.2 Pengolahan data

Pengolahan data adalah proses menganalisis data menjadi lebih bermakna yg digunakan untuk membuktikan hipotesis suatu penelitian, ketepatan dalam melakukan teknik analisis menjadi kunci di tahap ini.

Ada beberapa tahap yang dilakukan dalam pengolahan data, antara lain:

1. Editing, editing bertujuan untuk memeriksa kelengkapan data
2. Coding atau pemberian kode, pemberian kode ini bertujuan memudahkan data di analisis
3. Entry, pada tahap entry data yang sudah dicoding dimasukkan kedalam aplikasi pengolahan data komputer (SPSS, STATA dll)
4. Cleaning, pada tahap ini data dilakukan pengecekan kembali untuk memastikan tidak ada kesalahan pada data yang akan dianalisis.

1.5.3 Penyajian data

Setelah data dianalisis, biasanya output yang didapat masih sulit dimengerti dan susunannya kurang teratur, pada tahap penyajian data ini bertujuan menjadikan data hasil analisis lebih tertata dan mudah dipahami.

1.5.4 Analisis

Setelah data benar-benar siap akan dilanjutkan dengan proses analisis untuk mendapatkan Informasi yang diinginkan. Pemilihan teknik analisis mengikuti skala data dari variabel yang diukur, secara umum ada dua Teknik statistik yang biasa digunakan, yaitu statistik deskriptif dan inferensial. Analisis yang dilakukan mengacu pada tujuan penelitian.

1.5.5 Interpretasi data

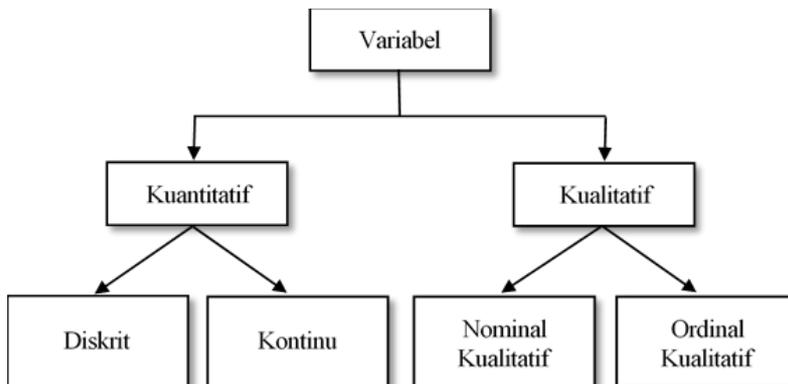
Tahap terakhir adalah melakukan interpretasi terhadap data yang dianalisis, interpretasi data hendaknya melihat rumusan masalah dan hipotesis dari suatu penelitian.

1.6 Variabel

Variabel adalah sesuatu yang cenderung bervariasi atau sesuatu yang dapat bervariasi. Dengan kata lain variabel juga bisa dapat dikatakan kumpulan nilai dari suatu objek (Bajwa, 2017). Variabel bisa dikatakan suatu objek atau fenomena yang menjadi pusat pengamatan dalam suatu penelitian, variabel inilah yang nantinya menjadi objek penelitian yang akan dilakukan pengukuran dengan berbagai macam cara sehingga didapatkan data yang akan dilanjutkan pada analisis sehingga didapatkan Informasi.

Variabel juga sering disebut item data. Contoh variabel pada manusia: tinggi badan, berat badan, lingkaran lengan, kadar gula darah, kadar kolesterol, tekanan darah, golongan darah, jumlah gaji, status ekonomi, tipe kepribadian, kualitas hidup, kualitas tidur, status perkawinan, jumlah anak, dll. Variabel inilah yang nantinya akan dilakukan pengukuran oleh peneliti, variabel yang menghasilkan data numerik atau berupa angka disebut variabel kuantitatif, sedangkan variabel yang menghasilkan data atau Informasi non numerik disebut variabel kualitatif atau kategorikal.

Dari contoh diatas kalau kita kelompokkan variabel kuantitatif adalah: tinggi badan, berat badan, lingkaran lengan, kadar gula darah, kadar kolesterol, tekanan darah, jumlah gaji dan jumlah anak, sedangkan variabel kategorikal adalah: golongan darah, status ekonomi, tipe kepribadian, kualitas hidup, kualitas tidur dan status perkawinan.



Gambar 1.2: Variabel

1.6.1 Variabel kualitatif

Variabel penelitian kualitatif atau sering juga disebut variabel kategorik adalah variabel yang berupa nama atau label terhadap objek pengukuran. Contoh variabel kualitatif: nama kabupaten kota yang ada di Bali (Denpasar, Tabanan, Jembrana, Singaraja, Bangli, Klungkung, Badung dan Karangasem), status perkawinan (menikah, tidak menikah, janda dan duda), jenis kelamin (laki-laki dan perempuan), jenis-jenis anjing (ras, lokal dll), nilai mata kuliah statistik (A, B, C dan D)

Secara umum variabel kualitatif dapat dibedakan menjadi dua:

1. Nominal kualitatif: tidak dapat diurutkan atau mempunyai posisi atau kedudukan yang sama. Contohnya jenis kelamin, dan Status imunisasi atau vaksinasi dasar.
2. Ordinal kualitatif: dapat diurutkan dan berupa tingkatan. Contohnya nilai mata kuliah statistik A, B dan C.

1.6.2 Variabel kuantitatif

Variabel kuantitatif atau juga sering disebut variabel numerik adalah variabel yang berupa angka atau nomor. Contohnya kita berbicara tentang nominal gaji pegawai negeri sipil pada dinas kesehatan, atau membicarakan jumlah kasus baru covid 19 di suatu negara atau wilayah, suhu badan seseorang, panjang dari suatu objek. Kita juga bisa mengatakan variabel kuantitatif adalah suatu variabel yang bisa diukur dengan hasil ukur berupa angka atau nomor (Riffenburgh and Gillen, 2020).

Variabel kuantitatif dibagi menjadi dua yaitu: diskrit dan kontinu.

1. Variabel diskrit adalah variabel yang datanya berupa bilangan bulat, nilai pada variabel diskrit hanya digunakan sebagai kode atau label dan tidak mempunyai nilai yang sebenarnya. Contoh memberikan kode 1 dan 0 pada Jenis kelamin (laki-laki=1, dan perempuan=0), variabel diskrit sering digunakan untuk membedakan dua jenis data atau dikotomi, sering juga disebut nominal dikotomi. Data dikotomi juga bisa kita temukan pada penelitian-penelitian kasus kontrol, kode 1 sering diberikan pada kasus dan kode 0 pada kontrol.

2. Variabel kontinu: data dari variabel yang didapatkan melalui pengukuran, data variabel kontinu bisa berupa pecahan ataupun desimal, pada dasarnya data dari variabel kontinu bisa dilakukan penghitungan matematika. Contoh data variabel kontinu: suhu ruangan, tinggi badan, lama waktu tidur, lama waktu istirahat, jumlah gaji, jumlah kekayaan seseorang, jumlah kendaraan yang dimiliki dan lain sebagainya. Data variabel kontinu secara umum dapat dibedakan menjadi tiga: ordinal, interval dan rasio (Bruce and Bruce, 2017).

1.7 Skala Data variabel

Ada dua tipe dasar dari variabel yaitu kategorikal(nominal dan ordinal) dan numerik(interval dan rasio), masing-masing akan memerlukan cara pengukuran dan analisis statistik yang berbeda tergantung jenis skala data dari masing-masing variabel yang digunakan. Ada empat jenis skala data yang umum digunakan dalam analisis statistik: nominal, ordinal, interval dan rasio.

1.7.1 Skala Nominal

Nominal adalah bentuk skala pengukuran paling sederhana yang digunakan peneliti. Kata Nominal memiliki arti “berhubungan dengan nama”. Pengukuran yang dilakukanpun hanya memberi nama terhadap variabel yang diukur. Nominal juga bisa diasumsikan sekumpulan kategori yang mempunyai nama berbeda. Pengukuran pada skala nominal bertujuan untuk memberi label dan mengkategorikan pengamatan tetapi tidak membuat perbedaan kuantitatif di antara mereka. Antara kategori yang satu dengan yang lainnya mempunyai tingkatan yang sama, dalam artian antara kategori yang satu dengan yang lain tidak mempunyai kedudukan yang berbeda. Pada penggolongan tingkatan data, skala nominal sering disebut skala data yang paling rendah tingkatannya di antara skala data yang lain.

Contoh: jenis kelamin terdiri dari 2 kategori (laki-laki dan perempuan), dapat dikatakan antara laki-laki dan perempuan memiliki makna yang sejajar atau sama (laki-laki tidak memiliki arti lebih tinggi dari perempuan, begitupun sebaliknya) contoh lain: seorang peneliti mengamati perilaku sekelompok bayi

baru lahir mungkin mengkategorikan tanggapan seperti langsung menangis, diam, hanya bergerak tanpa menangis.

1.7.2 Skala Ordinal

Ordinal adalah skala kategori yang memiliki tingkatan pada masing-masing kategori dari data pengukuran. Skala ordinal biasanya disusun atau diurutkan berdasarkan peringkat, contohnya seorang peneliti melakukan pengukuran tingkat kepuasan terhadap pelayanan suatu rumah sakit, maka hasil kategori dari pengukuran akan muncul “sangat puas” “puas” “tidak puas” “sangat tidak puas”, di antara masing-masing kategori ini pastinya adalah sebuah tingkatan, puas memiliki nilai lebih tinggi dari tidak puas, dan biasanya data kategori ordinal sebelumnya berupa angka yang kemudian dikategorikan.

Contoh lain data ordinal: nilai ujian mata kuliah statistik pada anak semester dua dikategorikan “A” “A-” “B+” “B”, nilai kategori seperti ini dapat digolongkan kedalam skala ordinal walaupun nilai awal berupa skor, dikategorikan A jika skor 85-100, dikategorikan A- jika mendapatkan skor 80-84.99.

1.7.3 Skala Interval

Skala interval merupakan salah satu jenis data numerik, dikatakan numerik berarti data berupa nomor atau angka, jika dilihat dari peringkat di antara skala data, skala numerik dikatakan mempunyai tingkatan yang lebih baik dari skala ordinal jika dilakukan analisis statistik, skala interval memiliki semua sifat atau karakteristik skala ordinal. Skala interval mempunyai jarak yang sama antara data satu dengan yang lainnya, sebagai contoh suhu tubuh “36°C” “37°C” “38°C” “39°C” jarak atau interval suhu 36°C ke 37°C sama dengan jarak dari 37°C ke 38°C. Skala interval dikatakan tidak memiliki 0 mutlak atau tidak memiliki titik awal. Titik 0 pada skala interval tidak berarti ketiadaan dari apa yang sedang diukur, contoh kita mengukur suhu ruang pendingin 0°C, bukan berarti ruangan tersebut tidak ada suhu, bahkan sebaliknya suhu di ruangan tersebut sangat dingin. Karakteristik lain dari skala interval adalah bisa dilakukan penghitungan aritmatika tetapi tidak mempunyai makna yang sebenarnya, sebagai contoh: suhu air di gelas A 50°C ditambah dengan suhu air di gelas B 40°C jika dilakukan penghitungan matematika maka $50^{\circ}\text{C} + 40^{\circ}\text{C} = 90^{\circ}\text{C}$ apakah 90°C mempunyai makna sebenarnya jika kita lakukan pengukuran suhu ulang? tentu tidak.

1.7.4 Skala Rasio

Rasio adalah skala data yang mempunyai tingkatan paling tinggi bila dibandingkan tiga skala lainnya, skala rasio memiliki semua ciri dari skala interval akan tetapi memiliki titik nol mutlak atau absolut. Contoh skala rasio adalah berat suatu benda 20Kg 30Kg, gaji juga bisa kita golongkan skala rasio. Banyak di antara kita mungkin agak kebingungan membedakan skala interval dengan rasio, ada beberapa perbedaan mendasar antara skala interval dan rasio, yang pertama adalah 0 mutlak atau absolut, maksud 0 mutlak ini adalah jika kita mengatakan uang di dompet saya Rp. 0.- berarti saat ini dompet saya tidak berisi uang, 0 mempunyai arti sebenarnya. Perbedaan lain dari data rasio bisa dilakukan perhitungan aritmatika dan mempunyai makna sebenarnya, contoh gaji amir dalam sebulan Rp. 5.000.000,- gaji anton Rp. 4.500.000,- bila gaji amir dan anton ditambahkan menjadi Rp. 9.500.000,- jumlah hasil penambahan ini mempunyai makna sebenarnya.

Bab 2

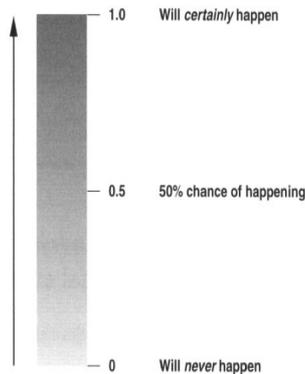
Probabilitas

2.1 Pendahuluan

Pemahaman tentang probabilitas sangat penting untuk memahami statistik. Kompetensi dalam probabilitas diperlukan untuk memahami p-value, membaca tabel tabulasi silang, dan memahami distribusi frekuensi, yang digunakan dalam penelitian kesehatan. Selain itu, beberapa uji statistik yang digunakan untuk menentukan nilai signifikansi suatu hasil bergantung pada distribusi probabilitas seperti distribusi normal (Plichta, Stacey B; Garzon, 2009). Kita sering mendengar berapakah probabilitas munculnya mata dadu bergambar kepala jika kita melempar koin tersebut? Kemudian, berapakah probabilitas seorang anak yang sudah diberikan vaksin BCG akan mengalami penyakit TBC? Dalam alam tidak ada kepastian, begitu juga dengan keberadaan statistik sering dianggap karena sebuah ketidakpastian, sehingga diperlukan membuat suatu kesimpulan (Sabri, Luknis; Hastono, 2018). Ketika kita berbicara tentang probabilitas atau peluang suatu peristiwa, maka kita mengacu pada pernyataan tentang seberapa besar kemungkinan (apa kemungkinannya) bahwa peristiwa itu akan terjadi? (Stommel and Katherine, 2014). Dalam bab ini akan dibahas mengenai probabilitas dan asas perhitungan probabilitas tersebut.

2.2 Pengertian Probabilitas

Dalam hidup kita dihadapkan pada ketidaktahuan. Kita merenung mengenai kemacetan lalu lintas dan cuaca besok pagi, harga saham minggu depan, atau pemilihan umum yang akan datang. Kita seringkali kita tidak mengetahui suatu dengan pasti, sebaliknya kita dipaksa untuk menebak atau memperkirakan. Probabilitas adalah ilmu ketidakpastian (Selvamuthu and Das, 2018). Teori statistik didasarkan pada teori probabilitas yang pada awalnya digunakan untuk menyelidiki pola dalam permainan judi dengan menggunakan kartu dan dadu (Peacock and Peacock, 2010). Teori probabilitas membuat suatu prediksi mengenai eksperimen yang hasilnya tergantung pada kebetulan (Selvamuthu and Das, 2018). Probabilitas adalah cara untuk mengukur ketidakpastian suatu peristiwa. Misalnya, ada kemungkinan kecil kita memenangkan sebuah undian namun ada peluang yang jauh lebih besar bahwa kita tidak akan memenangkan undian tersebut. Seseorang yang membeli tiket undian pada dasarnya bersedia membayar ketidakpastian (kemungkinan) untuk menang (Plichta, Stacey B; Garzon, 2009).



Gambar 2.1: Skala Probabilitas (Stewart, 2018)

Dalam teori probabilitas, percobaan atau eksperimen tidak selalu merupakan percobaan yang rumit tetapi seringkali menggunakan percobaan sederhana dengan menggunakan alat yang sederhana. Misalnya, kita ingin memproduksi sekrup mesin maka akan ada kemungkinan beberapa di antara sekrup tersebut ada yang rusak sehingga kemungkinan hasil yang diperoleh adalah rusak atau tidak rusak (Setiawan, 2015).

Probabilitas digunakan untuk memprediksi seberapa besar kemungkinan suatu peristiwa akan terjadi. Probabilitas diukur pada skala dari 0 hingga 1.0 seperti yang ditunjukkan pada Gambar 2.1. Misalnya, ketika seseorang melempar koin, ada kemungkinan 50% untuk mendapatkan gambar A. Probabilitas biasanya dinyatakan dalam format desimal (- 50% menjadi 0,5; 10% menjadi 0,1; dan 5% menjadi 0,05). Oleh karena itu, kemungkinan mendapatkan gambar A ketika sebuah koin dilemparkan adalah 0,5. Probabilitas tidak pernah bisa lebih dari 1,0 dan juga tidak bisa negatif (Stewart, 2018).

2.2 Konsep Probabilitas

Dalam statistik, terdapat beberapa konsep yang kesemuanya digunakan dalam memahami arti probabilitas, yaitu (Sabri, Luknis; Hastono, 2018):

1. Pandangan klasik/ intuitif

Menurut pandangan ini, probabilitas menunjukkan seberapa besar kemungkinan suatu peristiwa terjadi, di antara keseluruhan peristiwa yang terjadi.

Rumus yang digunakan adalah:

$$P(E) = X / N$$

P = Probabilitas

E = Event (kejadian)

X = Jumlah kejadian yang diinginkan (peristiwa)

N = Keseluruhan kejadian yang diinginkan

Contoh:

- a. Sebuah koin mempunyai dua buah sisi yaitu sisi H dan T. Jika koin tersebut dilempar ke atas satu kali, maka peluang untuk keluar sisi H adalah $\frac{1}{2}$
- b. Sebuah dadu jika dilemparkan ke atas satu kali maka peluang munculnya mata 'enam' adalah $\frac{1}{6}$.

2. Pandangan empiris

Menurut pandangan ini, probabilitas adalah berdasarkan pada pengamatan, pengalaman, atau suatu peristiwa yang telah terjadi.

Contoh:

- a. Sebuah koin dilemparkan ke atas 100 kali, jika kejadian yang diinginkan adalah 59 kali keluar sisi H, maka $P(H) = 59\%$
 - b. Jumlah hasil produksi tekstil adalah 10.000 dan 100 rusak. Maka, $P(\text{rusak}) = 100/10.000 = 1\%$
- ## 3. Pandangan subjektif

Menurut pandangan ini, probabilitas ditentukan oleh orang yang membuat pernyataan. Misalnya, seseorang meyakini bahwa jika ada kesempatan untuk melanjutkan pendidikan di sebuah universitas, maka yang diterima adalah dirinya (misalnya dia meyakini $95\% = 0.95$).

2.3 Unsur-Unsur Probabilitas

Dalam statistik kita perlu memahami ruang sampel, titik sampel, dan distribusi probabilitas. Berikut adalah beberapa definisi yang sering digunakan dalam eksperimen (Plichta, Stacey B; Garzon, 2009; Sabri, Luknis; Hastono, 2018):

1. Ruang sampel adalah kumpulan dari semua kemungkinan hasil suatu eksperimen. Misalnya, jika sekeping koin dibalik, maka ruang sampel memiliki dua kemungkinan hasil yaitu akan terlihat koin bagian kepala dan ekor. Jika sebuah dadu dilemparkan ke atas, maka ruang sampel memiliki enam kemungkinan hasil. Demikian pula, ruang sampel untuk gender memiliki dua hasil yaitu perempuan dan laki-laki. Ruang sampel ditulis dengan lambang S. Jika suatu eksperimen adalah $a_1, a_2, a_3, a_4, \dots, a_n$ menunjukkan semua hasil yang terjadi, maka ruang sampel dituliskan:

$$S = (a_1, a_2, a_3, a_4, \dots, a_n)$$

2. Titik sampel merupakan semua elemen yang ada dalam ruang sampel, yaitu $a_1, a_2, a_3, a_4, \dots, a_n$
3. Distribusi probabilitas adalah sekumpulan probabilitas yang terkait dengan setiap kemungkinan kejadian di ruang sampel. Distribusi probabilitas suatu variabel dapat dinyatakan dengan tabel, grafik, atau rumus.
4. Peristiwa/ Kejadian/ Event adalah himpunan bagian dari suatu ruang sampel, ditulis dengan huruf besar A, B , dan seterusnya dan dituliskan peristiwa yang mungkin muncul dalam hasil.

Misalnya a_1, a_2 adalah hasil dari peristiwa, maka yang dituliskan $A =$ hasil yang diterima $\{ a_1, a_2 \}$

Contoh:

Jika sebuah dadu dilemparkan ke atas, maka hasil adalah mata dadu yang muncul. Ruang sampel adalah $S = (1, 2, 3, 4, 5, 6)$.

Suatu peristiwa adalah:

$A =$ Bilangan ganjil yang muncul $\{1, 3, 5\}$

$B =$ Bilangan genap yang muncul $\{2, 4, 6\}$

Kejadian A adalah kejadian memperoleh mata dadu yang merupakan bilangan ganjil sedangkan kejadian B adalah kejadian memperoleh mata dadu yang merupakan bilangan genap (Setiawan, 2015).

2.4 Asas Perhitungan Probabilitas

Asas dalam perhitungan probabilitas terdiri dari dua jenis perhitungan yaitu pertambahan dan perkalian. Dalam hukum pertambahan, jika diperlukan probabilitas dalam dua peristiwa maka digunakan kata kunci 'atau' pada pernyataan tersebut. Contoh, probabilitas untuk keluar mata dadu dua atau lima pada pelemparan satu kali mata dadu. Dan untuk hukum perkalian digunakan kata kunci 'dan'. Contoh, sebuah dadu dan koin dilemparkan secara bersamaan, maka berapakah peluang terjadinya hasil lemparan sisi koin bergambar H dan mata dadu tiga (Sabri, Luknis; Hastono, 2018). Probabilitas

dapat dinyatakan dengan atau tanpa kondisi yang menyertainya. Misalnya, kita ingin mengetahui kemungkinan bahwa penduduk dalam suatu negara akan tertular flu selama periode dari Januari hingga April di tahun 2021. Atau mungkin kita ingin mengetahui kemungkinan bersyarat (conditional probability) yaitu mengingat bahwa seseorang telah menerima vaksin flu pada bulan oktober atau november, seberapa besar kemungkinan orang tersebut akan terserang gejala flu yang parah antara Januari dan April? (Stommel and Katherine, 2014).

Berikut terdapat beberapa formula dalam menghitung probabilitas (Plichta, Stacey B; Garzon, 2009; Stewart, 2018):

1. Menghitung probabilitas (P) dari satu peristiwa (A)

Misalnya untuk mengetahui probabilitas dengan melempar sebuah dadu

$$P(A) = \frac{\text{jumlah kemungkinan dari satu peristiwa}}{\text{jumlah kemungkinan hasil yang sama}}$$

$$P(A) = \frac{\text{jumlah enam dalam dadu}}{\text{jumlah sisi pada dau}}$$

$$P(A) = \frac{1}{6} = 0.1667 = 16.67\%$$

2. Menghitung probabilitas dari satu peristiwa (A) dan peristiwa (B) terjadi (peristiwa independen/ bersyarat). Peristiwa ini disebut independen apabila kejadian atau ketidakjadian dari satu peristiwa tidak memengaruhi peristiwa lain.

Misalnya, jika Anda memiliki dua kartu yaitu kartu A dan kartu B, maka berapakah probabilitas untuk menarik kartu As dari kedua kartu tersebut?

$$\text{Rumus} = P(A) \times P(B)$$

$$P(\text{paket A}) = 1 \text{ kartu (1 paket kartu 52)} = 1/52 = 0.0192$$

$$P(\text{paket B}) = 1 \text{ kartu (1 paket kartu 52)} = 1/52 = 0.0192$$

$$P(A) \times P(B) = 0.0192 \times 0.0192 = 0.00037$$

Ini disebut aturan perkalian, contoh diatas menjelaskan bahwa peristiwa A dan B tidak bergantung satu sama lain. Ini berarti bahwa satu peristiwa

yang terjadi terlepas dari yang lain, dan hasilnya tidak berhubungan dengan yang lain.

3. Menghitung probabilitas dari suatu peristiwa (A) dan peristiwa (B) (conditional probability/ peristiwa bersyarat). Peristiwa ini disebut bersyarat apabila kejadian atau ketidakjadian dari satu peristiwa akan berpengaruh terhadap peristiwa yang lain.

Berapa probabilitas penarikan kartu As dan King secara berurutan dari satu kartu?

$$\text{Rumus} = P(A) \times P(B | A)$$

Berdasarkan contoh nomor 2, probabilitas penarikan kartu As dari satu kartu (berisi 52 kartu) adalah $1/52 = 0,0192$, maka $P(A) = 0,0192$.

Peluang untuk menarik kartu King sedikit lebih tinggi, karena sisa kartu menjadi 51 kartu, maka probabilitas $P(B | A) = 1/51 = 0,0196$.

Maka, $P(A) \times P(B | A) = (1/52) \times (1/51) = 0,0192 \times 0,0196 = 0,0004$

4. Menghitung kemungkinan terjadinya peristiwa (A) atau peristiwa (B). Ini disebut juga dengan kejadian mutually exclusive (saling terpisah/ saling meniadakan/ disjoint).

Contoh :

- Permukaan sebuah koin
- Permukaan dadu
- Kelahiran anak perempuan atau laki-laki pada ibu yang hamil tunggal

Contoh, jika kita melemparkan sebuah dadu, maka berapakah probabilitas untuk muncul dadu enam atau dadu lima?

$$\text{Rumus: } P(A) + P(B)$$

$$P(A) = 0,1667$$

$$P(B) = 0,1667$$

$$P(A) + P(B) = 0,1667 + 0,1667 = 0,333 \text{ (atau } 33,3\%)$$

5. Menghitung kemungkinan terjadinya peristiwa (A) atau peristiwa (B) (Non mutually exclusive/ joint/ terjadi secara bersama).

Contoh dari peristiwa ini adalah:

- a. Penarikan karu As dan King
- b. Seorang dosen dan guru

Misalkan sebuah penelitian menemukan bahwa 90% orang yang berusia di atas 60 tahun di suatu daerah menderita setidaknya satu kali flu dalam periode satu tahun, dan 20% menderita diare setidaknya satu kali. Berapa probabilitas seseorang yang berusia di atas 60 tahun mengalami flu biasa atau diare?

Diasumsikan bahwa flu dan diare terjadi secara independen.

Rumus = $P(A) + P(B) - P(A \text{ dan } B)$

$$P(A) = 0.9 \text{ (flu)}$$

$$P(B) = 0.2 \text{ (diare)}$$

$$P(A \text{ dan } B) = 0.9 \times 0.2 = 0.18$$

$$\begin{aligned} \text{Maka, } P(A) + P(B) - P(A \text{ dan } B) &= (0.9 + 0.2) - 0.18 \\ &= 1.1 - 0.18 \\ &= 0.92 \text{ (or } 92\%) \end{aligned}$$

Sehingga, ada kemungkinan bahwa 0,92 (92%) dari orang yang berusia di atas 60 tahun di suatu daerah akan mengalami flu atau diare dalam periode satu tahun.

6. Joint Probability

Joint probability/ probabilitas gabungan adalah terjadinya dua peristiwa atau lebih secara bersamaan. Kunci untuk memahami joint probability adalah menggunakan 'keduanya' dan 'dan'. Joint probability ini biasa disusun dalam tabel silang. Dalam tabel berikut akan dijelaskan mengenai jumlah perawat dalam suatu rumah sakit berdasarkan kelompok usia dan jenis kelamin.

Tabel 2.1: Jumlah perawat berdasarkan kelompok usia dan jenis kelamin

Usia	Jenis Kelamin		jumlah
	Wanita	Laki-laki	
<40 tahun	50	30	80
>40tahun	20	10	30
Jumlah	70	40	110

- Probabilitas perawat wanita adalah $70/110 = 0.63$. Ini disebut dengan probabilitas marginal.
- Probabilitas perawat berumur < 40 tahun $80/110 = 0.72$. Ini disebut dengan probabilitas marginal.
- Probabilitas perawat wanita dan berumur < 40 tahun adalah $50/110 = 0.45$. Ini disebut dengan joint probabilitas.

Bab 3

Populasi dan Sampling

3.1 Pendahuluan

Penelitian adalah pekerjaan ilmiah yang bertujuan untuk mengungkapkan rahasia ilmu secara objektif, dengan berdasarkan bukti-bukti yang lengkap. Penelitian merupakan proses kreatif untuk mengungkapkan suatu gejala melalui cara tersendiri sehingga diperoleh suatu informasi. Pada dasarnya informasi tersebut merupakan jawaban atas masalah-masalah yang dipertanyakan sebelumnya. Oleh karena itu, penelitian juga dapat dipandang sebagai usaha mencari tahu tentang berbagai masalah yang dapat merangsang pikiran atau kesadaran seseorang. Penelitian ilmiah pada dasarnya merupakan usaha memperkecil interval dugaan peneliti melalui pengumpulan dan penganalisaan data atau informasi yang diperolehnya. Salah satu bagian dalam langkah-langkah penelitian adalah menentukan populasi dan sampel penelitian.

Untuk dapat melaksanakan penelitian dengan baik, seorang peneliti harus memahami konsep populasi dan sampel. Populasi merupakan keseluruhan objek atau subjek penelitian, sedangkan sampel merupakan Sebagian atau wakil yang memiliki karakteristik representasi dari populasi. Untuk dapat menentukan atau menetapkan sampel yang tepat, maka diperlukan pemahaman yang baik dari peneliti mengenai sampling, baik penentuan jumlah maupun dalam menentukan sampel mana yang diambil. Kesalahan

dalam menentukan populasi akan berakibat tidak tepatnya data yang dikumpulkan sehingga hasil penelitian tidak memiliki kualitas yang baik, tidak representatif, dan tidak memiliki generalisasi yang baik.

Kegiatan penelitian banyak dilakukan dengan penarikan sampel dari populasi, karena metode penarikan sampel lebih praktis, biaya yang dikeluarkan lebih hemat, dan memerlukan waktu serta tenaga yang lebih sedikit dibandingkan dengan metode sensus. Penentuan sampel dari suatu populasi disebut dengan penarikan sampel (Sukmadinata, 2011). Penelitian yang memakai sampel untuk meneliti karakteristik objek penelitian, dilakukan dengan beberapa alasan antara lain objek yang diteliti sifatnya mudah rusak, objek yang diteliti bersifat homogen, tidak mungkin meneliti secara fisik seluruh objek dalam populasi, bertujuan untuk menghemat biaya, menghemat waktu dan tenaga, serta keakuratan hasil sampling.

Di dalam sebuah penelitian yang menggunakan sampel sebagai sebuah unit analisis, baik pada penelitian kuantitatif dan kualitatif, setidaknya terdapat dua hal yang menjadi masalah yang dihadapi yaitu persoalan sampling dan pengambilan sampel. Persoalan sampling adalah proses untuk mendapatkan sampel dari suatu populasi. Sampel harus bisa mencerminkan keadaan populasi, artinya kesimpulan hasil penelitian yang diangkat dari sampel harus merupakan kesimpulan atas populasi. Sehingga masalah yang dihadapi adalah bagaimana cara memperoleh sampel yang representatif, yaitu sampel yang dapat mewakili elemen lain dalam populasi atau mencerminkan keadaan populasi. Masalah lain yang dihadapi dalam penelitian yang menggunakan sampel sebagai unit analisis adalah bagaimana proses pengambilan sampel dan berapa banyak unit analisis yang akan diambil sebagai sampel (Mamik, 2015).

3.2 Populasi

Penelitian bertujuan untuk memperoleh kesimpulan tentang kelompok yang besar dalam lingkup wilayah yang luas, tetapi hanya dengan meneliti kelompok kecil dalam daerah yang lebih sempit. Lingkup wilayah bisa mencakup seluruh wilayah negara, satu propinsi ataupun suatu kota atau kabupaten. Kelompok besar dan wilayah yang menjadi lingkup penelitian disebut populasi (Sukmadinata, 2012).

Populasi adalah wilayah generalisasi yang terdiri atas objek atau subjek yang mempunyai kuantitas dan karakteristik tertentu yang ditetapkan oleh peneliti untuk dipelajari dan kemudian ditarik kesimpulannya (Sugiyono, 2001). Populasi adalah seluruh data yang menjadi perhatian kita dalam suatu ruang lingkup dan waktu yang kita tentukan. Jadi populasi berhubungan dengan data, bukan manusianya. Kalau setiap manusia memberikan suatu data, maka banyaknya atau ukuran populasi akan sama dengan banyaknya manusia. Populasi adalah keseluruhan subjek penelitian (Arikunto, 2006).

Populasi merupakan semua anggota kelompok orang, kejadian, atau objek yang telah dirumuskan secara jelas (Furchan, 2004). Populasi merupakan kumpulan dari individu dengan kualitas serta ciri-ciri yang telah ditetapkan. Kualitas atau ciri tersebut dinamakan variabel. Sebuah populasi dengan jumlah individu tertentu dinamakan populasi finit. Sedangkan apabila jumlah individu dalam kelompok tidak mempunyai jumlah yang tetap, maupun jumlahnya tidak terhingga disebut populasi infinit. Misalnya, jumlah petani dalam sebuah desa adalah finit. Sebaliknya, jumlah pelemparan mata dadu yang terus-menerus merupakan populasi infinit (Nazir, 2005).

Dalam penelitian, populasi ini dibedakan antara populasi secara umum dengan populasi target atau disebut dengan "target population". Populasi target adalah populasi yang menjadi sasaran keberlakuan kesimpulan penelitian. Populasi adalah seluruh data yang menjadi perhatian dalam suatu ruang lingkup dan waktu yang ditentukan. Jadi, populasi berhubungan dengan data, bukan manusianya. Jika manusia memberikan suatu data, maka banyaknya atau ukuran populasi akan sama banyaknya dengan ukuran manusia. Populasi memiliki parameter yakni besaran terukur yang menunjukkan ciri populasi tersebut. Besaran-besaran yang kita kenal antara lain yaitu rata-rata, rata-rata simpangan, variansi, simpangan baku sebagai parameter populasi. Parameter suatu populasi adalah tetap nilainya, jika nilainya berubah, maka populasinya pun berubah (S. Margono, 2004).

Pengertian lain menyebutkan bahwa populasi adalah keseluruhan objek penelitian yang terdiri dari manusia, benda-benda, hewan, tumbuh-tumbuhan, gejala-gejala, nilai tes, atau peristiwa-peristiwa sebagai sumber data yang memiliki karakteristik tertentu di dalam suatu penelitian. Data yang digunakan dalam penelitian (bahan penelitian), dapat berupa populasi (universe) atau sampel.

Populasi dapat dibedakan sebagai berikut:

1. Populasi terbatas atau populasi terhingga, yakni populasi yang memiliki batas kuantitatif secara jelas karena memiliki karakteristik yang terbatas. Misalnya 5.000 orang dokter pada awal tahun 1995, dengan karakteristik: masa kerja 2 tahun, lulusan program pendidikan dokter, dan lain-lain.
2. Populasi tak terbatas atau populasi tak terhingga, yakni populasi yang tidak dapat ditemukan batas-batasnya, sehingga tidak dapat dinyatakan dalam bentuk jumlah secara kuantitatif. Misalnya dokter di Indonesia, yang berarti harus dihitung jumlahnya sejak dokter pertama ada sampai sekarang dan yang akan datang. Dalam keadaan seperti itu jumlahnya tidak dapat dihitung, hanya dapat digambarkan suatu jumlah objek secara kualitas dengan karakteristik yang bersifat umum yaitu orang-orang, dahulu, sekarang, dan yang akan menjadi dokter. Populasi ini disebut juga parameter.

Selain itu, populasi dapat dibedakan ke dalam hal berikut ini:

1. Populasi teoritis (Theoretical Population), yakni sejumlah populasi yang batas-batasnya ditetapkan secara kualitatif. Kemudian agar hasil penelitian berlaku juga bagi populasi yang lebih luas, maka ditetapkan terdiri dari dokter, berumur 27 tahun sampai 40 tahun, program pendidikan dokter, dan lain-lain.
2. Populasi yang tersedia (Accessible population), yakni sejumlah populasi yang secara kuantitatif dapat dinyatakan dengan tegas. Misalnya, dokter sebanyak 250 di kota Bandung terdiri dari dokter yang memiliki karakteristik yang telah ditetapkan dalam populasi teoritis.

Di samping itu persoalan populasi bagi suatu penelitian harus dibedakan ke dalam sifat berikut ini:

1. Populasi yang bersifat homogen, yaitu populasi yang unsur-unsurnya memiliki sifat yang sama, sehingga tidak perlu dipermasalahkan jumlahnya secara kuantitatif. Misalnya, seorang dokter yang akan melihat golongan darah seseorang, maka ia cukup mengambil setetes

darah saja. Dokter itu tidak perlu mengambil satu botol darah, karena baik setetes maupun satu botol hasilnya akan sama saja.

2. Populasi yang bersifat heterogen, yaitu populasi yang unsur-unsurnya memiliki sifat atau keadaan yang bervariasi, sehingga perlu ditetapkan batas-batasnya, baik secara kualitatif maupun kuantitatif. Penelitian dibidang sosial yang objeknya manusia atau gejala-gejala dalam kehidupan manusia menghadapi populasi yang heterogen.

Meskipun banyak anggotanya terbatas jumlahnya, seperti jumlah mobil di Jakarta, jumlah mahasiswa di Universitas Islam Negeri Jakarta, di mana keduanya sebenarnya dapat dihitung namun karena hal tersebut sulit dilakukan maka dianggap tidak terbatas. Metode penarikan/ pengambil data dengan jelas mengawali/ melibatkan seluruh anggota populasi disebut sensus.

Seorang peneliti meskipun mengetahui bahwa metode sensus ini akan banyak memerlukan pemikiran, memakan waktu yang lama serta relatif mahal, namun tetap melakukan sensus dengan alasan sebagai berikut:

1. Untuk ketelitian

Suatu penelitian sering meminta ketelitian dan kecermatan yang tinggi, sehingga memerlukan data-data yang besar jumlahnya. Apabila unsur ketelitian dan kecermatan ini harus diprioritaskan maka harus digunakan metode sensus.

2. Sumber bersifat heterogen

Apabila menghadapi sumber informasi yang bersifat heterogen di mana sifat dan karakteristik masing-masing sumber sulit untuk dibedakan, maka lebih baik digunakan metode sensus.

Karena populasi merupakan keseluruhan subjek penelitian, maka jika seseorang meneliti semua elemen dia harus meneliti semua populasi. Penelitian populasi dilakukan dengan cara sensus. Cara sensus yang baik dilakukan jika sesuai dengan hal-hal berikut:

1. Tingkat presisi karakteristik subjek penelitian sangat diutamakan (seperti jumlah, jenis, waktu dan ukuran). Misalnya, pada kegiatan sensus penduduk, sensus ekonomi, dan lain-lain.

2. Ukuran populasi sangat kecil. Bila jumlah populasi sedikit, sempit, sebentar maka cara sensus tepat diterapkan. Misalnya, pada penelitian kelas atau penilaian diri bagi para pembuat kebijakan bagi lingkungan kantor.

Pada dasarnya, penelitian dengan cara sensus lebih baik daripada sampling, karena cara sensus lebih mempresentasikan populasinya. Meskipun demikian, seperti yang dikemukakan di atas, pada hal-hal tertentu cara sampling bisa lebih efektif dan efisien daripada cara sensus (Alkaf, 2009).

3.3 Sampel

Sampel adalah sebagian dari populasi, sebagai contoh populasi yang diambil dengan menggunakan cara-cara tertentu (S. Margono, 2004). Sampel adalah Sebagian atau wakil populasi yang diteliti (Furchan, 2004). Pendapat lain tentang pengertian dari sampel adalah Sebagian dari jumlah dan karakteristik yang dimiliki oleh populasi. Apabila populasi besar, dan peneliti tidak mungkin mempelajari semua yang ada pada populasi, misalnya karena keterbatasan dana, tenaga dan waktu, maka peneliti dapat menggunakan sampel yang diambil dari populasi tersebut. Apa yang dipelajari dari sampel tersebut, kesimpulannya akan diberlakukan untuk populasi. Sehingga sampel yang diambil dari populasi harus benar-benar representatif (Sugiyono, 2001).

Masalah sampel dalam suatu penelitian muncul dikarenakan hal berikut ini:

1. Penelitian bermaksud mereduksi objek penelitian sebagai akibat dari besarnya jumlah populasi sehingga harus meneliti sebagian saja dari populasi.
2. Penelitian bermaksud mengadakan generalisasi dari hasil-hasil penelitiannya, dalam arti menegakkan kesimpulan-kesimpulan kepada objek, gejala atau kejadian yang lebih luas.

Adapun alasan-alasan penelitian dilakukan dengan mempergunakan sampel adalah sebagai berikut ini:

1. Ukuran populasi

Dalam hal populasi tak terbatas (tak terhingga) berupa parameter yang jumlahnya tidak diketahui dengan pasti, pada dasarnya bersifat konseptual. Karena itu sama sekali tidak mungkin mengumpulkan data dari populasi seperti itu. Demikian juga dalam populasi terbatas (terhingga) yang jumlahnya sangat besar, misalnya tidak praktis untuk mengumpulkan data dari populasi 50 juta murid sekolah dasar yang tersebar di seluruh pelosok Indonesia.

2. Masalah biaya

Besar ataupun kecilnya biaya tergantung juga dari banyak sedikitnya objek yang diselidiki. Semakin besar jumlah objek, maka semakin besar biaya yang diperlukan, terlebih lagi apabila objek tersebar di wilayah yang cukup luas. Oleh karena itu, sampling adalah satu cara untuk mengurangi biaya.

3. Masalah waktu

Penelitian sampel selalu memerlukan waktu yang lebih sedikit daripada penelitian populasi. Sehubungan dengan hal tersebut, apabila waktu yang tersedia terbatas, dan kesimpulan diinginkan dengan segera, maka dalam hal ini penelitian sampel lebih cepat.

4. Percobaan yang sifatnya merusak

Banyak penelitian yang tidak dapat dilakukan pada seluruh populasi karena dapat merusak atau merugikan. Misalnya, tidak mungkin mengeluarkan semua darah dari tubuh seseorang pasien yang akan dianalisis keadaan darahnya, juga tidak mungkin mencoba seluruh neon untuk diuji kekuatannya. Karena itu penelitian harus dilakukan hanya pada sampel (S. Margono, 2004).

5. Masalah ketelitian

Adalah salah satu segi yang diperlukan agar kesimpulan cukup dapat dipertanggungjawabkan. Ketelitian dalam hal ini meliputi pengumpulan, pencatatan, dan analisis data. Penelitian terhadap populasi belum tentu terjadi ketelitian. Boleh jadi peneliti akan

menjadi bosan dalam melaksanakan tugasnya. Untuk menghindarkan itu semua, penelitian terhadap sampel memungkinkan ketelitian dalam suatu penelitian.

6. Masalah ekonomis

Pertanyaan yang harus selalu diajukan oleh seseorang peneliti yaitu apakah kegunaan dari hasil penelitian sepadan dengan biaya, waktu, dan tenaga yang telah dikeluarkan? Jika tidak, mengapa harus dilakukan penelitian? Dengan kata lain penelitian sampel pada dasarnya akan lebih ekonomis daripada penelitian populasi.

Selanjutnya, mengenai penetapan besar kecilnya sampel tidaklah ada suatu ketentuan yang mutlak, artinya tidak ada suatu ketentuan berapa persen suatu sampel harus diambil. suatu hal yang perlu diperhatikan adalah keadaan homogenitas dan heterogenitas populasi. Jika keadaan populasi homogen, jumlah sampel hampir tidak menjadi persoalan, sebaliknya jika keadaan populasi heterogen, maka pertimbangan pengambilan sampel harus memperhatikan hal:

1. Harus diselidiki kategori-kategori heterogenitas.
2. Besarnya populasi dalam tiap kategori.

Informasi tentang populasi perlu dikejar seberapa jauh dapat diusahakan. Penetapan jumlah sampel yang lebih banyak selalu lebih baik dari pada kurang (*oversampling is always better than undersampling*). Namun demikian, terdapat cara untuk memperoleh sampel minimal yang harus diselidiki dengan menggunakan rumus sebagai berikut:

$$n \geq \frac{pqz^2}{2ab}$$

keterangan:

- | | |
|------------------|---|
| n | = jumlah sampel |
| ≥ | = sama dengan atau lebih besar |
| p | = proporsi populasi persentase kelompok pertama |
| q | = proporsi sisa di dalam populasi |
| z ^{1/2} | = derajat koefisien konfidensi pada 99%, 95 % |

b = persentase perkiraan kemungkinan membuat kekeliruan dalam menentukan sampel.

Contoh:

Jika diketahui jumlah populasi perawat lulusan D3 di Jawa Tengah adalah 400.000 orang. Di antara mereka yang tinggal di daerah pedesaan (luar kota) sebanyak 50.000 orang. Berapa sampel yang perlu diselidiki dalam rangka mengungkapkan hambatan penanaman disiplin di sekolah di wilayah masing-masing.

Perhitungan:

$$F = \frac{50.000}{400.000} \times 100 \% = 12,5 \% \text{ atau } P = 0,125$$

$$q = 1,00 - 0,125 = 0,875$$

$$z_{1/2} = 1,96 \text{ (pada derajat kepastian 99% atau 0,05)}$$

$$b = 5 \% \text{ atau } 0,05$$

Dimasukkan ke dalam rumus sebagai berikut:

$$n \geq \frac{0,125 \times 0,875 \times 1,96^2}{0,05}$$

$n < 168,05$ dibulatkan 169 orang.

Jika penelitian kurang puas dengan jumlah sampel minimal itu, maka dapat dilakukan peningkatan jumlah sampel dengan meningkatkan jumlah sampel dengan sebesar 2,58. Demikian juga ukuran sampel dapat diperbesar lagi dengan memperkecil perkiraan persentase kemungkinan membuat kesalahan dalam penarikan sampel, misalnya sebesar 2% atau $b = 0,02$. Dari contoh itu, maka sampel minimum menjadi:

Dimasukkan ke dalam rumus sebagai berikut:

$$n \geq \frac{0,125 \times 0,875 \times 2,58^2}{0,02}$$

$n > 1.740,21$ dibulatkan 1.740 orang.

Apabila proporsi di dalam populasi yang tersedia tidak diketahui maka variasi p dan q dapat mengganti dengan harga maksimum, yakni $(0,50 \times 0,50 = 0,25)$ ukuran sampel yang harus diselidiki:

2

$$n \geq \frac{1,96}{0,05}$$

$$n \geq 384.$$

Sampel yang baik adalah sampel yang memiliki populasi atau yang representatif, artinya yang menggambarkan keadaan populasi atau mencerminkan populasi secara maksimal walaupun mewakili sampel bukan merupakan duplikat dari populasi.

Pada umumnya masalah sampling timbul apabila penelitian bermaksud untuk:

1. Mereduksi objek penyelidikannya. Karena suatu alasan seringkali seseorang melakukan penelitian tidak meneliti semua objek, semua gejala, semua kejadian atau peristiwa, melainkan hanya sebagian saja dari objek gejala atau kejadian yang dimaksudkan.
2. Ingin mengadakan generalisasi dari hasil-hasil penelitiannya. Mengadakan generalisasi berarti mengesahkan kesimpulan-kesimpulan kepada objek-objek, gejala-gejala, dan kejadian-kejadian yang diteliti.

Mahasiswa yang baru belajar metodologi penelitian di tingkat awal harus menyadari bahwa sampel bukan merupakan duplikat populasi. Oleh karena itu, tidak boleh menyimpulkan bahwa suatu sampel jika telah ditetapkan dengan cara-cara tertentu dapat menjadi cermin yang sempurna bagi populasi, artinya tidak boleh meyakini bahwa sampel tidak mengalami kesesatan walaupun pengambilannya sudah menggunakan metode-metode statistik tertentu.

Berikut ini merupakan petunjuk untuk mengambil sampel:

1. Daerah generalisasi

Yang penting disini adalah menentukan terlebih dahulu luas populasi nya sebagai daerah generalisasi, selanjutnya baru menentukan sampel sebagai daerah penelitiannya. Di sampling itu, yang penting adalah: “apabila yang diselidiki hanya satu kelas saja, jangan diperluas

sampai kelas-kelas lainnya apalagi menyimpulkan untuk sekolah-sekolah lain”.

2. Pengesahan sifat-sifat populasi dan ketegasan batas-batasnya
Apabila luas populasinya telah ditetapkan, harus segera diikuti penegasan tentang sifat-sifat populasinya. Penegasan ini sangat penting jika menginginkan adanya validitas dan reliabilitas bagi penelitiannya. Oleh karena itu, haruslah ditentukan terlebih dahulu luas dan sifat-sifat populasi, dan memberikan batas-batas yang tegas, kemudian menentukan sampelnya. Jangan terjadi kebalikannya, yaitu menetapkan populasilah yang lebih dahulu baru kemudian sampelnya.
3. Sumber-sumber informasi tentang populasi
Untuk mengetahui ciri-ciri populasinya secara terperinci dapat diperoleh melalui bermacam-macam sumber informasi tentang populasi tersebut. Misalnya, sensus penduduk dokumen-dokumen yang disusun oleh instansi-instansi dan organisasi-organisasi, seperti pengadilan, kepolisian, kantor kelurahan, dan sebagainya. Meskipun demikian, haruslah diteliti kembali apakah informasi tersebut telah menunjukkan validitasnya (kesahihan). Hal tersebut perlu, karena jangan sampai terjadi data tahun 1954 masih dipakai sebagian sumber untuk tahun 1965, misalnya bila tahun 1954 tercatat jumlah anak rata-rata dalam setiap keluarga 4 orang, maka pada tahun 1965 jumlah anak rata-rata mungkin tidak seperti itu (4 orang).
4. Menetapkan besar kecilnya sampel
Mengenai berapa besar kecilnya sampel yang harus diambil untuk sebuah penelitian, memang tidak ada ketentuan yang pasti.
5. Menetapkan teknik sampling
Dalam masalah sampel, ada yang disebut *biased sample*, yaitu sampel yang tidak mewakili populasi atau disebut juga dengan sampel yang menyeleweng. Pengambilan sampel yang menyeleweng disebut: *biased sampling*. *Biased sampling* adalah pengambilan sampel yang tidak dari seluruh populasi, tetapi hanya dari salah satu golongan populasi saja, tetapi generalisasinya dikenakan kepada seluruh

populasi. Contoh: Penelitian tentang penghasilan rata-rata orang Indonesia hanya diambil sampel yang kaya raya saja, ataupun hanya yang miskin saja. Dengan sendirinya akan mengakibatkan adanya kesimpulan yang menyeleweng atau disebut *biased conclusion* (Haryono, 1998).

3.4 Teknik Sampling

Teknik sampling adalah cara untuk menentukan sampel yang jumlahnya sesuai dengan ukuran sampel yang akan dijadikan sumber data sebenarnya, dengan memperhatikan sifat-sifat dan penyebaran populasi agar diperoleh sampel yang representatif. Secara umum, ada dua jenis teknik pengambilan sampel yaitu, sampel acak atau *random sampling/ probability sampling* dan sampel tidak acak atau *nonrandom sampling/ non probability sampling* (S. Margono, 2004).

Random sampling adalah cara pengambilan sampel yang memberikan kesempatan yang sama untuk diambil kepada setiap elemen populasi. Artinya jika elemen populasinya ada 100 dan yang akan dijadikan sampel adalah 25, maka setiap elemen tersebut mempunyai kemungkinan $25/100$ untuk bisa dipilih menjadi sampel. Sedangkan yang dimaksud dengan *non random sampling* atau *non probability sampling*, setiap elemen populasi tidak mempunyai kemungkinan yang sama untuk dijadikan sampel. Lima elemen populasi dipilih sebagai sampel karena letaknya dekat dengan rumah peneliti, sedangkan yang lainnya, karena jauh, tidak dipilih; artinya kemungkinannya 0 (nol).

3.4.1 Random Sampling/ Probability Sampling

Teknik *random sampling* adalah teknik pengambilan sampel di mana semua individu dalam populasi, baik secara individu maupun kelompok memiliki kesempatan yang sama untuk menjadi sampel. Teknik ini tidak pilih-pilih dan didasarkan atas prinsip-prinsip matematis yang telah diuji dalam praktek.

Berikut ini merupakan jenis-jenis dari teknik random sampling/ probability sampling:

1. Simple random sampling atau sampel acak sederhana
Teknik untuk mendapatkan sampel yang langsung dilakukan pada unit sampling. Dengan demikian setiap unsur populasi harus mempunyai kesempatan sama untuk bisa dipilih menjadi sampel. Cara demikian dilakukan apabila anggota populasi dianggap homogen. Teknik ini dapat dipergunakan bilamana jumlah unit sampling di dalam suatu populasi tidak terlalu besar. Misalnya, populasi terdiri dari 500 orang mahasiswa program S1 (unit sampling). Untuk memperoleh sampel sebanyak 150 orang dari populasi tersebut, digunakan Teknik ini, baik dengan cara undian, ordinal, maupun tabel bilangan random (S. Margono, 2004).
2. Stratified random sampling atau sampel acak distratifikasikan
Teknik ini biasa digunakan pada populasi yang mempunyai susunan bertingkat atau berlapis-lapis. Teknik ini digunakan apabila populasi mempunyai anggota yang tidak homogen. Dan berstrata secara proporsional. Misalnya sekolah, terdapat beberapa tingkatan kelas. Jika tingkatan dalam populasi diperhatikan, mula-mula harus dipastikan strata yang ada, kemudian tiap strata diwakili sampel penelitian. Untuk lebih jelasnya adalah sebagai berikut: misalnya jumlah pegawai yang lulus $S_1=45$, $S_2=30$, $STM=900$, $SMEA=400$, $SD=300$. Jumlah sampel yang harus diambil meliputi strata pendidikan tersebut yang diambil secara proporsional jumlah sampel (Margono, 2004).
3. Disproportionate Stratified Random Sampling
Teknik ini digunakan untuk menentukan jumlah sampel apabila populasinya berstrata tetapi kurang proporsional. Misalnya, pegawai PT tertentu mempunyai 3 orang lulusan S3, 4 orang lulusan S2, 90 orang lulusan S1, 800 orang lulusan SMU, 700 orang lulusan SMP, maka 3 orang lulusan S3 dan 4 orang lulusan S2 itu diambil semuanya sebagai sampel. Karena dua kelompok tersebut terlalu

kecil apabila dibandingkan dengan kelompok S1, SMU, dan SMP (Sugiyono, 2004).

4. Cluster sampling atau sampel gugus

Teknik ini digunakan jika populasi tidak terdiri dari individu-individu, melainkan terdiri dari kelompok atau cluster. Teknik sampling daerah digunakan untuk menentukan sampel apabila objek yang akan diteliti atau sumber data sangat luas, misalnya penduduk dari suatu negara, provinsi, atau kabupaten. Untuk menentukan penduduk mana yang akan dijadikan sebagai sumber data, maka pengambilan sampelnya berdasarkan daerah populasi yang telah ditetapkan (S. Margono, 2004). Misalnya, di Indonesia terdapat 27 propinsi, dan sampelnya akan menggunakan 10 provinsi, maka pengambilan 10 provinsi tersebut dilakukan secara random. Tetapi perlu diingat bahwa propinsi-propinsi di Indonesia itu berstrata, maka pengambilan sampelnya perlu menggunakan stratified random sampling (Sugiyono, 2001). Dan contoh lain yaitu penelitian dilakukan terhadap populasi pelajar SMU di suatu kota. Untuk itu random tidak dilakukan secara langsung pada semua pelajar, tetapi pada sekolah/ kelas sebagai kelompok atau cluster (S. Margono, 2004).

3.4.2 Non Random Sampling/ Non Probability Sampling

Seperti telah diuraikan sebelumnya, jenis sampel ini tidak dipilih secara acak. Tidak semua unsur atau elemen populasi mempunyai kesempatan sama untuk bisa dipilih menjadi sampel. Unsur populasi yang terpilih menjadi sampel bisa disebabkan karena kebetulan atau karena faktor lain yang sebelumnya sudah direncanakan oleh peneliti.

Berikut ini merupakan jenis-jenis dari teknik non random sampling/ non probability sampling:

1. Convenience sampling atau sampel yang dipilih dengan pertimbangan kemudahan

Dalam memilih sampel, peneliti tidak mempunyai pertimbangan lain kecuali berdasarkan kemudahan saja. Seseorang diambil sebagai

sampel karena kebetulan orang tadi ada di situ atau kebetulan dia mengenal orang tersebut. Oleh karena itu ada beberapa penulis menggunakan istilah *accidental sampling* – tidak disengaja – atau juga *captive sample* (*man-on-the-street*). Jenis sampel ini sangat baik jika dimanfaatkan untuk penelitian penjajagan, yang kemudian diikuti oleh penelitian lanjutan yang sampelnya diambil secara acak (*random*). Beberapa kasus penelitian yang menggunakan jenis sampel ini, hasilnya ternyata kurang obyektif.

2. Purposive sampling

Sesuai dengan namanya, sampel diambil dengan maksud atau tujuan tertentu. Seseorang atau sesuatu diambil sebagai sampel karena peneliti menganggap bahwa seseorang atau sesuatu tersebut memiliki informasi yang diperlukan bagi penelitiannya. Sampel dipilih berdasarkan penilaian peneliti bahwa dia adalah pihak yang paling baik untuk dijadikan sampel penelitiannya. Misalnya, untuk memperoleh data tentang bagaimana keadaan atau karakteristik suatu sekolah, maka kepala sekolah merupakan orang yang terbaik untuk bisa memberikan informasi. Jadi, *judgment sampling* umumnya memilih sesuatu atau seseorang menjadi sampel karena mereka mempunyai “*information rich*” (Hadi, 1998).

3. Quota sampling

Teknik sampel ini adalah bentuk dari sampel distratifikasikan secara proporsional, namun tidak dipilih secara acak melainkan secara kebetulan saja. Dalam teknik ini jumlah populasi tidak diperhitungkan akan tetapi diklasifikasikan dalam beberapa kelompok. Sampel diambil dengan memberikan jatah atau *quorum* tertentu pada setiap kelompok. Pengumpulan data dilakukan langsung pada unit sampling. Setelah jatah terpenuhi, pengumpulan data dihentikan.

4. Snowball sampling atau sampel bola salju

Teknik ini adalah teknik penentuan sampel yang mula-mula jumlahnya kecil, kemudian membesar. Ibarat bola salju yang menggelinding yang lama-lama menjadi besar. Teknik ini banyak

dipakai ketika peneliti tidak banyak tahu tentang populasi penelitiannya. Dia hanya tahu satu atau dua orang yang berdasarkan penilaiannya bisa dijadikan sampel. Karena peneliti menginginkan lebih banyak lagi, lalu dia minta kepada sampel pertama untuk menunjukan orang lain yang kira-kira bisa dijadikan sampel.

5. Systematic sampling atau sampel sistematis

Jika peneliti dihadapkan pada ukuran populasi yang banyak dan tidak memiliki alat pengambil data secara random, cara pengambilan sampel sistematis dapat digunakan. Cara ini menuntut kepada peneliti untuk memilih unsur populasi secara sistematis, yaitu unsur populasi yang bisa dijadikan sampel adalah yang “keberapa” (Alkaf, 2009). Misalnya, setiap unsur populasi yang keenam, yang bisa dijadikan sampel. Soal “keberapa”-nya satu unsur populasi bisa dijadikan sampel tergantung pada ukuran populasi dan ukuran sampel. Misalnya, dalam satu populasi terdapat 5000 rumah. Sampel yang akan diambil adalah 250 rumah dengan demikian interval di antara sampel kesatu, kedua, dan seterusnya adalah 25.

6. Area sampling atau sampel wilayah

Teknik ini dipakai ketika peneliti dihadapkan pada situasi bahwa populasi penelitiannya tersebar di berbagai wilayah. Misalnya, dalam penelitian pendidikan mengadakan penelitian acak terhadap wilayah-wilayah pendidikan dari suatu populasi atau kabupaten, kemudian terhadap sekolah-sekolah, lalu kelas-kelas dan akhirnya para siswa.

Penentuan sampel perlu memperhatikan sifat dan penyebaran populasi. Berkenaan dengan hal tersebut, dikenal beberapa kemungkinan dalam menentukan sampel dari suatu populasi yaitu sebagai berikut:

1. Sampel proporsional

Sampel proporsional menunjuk pada perbandingan penarikan sampel dari beberapa sub populasi yang tidak sama jumlahnya. Dengan kata lain unit sampling pada setiap subsample sebanding jumlahnya dengan unit sampling dalam setiap subpopulasi, misalnya penelitian dengan menggunakan murid SLTA Negeri sebagai unit sampling yang terdiri dari 3.000 murid SMA Negeri dan 1.500 murid STM

Negeri. Dengan demikian perbandingan sub populasi adalah 2:1. Dari populasi tersebut akan diambil sebanyak 150 murid. Sesuai dengan proporsi setiap subpopulasi, maka harus diambil sebanyak 100 murid SMA Negeri dan 50 murid STM Negeri sebagai sampel.

2. Area sampel

Area ini memiliki kesamaan dengan proporsional sampel. Perbedaannya terletak pada sub populasi yang ditetapkan berdasarkan daerah penyebaran populasi menurut daerah penelitian dijadikan dasar dalam menentukan ukuran setiap subsampel. Misalnya, penelitian yang menggunakan guru SMP Negeri sebagai unit sampling yang tersebar pada lima kota kabupaten. Setiap kabupaten memiliki populasi guru sebanyak 500, 400, 300, 200, dan 100. Melihat populasi tersebut, maka perbandingan yang digunakan adalah 5:4:3:2:1. Jumlah sampel yang akan diambil 150. Dengan demikian setiap kabupaten harus diambil sampel sebesar 50, 40, 30, 20, dan 10 orang guru.

3. Sampel ganda

Penarikan ganda atau sampel kembar dilakukan dengan tujuan menanggulangi kemungkinan sampel minimum yang diharapkan tidak masuk seluruhnya. Untuk itu jumlah atau ukuran sampel ditetapkan dua kali lebih banyak dari yang ditetapkan. Penentuan sampel sebanyak dua kali lipat itu dilakukan terutama apabila alat pengumpul data yang digunakan adalah kuesioner pada dua unit sampling yang memiliki persamaan, maka dapat diharapkan salah satu di antaranya akan dikembalikan, sehingga jumlah atau ukuran sampel yang telah ditetapkan terpenuhi.

4. Sampel majemuk (multiple samples)

Sampel majemuk merupakan perluasan dari sampel ganda. Pengambilan sampel dilakukan lebih dari dua kali lipat, tetap memiliki kesamaan dengan unit sampling yang pertama. Dengan sampel multiple ini kemungkinan masuknya data sebanyak jumlah sampel yang telah ditetapkan tidak diragukan lagi. Penarikan sampel

majemuk ini hanya dapat dilakukan apabila jumlah populasi cukup besar.

Dalam setiap penelitian, populasi yang dipilih erat hubungannya dengan masalah yang ingin dipelajari. Misalnya dalam penelitian fertilitas, suatu sampel biasanya dipilih dari populasi wanita usia subur yaitu umur 15 – 49 tahun yang pernah menikah, dalam penelitian tenaga kerja dipilih populasi penduduk usia kerja, dalam penelitian transmigrasi para transmigran yang menjadi populasi sasaran, dan dalam penelitian memakai alat kontrasepsi pada para akseptor yang menjadi sasaran peneliti.

Unsur-unsur yang diambil sebagai sampel disebut unsur sampling. Unsur sampling diambil dengan menggunakan kerangka sampling (sampling frame). Kerangka sampling adalah daftar dari semua unsur sampling dalam populasi sampling. Kerangka sampling dapat berupa daftar mengenai jumlah penduduk, jumlah bangunan, mungkin pula sebuah peta yang unit-unitnya tergambar secara jelas.

Sebuah kerangka sampling yang baik harus memenuhi syarat-syarat sebagai berikut:

1. Harus meliputi seluruh unsur sampel.
2. Tidak ada unsur sampel yang dihitung dua kali.
3. Harus up to date.
4. Batas-batasnya harus jelas, misalnya wilayah, rumah tangga (siapa-siapa yang menjadi anggota rumah tangga).
5. Harus dapat dilacak di lapangan, jadi hendaknya tidak terdapat beberapa desa dengan nama yang sama (S. Margono, 2004).

Bab 4

Uji Hipotesis

4.1 Pendahuluan

Setiap penelitian kuantitatif dan kualitatif tentunya dimulai dengan hipotesis penelitian. Pengujian hipotesis dapat berguna untuk membantu pengambilan keputusan tentang apakah suatu hipotesis yang diajukan, seperti perbedaan atau hubungan cukup meyakinkan untuk diterima atau ditolak (Hastono, 2016).

Secara sederhana kita dapat mengartikan hipotesis statistik sebagai pernyataan atau dugaan mengenai satu atau lebih populasi. Pengujian hipotesis ada kaitannya dengan penerimaan atau penolakan suatu hipotesis. Kebenaran suatu hipotesis tidak akan pernah diketahui dengan pasti kecuali kita memeriksa seluruh populasi. Jika kita memeriksa seluruh populasi tentunya akan menghabiskan waktu yang cukup lama dan biaya yang tidak sedikit. Apa yang bisa kita lakukan jika tidak memeriksa seluruh populasi untuk memastikan kebenaran dari suatu hipotesis. Kita dapat mengambil sampel secara acak dan menggunakan informasi dari sampel itu untuk menerima atau menolak hipotesis.

Dalam bidang kesehatan pengujian hipotesis diawali dengan pertanyaan penelitian dan hipotesis untuk memandu analisis. Dimulai dengan pertanyaan penelitian secara luas, kemudian peneliti mengidentifikasi celah dalam

penelitian yang sedang dilakukan (Shreffler J, 2020). Masalah penelitian atau pernyataan apapun didasarkan pada pemahaman yang lebih baik tentang hubungan antara dua variabel atau lebih.

Contoh pertanyaan penelitian: Apakah merokok berhubungan dengan Covid-19. Untuk menjawab pertanyaan ini maka perlu dilakukan pengujian hipotesis. Dengan pengujian hipotesis akan diperoleh suatu kesimpulan secara probabilistik apakah merokok menjadi sumber penularan Covid-19 atau malah tidak.

Pertanyaan penelitian tidak secara langsung menyiratkan tebakan atau prediksi tertentu. Kita harus merumuskan hipotesis penelitian. Hipotesis adalah pernyataan yang ditentukan sebelumnya mengenai pertanyaan penelitian di mana peneliti membuat tebakan yang tepat tentang hasil studi. Ini kadang-kadang disebut hipotesis alternatif dan pada akhirnya peneliti untuk mengambil sikap berdasarkan pengalaman atau wawasan dari berbagai literatur.

- Hipotesis penelitian: Merokok terbukti berhubungan dengan penularan Covid-19.
- Hipotesis nol (H_0): Tidak terdapat hubungan merokok dengan penularan Covid-19
- Hipotesis alternatif (H_a): Ada hubungan antara merokok dengan penularan Covid-19

Sebelum menyimpulkan hipotesis, peneliti harus menyadari rekomendasi jurnal dalam mempertimbangkan nilai p value. Jika jumlah individu yang terdaftar dalam suatu penelitian meningkat maka kemungkinan untuk menemukan hubungan yang signifikan secara statistik meningkat. Dengan ukuran sampel yang sangat besar, nilai p value bisa menjadi perbedaan signifikan yang sangat rendah dalam menentukan hubungan antara merokok dengan Covid-19.

Hipotesis nol (H_0) dianggap benar sampai sebuah penelitian menyajikan data yang signifikan untuk mendukung penolakan hipotesis nol (H_0). Berdasarkan hasil, peneliti akan menolak hipotesis nol (H_0) jika mereka menemukan perbedaan atau hubungan yang signifikan atau mereka akan gagal untuk menolak hipotesis nol (H_0) jika mereka tidak dapat memberikan bukti bahwa terdapat perbedaan atau hubungan yang signifikan.

Untuk menguji hipotesis, peneliti harus memperoleh data tentang sampel yang representatif untuk menentukan apakah menolak atau gagal untuk menolak hipotesis nol (H_0). Dalam kebanyakan penelitian tidak mungkin atau sulit mendapatkan data untuk seluruh populasi.

Kesimpulan yang didapat dari hasil pengujian hipotesis ada dua kemungkinan yaitu menolak hipotesis (gagal menolak hipotesis) (Hastono, 2016). Dalam uji hipotesis jika kesimpulannya menerima hipotesis artinya bukan berarti kita telah membuktikan bahwa hipotesis tersebut benar karena benar atau tidaknya suatu hipotesis hanya dapat dibuktikan dengan melakukan observasi ke seluruh populasi. Menerima hipotesis maksudnya kita tidak cukup bukti untuk menolak hipotesis.

4.2 Defenisi Hipotesis

Secara bahasa hipotesis berasal dari dua kata yaitu hypo dan thesis. Hypo artinya sementara/lemah kebenarannya dan thesis artinya pernyataan. Dengan demikian hipotesis dapat diartikan sebagai suatu pernyataan yang perlu diuji kebenarannya. Sedangkan menurut (Solution, 2013) bahwa uji hipotesis adalah sebuah metode statistik yang digunakan untuk membuat keputusan statistik dari data eksperimen.

Dalam pengujian hipotesis sering kita jumpai ada dua jenis hipotesis yaitu hipotesis nol (H_0) dan hipotesis alternatif (H_a) (Banerjee, A., Jadhav, S. L., & Bhawalkar, 2009) . Berikut penjelasannya:

Hipotesis nol (H_0) adalah hipotesis yang akan diuji. Dilambangkan dengan simbol H_0 . Dikenal sebagai hipotesis yang tidak ada perbedaan. Hipotesis ini dibuat dengan tujuan untuk menolaknya. Dalam pengujian hipotesis, hipotesis nol (H_0) dapat ditolak atau diterima. Jika tidak ditolak (diterima) artinya bahwa data tersebut tidak cukup bukti untuk menyebabkan penolakan. Jika proses pengujian menolak hipotesis nol (H_0), kesimpulannya bahwa data yang tersedia tidak sesuai dengan hipotesis nol (H_0) sehingga kita bisa menerima hipotesis alternatif (H_a).

Contoh H_0 :

Tidak ada hubungan antara perilaku penggunaan APD dengan kejadian Kecelakaan Kerja

Hipotesis alternatif (H_a) ditetapkan dengan simbol H_A . Hipotesis ini menyatakan ada perbedaan suatu kejadian antara kedua kelompok atau hipotesis yang menyatakan ada hubungan antara satu variabel dengan variabel lainnya.

Contoh H_a :

Ada hubungan antara perilaku penggunaan APD dengan kejadian kecelakaan kerja

Karena kita membuat keputusan berdasarkan sampel yang terbatas, sehingga ada kemungkinan kita akan melakukan kesalahan. Hasil yang mungkin didapat adalah:

Tabel 4.1: Kesalahan dalam Membuat Hipotesis (Paiva, 2010)

	H_0 Benar	H_a Benar
Tidak Menolak H_0	Keputusan tepat	Kesalahan tipe II
Menolak H_0	Kesalahan tipe I	Keputusan tepat

Menerima H_a ketika H_0 benar disebut kesalahan tipe I. Probabilitas melakukan kesalahan tipe I disebut tingkat signifikansi yang dilambangkan dengan α . Contoh kasus di pengadilan seperti menghukum terdakwa yang tidak bersalah.

Tingkat signifikansi yang lebih rendah dari α . Semakin kecil kemungkinan kita untuk melakukan kesalahan tipe I jika nilai α 0,05 atau lebih kecil dari 0,05.

Kegagalan menolak H_0 ketika H_a benar disebut dengan kesalahan tipe II. Kemungkinan untuk melakukan kesalahan tipe II disimbolkan dengan β .

Nilai α dan β selalu berkebalikan. Jika nilai α tinggi maka nilai β rendah dan sebaliknya (Sadik, 2015).

4.3 Jenis Hipotesis

Jenis uji hipotesis dibagi menjadi 2 yaitu hipotesis komparatif/asosiatif dan hipotesis korelatif (Anwar, 2005). Untuk dapat membedakannya perhatikan pertanyaan berikut:

1. Pertanyaan penelitian untuk hipotesis komparatif/asosiatif: apakah terdapat perbedaan kadar gula darah antara kelompok yang diberikan pengobatan glibenklamid dan kelompok plasebo?
 - Apakah terdapat hubungan antara kadar gula darah dengan jenis pengobatan yang diterima (glibenklamid dan plasebo)?
 - Apakah terdapat perbedaan terjadinya kanker paru antara perokok dan bukan perokok?
 - Apakah terdapat hubungan antara perilaku merokok dengan terjadinya kanker paru?
2. Pertanyaan untuk hipotesis korelasi
Berapa besar korelasi antara kadar trigliserida dan kadar gula darah?
Dari beberapa contoh pertanyaan diatas, kita bisa membedakan jenis pertanyaan penelitian untuk masing-masing jenis hipotesis.

4.4 Pemilihan Jenis Uji Parametrik atau Non Parametrik

Dalam pengujian hipotesis sangat berhubungan dengan distribusi data populasi yang akan diuji (Hastono, 2016). Bila distribusi data populasi yang akan diuji berbentuk normal/simetris/Gauss, maka proses pengujian dapat digunakan dengan pendekatan uji statistik parametrik. Sedangkan jika distribusi data populasinya tidak normal maka dapat digunakan uji statistik non parametrik.

Perbedaan uji parametrik dan uji non parametrik dapat pula dilihat dari jenis variabelnya. Uji parametrik salah satu syarat variabelnya harus menggunakan variabel numerik sedangkan uji non parametrik menggunakan variabel

kategorik atau variabel numerik yang tidak memenuhi syarat untuk dilakukan uji parametrik (seperti sebaran datanya tidak normal).

Untuk lebih mengerti tentang perbedaan jenis uji parametrik dan non parametrik dalam menentukan jenis hipotesis dapat dilihat pada tabel berikut:

Tabel 4.2: Tabel uji hipotesis alur menuju pemilihan uji hipotesis komparatif variabel numerik (Anwar, 2005)

Skala Pengukuran Variabel	Jenis Hipotesis			
	Komparatif/Asosiatif			
	2 Kelompok		>2 Kelompok	
	Berpasangan	Tidak Berpasangan	Berpasangan	Tidak Berpasangan
Ordinal	Wilcoxon	Mann Whitney	Friedman	Kruskal Wallis
Numerik (Interval dan rasio)	Uji t berpasangan	Uji t tidak berpasangan	Anova	Anova

Tabel 4.3: Tabel uji hipotesis: alur menuju pemilihan hipotesis variabel kategorikal (Anwar, 2005)

Skala Pengukuran Variabel	Jenis Hipotesis			
	Komparatif/Asosiatif			
	2 Kelompok		>2 Kelompok	
	Berpasangan	Tidak Berpasangan	Berpasangan	Tidak Berpasangan
Nominal	McNemar Marginal Homogeneity	Chi Square Fisher Kolmogorov Smirnov	Cochran	Chi Square Fisher Kolmogorov Smirnov

Ordinal	McNemar Marginal Homogeneity	Chi Square Fisher Kolmogorov Smirnov	Cochran	Chi Square Fisher Kolmogorov Smirnov
---------	------------------------------------	---	---------	---

Dengan penjelasan tabel diatas maka kita bisa lebih mudah menentukan uji hipotesis apakah menggunakan uji parametrik (variabel numerik) atau menggunakan uji non parametrik (variabel kategorik dan numerik yang sebaran datanya tidak normal).

4.5 Langkah Uji Hipotesis

Sebelum melakukan uji hipotesis ada beberapa langkah yang harus dilakukan di antaranya:

1. Menetapkan hipotesis

Langkah pertama yang dilakukan dalam menguji hipotesis adalah menentukan hipotesis nol (H_0) dan hipotesis alternatif (H_a). Dari hipotesis alternatif (H_a) akan diketahui apakah uji statistik menggunakan satu arah (one tail) atau dua arah (two tail) (Hastono, 2016).

Bentuk hipotesis satu arah (one tail) adalah hipotesis yang dinyatakan dengan jelas arah hubungan atau perbedaan nilai/tingkat. Contohnya: semakin tinggi pengetahuan Ibu maka semakin rendah proporsi terjadinya diare pada balita. Sedangkan hipotesis dua arah (two tail) adalah hipotesis yang dinyatakan dengan tidak ada arah hubungan atau tidak ada perbedaan nilai suatu tingkat (Manglapy, 2014). Contohnya ada hubungan antara merokok dengan kanker paru.

2. Penentuan uji statistik yang sesuai

Ada beberapa uji statistik yang dapat digunakan. Setiap uji statistik mempunyai persyaratan tertentu yang harus dipenuhi. Oleh karena itu harus digunakan uji statistik yang sesuai dengan data yang diuji. Jenis uji statistik sangat tergantung dari (Hastono, 2016):

- Jenis variabel yang akan dianalisis
 - Jenis data apakah dependen atau independen
 - Jenis distribusi data populasinya apakah mengikuti distribusi normal atau tidak.
3. Menentukan batas atau tingkat kemaknaan
- Batas/tingkat kemaknaan, sering juga disebut dengan nilai α . Penentuan nilai α tergantung dari tujuan dan kondisi penelitian. Nilai α yang sering digunakan adalah 0,01, 0,05 dan 0,1. Untuk bidang kesehatan masyarakat menggunakan nilai alpha 0,05. Sedangkan untuk pengujian obat-obatan digunakan batas toleransi yang lebih kecil misalnya 0,01 karena mengandung resiko fatal. Sam halnya dengan penelitian eksperimental obat atau vaksin lainnya dengan menggunakan nilai α 0,01 karena peneliti tersebut tidak mau mengambil resiko bahwa ketidakberhasilan obat atau vaksin karena berhubungan dengan nyawa manusia.
4. Perhitungan uji statistik
- Perhitungan uji statistik yaitu menghitung data sampel ke dalam uji hipotesis yang sesuai. Misalnya jika ingin menguji perbedaan mean antara dua kelompok berpasangan, maka data hasil pengukuran dimasukkan ke dalam rumus uji t berpasangan. Jika ingin menguji antara dua kelompok tidak berpasangan maka data hasil pengukuran dimasukkan ke dalam uji t tidak berpasangan. Jika ingin menguji > 2 kelompok baik berpasangan atau tidak berpasangan maka data hasil pengukuran dimasukkan ke dalam uji Anova. Hasilnya pengujian diketahui apakah ada hipotesis ditolak atau gagal menerima hipotesis.
5. Keputusan uji statistik
- Hasil pengujian statistik akan menghasilkan dua kemungkinan yaitu menolak atau menerima hipotesis nol (H_0). Setiap kita melakukan uji statistik lewat SPSS maka yang akan kita cari adalah nilai p (p value). Dengan nilai p ini kita dapat menggunakan untuk keputusan uji statistik dengan cara membandingkan nilai p dengan α (alpha) (Hastono, 2016). Ketentuan yang berlaku adalah:
- Bila nilai $p \leq \alpha$ maka keputusannya adalah H_0 ditolak

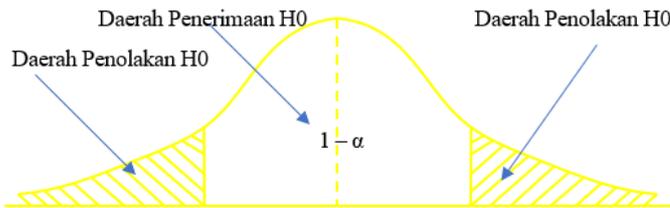
- Bila nilai $p \geq \alpha$ maka keputusannya adalah H_0 gagal ditolak

Cara lain dapat dilakukan untuk penarikan kesimpulan dengan cara probabilitas. Cara ini dilakukan dengan menggunakan rumus:

$$P = (0,5 - \text{luas daerah yang diblok pada kurva normal standar}) \times 2$$

Untuk uji two tail dengan $\alpha = 0,05$ maka nilai $Z_{95\%} = 1,96$

Z tabel = $(1 - \alpha)/2$ (luas daerah yang tidak diarsir pada gambar dibawah ini)



Gambar 4.1: Kurva Normal Standar (Manglapy, 2014)

Berdasarkan Z tabel dengan $Z = 1,96$ adalah $0,475$. Sehingga nilai $p = (0,5 - 0,4938) \times 2$

$$P = 0,0062 \times 2$$

$P = 0,012$ hasil ini kemudian kita bandingkan dengan nilai $\alpha = 0,05$ maka nilai $p < \alpha$ kesimpulannya adalah H_0 gagal diterima.

Contoh Kasus:

Apakah terdapat perbedaan rerata *body mass index* (bmi) antara kelompok status ekonomi tinggi dibandingkan kelompok ekonomi rendah?

Untuk menjawab contoh kasus diatas dapat dilakukan langkah-langkah sebagai berikut:

1. Menentukan variabel yang diuji: Body mass index
2. Menentukan skala pengukuran variabel: Body mass index adalah variabel dengan skala pengukuran numerik. Uji yang mungkin bisa dilakukan dengan t tes berpasangan, t tes tidak berpasangan, anova atau pearson

3. Menentukan jenis hipotesis: hipotesis komparatif. Uji yang mungkin bisa dilakukan dengan t tes berpasangan, t tes tidak berpasangan, anova, atau pearson
4. Menentukan jumlah kelompok: jumlah kelompok yang diuji adalah 2 kelompok (kelompok ekonomi rendah dan ekonomi tinggi). Uji yang mungkin bisa dilakukan adalah dengan t tes berpasangan atau t tes tidak berpasangan
5. Menentukan berpasangan atau tidak: pada kasus ini kedua kelompok tidak berpasangan sehingga uji yang dilakukan adalah menggunakan t tes tidak berpasangan (uji parametrik). Jika memenuhi syarat, jika tidak memenuhi syarat maka yang digunakan adalah uji Mann-Whitney (uji non parametrik).

Bab 5

Statistik Deskriptif

5.1 Pendahuluan

Statistika merupakan ilmu yang mempelajari tentang tata cara mengumpulkan data, menyajikan data, proses analisis dan interpretasi data untuk pengambilan keputusan yang efektif (Arifin, 2014). Statistik biasanya menggunakan data digital, dari data sensus atau pengukuran, dan data digital diperoleh dengan menggunakan data rahasia yang dikategorikan sesuai dengan standar tertentu. Selanjutnya hasil akan dicatat dan dikelompokkan ke dalam bentuk numerik ataupun kategorik yang dikenal dengan nama pengamatan. Prosedur statistik pada dasarnya adalah metode untuk mengolah informasi kuantitatif, sehingga informasi memiliki arti. Statistika menurut fungsinya dapat dikategorikan menjadi statistik deskriptif dan statistik inferensial (Kariadinata and Abdurrahman, 2012).

Proses statistik deskriptif merupakan proses yang memungkinkan untuk mengelola, meringkas dan mendeskripsikan pengamatan. Statistik deskriptif dilakukan dengan tujuan untuk menyederhanakan data (Lotka and Odell, 1926). Data yang dapat dianalisis secara deskriptif adalah semua skala data termasuk skala data nominal, ordinal, interval dan rasio. Tugas statistik deskriptif meliputi informasi statistik yang berhubungan dengan proses mengumpulkan, menyajikan serta melakukan perhitungan statistik. Kemudian

untuk penyajian data statistik deskriptif dapat dilakukan dengan menggunakan tabel, grafik, ukuran rata-rata, diagram, kurva dan lain-lain.

5.2 Pengertian Statistik Deskriptif

Statistik deskriptif adalah sebuah metode yang digunakan dalam mengorganisasikan, menjumlahkan, serta menyajikan data dengan cara yang tepat dan informatif, dengan menggunakan teknik grafik maupun penghitungan statistik. Menggambarkan proses analisa data, namun tidak dapat mengambil sebuah kesimpulan terhadap hasil analisa data tersebut merupakan konsep dari statistik deskriptif. Upaya mengumpulkan data, menyederhanakan angka hasil pengamatan, dan melakukan pengukuran pemusatan untuk mendapatkan informasi yang bermanfaat dan mudah dipahami serta menarik dapat dilakukan dengan statistika deskriptif. Statistik deskriptif akan mampu mengarahkan peneliti untuk dapat memahami karakteristik data yang dimiliki (Dahlan, 2008).

Penyajian data dalam statistik deskriptif biasanya dilakukan dengan membuat grafik dan penyajiannya dalam grafik menjadi tabel, atau memiliki ciri konsentrasi dan distribusi. Contoh penerapan statistik deskriptif yaitu: dampak dari kekurangan gizi pada balita adalah kejadian stunting pada balita. Data yang dikumpulkan setiap tahun misalnya tahun 2017 angka stunting mencapai 8 juta, pada tahun 2018 angka stunting mencapai angka 8,9 juta, dan pada tahun 2019 angka stunting mencapai 9,2 juta. Dalam contoh ini, statistik deskriptif hanya mendeskripsikan masalah yang terjadi namun tidak dilakukan penarikan kesimpulan apapun namun informasi yang disajikan akan lebih menarik dan bermanfaat. Data statistik deskriptif akan memberikan informasi berupa ukuran tendensi sentral, sebaran data dan kecenderungan gugus sebuah data.

5.3 Manfaat Statistik Deskriptif

Dalam statistik deskriptif, data pengamatan akan disederhanakan serta dilakukan pengukuran pemusatan dan distribusi data agar informasi yang

didapat lebih menarik, bermanfaat dan mudah untuk dipahami (Arifin, 2014). Adapun manfaat dari statistik deskriptif yaitu:

Dengan statistik deskriptif, data yang disajikan menjadi lebih ringkas, rapi dan akan memberikan informasi yang tepat.

Dengan statistika deskriptif akan memungkinkan peneliti untuk menampilkan dan menjelaskan data penelitiannya dengan penyajian berupa grafik ataupun numerik

Dengan statistika deskriptif akan memberikan kesempatan peneliti untuk melakukan pengukuran dua karakteristik setiap respondennya kemudian meneliti hubungan dari kedua karakteristik tersebut

Statistik deskriptif memiliki peran yang penting dalam proses persiapan analisa data di mana analisa ini dilakukan oleh peneliti sebelum penerapan statistik inferensial untuk data penelitiannya (Arifin, 2014).

5.4 Metode Dasar Statistik Deskriptif

Metode dasar dari statistik deskriptif yaitu:

1. Metode Numerik

Metode ini dilakukan untuk menghitung nilai statistik dari sekelompok data misalnya untuk menghitung nilai mean dan standar deviasi. Metode numerik dapat memberikan informasi tentang rata-rata maupun informasi distribusi sebuah data.

2. Metode Grafik

Grafik memiliki kelebihan jika digunakan dalam penyajian data berhubungan dengan pola-pola tertentu di dalam data atau dapat dikatakan lebih tepat dan lebih objektif untuk penyajian sebuah data bisa dibandingkan dengan metode grafik. Akan tetapi, metode grafik maupun numerik akan saling melengkapi dalam proses penyajian data penelitian, sehingga kita dapat menerapkan kedua metode tersebut dalam penyajian data penelitian.

5.4.1 Distribusi Frekuensi/ Distribusi Data

Proses Menyusun, meringkas dan membuat tabel akan sangat membantu dalam penyajian jumlah data yang besar. Nilai data bisa saja berbeda dari daftar yang ada di dalam tabel beserta frekuensinya (data tunggal atau berkelompok) termasuk nilai frekuensinya. Nilai frekuensi mendeskripsikan banyaknya kejadian atau kemunculan nilai suatu data dengan kategori tertentu dikenal dengan distribusi frekuensi. Distribusi frekuensi merupakan tabel yang berisi data yang diringkas yang menunjukkan seberapa banyak objek dalam suatu kelas yang bertujuan untuk mendapatkan informasi mendalam dan tepat terhadap data yang dimiliki (Maiti and Bidinger, 1981).

Penyajian data distribusi frekuensi dapat dilakukan dengan penyajian tabel atau grafik dan dapat disajikan dengan nilai persentase. Dengan menggunakan grafik, akan membantu memudahkan penyajian karakteristik serta kecenderungan tertentu dari sekelompok data. Selain itu, pola data dapat kita lihat dengan mudah dengan distribusi frekuensi namun informasi dari nilai individunya akan hilang.

5.4.2 Ukuran Pemusatan (Central Tendency)

Tendensi sentral merupakan salah satu komponen yang penting dalam menggambarkan distribusi dari suatu data. Ukuran tendensi sentral merupakan pengukuran aritmetika yang ditujukan untuk menggambarkan sebuah nilai yang mewakili nilai pusat maupun nilai sentral dari suatu gugus data (himpunan pengamatan). Ukuran tendensi sentral yang sering digunakan meliputi: Mean, Median, dan Modus.

Ukuran dari nilai pusat (average) merupakan nilai yang mewakili distribusi sebuah data, oleh karena itu wajib memiliki karakteristik seperti:

- Seluruh kelompok data harus dipertimbangkan
- Ukuran pemusatan tidak dipengaruhi nilai ekstrim
- Data dari sampel harus stabil
- Harus dapat dimanfaatkan dalam analisis statistik lebih lanjut

1. Mean

Nilai Mean adalah cara yang paling dipakai untuk menyajikan nilai ukuran dari central tendency. Nilai Mean dihitung dengan cara mengakumulasi atau menjumlahkan semua data dalam pengamatan

selanjutnya dibagi dengan banyaknya data. Nilai ekstrim berpengaruh terhadap nilai Mean. Mean ditulis menggunakan symbol μ (dibaca: "miu") untuk menggambarkan rata-rata populasi, (dibaca: x bar) untuk menggambarkan rata-rata sampel.

Rumus Rata-rata populasi:

$$\mu = \frac{\sum_{i=1}^N X_i}{N}$$

N merupakan banyaknya populasi

Rumus Rata-rata sampel:

$$\bar{X} = \frac{\sum_{i=1}^n X_i}{n}$$

n merupakan banyaknya sampel

Rata-rata (Mean) untuk data tunggal

Contoh:

Terdapat 11 pohon jeruk yang dipanen dan menghasilkan buah dengan berat (dalam Kg):

330 284 326 268 236 346 326 402 374 292 380

Apabila nilai rata-rata jumlah buahnya dihitung, maka:

$$\begin{aligned} \bar{X} &= \frac{\sum_{i=1}^{11} X_i}{11} \\ &= \frac{330 + 284 + \dots + 380}{11} \\ &= \frac{3564}{11} = 324 \end{aligned}$$

Sehingga rata-rata jumlah buahnya yaitu 324 Kg

Rata-rata (Mean) untuk data berkelompok.

Rumus rata-rata untuk data yang disajikan dalam data berkelompok contohnya bentuk tabel distribusi frekuensi di mana hasil pengamatan-pengamatan dikelompokkan ke dalam kelas-kelas yaitu:

$$\bar{X} = \frac{\sum_{i=1}^K f_i M_i}{\sum_{i=1}^K f_i}$$

dimana:

M_i adalah nilai tengah kelompok data ke-i

f_i adalah frekuensi atau banyaknya observasi pada kelompok data ke-i

K adalah banyaknya kelompok data (kelas).

Contoh: Hitunglah nilai median pada nilai Mata Kuliah Biostatistik pada tabel di bawah:

Tabel 5.1 Data Nilai Mata Kuliah Biostatistik Mahasiswa

Nilai	Banyaknya Mahasiswa
30-39	3
40-49	5
50-59	8
60-69	14
70-79	10
80-89	7
90-99	3
Jumlah	50

Solusi

Nilai	f_i	M_i	$f_i M_i$
30-39	3	34,5	103,5
40-49	5	44,5	222,5
50-59	8	54,5	463,0
60-69	14	64,5	903,0
70-79	10	74,5	745,0
80-89	7	84,5	591,5
90-99	3	94,5	283,5
Jumlah	50	-	3.285,0

Hasil: Nilai Rata-rata (Mean) Mata Kuliah Biostatistik Mahasiswa adalah 65,7

2. Median

Median adalah nilai tengah. Nilai Median merupakan nilai tengah dari pengamatan sekelompok data setelah data tersebut diurut. Apabila jumlah pengamatan (n) nilainya ganjil, maka nilai median berada tepat di tengah kelompok data, namun apabila jumlah pengamatan (n) nilainya genap, maka nilai median diperoleh melalui interpolasi yaitu cara dengan menghitung rata-rata dua data yang ada di tengah kelompok data. Nilai ekstrem tidak memengaruhi Median. Dalam menentukan median, data diurutkan dari nilai paling kecil hingga yang paling besar terlebih dahulu. Dapat dikatakan median adalah nilai ke.

Nilai tengah data tunggal

$$\frac{(n + 1)}{2}$$

Contoh data buah pohon jeruk:

236 268 284 292 326 326 330 346
 374 380 402

Apabila contoh data jumlah buah pohon jeruk di atas dicari nilai tengahnya, maka setelah datanya diurut, sehingga akan didapat mediannya adalah 326, artinya lima pohon jeruk memiliki jumlah buah di bawah 325 kg dan 5 pohon jeruk memiliki buah di atas 326 kg.

Nilai tengah data berkelompok

Rumus untuk menentukan nilai tengah dari data berkelompok yaitu:

$$\text{Median} = B_m + I \cdot \left(\frac{\frac{n}{2} - (\sum f_i)_0}{f_m} \right)$$

dimana:

B_m = bonderi bawah dari kelas median

$(\sum f_i)_0$ = jumlah frekuensi dari kelas-kelas sebelum kelas median

f_m = frekuensi dari kelas median

n = banyaknya seluruh observasi (jumlah semua frekuensi)

I = interval kelas median.

Kelas Median

Kelas median merupakan kelas yang terdapat nilai median di dalamnya. Dalam menentukan kelas median, semua pengamatan dibagi dua yang artinya 50% dari seluruh pengamatan terletak sebelum dan setelah median.

Pada tabel 5.1, nilai mediannya terletak pada pengamatan ke $(50/2)$ yaitu pada pengamatan ke 25. Jumlah nilai tiga frekuensi pertama ($f_1+f_2+f_3$) adalah $3+5+8=16$. Untuk memperoleh 25 pengamatan diperlukan 9 pengamatan lagi. 9 pengamatan tersebut dapat terpenuhi oleh frekuensi keempat (f_4). Hal ini disebabkan karena jumlah pengamatan pada f_4 sebanyak 14 pengamatan. oleh karena itu, Median terletak pada kelas keempat (60-69).

Contoh: cara untuk menghitung median dari data kelompok nilai Mata Kuliah Biostatistik Mahasiswa pada Tabel 5.1 yaitu:

Solusi:

$$\begin{aligned} B_m &= 59,5 \\ (\sum f_i)_0 &= 3 + 5 + 8 = 16 \\ f_m &= 14 \\ n &= 50 \\ I &= 59,5 - 69,5 = 10 \end{aligned}$$

sehingga Mediannya:

$$\text{Median} = 59,5 + 10 \left\{ \frac{\frac{50}{2} - 16}{14} \right\} = 66,33$$

Melihat hasil analisa di atas, nilai Median yang sesungguhnya belum tentu 66,33 bila kita mengetahui data aslinya, sebab penghitungan ini merupakan suatu interpolasi dari data asli yang tidak diketahui dan data telah dibuat menjadi data berkelompok dalam tabel frekuensi.

3. Modus

Modus adalah suatu data yang paling sering terjadi atau data yang paling sering muncul. Langkah-langkah untuk menentukan nilai modus adalah dengan cara menyusun data terlebih dahulu, menghitung nilai frekuensi dan dilihat nilai mana yang paling besar dan paling banyak muncul, itulah yang disebut modus. Penghitungan nilai modus bagus digunakan untuk menghitung data katagorik maupun numerik. Nilai ekstrem tidak berpengaruh terhadap modus. Modus biasanya digunakan untuk data dengan variasi dan ukuran yang besar, misalnya untuk melihat banyaknya produk yang terjual kepada pembeli sehingga dapat dianalisis kebutuhan pasar dari sebuah produk. Distribusi dari sekelompok data bisa saja tidak memiliki modus atau bisa juga memiliki lebih dari satu modus. Distribusi yang memiliki satu modus dikenal dengan Unimodus, memiliki dua modus dikenal dengan Bimodus dan memiliki lebih dari dua modus dikenal dengan Multimodus.

Modus data tunggal

Sekelompok data disajikan seperti di bawah ini, untuk menghitung modulusnya maka:

- | | | | | | | | | |
|----|---|----|----|---|---|---|---|---|
| a. | 3 | 5 | 7 | 8 | 9 | | | |
| b. | 3 | 5 | 7 | 8 | 8 | 8 | 9 | 9 |
| | | 10 | 11 | | | | | |
| c. | 2 | 3 | 3 | 3 | 4 | 5 | 5 | 5 |
| | | 7 | 7 | 9 | | | | |

Jawaban untuk data:

1. Tidak memiliki Modus sebab semua nilai frekuensinya sama
2. Modusnya adalah 9, sebab nilai pengamatan ini memiliki nilai frekuensi terbanyak
3. Modusnya ada 2 yaitu 3 dan 5, kedua nilai pengamatan memiliki frekuensi sama dan paling banyak

Modus data berkelompok

Rumus untuk menghitung modus bagi data yang telah dikelompokkan dan disajikan dalam tabel distribusi frekuensi yaitu:

$$Modus = B_{mod} + I \cdot \left(\frac{f_1}{f_1 + f_2} \right)$$

dimana,

B_{mod} = Banderi bawah dari kelas modus

f_1 = selisih frekuensi kelas modus dengan frekuensi kelas sebelumnya.

f_2 = selisih frekuensi kelas modus dengan frekuensi kelas sesudahnya.

I = interval kelas modus.

Kelas modus merupakan nilai modus yang terdapat dalam kelas. Misalnya apabila kita ingin menghitung modus dari tabel 1, maka solusinya adalah:

Kelas Modus adalah kelas pada kelompok (60-69), sebab nilai frekuensi paling banyak terdapat pada kelompok ini.

$$B_{mod} = 59,5$$

$$f_1 = 14 - 8 = 6$$

$$f_2 = 14 - 10 = 4$$

$$I = 59,5 - 69,5 = 10$$

Jadi,

$$Modus = 59,5 + 10 \left(\frac{6}{6 + 4} \right) = 65,5$$

5.4.3 Dispersion

Ukuran dispersion atau ukuran penyimpangan adalah ukuran yang menyatakan seberapa jauh nilai-nilai data mengalami penyimpangan dari pusatnya atau dengan kata lain merupakan ukuran yang menyatakan seberapa banyak nilai-nilai dari data berbeda dengan nilai pusatnya. Ukuran disperse ini merupakan pelengkap dari ukuran nilai pusat dalam menggambarkan sekumpulan data. Oleh karena itu, data dispersi merupakan gambaran sekelompok data sehingga hasil menjadi lebih jelas dan tepat.

Bentuk-bentuk ukuran dispersi meliputi:

1. Range

Range (jarak) merupakan selisih nilai terbesar dan terkecil dari data. Nilai Range dapat dibedakan menjadi dua yaitu Range data tunggal dan data berkelompok.

Range data tunggal

$$\text{Range} = \text{Nilai maksimal} - \text{Nilai minimal} (R = X_{\max} - X_{\min})$$

Contoh soal Range data tunggal:

Kelompok data A: 1, 3, 4, 7, 9, 12 $\square R = 12 - 1 = 10$

Kelompok data B: 8, 8, 8, 8, 8, 8 $\square R = 8 - 8 = 0$

Kelompok data C: 20, 25, 35, 45, 50 $\square R = 50 - 20 = 30$

Range data ber kelompok

Range data berkelompok bisa diperoleh dengan nilai tengah atau tepi kelas.

Contoh soal Range data berkelompok: berat badan 30 siswa

Berat Badan (Kg)	Frekuensi
40-44	4
45-49	8
50-54	11

55-59	5
60-64	2
Jumlah	30

Jawaban:

Titik tengah kelas paling rendah: 42

Titik tengah kelas paling rendah: 62

Tepi bawah kelas paling rendah: 39.5

Tepi atas kelas paling tinggi: 64.5

$$1) R = 62 - 42 = 20$$

$$2) R = 64.5 - 39.5 = 25$$

2. Mean Deviation (Simpangan rata-rata)

Mean deviation (MD) merupakan nilai rata-rata hitung dari harga mutlak simpangan-simpangannya. Penghitungan nilai MD menggunakan nilai mutlak saja, yaitu menggunakan nilai-nilai bertanda positif dan tidak memperhitungkan nilai yang bersifat negatif. Mean deviation dibedakan atas data tunggal dan data berkelompok.

3. Varian (Variance)

Varian merupakan nilai tengah kuadrat simpangan dari nilai tengah atau simpangan rata-rata kuadrat. Varian sampel digunakan untuk sampel simbolnya adalah S^2 , sedangkan varian populasi untuk populasi simbolnya adalah σ^2 (dibaca: sigma kuadrat).

4. Standar Deviasi (Simpangan baku)

Simpangan Baku adalah statistik untuk mengukur sekelompok data relative dari Mean dan dihitung menggunakan akar kuadrat dari varian. Standar deviasi untuk sampel disimbolkan dengan S^2 , dan standar deviasi untuk populasi disimbolkan dengan simbol sigma.

5.5 Penyajian Data Statistika Deskriptif

Hasil penelitian statistik deskriptif, data dapat disajikan dengan grafik maupun numerik. Penyajian data dalam statistika deskriptif bertujuan agar mudah dimengerti dan dipahami oleh pembaca (Sataloff, Johns and Kost, 2013).

Penyajian data statistik deskriptif dalam bentuk grafik di antaranya:

1. Histogram
2. Pie Chart
3. Ogive
4. Polygon
5. Diagram Batang Daun (Stem and Leaf)

Data statistik deskriptif dapat disajikan secara numerik dengan:

1. Tendensi sentral
2. Fractile
3. Skewness
4. Pengukuran keruncingan
5. Dispersion

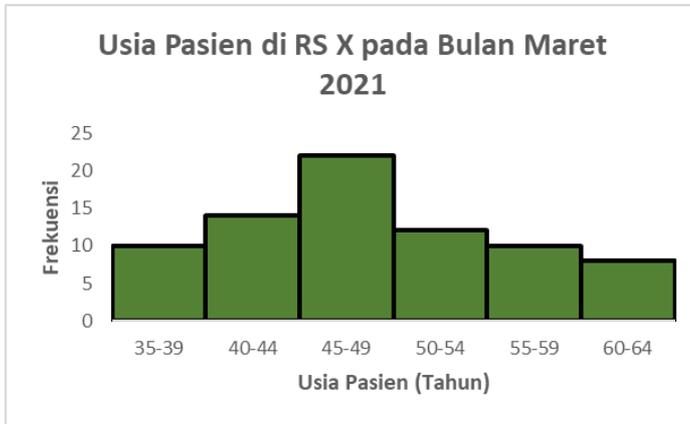
5.5.1 Penyajian Data dalam Bentuk Grafik

Grafik statistik adalah alat yang dapat membantu untuk mempelajari bentuk atau distribusi sampel atau populasi. Grafik bisa menjadi cara yang lebih efektif untuk menyajikan data daripada kumpulan angka karena kita dapat melihat di mana cluster data dan di mana hanya ada beberapa nilai data. Koran dan internet menggunakan grafik untuk menunjukkan tren dan memungkinkan pembaca membandingkan fakta dan angka (Bridges, Weiss and Hassett, 1992).

1. Histogram

Histogram merupakan suatu grafik distribusi frekuensi dari sebuah variabel. Histogram pada umumnya berbentuk balok, dan disajikan dalam dua sumbu utama dengan sudut 90° di mana sebagai aksisnya sumbu X dan ordinatnya sumbu Y. Jarak dari kelas interval dalam

grafik histogram ditunjukkan dengan lebar balok, kemudian besarnya frekuensi data ditunjukkan oleh tinggi balok.



Gambar 5.1: Histogram

2. Pie Chart

Diagram pie atau pie chart adalah suatu diagram berupa lingkaran yang terbagi menjadi beberapa bagian. Bagian-bagian dari diagram menunjukkan besaran persentase dari tiap-tiap kelompok.

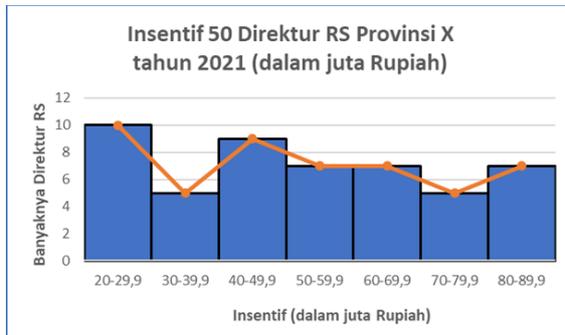


Gambar 5.2: Pie Chart

3. Poligon

Poligon merupakan sebuah grafik untuk menyajikan distribusi frekuensi yang termasuk dalam variabel. Grafis poligon adalah

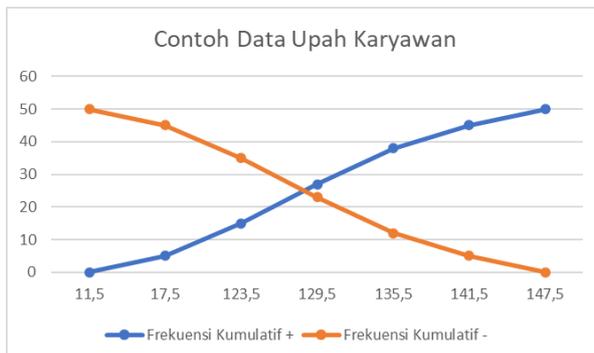
berupa garis patah-patah yang diperoleh melalui upaya menghubungkan puncak pada masing-masing nilai tengah dari kelas. Grafis poligon bagus digunakan untuk menampilkan bentuk perbandingan data dari dua distribusi.



Gambar 5.3: Poligon

4. Ogive

Ogive merupakan sebuah bentuk gambar distribusi frekuensi kumulatif pada suatu variabel. Ogive disajikan dengan ogive positif dan ogive negatif. Ogive positif disajikan untuk data yang disusun dalam bentuk tabel distribusi frekuensi kumulatif kurang dari sedangkan ogive negatif disajikan untuk data yang disusun dalam bentuk tabel distribusi frekuensi kumulatif lebih dari.



Gambar 5.4: Ogive

5. Diagram Batang Daun (Stem and Leaf)

Diagram Stem and Leaf biasanya menampilkan nilai hasil pengamatan asli agar informasi yang didapatkan lebih baik.

Contoh diagram Batang Daun:

Data: 18 15 29 36 62 49 44 45 46 56 57 50 79 73 89 90 99 98 82

Stem	Leaf
1	5 8
2	9
3	6
4	4 5 6 9
5	6 7
6	2
7	3 9
8	2 9
9	8 9

5.5.2 Penyajian Data Numerik

Dalam statistik deskriptif, selain dengan grafik, data juga dapat disajikan dengan cara numerik. Pendekatan ini biasanya digunakan untuk menghitung nilai statistik dari sekelompok data misalnya mean dan standar deviasi. Secara numerik, data dapat disajikan dengan:

1. Tendensi Sentral
2. Dispersion
3. Fractile
4. Skewness
5. Pengukuran keruncingan

Bab 6

Uji Beda Berpasangan (Kategorik Dan Numerik)

6.1 Pendahuluan

Adapun tujuan dari uji t berpasangan atau paired t test atau uji t dependen atau uji beda dua mean dependen atau satu kelompok adalah untuk membandingkan (membedakan) apakah kedua mean tersebut sama atau berbeda. Gunanya adalah untuk menguji kemampuan generalisasi (signifikansi hasil penelitian yang berupa perbandingan keadaan variabel dari dua rata-rata satu sampel/ kelompok/ berpasangan).

6.2 Uji beda Berpasangan/ Uji Beda Dua Mean Dependen (Uji t Dependen)

Tujuan pengujian ini adalah untuk menguji perbedaan mean antara dua kelompok data dependen (subjeknya sama diukur dua kali),

Contoh:

1. Apakah ada perbedaan rata-rata berat kadar Hb ibu antara sebelum dan sesudah melahirkan
2. Apakah ada perbedaan kadar tekanan darah sistol antara sebelum dan sesudah melakukan aktivitas naik tangga

Syarat yang harus dipenuhi :

1. Distribusi data harus normal
2. Kedua kelompok sama
3. Variabel yang dihubungkan adalah kategorik dengan numerik (hanya dua kelompok)

Uji yang digunakan adalah uji t :

$$t \text{ hitung} = \frac{d}{S/\sqrt{n}}$$

$$df = n - 1$$

d (debar) = rata-rata selisih / deviasi pengukuran pertama dan kedua

S = Standar deviasi dari nilai d

n = Jumlah sampel

Contoh soal:

Suatu penelitian yang dilakukan kepada pegawai perguruan tinggi A untuk mengetahui perbedaan berat badan antara sebelum dan sesudah diet. Kemudian dilakukan penelitian dengan mengambil sampel lima orang relawan berbadan gemuk didapatkan data berat badan sebagai berikut: (asumsi distribusi data normal)

Sebelum diet =	70	65	80	75	65
Sesudah diet =	65	66	60	72	61

Berdasarkan data tersebut, buktikan apakah ada perbedaan berat badan karyawan perguruan tinggi antara sebelum dan sesudah diet.

Langkah-Langkah Menjawab:

Langkah 1. Membuat H_a dan H_o dalam bentuk kalimat:

H_a : Terdapat perbedaan yang signifikan antara berat badan karyawan sebelum dan sesudah diet

Ho: Tidak terdapat perbedaan yang signifikan antara berat badan karyawan sebelum dan sesudah diet.

Dengan H_a seperti di atas berarti ujinya dengan two tail (dua arah)

Langkah 2. Membuat H_a dan H_o model statistik:

$$H_a : \mu_1 \neq \mu_2$$

$$H_o : \mu_1 = \mu_2$$

Langkah 3. Mencari selisih/ deviasi antara nilai sebelum dan sesudah diet dan rata-rata deviasinya.

Sebelum diet	=	70	65	80	75	65	
Sesudah diet	=	65	66	60	72	61	
Selisih	=	5	-1	20	3	4	=
		31					

$$d = 31 / 5 = 6,2$$

Langkah 4. Mencari nilai standar deviasi dari selisih nilai data

$$S = \sqrt{\frac{(5-6,2)^2 + (-1-6,2)^2 + (20-6,2)^2 + (3-6,2)^2 + (4-6,2)^2}{5-1}}$$

$$S = \sqrt{\frac{(1,2)^2 + (7,2)^2 + (13,8)^2 + (-3,2)^2 + (-2,2)^2}{4}}$$

$$S = \sqrt{\frac{1,44 + 51,84 + 190,44 + 10,24 + 4,4}{4}} = \sqrt{\frac{258,36}{4}}$$

$$S = \sqrt{64,59} = 8,04$$

Langkah 5. Menghitung dengan rumus uji t

$$t \text{ hitung} = \frac{6,2}{8,04/\sqrt{5}} = \frac{6,2}{8,04/2,24} = \frac{6,2}{3,59} = 1,73$$

$$df = 5 - 1 = 4$$

Cara I:

Untuk mengetahui cara mencari nilai p value pada $df = 4$, ikuti ilustrasi sebagai berikut:

Pada soal di atas nilai t tabel adalah 2,776 sedangkan t hasil = 1,73. Dengan demikian t hasil < t tabel, maka H_0 gagal ditolak artinya tidak ada perbedaan yang signifikan.

Cara II:

Dengan $df = 4$, maka nilai t hasil = 1,73 tersebut terletak di sebelah kiri dari nilai 2,132 atau kalau kita seajarkan dengan nilai α value Uji Dua Pihak, maka terletak di sebelah kiri dari nilai $p = 0,1$. Maka kesimpulannya $p > \alpha (0,05)$ artinya H_0 gagal ditolak dan tidak ada perbedaan yang signifikan.

Langkah 6. Keputusan Uji Statistik

Hasil perhitungan didapatkan nilai t hasil < t table atau p value > dari pada nilai $\alpha (0,05)$, maka dapat disimpulkan H_0 gagal ditolak (diterima). Dengan menggunakan $\alpha 5\%$ dapat disimpulkan bahwa secara statistic tidak terdapat perbedaan yang signifikan rata-rata berat badan karyawan antara sebelum dan sesudah diet.

6.3 Uji Beda Dua Mean Dependen (Uji t Wilcoxon)

Bila ingin mengetahui perbedaan rata-rata pre dan post dari suatu perlakuan, kemudian distribusi data tersebut tidak normal sebaiknya jangan menggunakan uji t dependen, tetapi yang tepat menggunakan uji non parametric yaitu uji Wilcoxon.

$$Z = \frac{T - \mu_T}{\tau_T}$$

Keterangan :

T = Jumlah jenjang/ rangking yang terkecil

$$\mu_T = \frac{n(n+1)}{4}$$

$$\tau_T = \sqrt{\frac{n(n+1)(2n+1)}{24}}$$

Keputusan uji:

Jika $T \text{ terkecil} \leq T \text{ tabel Wilcoxon}$, maka H_0 ditolak

Jika $T \text{ terkecil} > T \text{ tabel Wilcoxon}$, maka H_0 gagal ditolak

Contoh kasus :

Suatu penelitian ingin mengetahui pengaruh terapi obat X untuk menaikkan tekanan darah diastol, kemudian melakukan penelitian terhadap 10 orang relawan. Hasil pengukuran tekanan darah diastole sebelum dan sesudah diberikan obat X adalah:

Sebelum :

100 98 76 90 87 89 77 92 78 82

(mmHg)

Sesudah :

105 94 78 98 90 85 86 87 80 83

(mmHg)

Berdasarkan data tersebut buktikan dugaan peneliti, dengan alpha 5% (asumsi distribusi data tidak normal)

Langkah-Langkah Menjawab

Langkah 1. Membuat H_a dan H_0 dalam bentuk kalimat:

H_a : Obat X memengaruhi tekanan darah diastol

H_0 : Obat X tidak memengaruhi tekanan darah diastol

Dengan H_a seperti di atas berarti ujinya dengan two tail (dua arah)

Langkah 2. Membuat H_a dan H_0 model statistik:

$H_a : \mu_1 \neq \mu_2$

$H_0 : \mu_1 = \mu_2$

Langkah 3. Penyajian data

Tabel 6.1: Penyajian data

No	Diastol sebelum	Diastol sesudah	Selisih sesudah-sebelum	Rangking	+	-
1	100	105	5	7,5	7,5	-
2	98	94	-4	5,5	-	5,5
3	76	78	2	2,5	2,5	-
4	90	98	8	9	9	-
5	87	90	3	4	4	-
6	89	85	-4	5,5	-	5,5
7	77	86	9	10	10	-
8	92	87	-5	7,5	-	7,5
9	78	80	2	2,5	2,5	-
10	82	83	1	1	1	-
	Jumlah				T=36,5	T=18,5

Tabel 6.2: Bantuan Rangking

No	Selisih diurutkan dari yang terkecil	Rangking
1	1	1
2	2	2,5
3	2	2,5
4	3	4
5	4	5,5
6	4	5,5

7	5	7,5
8	5	7,5
9	8	9
10	9	10

Langkah 4. Keputusan Uji

Berdasarkan tabel Wilcoxon di lampiran dengan $n = 10$ taraf kesalahan 5% (0,05) pada uji dua pihak, maka T tabel = . Sehingga T terkecil (18,5) > T tabel (8), kesimpulan H_0 gagal ditolak, sehingga keputusannya tidak ada perbedaan yang bermakna tekanan darah diastole sebelum dan sesudah diberikan obat X (obat X tidak memengaruhi tekanan darah sistol).

Jika sampel lebih besar dari 25, maka distribusinya akan mendekati normal, sehingga digunakan rumus Z dalam pengujiannya:

Menghitung dengan rumus uji Z

$$Z = \frac{T - \mu_T}{\tau_T}$$

Keterangan:

T = jumlah jenjang/ rangking yang terkecil kasus di atas = 18,5

$$\mu_T = \frac{n(n+1)}{4} = \frac{10(10+1)}{4} = 27,5$$

$$\tau_T = \sqrt{\frac{n(n+1)(2n+1)}{24}} = \sqrt{\frac{10(10+1)[(2 \times 10) + 1]}{24}}$$

$$\tau_T = \sqrt{\frac{2310}{24}} = \sqrt{96,25} = 9,8$$

$$Z = \frac{T - \mu_T}{\tau_T} = \frac{18,5 - 27,5}{9,8} = -0,918$$

Cara klasik:

Untuk mengetahui cara mencari nilai p value pada tabel kurva normal, ikuti ilustrasi berikut:

Pada soal di atas nilai Z hitung = 0,918, sedangkan nilai Z kurva = 1,96. Dengan demikian Z hitung < Z tabel, maka H_0 gagal ditolak artinya tidak ada perbedaan yang signifikan.

Cara probabilistik :

Dengan nilai Z = 0,918, jika dibulatkan 0,92, maka peluang tabel = 0,3212. Berarti nilai p-nya = $0,5 - 0,3212 = 0,1788$ karena ujinya two tail maka dikali dua, maka $p = 2 \times 0,1788 = 0,3576$.

Dengan demikian $p > \alpha (0,05)$, maka H_0 gagal ditolak, sehingga keputusannya tidak ada perbedaan yang bermakna tekanan darah diastole sebelum dan sesudah diberikan obat X (obat X tidak memengaruhi tekanan darah diastol).

Langkah 5. Keputusan Uji Statistik

Hasil perhitungan didapatkan nilai t hasil < t tabel atau p value > dari pada nilai alpha (0,05), maka dapat disimpulkan H_0 gagal ditolak (diterima). Dengan menggunakan alpha 5% dapat disimpulkan bahwa secara statistik tidak terdapat perbedaan yang signifikan rata-rata tekanan darah diastol antara sebelum dan sesudah diberikan obat X.

Latihan Kasus :

Seorang peneliti ingin mengetahui hubungan antara pemakaian kontrasepsi oral dengan tekanan darah sistol pada wanita. Kemudian dilakukan eksperimen terhadap delapan orang wanita usia subur yang belum pernah menggunakan kontrasepsi oral. Hasil penelitian sebelum dan sesudah menggunakan kontrasepsi oral selama satu tahun didapatkan data sebagai berikut:

Sebelum :

115 115 104 112 105 107 126 119

(mmHg)

Sesudah :

117 128 102 120 112 115 130 120

(mmHg)

Berdasarkan data di atas, buktikan dugaan peneliti dengan alpha 5% (asumsi distribusi data tidak normal).

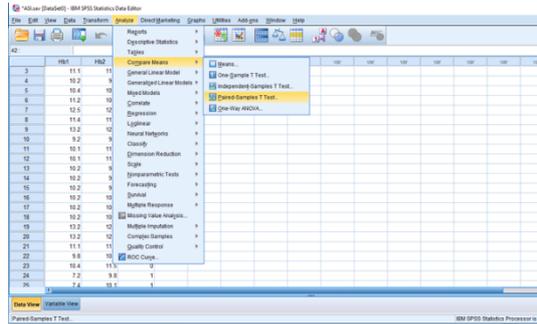
6.4 Uji T Dependen dengan menggunakan SPSS

Uji T dependen seringkali disebut uji T Paired/Related atau pasangan. Uji T dependen sering digunakan pada analisis data penelitian eksperimen. Seperti sudah dijelaskan sebelumnya bahwa disebut kedua sampel bersifat independen kalau kedua kelompok sampel yang dibandingkan mempunyai subjek yang sama. Dengan kata lain disebut dependen bila responden diukur dua kali/diteliti dua kali, sering orang mengatakan penelitian pre dan post. Misalnya kita ingin membandingkan Hb antara sebelum dan setelah menyusui ASI Eksklusif.

Untuk ini akan dilakukan uji beda rata-rata kadar Hb antara kadar Hb pengukuran pertama dengan kadar Hb pengukuran kedua pada ibu yang menyusui ASI Eksklusif, ingin diketahui apakah ada perbedaan kadar Hb antara pengukuran pertama dengan pengukuran kedua. Disini terlihat sampelnya dependen karena orangnya sama diukur dua kali.

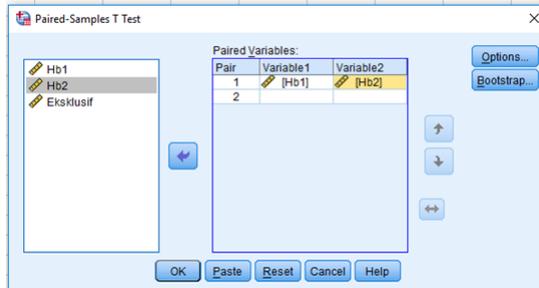
Adapun langkahnya:

1. Pastikan anda berada di file “ASI.SAV”, jika belum aktifkan atau buka
2. Dari menu utama SPSS, pilih menu “ Analyze”, kemudian pilih sub menu “Compare Means”, lalu pilih “Paired-Samples T Test”



Gambar 6.1: Tampilan SPSS

1. Klik Hb1
2. Klik Hb2
3. Klik tanda panah sehingga kedua variabel masuk kotak sebelah kanan



Gambar 6.2: Uji Coba SPSS

1. Klik “OK”, hasilnya terlihat sebagai berikut
T-Test

Paired Samples Statistics

		Mean	N	Std. Deviation	Std. Error Mean
Pair 1	Hb1	10.346	50	1.3835	.1957
	Hb2	10.860	50	1.0558	.1493

Paired Samples Correlations

		N	Correlation	Sig.
			n	

Pair 1	Hb1 & Hb2	50	.707	.000
-----------	-----------------	----	------	------

Paired Samples Test

	Paired Differences					t	d f	Sig. (2- taile d)
	Me an	Std. Deviat ion	Std. Err or Me an	95% Confidence Interval of the Difference				
				Low er	Upp er			
Pa ir 1 H b1 - H b2	- .51 40	.9821	.13 89	- .793 1	- .234 9	- 3.7 01	4 9	.001

Pada tabel pertama terlihat statistik deskriptif berupa rata-rata dan standar deviasi kadar Hb antara pengukuran pertama dan pengukuran kedua. Rata-rata kadar Hb pada pengukuran pertama (Hb1) adalah 10,346 gr% dengan standar deviasi 1,38 gr%. Pada pengukuran kedua (Hb2) didapatkan rata-rata kadar Hb adalah 10,860 gr% dengan standar deviasi 1,05 gr%.

Uji T berpasangan dilaporkan pada tabel kedua, terlihat nilai mean perbedaan antara pengukuran pertama dan kedua adalah 0,514 dengan standar deviasi 0,982. Perbedaan ini diuji dengan uji T berpasangan menghasilkan nilai p yang dapat dilihat pada kolom "Sig (2-tailed)". Pada contoh diatas didapatkan nilai $p= 0,001$, maka dapat disimpulkan ada perbedaan yang signifikan kadar Hb antara pengukuran pertama dengan pengukuran kedua.

6.5 Penyajian dan Interpretasi di Laporan Penelitian

Dari hasil yang didapat di atas kemudian angka-angka disusun dalam tabel yang disajikan dalam laporan penelitian. Bentuk penyajian dan interpretasinya sebagai berikut:

Tabel 6.3: Distribusi Rata-Rata Kadar Hb Responden Menurut Pengukuran Pertama dan Kedua di... Tahun....

Variabel	Mean	SD	SE	P value	N
Kadar Hb					
Pengukuran I	10,346	1,38	0,19	0,001	50
Pengukuran II	10,860	1,05	0,14		

Rata-rata kadar Hb pada pengukuran pertama adalah 10,346 gr% dengan standar deviasi 1,38 gr%. Pada pengukuran kedua didapatkan rata-rata kadar Hb adalah 10,860 gr% dengan standar deviasi 1,05 gr%. Terlihat nilai mean perbedaan antara pengukuran pertama dan kedua adalah 0,514 dengan standar deviasi 0,982. Hasil uji statistic didapatkan nilai 0,001 maka dapat disimpulkan ada perbedaan yang signifikan antara kadar Hb pengukuran pertama dan kedua.

Bab 7

Uji Beda Dua Kelompok

7.1 Pendahuluan

Statistik parametrik dan non parametrik adalah alat bantu untuk menganalisis data. Statistik parametrik digunakan untuk data yang sudah memenuhi syarat berdistribusi normal, diambil secara random dan data memiliki skala interval atau rasio. Uji statistik non parametrik tidak memiliki syarat seperti uji statistik parametrik, sehingga asumsi yang dibuat lebih lemah dan kurang teliti karena pengamatannya bebas dan variabel yang diamati kontinu (Astutik, 2011). Uji statistik parametrik dan non parametrik dapat digunakan untuk menguji dua sampel dalam kelompok atau lebih, pada bab ini akan dibahas lebih lanjut mengetahui uji beda dua kelompok.

Uji beda dua kelompok merupakan uji untuk mengetahui adanya perbedaan rata-rata dua kelompok, di mana pengujian ini bisa dilakukan secara parametrik dan non parametrik. Statistik parametrik memiliki syarat yaitu populasi harus berdistribusi normal, variannya homogen, sampel dipilih secara random (Sunjoyo et al., 2013; Tyastirin and Hidayati, 2017). Berbeda halnya dengan statistik parametrik, statistik non parametrik tidak perlu memiliki syarat populasi berdistribusi normal, serta variannya tidak perlu homogen. Beberapa keuntungan yang didapat dengan menggunakan uji non parametrik adalah data yang digunakan tidak harus berdistribusi normal, bisa dilakukan pada level

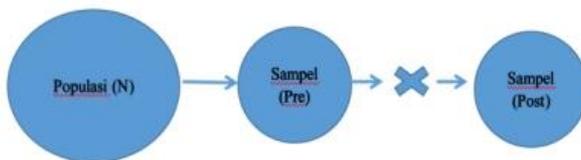
binomial atau ordinal serta prosesnya lebih sederhana dan mudah dimengerti (Sunjoyo et al., 2013).

7.2 Parametrik

Uji parametrik harus memenuhi beberapa asumsi yaitu memiliki data yang berdistribusi normal, sampel yang homogen, skala data interval atau rasio, serta jumlah sampel yang digunakan lebih besar. Kelebihan dari statistik parametrik adalah hasil atau kesimpulan dari penelitian yang didapat lebih kuat, sedangkan kelemahannya adalah cenderung lebih rumit untuk dilakukan (Sunjoyo et al., 2013). Uji parametrik ini bisa dibagi menjadi uji parametrik tidak berpasangan atau independent menggunakan uji t. Sedangkan pada kelompok berpasangan atau dependent menggunakan uji t-paired.

7.2.1 Berpasangan atau Dependent (Uji t-paired)

Uji t-paired merupakan uji pada kelompok berpasangan, di mana pasangan yang dimaksud adalah saling terkait. Asumsi yang harus dipenuhi untuk uji ini adalah, skala pengukuran rasio atau interval, sampel diambil secara acak, data berdistribusi normal, data independent, ukuran sampel digunakan cukup besar dan homogenitas pada variannya (Lotka and Odell, 1926). Misalnya sebuah kelompok diuji tekanan darahnya sebelum diberikan sebuah terapi. Maka tekanan darah pada kelompok tersebut akan diuji berulang yaitu sebelum dan sesudah diberikan terapi sehingga mendapatkan hasil dua pengukuran pada sampel yang sama (Tyastirin and Hidayati, 2017).



Gambar 7.1: Ilustrasi Sampel Berpasangan

Uji t-paired dapat dilakukan dengan cara manual dan SPSS. Berikut rumus yang digunakan untuk melakukan uji t berpasangan secara manual:

Rumus uji t-berpasangan

$$T = \frac{\bar{d}}{SD_d/\sqrt{n}} \quad df = n - 1$$

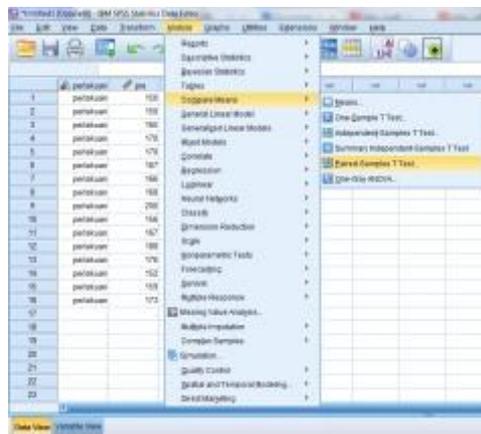
Keterangan:

\bar{d} = rata-rata deviasi/ selisih sampel 1 dan sampel 2

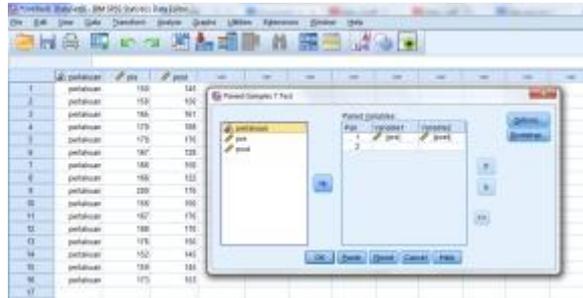
SD_d = Standar deviasi dari deviasi/selisih sampel 1 dan sampel 2

Tahapan uji t-paired dengan SPSS:

1. Pertama buatlah variabel pada window variable view ke dalam SPSS
2. Masukkan data tersebut ke window data view
3. Lakukan analisis dengan memilih Analyze → compare means → paired sample T test. Data akan terlihat seperti ini:



4. Variabel yang akan dianalisis sebelum (pre) dan sesudah (post) pindahkan ke paired variables (data yang dimasukkan pastikan berpasangan). Kemudian tekan OK



5. Hasil dapat dilihat pada output

T-Test

Paired Samples Statistics

Paired Samples	Mean	Std. Deviation	Std. Error Mean
pre	187.31	18.120	3.500
post	183.94	18.140	3.685

Paired Samples Correlations

Paired Samples	Correlation	Sig.
pre & post	.828	.000

Paired Samples Test

Paired Samples	Mean	Std. Deviation	Mean Difference	95% Confidence Interval of the Difference	t	df	Sig. (2-tailed)
pre - post	3.370	18.000	3.370	0.384 - 6.356	0.584	18	.578

Tabel pertama menunjukkan nilai rata-rata setiap variabel, sedangkan tabel akhir menunjukkan nilai akhir sig (2-tailed).

7.2.2 Tidak Berpasangan atau Independent (Uji t)

Uji t independent ini digunakan pada variabel yang tidak berkaitan, misalnya ingin menguji tekanan darah pasien setelah diberikan terapi komplementer bekam kering dengan tekanan darah pasien yang tidak diberikan terapi komplementer bekam kering (Tyastirin and Hidayati, 2017)



Gambar 7.2: Ilustrasi Sampel Tidak Berpasangan

Uji normalitas tetap harus dilakukan pada uji t independent, yang penting perlu diperhatikan juga adalah hasil uji homogenitas untuk mengetahui perbandingan variasi antar kelompok. Hasil uji homogenitas ini yang akan digunakan untuk melihat variasi antar kelompok serta memengaruhi standar error dari hasil uji hipotesis. Uji t independent dapat dilakukan secara manual dan SPSS. Rumus dari uji t tidak berpasangan (Tyastirin and Hidayati, 2017).

Uji-T tidak berpasangan untuk variasi sama

$$T = \frac{X_1 - X_2}{Sp \sqrt{(1/n_1) + (1/n_2)}} \quad \text{dan} \quad Sp^2 = \frac{(n_1 - 1)S_1^2 + (n_2 - 1)S_2^2}{n_1 + n_2 - 2}$$

Uji-T tidak berpasangan untuk variasi berbeda

$$T = \frac{X_1 - X_2}{\sqrt{(S_1^2/n_1) + (S_2^2/n_2)}}$$

$$df = \frac{[(S_1^2/n_1) + (S_2^2/n_2)]^2}{\left[\frac{(S_1^2/n_1)^2}{(n_1 - 1)} + \frac{(S_2^2/n_2)^2}{(n_2 - 1)} \right]}$$

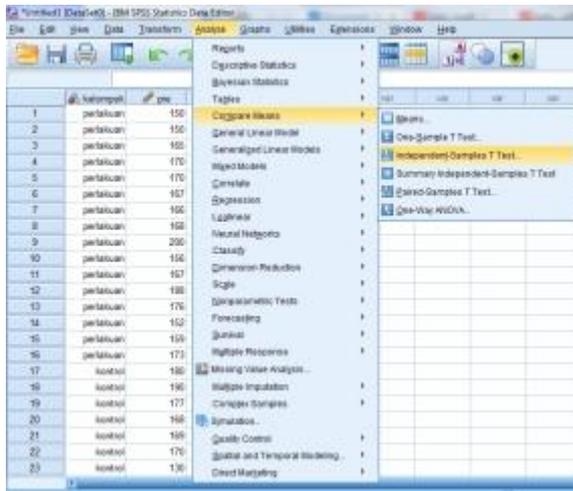
Keterangan

n_1 dan n_2 adalah jumlah sampel pada kelompok 1 dan 2

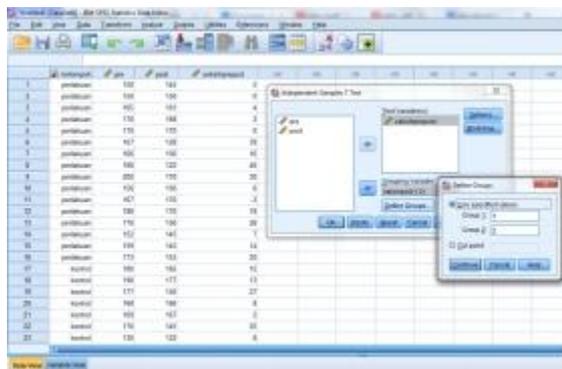
S_1 dan S_2 adalah standar deviasi pada kelompok 1 dan 2

Langkah uji t independent dengan SPSS:

1. Data dibuat pada windows variable view, memasukkan variable independent dan variable dependent.
2. Masukkan data pada windows data view
3. Klik Analyze \rightarrow compare means \rightarrow Independent-Sampel T test



4. Masukkan variable independent pada test variable, dan variable dependent pada grouping variable.



5. Aktifkan kotak define group, karena variable independent 2 maka angka 1 untuk kotak group 1, angka 2 untuk kotak group 2.
6. Prosedur sudah selesai, lanjutkan dengan klik continue dan klik OK
7. Hasil akan dibaca pada tabel

T-Test

Group Statistics

Group	N	Mean	Std. Deviation	Std. Error
Kontrol	14	14.36	14.051	3.694
Eksperimen	14	17.43	11.881	3.083

Independent Samples Test

	Levene Test for Equality of Variances	F	Sig.	t	Sig. (2-tailed)	95% Confidence Interval of the Difference		Lower	Upper	
						Equal Variances Assumed	Equal Variances Not Assumed			
Kontrol - Eksperimen	1.438	.241	.625	1.46	.154	-3.884	8.889	-2.240	6.240	
Kontrol - Eksperimen				-1.00	.320	0.117	-3.882	4.649	-4.280	3.450

8. Hasil pertama akan disimpulkan dulu pada tabel Levene's test (uji hipotesis untuk menguji varian atau uji homogenitas), apabila nilai $p < 0.05$ maka varian berbeda yang dibaca pada hasil uji t tidak berpasangan adalah pada baris equal varian not assumed, nilai $p > 0.05$ maka varian sama yang dibaca pada hasil uji t tidak berpasangan adalah pada baris equal varian assumed.

7.3 Non Parametrik

Uji non parametrik digunakan pada penelitian dengan tipe data dan hasil uji normalitas yang berbanding terbalik dengan uji parametrik. Tipe data pada uji non parametrik adalah data dengan skala kategorik dan data berdistribusi tidak normal, uji non parametrik menjadi pilihan untuk penelitian kuantitatif apabila syarat menggunakan uji parametrik tidak terpenuhi. Uji non parametrik cenderung lebih mudah dan sederhana untuk dilakukan dibandingkan dengan uji parametrik (Sunjoyo et al., 2013). Uji non parametrik sama halnya dengan uji parametrik, dapat menguji hipotesis pada kelompok berpasangan (dependent) atau tidak berpasangan (independent).

7.3.1 Berpasangan (Dependent)

Uji beda dua kelompok berpasangan pada uji non parametrik merupakan uji yang digunakan pada skala data kategorik. Beberapa syarat yang tidak terpenuhi menggunakan uji parametrik maka uji non parametrik menjadi pilihan. Sampel yang saling terkait merupakan salah satu syarat untuk dapat

menggunakan uji non parametrik pada kelompok berpasangan. Beberapa uji non parametrik yang bisa digunakan pada kelompok berpasangan adalah:

7.3.2 Uji Sign

Uji non parametrik uji Sign, merupakan uji yang hampir mirip dengan uji Wilcoxon, membedakan rata-rata dua kelompok. Hasil dari uji ini bukan nilai signifikan melainkan nilai positif dan negatif. Di antara nilai positif dan negatif tersebut, dihitung mana yang lebih banyak (Artaya, 2019). Contohnya: petugas kesehatan ingin melihat hasil pelatihan yang diberikan kepada peserta, pertama melakukan pelatihan secara mandiri dan kedua melakukan pelatihan secara berkelompok.

Rumus uji Sign:

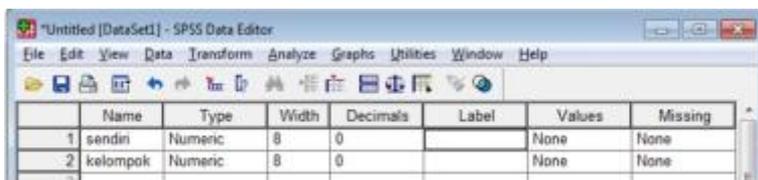
Uji Tanda dilambangkan dengan khai-kuadrat (χ^2). Formula uji Tanda :

$$\chi^2 = \frac{[(n_1 - n_2) - 1]^2}{n_1 + n_2}$$

Dimana : χ^2 = Hasil perhitungan, n_1 = jumlah selisih positif, n_2 = jumlah selisih negatif

Langkah SPSS uji Sign:

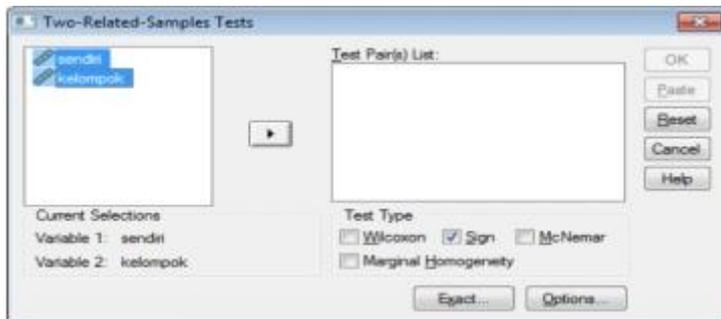
1. Pilih file new
2. Data variable diisi sesuai dengan yang dibutuhkan



3. Masukkan data yang dimiliki pada data view, kemudian save data yang sudah dimasukkan

	sendiri	kelompok	var
1	70	68	
2	72	75	
3	74	72	
4	76	68	
5	70	74	
6	74	79	
7	73	73	
8	71	71	
9	72	80	
10	70	68	
11	73	75	
12	74	72	
13	72	76	
14	71	81	
15	70	70	
16			

4. Klik Analyze → non parametric test → 2 related sample



5. Pindahkan variable sebelum dan sesudah pada test pair (s) list, kemudian untuk test type klik Sign.
6. Hasilnya dapat dilihat pada output

Frequencies

		N
kelompok – sendiri	Negative Differences(a)	5
	Positive Differences(b)	7
	Ties(c)	3
	Total	15

a kelompok < sendiri

b kelompok > sendiri

c kelompok = sendiri

Test Statistics(b)

	kelompok - sendiri
Exact Sig. (2- tailed)	.774(a)

a Binomial distribution used.

b Sign Test

Hipotesanya: H_0 median populasi beda-beda adalah sama dengan nol artinya tidak ada perbedaan, H_a median populasi beda-beda tidak sama dengan nol, artinya ada perbedaan. Interpretasi dari hasil output dapat dilihat berdasarkan probabilitas, jika nilai probabilitas >0.05 maka H_0 diterima atau gagal ditolak, jika nilai probabilitas <0.05 maka H_0 ditolak.

7.3.2 Wilcoxon

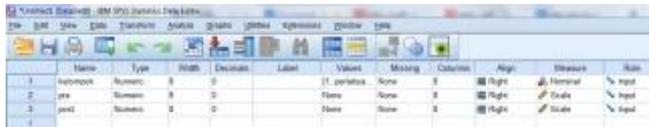
Wilcoxon adalah uji beda kelompok berpasangan, uji ini digolongkan kedalam uji beda non parametrik, di mana skala data yang diuji berupa kategorik (ordinal). Uji Wilcoxon juga bisa digunakan untuk melihat beda proporsi data skala numerik yg tidak berdistribusi normal (Armitage, Berty and Matthews, 2002). Misalnya: peneliti ingin memberikan terapi komplementer akupresure pada penderita hipertensi. Subjek tersebut diukur tekanan darahnya sebelum dan sesudah diberikan terapi komplementer akupresure. Hasilnya dilihat apakah ada perbedaan tekanan darah sebelum dan sesudah diberikan terapi komplementer akupresure.

Rumus statistik uji Wilcoxon:

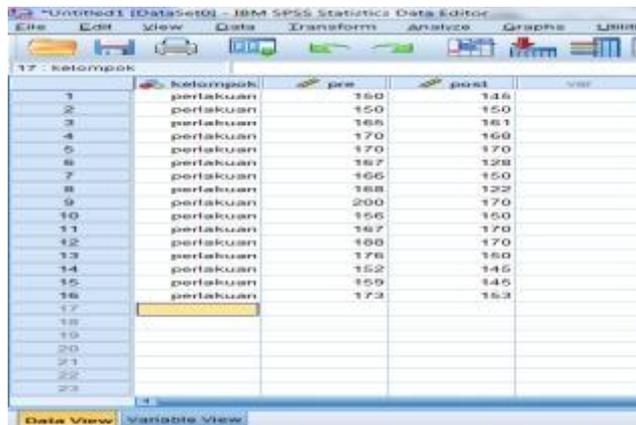
$$Z = \frac{T - \left[\frac{1}{4N(N+1)} \right]}{\sqrt{\frac{1}{24(N)(N+1)(2N+1)}}$$

Langkah-langkah uji Wilcoxon dengan SPSS:

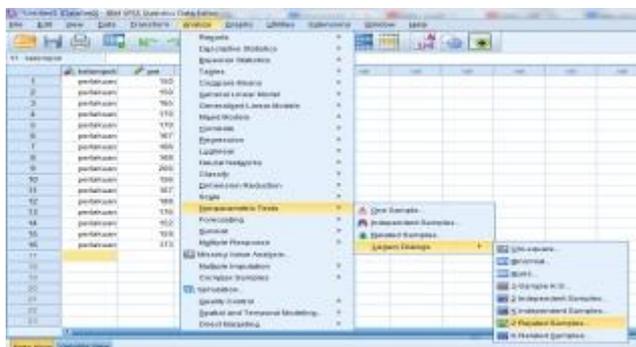
1. Pilih file new
2. Data variable diisi sesuai dengan yang dibutuhkan



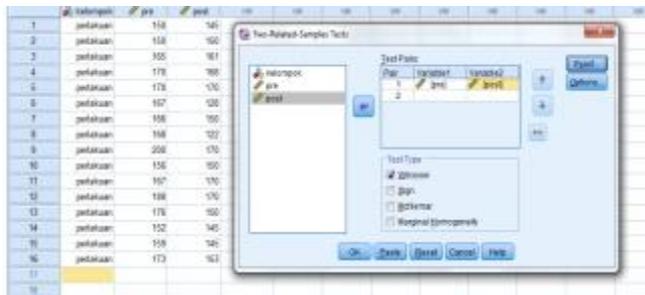
3. Masukkan data yang dimiliki pada data view, kemudian save data yang sudah dimasukkan



4. Klik Analyze → non parametric test → Legacy dialogs → 2 related sample



- Pindahkan sebelum ke variable 1 dan sesudah ke variable 2 pada test pair (s) list, kemudian untuk test type pilih Wilcoxon pilih OK



- Hasilnya dapat dilihat pada output

Ranks			
	LL	Median Rank	Number of Ranks
sebelum - sesudah	17.500	7.000	40 (100%)
sesudah - sebelum	17.500	7.000	40 (100%)
Total	35.000		

Test Statistics ^a	
sebelum - sesudah	0.002
Asymp. Sig. (2-tailed)	
a. Wilcoxon Signed Ranks Test	
b. Ranksum for probability ranks.	

Hasil uji Wilcoxon, ditemukan hasil sig (2 tailed) dapat disimpulkan bahwa H_0 ditolak dengan p value $0.002 < \alpha 0.05$ artinya ada perbedaan sebelum dan sesudah terapi.

7.3.3 Mc.Nemar

Uji Mc Nemar berbeda halnya dengan uji Sign dan Wilcoxon, di mana skala pengukuran data harus nominal atau binomial atau kategori binari. Mc Nemar menguji data berpasangan yang memiliki skala binomial (Riffenburgh and Gillen, 2020). Contohnya: peneliti ingin melihat apakah kontrasepsi hormonal menyebabkan terjadinya kanker payudara, kemudian peneliti memasang pasien dengan kelompok kontrol yang tanpa penyakit. Memasang sampel harus dengan karakteristik yang sama atau hampir mirip dengan pasien atau matching. Uji Mc. Nemar harus menggunakan tabel kontingensi 2×2 , sampel diambil secara acak dari populasi yang mewakili, skala pengukuran dikotom

dan bersifat mutually exclusive, jumlah sampel harus besar apabila jumlah sampel kecil harus menggunakan correction of continuity untuk menjamin hasil yang akurat dari estimasi chi square hitung pada distribusi binomial (Heryana, 2020).

Tabel 7.1: Model Uji Mc.Nemar (Heryana, 2020)

		Kondisi 2 (Post-Test)		Penjumlahan Baris
		Respon kategori 1	Respon kategori 2	
Kondisi 1 (Pre-Test)	Respon kategori 1	<i>a</i>	<i>b</i>	<i>a+b = n₁</i>
	Respon kategori 2	<i>b</i>	<i>d</i>	<i>c+d = n₂</i>
	Penjumlahan Kolom	<i>a+c</i>	<i>b+d</i>	<i>n = n₁ + n₂</i>

Rumus Statistik Mc.Nemar berdasarkan distribusi Chi Square

$$\chi^2 = \frac{(f_o - f_e)^2}{f_e}$$

Pengamatan berdasarkan sel A dan D, maka rumus statistik Mc.Nemar menjadi:

$$\chi^2 = \sum_{i=1}^k \frac{(O_i - E_i)^2}{E_i} = \frac{(A - \frac{A+D}{2})^2}{\frac{A+D}{2}} + \frac{(D - \frac{A+D}{2})^2}{\frac{A+D}{2}} = \frac{(A - D)^2}{A + D}$$

Keterangan:

O_i: Banyak kasus yang diobservasi dalam kategori ke-i

E_i: Banyak kasus yang diharapkan dibawah H₀ dalam kategori ke-i

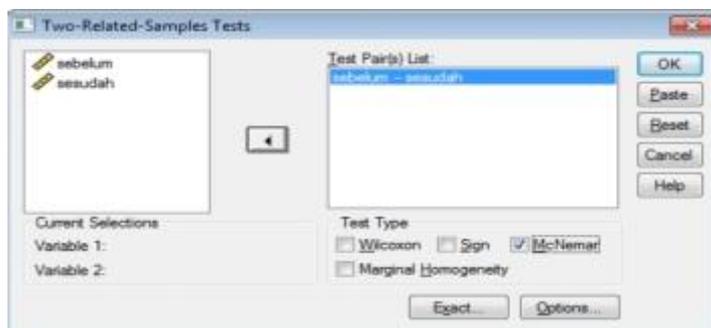
A: Banyak kasus yang diobservasi dalam “sel A”

B: Banyak kasus yang diobservasi dalam “sel D”

Langkah uji menggunakan SPSS:

1. Pilih file new
2. Data variable diisi sesuai dengan yang dibutuhkan
3. Masukkan data yang dimiliki pada data view, kemudian save data yang sudah dimasukkan
4. Klik Analyze → non parametric test → Legacy dialogs → 2 related sample

- Pindahkan sebelum ke variable 1 dan sesudah ke variable 2 pada test pair (s) list, kemudian untuk test type klik Mc.Nemar pilih OK



Hasilnya dapat dilihat pada output

NPar Tests
[Dataset0] C:\Users\A302\Desktop\Nofemar .sav

McNemar Test

Crosstabs

Sebelum & Sesudah

	Sesudah	
Sebelum	Jam Tidak Normal	Jam Tidak Tidak Normal
Jam Tidak Normal	20	0
Jam Tidak Tidak Normal	16	0

Test Statistics^a

	Sebelum & Sesudah
N	36
Exact Sig. (2-tailed)	.027 ^b

a. McNemar Test
b. Binomial distribution used.

Nilai sig (2-tailed) 0.027 < taraf signifikan 0.05 menunjukkan H₀ ditolak dan H_a diterima.

7.3.4 Tidak Berpasangan (Independent)

Uji non parametrik tidak berpasangan merupakan uji yang digunakan pada dua sampel yang memiliki data ordinal serta tidak berdistribusi normal. Beberapa uji dua sampel tidak berpasangan yang menjadi pilihan dalam menguji hipotesis apabila data tidak memenuhi syarat dilakukan uji parametrik adalah:

7.3.5 Mann-Whitney

Uji Mann-Whitney merupakan pengujian rata-rata dua sampel bebas yang memiliki perlakuan yang berbeda antar kelompoknya dan tidak berdistribusi

normal (Sunjoyo et al., 2013). Asumsi yang harus terpenuhi untuk melakukan uji Mann Whitney adalah bentuk dan sebaran data sama serta harus memiliki varian yang homogen (Liese and Miescke, 2008). pengamatan yang dilakukan pada uji ini apabila memiliki data dalam bentuk interval, harus diubah dulu ke dalam data ordinal. Ukuran sampel tidak harus sama karena sampel tidak berpasangan sehingga kedua sampel tidak saling memengaruhi (Astutik, 2011). Misalnya peneliti akan memberikan senam kaki diabetik pada penderita diabetes mellitus. Beda rata-rata dari kelompok perlakuan yang diberikan senam kaki diabetik akan dilihat perbedaannya dengan kelompok kontrol yang tidak diberikan senam kaki diabetik.

Rumus uji Mann-Whitney untuk sampel kecil ($n < 20$):

$$U_1 = n_1 n_2 + \frac{n_1(n_1+1)}{2} - R_1$$

$$U_2 = n_1 n_2 + \frac{n_2(n_2+1)}{2} - R_2$$

Rumus uji Mann Whitney untuk sampel ($n > 20$):

$$Z_{hit} = \frac{U - \frac{n_1 n_2}{2}}{\sqrt{\frac{n_1 n_2 (n_1 + n_2 + 1)}{12} - \frac{n_1 n_2 (\sum t^3 - \sum t)}{12(n_1 + n_2)(n_1 + n_2 - 1)}}$$

Keterangan:

n_1 : Jumlah sampel 1

n_2 : Jumlah sampel 2

U_1 : Jumlah peringkat 1

U_2 : Jumlah Peringkat 2

R_1 : Jumlah ranking pada sampel 1

R_2 : Jumlah ranking pada sampel 2

Langkah uji Mann Whitney dengan SPSS:

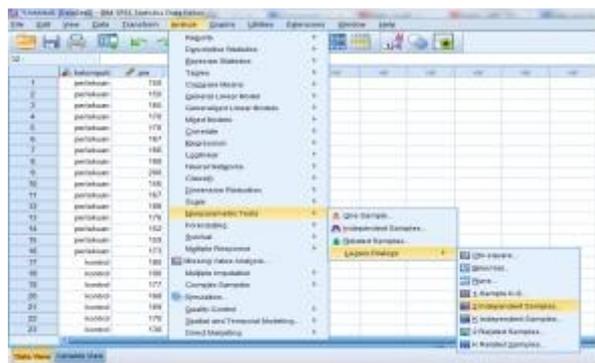
1. Pilih file new
2. Data variable diisi sesuai dengan yang dibutuhkan



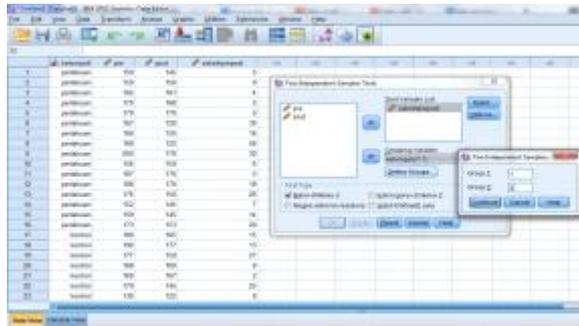
3. Penulisan variabel kelompok dilengkapi perlakuan dan kontrol

	kelompok	pre	post	selisihpost
1	perlakuan	150	145	5
2	perlakuan	150	150	0
3	perlakuan	165	161	4
4	perlakuan	170	168	2
5	perlakuan	170	170	0
6	perlakuan	167	128	39
7	perlakuan	166	150	16
8	perlakuan	168	122	46
9	perlakuan	200	170	30
10	perlakuan	156	150	6
11	perlakuan	167	170	-3
12	perlakuan	188	170	18
13	perlakuan	176	150	26
14	perlakuan	162	145	17
15	perlakuan	169	145	24
16	perlakuan	173	152	21
17	kontrol	180	188	-8
18	kontrol	180	177	3
19	kontrol	177	150	27
20	kontrol	168	160	8
21	kontrol	169	167	2
22	kontrol	170	145	25
23	kontrol	130	122	8

1. Masukkan data yang dimiliki pada data view, kemudian save data yang sudah dimasukkan
2. Klik Analyze → non parametric test → Legacy dialogs → 2 independent samples



3. Memasukkan variabel pada test variable list, dan grouping variable untuk variabel pada kelompok kontrol dan perlakuan, kemudian untuk test type klik Mann-Whitney pilih OK



4. Hasilnya dapat dilihat pada output



7.3.6 Kolmogorov Smirnov

Uji dua sampel tidak berpasangan dengan skala kategori juga bisa digunakan dengan uji Kolmogorov Smirnov, di mana uji ini melihat adanya perbedaan rata-rata dua sampel independen bila datanya ordinal dan telah tersusun pada distribusi frekuensi kumulatif dengan menggunakan klas-klas interval (Astutik, 2011). Tujuan dari uji ini ingin melihat apakah dua sampel independen berasal dari populasi yang sama.

Rumus uji Statistik Kolmogorov Smirnov untuk sampel ($n < 20$):

$$D = \text{maksimum} [s_{n_1}(X) - s_{n_2}(X)]$$

Keterangan:

$s_{n_1}(X)$: Proporsi kumulatif sampel pertama

$s_{n_2}(X)$: Proporsi kumulatif sampel kedua

Rumus uji statistik Kolmogorov Smirnov untuk sampel besar:

$$K_D = 1,36 \sqrt{\frac{n_1 + n_2}{n_1 n_2}}$$

Kriteria dari hasil uji Kolmogorov Smirnov adalah H_0 diterima jika K_D Hitung $\leq K_D$ tabel.

Langkah-langkah uji menggunakan SPSS:

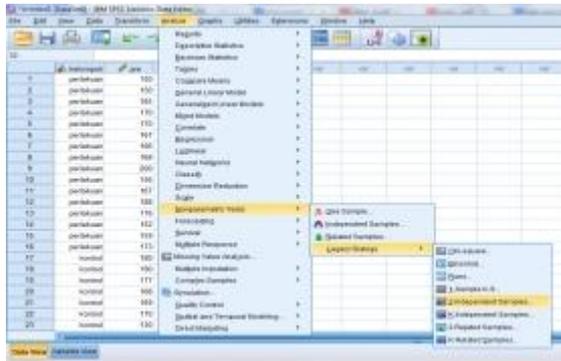
1. Pilih file new
2. Data variable diisi sesuai dengan yang dibutuhkan



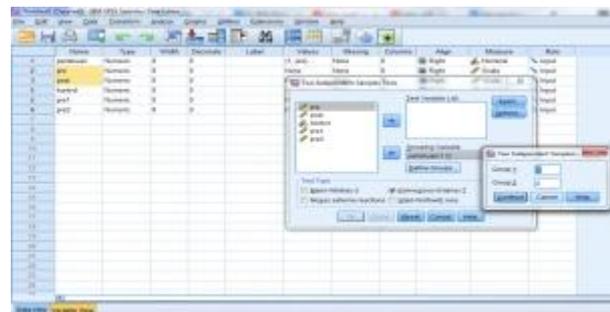
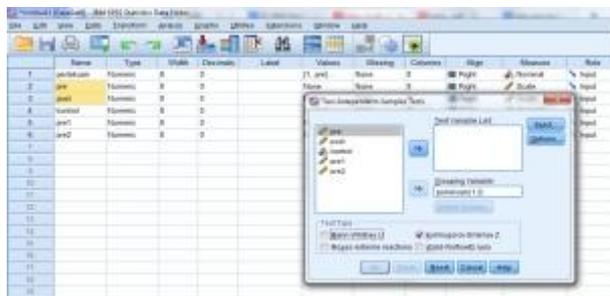
3. Penulisan variabel kelompok dilengkapi perlakuan dan kontrol

	kelompokpost	pre	post	kelompokpost	pre
1	post	10,0	14,0	0	0
2	post	10,0	10,0	0	0
3	post	10,0	10,1	0	0
4	post	17,0	10,0	0	0
5	post	17,0	17,0	0	0
6	post	10,7	12,0	0	0
7	post	10,0	15,0	0	0
8	post	10,0	12,0	0	0
9	post	20,0	17,0	0	0
10	post	10,0	10,0	0	0
11	post	10,0	17,0	0	0
12	post	10,0	17,0	0	0
13	post	17,0	10,0	0	0
14	post	15,0	14,0	0	0
15	post	10,0	14,0	0	0
16	post	17,0	10,0	0	0
17	kontrol	10,0	10,0	0	0
18	kontrol	10,0	17,0	0	0
19	kontrol	17,0	10,0	0	0
20	kontrol	10,0	10,0	0	0
21	kontrol	10,0	10,0	0	0
22	kontrol	17,0	14,0	0	0
23	kontrol	17,0	10,0	0	0

4. Masukkan data yang dimiliki pada data view, kemudian save data yang sudah dimasukkan
5. Klik Analyze \rightarrow non parametric test \rightarrow Legacy dialogs \rightarrow 2 independent samples



- Memasukkan variabel ke test variabel dan grouping variabel, kemudian klik Kolmogorov Smirnov, klik OK.



Hasil dapat dilihat di output

Test Statistics(a)

		Salesman
Most Extreme Differences	Absolute Positive	,933
	Negative	,000
Kolmogorov-Smirnov Z		2,286
Asymp. Sig. (2-tailed)		,000

a. Grouping Variable: kelompok

Analisa hipotesis dari uji ini adalah H_0 kedua populasi identik, H_a kedua populasi tidak identik. Jika probabilitas >0.05 H_0 diterima, jika probabilitas <0.05 maka H_0 ditolak.

7.3.7 Run Wald-Wol Fowitz

Uji Run Wald-Wol Fowitz merupakan uji non parametrik menggunakan dua sampel tidak berpasangan dengan data ordinal dengan tujuan melihat apakah dua sampel tersebut berasal dari populasi yang sama. Uji ini bisa digunakan untuk mengetahui sembarang perbedaan, menguji sekumpulan hipotesis pengganti dengan sebaran data yang digunakan kontinu. Uji ini menggunakan data yang disusun dalam bentuk run dalam melihat perbedaan dari dua sampel (Astutik, 2011).

Rumus uji statistik Run Wald-Wol Fowitz untuk sampel besar n_1 dan n_2 masing-masing lebih dari atau sama dengan 20:

$$z = \frac{\left| r - \left(\frac{2n_1n_2}{n_1 + n_2} + 1 \right) \right| - 0.5}{\sqrt{\frac{2n_1n_2(2n_1n_2 - n_1 - n_2)}{(n_1 + n_2)^2(n_1 + n_2 - 1)}}$$

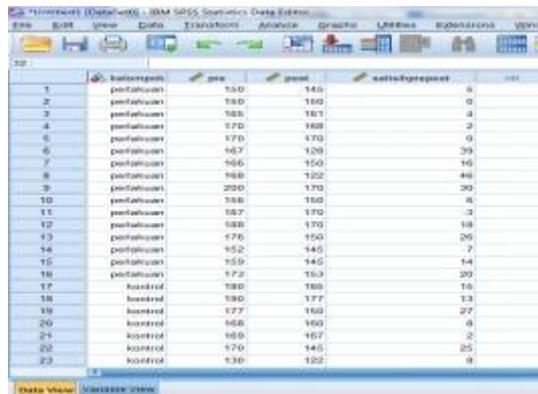
Kriteria hasil uji dari Run Wald-Wol Fowitz pada sampel (≤ 20) H_0 diterima jika run hitung \geq run tabel, jika ($n > 20$) maka H_0 diterima dengan taraf signifikan yaitu Asymp Sig $> 5\%$.

Langkah-langkah uji menggunakan SPSS:

1. Pilih file new
2. Data variable diisi sesuai dengan yang dibutuhkan

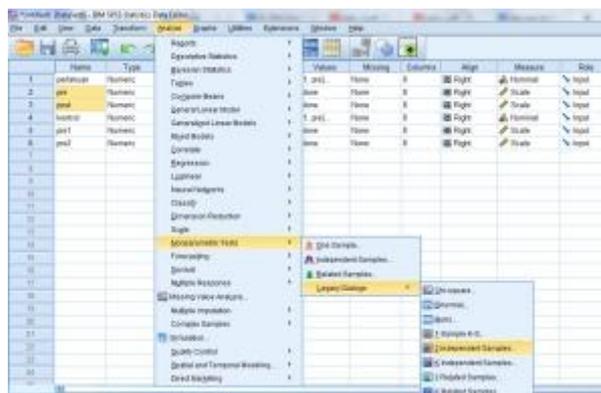


3. Penulisan variabel kelompok dilengkapi perlakuan dan kontrol



4. Masukkan data yang dimiliki pada data view, kemudian save data yang sudah dimasukkan

5. Klik Analyze → non parametric test → Legacy dialogs → 2 independent samples



6. Memasukkan variabel ke test variabel dan grouping variabel, kemudian klik Run Wald-Wol Fowitz, klik OK.



7. Hasil dapat dilihat di output

Frequencies

	Kelompok	N
Salesman	Pelatihan tanpa pelatihan	15
	Total	25

Test Statistics(b,c)

		Number of Runs	Z	Exact Sig. (1-tailed)
Salesman	Minimum Possible	2(a)	-4,477	,000
	Maximum Possible	4(a)	-3,624	,000

a. There are 1 inter-group ties involving 2 cases.

b. Wald-Wolfowitz Test

c. Grouping Variable: kelompok

Analisa hipotesa dari uji ini adalah H_0 kedua populasi identik, H_a kedua populasi tidak identik. Jika probabilitas >0.05 H_0 diterima, jika probabilitas <0.05 maka H_0 ditolak.

Bab 8

Uji Beda Rata-rata (MEAN) Lebih dari Dua Kelompok (ANOVA)

8.1 Pendahuluan

Analysis of Variance (ANOVA) adalah suatu uji statistik yang berfungsi untuk melihat perbedaan rata-rata (mean) dari minimal tiga kelompok sampel (Rusydi and Fadhli, 2018), (Sawyer, 2009). Menurut Kennedy dan Bush, ANOVA pertama kali dikembangkan dan dikenalkan oleh Sir R.A Fisher dengan prinsip membandingkan beberapa mean populasi dengan cara membandingkan variansnya (Setiawan, 2019). Menguji perbedaan rata-rata (mean) yang terdiri dari dua kelompok biasanya memakai uji T, namun kadang kita akan menjumpai penelitian yang terdiri dari 3 kelompok atau lebih. Sebenarnya bisa saja menggunakan uji T akan tetapi lebih baik untuk dihindari. Hal ini dikarenakan uji T yang dilakukan berulang kali akan berisiko meningkatkan probabilitas hasil yang keliru akibat adanya kenaikan dari inflasi nilai alpha (Hastono, 2016).

Contoh penelitian yang dapat menggunakan uji ini adalah studi untuk mengetahui apakah ada perbedaan rata-rata signifikan terkait tingkat kepuasan

pasien terhadap pelayanan kesehatan di 3 RS yang berbeda, di mana responden yang diambil di tiga RS tersebut diambil secara acak (random).

8.2 Jenis ANOVA

ANOVA seringkali digunakan untuk menguji desain penelitian eksperimen. Menurut Jusmiana and Herianto (2020) uji ANOVA selalu menggunakan uji dua sisi (two tailed), akan tetapi nilai alpha (α) tidak perlu dibagi dua. *Analysis of Variance* (ANOVA) terdiri dari berbagai jenis berdasarkan banyaknya jumlah variabel independen (variabel factor) dan jumlah variabel dependen/terikat yang diuji.

Berikut ini adalah jenis ANOVA:

1. ANOVA satu arah

ANOVA satu arah atau biasa disebut dengan One way ANOVA berfungsi untuk menguji hipotesis dari dua variabel yaitu satu variabel bebas (independen) dan satu variabel tergantungan (dependen). Variabel bebas dibedakan menjadi minimal tiga kelompok, sedangkan untuk variabel dependen harus berskala numerik (interval/rasio) dan berdistribusi normal.

Memakai ANOVA satu arah harus memenuhi beberapa asumsi, seperti yang dijelaskan oleh Irianto dalam (Rusydi and Fadhli, 2018) dan (Tabachnick and Fidell, 2007) yaitu:

a. Normalitas

Uji normalitas merupakan salah satu syarat dalam uji parametrik, jika data tidak terdistribusi secara normal maka tidak dapat memakai uji ANOVA. Dalam arti penelitian tersebut hendaknya memakai uji non parametrik seperti uji Kruskal Wallis. Menurut (Sawyer, 2009) normalitas data dapat diabaikan jika jumlah sampel besar. Uji normalitas dapat menggunakan berbagai alternatif seperti menggunakan uji Kolmogorov Smirnov, uji Saphiro Wilk, ataupun memakai metode deskriptif dengan membandingkan rasio skewness / melihat grafik histogram (Dahlan, 2008)

b. Homogenitas (kesamaan varians)

Tiap kelompok yang diuji sebaiknya berasal dari populasi yang memiliki varians sama. Jika jumlah sampel sama di tiap kelompok, maka uji kesamaan varians dapat diabaikan. Namun jika jumlah sampel pada tiap kelompok berbeda, maka kesamaan varians di populasi sangat dibutuhkan. Hal ini untuk menghindari terjadinya kesimpulan yang bias. Jika terdapat varians berbeda serta jumlah sampel tiap kelompok juga berbeda maka transformasi data dapat digunakan sebagai salah satu jalan keluarnya.

c. Pengamatan Bebas

Pengambilan sampel secara random hendaknya dilakukan dalam studi yang menggunakan uji ANOVA, hal ini bertujuan untuk memastikan setiap pengamatan/pengukuran merupakan informasi yang bebas.

2. ANOVA dua arah

ANOVA dua arah/ Two Way ANOVA berfungsi untuk menguji hipotesis penelitian yang terdiri dari dua variabel independen dan tiap variabel independen tersebut dibagi menjadi beberapa kelompok. Misalnya peneliti ingin menguji tiga macam obat hipertensi (obat A obat B, dan obat C), di mana tiap obat tersebut diuji cobakan di dua kelompok berdasarkan jenis kelamin (pria dan wanita). Berdasarkan desain tersebut maka peneliti memiliki kelompok sel data penelitian = 3×2 .

Asumsi yang harus dipenuhi dalam ANOVA dua arah seperti yang dijelaskan oleh Irianto dalam (Rusydi and Fadhli, 2018) yaitu:

- a. Setiap nilai dalam tiap kelompok sel harus terdistribusi normal. Jika jumlah sampel tiap kelompok banyak maka asumsi ini tidak mutlak dipenuhi.
- b. Homogenitas. Varians hendaknya sama pada setiap kelompok sel
- c. Nilai tiap sel independen dari pengaruh variabel lain yang tidak diteliti. Hal ini bisa dicapai jika pengambilan sampel dilakukan

secara random dari populasi yang telah dikelompokkan sesuai dengan sel yang ada.

8.3 Rumus ANOVA

Prinsip dari uji ANOVA adalah dengan melakukan pengujian variabilitas data menjadi dua variasi yakni variasi dalam kelompok (within) dan variasi antar kelompok (between). Jika variasi within sama dengan variasi between maka nilai mean di antara beberapa kelompok tersebut tidak ada perbedaan (Hastono, 2016).

ANOVA satu arah memakai Rumus F dalam penghitungannya, yakni sebagai berikut:



Keterangan:

Sb2 : Varians between

Sw2: Varians within

$$S_b^2 = \frac{n_1(\bar{x}_1 - \bar{x})^2 + n_2(\bar{x}_2 - \bar{x})^2 + \dots + n_n(\bar{x}_n - \bar{x})^2}{k - 1}$$

$$S_w^2 = \frac{(n_1 - 1)S_1^2 + (n_2 - 1)S_2^2 + \dots + (n_n - 1)S_n^2}{n - k}$$

$$\bar{X} = \frac{n_1 \cdot \bar{X}_1 + n_2 \cdot \bar{X}_2 + \dots + n_n \cdot \bar{X}_n}{k - 1}$$

Sedangkan derajat kebebasan atau degree of freedom (df) dari uji F adalah:

$$\text{Pembilang (numerator)} = k - 1$$



Keterangan:

Sb2 : Varians between (varians antar kelompok)

Sw2 : Varians within (varians dalam kelompok)

S2 : Varians

k : jumlah kelompok

n1, n2, nk : jumlah data dalam kelompok

n : jumlah keseluruhan data (n1 + n2 + n3 + ... + nk)

\bar{X} : rata-rata gabungan (grand mean)

8.4 Contoh Soal Uji ANOVA

Sebuah penelitian bertujuan untuk mengetahui apakah terdapat perbedaan yang signifikan pada rata-rata usia pasien penyakit jantung di 3 RS, di mana sampel telah diambil secara random.

Penelitian tersebut diperoleh data sebagai berikut:

Tabel 8.1: Contoh Soal uji anova

RS			Apabila diketahui ketiga populasi tersebut memiliki distribusi normal dan varian ketiganya sama, ujilah pada taraf $\alpha=5\%$.
A	B	C	
47	55	54	
53	58	50	
49	54	51	
50	61	51	
46	62	49	
$\bar{X}_1:$ 49	$\bar{X}_2:$ 56	$\bar{X}_3:$ 51	
S ₂₁ : 7,5	S ₂₂ : 12,5	S ₂₃ : 3,5	

Penyelesaian:

- Menentukan hipotesis nol (null hypothesis) dan hipotesis alternatif (ialternative hypothesis)
 - Ho: Tidak ada perbedaan usia pasien di tiga RS tersebut.
 - Ho: $\mu_1 = \mu_2 = \mu_3$
 - Ha: Minimal ada 1 perbedaan rata-rata usia pasien di tiga RS tersebut
- Menentukan alpha (tingkat kemaknaan) = 0,05
- Menentukan uji statistik dan menghitungnya
Dikarenakan membandingkan 3 rata-rata maka memakai uji Anova satu arah.

Penghitungan:

- Menghitung rata-rata gabungan

$$X = \frac{(5*49) + (5*56) + (5*51)}{15} = 52$$

- Menghitung varians between dan within

$$Sw2 = \frac{((5-1)*7,5) + ((5-1)*12,5) + ((5-1)*3,5)}{15-3} = 7,8$$

$$Sb2 = \frac{(5*(49-52)^2) + (5*(56-52)^2) + (5*(51-52)^2)}{3-1} = 65$$

- Menghitung uji F

F hitung = Sb2

$$= \frac{Sw2}{Sb2} = \frac{65}{7,8} = 8,3$$

- Melihat F tabel

Untuk mencari nilai F tabel diperlukan tabel distribusi F dan menentukan degree of freedom (df) nya, yaitu:

df numerator = k - 1 = 3 - 1 = 2

df denominator = n - k = 15 - 3 = 12

		PROBABILITAS DIBAWAH KURVA DISTRIBUSI-F							
		Numerator (k - 1)							
DENO									
(N - k)	VAL	VAL	2	3	4	5	6	7	
		VAL	Nilai-F						
10	0,001	0,001	21,0	14,9	12,6	11,3	10,5	9,9	9,5
12	0,1	0,1	3,2	2,8	2,6	2,5	2,4	2,3	2,3
12	0,05	0,05	4,7	3,89	3,5	3,3	3,1	3,0	2,9
12	0,025	0,025	6,6	5,1	4,5	4,1	3,9	3,7	3,6
12	0,010	0,010	9,3	6,9	6,0	5,4	5,1	4,8	4,6
12	0,005	0,005	11,8	8,5	7,2	6,5	6,1	5,8	5,5
				11,6					

Dengan melihat tabel F di nilai alpha = 0,05 maka diketahui F tabel = 3,89

4. Keputusan

Bandungkan antara F hitung dengan F tabel

F hitung: 8,3 > F tabel: 3,89

Sehingga keputusannya H_0 ditolak

5. Kesimpulan

Terdapat perbedaan signifikan rata-rata usia pasien penyakit jantung di 3 RS tersebut.

8.5 Uji Banding Ganda / Multiple Comparison (Post Hoc Test)

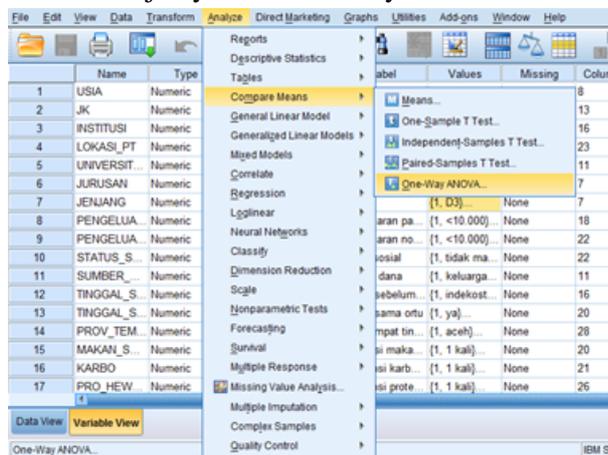
Jika ANOVA satu arah menyimpulkan nilai $p > \alpha$, maka H_0 gagal ditolak, sehingga dapat disimpulkan ada perbedaan nilai rata-rata antara kelompok pada tingkat kepercayaan tertentu. Akan tetapi jika nilai $p < \alpha$, maka H_0 ditolak/ H_a diterima maka berarti minimal ada satu perbedaan nilai rata-rata di antara kelompok. Sehingga perlu diketahui lebih lanjut nilai rata-rata kelompok manakah yang berbeda. Uji statistik untuk menjawab hal ini adalah uji banding ganda (Prasetyo and Ariawan, 2008). Uji banding ganda terdiri dari berbagai jenis, di antaranya adalah Bonferroni, Duncan, *Honestly Significant Different* (HSD), Scheffe, dan Tukey (Hastono, 2016).

8.6 Uji ANOVA dengan SPSS

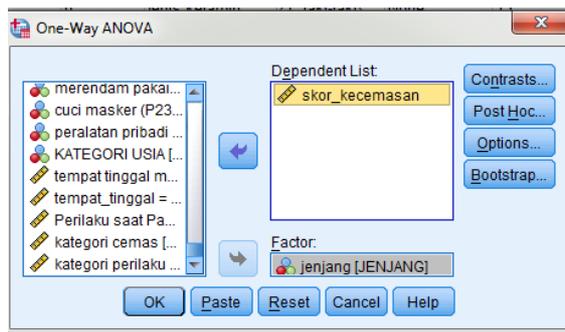
SPSS merupakan salah satu software yang umum digunakan dalam menguji statistik di berbagai bidang keilmuan, termasuk di antaranya di bidang kesehatan. Pada kesempatan ini akan dicontohkan tentang pengujian hubungan antara jenjang pendidikan mahasiswa dengan skor kecemasan. Variabel jenjang pendidikan terdiri dari 3 kelompok (D3, D4, dan S1) yang berskala kategorik sedangkan variabel kecemasan berskala numerik.

Adapun langkahnya adalah sebagai berikut:

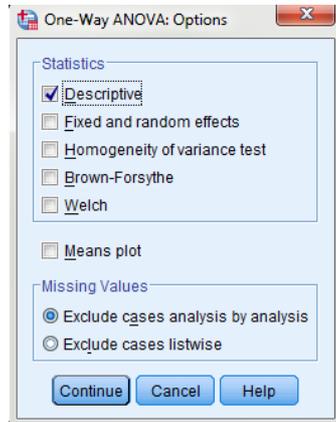
1. Membuka data file melalui aplikasi SPSS
2. Kemudian dari menu utama, klik “Analyze”, pilih “Compare means” selanjutnya klik “One way ANOVA”.



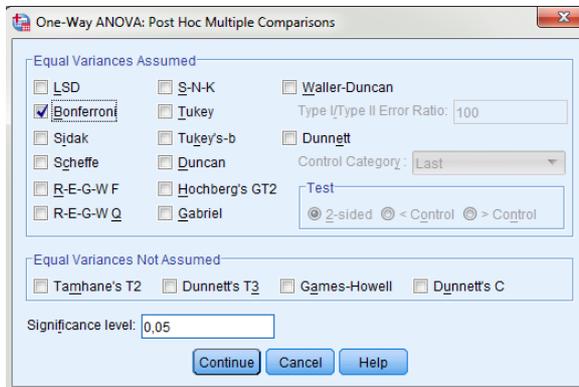
3. Akan muncul dialog box One way Anova, ada dua kotak yang perlu diisi. Kotak “dependent list” diisi dengan variabel numerik (kecemasan) dan kotak “factor” diisi dengan variabel kategorik (jenjang).



4. Klik tombol options dan check list (√) descriptive, dan klik continue



5. Klik tombol post hoc, dan check list bonferroni



6. Lanjutkan dengan klik continue dan OK

Pada menu output SPSS maka akan muncul hasil sebagai berikut:

Descriptives

skor_kecemasan

	N	Mean	Std. Deviation	Std. Error	95% Confidence Interval for Mean		Minimum	Maximum
					Lower Bound	Upper Bound		
D3	2169	43,6782	11,26757	,24194	43,2037	44,1526	13,00	65,00
D4	1226	43,7365	11,08645	,31663	43,1154	44,3577	13,00	65,00
S1	2529	42,3654	10,63265	,21143	41,9508	42,7800	13,00	65,00
Total	5924	43,1298	10,98085	,14267	42,8501	43,4095	13,00	65,00

ANOVA

skor_kecemasan

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	2581,488	2	1290,744	10,740	,000
Within Groups	711607,687	5921	120,184		
Total	714189,175	5923			

Post Hoc Tests**Multiple Comparisons**skor_kecemasan
Bonferroni

(I) jenjang	(J) jenjang	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval	
					Lower Bound	Upper Bound
D3	D4	-,05835	,39171	1,000	-,9964	,8797
	S1	1,31283*	,32083	,000	,5446	2,0811
D4	D3	,05835	,39171	1,000	-,8797	,9964
	S1	1,37118*	,38151	,001	,4576	2,2848
S1	D3	-1,31283*	,32083	,000	-2,0811	-,5446
	D4	-1,37118*	,38151	,001	-2,2848	-,4576

*. The mean difference is significant at the 0.05 level.

Dari hasil output di atas, membacanya adalah sebagai berikut:

1. Kotak Descriptive

Pada mahasiswa yang di jenjang pendidikan D3, rata-rata skor kecemasannya adalah 43,68 dengan standar deviasi 11,27; sedangkan pada mahasiswa D4 rata-rata skor kecemasannya adalah 43,74 dengan standar deviasi 11,09; mahasiswa S1 rata-rata skor kecemasannya adalah 42,37 dengan standar deviasi 10,63.

2. Kotak ANOVA

Nilai p dari uji ANOVA dapat dilihat di kolom Sig, di mana hasilnya adalah 0,000 (dapat ditulis menjadi 0,0005 atau 0,0001) yang berarti pada alpha 0,05 terdapat perbedaan rata-rata skor kecemasan mahasiswa berdasarkan jenjang pendidikannya.

3. Kotak Post Hoc Test

Dari kotak Post Hoc test dapat diketahui kelompok yang berbeda secara signifikan. Untuk mengetahuinya kita dapat melihat di kolom Sig. Jika hasilnya adalah kurang dari alpha (0,05) maka berarti berbeda secara signifikan. Terlihat bahwa yang berbeda secara signifikan adalah antara mahasiswa D3 dengan S1 dan antara mahasiswa D4 dengan S1. Sedangkan antara mahasiswa D3 dan D4 tidak terjadi perbedaan yang signifikan.

Hasil output di atas dapat dituliskan dan diinterpretasikan dalam laporan penelitian sebagai berikut:

Tabel 8.2: Distribusi Rata-rata Skor Kecemasan Mahasiswa Berdasarkan Jenjang pendidikan

Variabel	Mean	SD	Nilai p
Jenjang pendidikan			0,0005
D3	43,68	11,27	
D4	43,74	11,09	
S1	42,37	10,63	

Berdasarkan tabelmahasiswa yang di jenjang pendidikan D3, rata-rata skor kecemasannya adalah 43,68 dengan standar deviasi 11,27; sedangkan pada mahasiswa D4 rata-rata skor kecemasannya adalah 43,74 dengan standar deviasi 11,09; mahasiswa S1 rata-rata skor kecemasannya adalah 42, 37 dengan standar deviasi 10,63. Sedangkan dari uji F yang dilakukan menghasilkan nilai $p = 0,0005$ sehingga dapat disimpulkan bahwa terdapat perbedaan rata-rata skor kecemasan pada mahasiswa berdasarkan jenjang pendidikannya. Analisis lebih lanjut dilakukan menggunakan uji Bonferroni untuk melihat kelompok mana saja yang berbeda secara signifikan. Hasilnya adalah kelompok yang berbeda secara signifikan antara mahasiswa D3 dengan S1 dan antara mahasiswa D4 dengan S1.

Bab 9

Analisis Data Proporsi

9.1 Pendahuluan

Data merupakan sesuatu yang dikenal sebagai informasi yang memberikan gambaran tentang situasi atau masalah yang akan dipelajari. Misalnya ketika pemerintah ingin menyediakan lapangan pekerjaan. Maka dilakukan pengumpulan data penduduk berdasarkan usia, jenis kelamin, tingkat pendidikan, jenis pekerjaan dan lain – lain untuk mengetahui kondisi masyarakat.

9.1.1 Pengertian

Data didefinisikan sebagai fakta - fakta, pemikiran, atau pendapat yang tidak atau belum memiliki arti kegunaan. Berdasarkan jenisnya, data dibagi menjadi 2 yaitu:

1. Data kualitatif
Data yang tidak berbentuk angka (non numeric), seperti jenis kelamin dan warna
2. Data kuantitatif
Data yang berbentuk angka (numeric), seperti jumlah penjualan TV, umur, dan lain – lain

Data kuantitatif dibagi lagi menjadi 2 bagian yaitu data diskrit dan data kontinyu. Data diskrit adalah data yang dihasilkan dari perhitungan dan nilainya berupa bilangan bulat, biasanya juga disebut data komunikatif. Contohnya jumlah motor 0, 1, 2, dan seterusnya. Data kontinyu merupakan data yang merupakan hasil pengukuran dan nilainya bisa berupa bilangan pecahan atau bulat, yang disebut juga dengan data kuantitatif. Contohnya berat badan 60,5 kg atau 60kg.

Menurut sumbernya data dibagi menjadi data primer dan data sekunder. Data primer yaitu data yang didapatkan langsung dari lapangan atau objek yang dijadikan bahan penelitian, seperti data yang diperoleh menggunakan daftar pertanyaan. Hal lain yang dapat dijadikan data primer yaitu wawancara langsung dengan objek penelitian, melalui pihak ketiga (menggunakan pihak ketiga sebagai sumber informasi), mengikuti daftar pertanyaan melalui pos, email pada responden.

Data yang diperoleh dari pihak lain atau dari sumber lain disebut dengan data sekunder. Contoh sumber dari data sekunder berupa data yang telah dipublikasikan, majalah, koran, dan lain sebagainya.

9.1.2 Pengumpulan Data

Kegiatan statistika adalah metode yang digunakan untuk mempelajari pengumpulan, pengaturan, perhitungan, penggambaran, dan penganalisaan data, serta penarikan kesimpulan yang meyakinkan (valid) berdasarkan proses analisa yang dilakukan dan pembuatan keputusan yang rasional. Statistik dianggap sebagai gabungan bahan keterangan yang berupa angka atau bilangan; dapat pula diartikan sebagai deretan atau kumpulan angka yang menjelaskan tentang cabang kegiatan hidup tertentu. Statistik digunakan pula sebagai alat analisis, alat untuk membuat keputusan.

Data statistik harus mengandung data yang benar dan dapat dipercaya. Maka dari itu, proses pengumpulan data harus melihat objek penelitian, cara dalam pengumpulan data dilakukan, serta tujuan dari pengumpulan data.

Dalam mengerjakan analisis data perlu diketahui beberapa langkah yang harus dipahami, di antaranya:

1. Pengumpulan data

Hal ini dilakukan pertama kali saat dilakukan analisis data

2. Tahap penataan
Melakukan pemeriksaan terhadap transparansi dan kelengkapan pengisian instrumen pengumpulan data
3. Tahap pengkodean
Mengidentifikasi dan mengumpulkan semua pernyataan kelengkapan data berdasarkan variabel yang akan diteliti
4. Tahap pemeriksaan
Metode yang digunakan untuk pengecekan ulang kualitas data, baik dari prinsip dasar maupun kebutuhan alat pengumpulan data
5. Tahap menguraikan data
Proses ini dilakukan untuk menjelaskan dengan mentransmisikan data dalam bentuk tabel frekuensi atau diagram dengan berbagai macam ukuran. Tujuannya untuk memahami karakteristik data dari suatu penelitian
6. Tahap pemeriksaan hipotesis
Proses ini dilakukan untuk melihat apakah data penelitian dapat diterima atau ditolak, logis atau tidak. Dari proses ini akan menghasilkan suatu keputusan atau kesimpulan terhadap penelitian yang akan dijalankan

Objek penelitian merupakan komponen yang akan dilakukan pemeriksaan seperti individu, organisasi atau barang. Tujuan dari pengumpulan data yaitu digunakan sebagai contoh gambaran atau deskripsi mengenai suatu peristiwa serta proses penyelesaian masalah. Setelah mengetahui komponen yang diperiksa, maka perlu juga untuk mengetahui karakteristik dari hal tersebut. karakteristik adalah sifat atau ciri – ciri yang dimiliki oleh setiap elemen atau komponen. Alat ukur dari karakteristik suatu elemen adalah nilai variabel. Variabel sendiri memiliki arti nilai atau hal yang dapat berubah, contohnya pendapatan nasional, tinggi badan, hasil penjualan dan harga.

Metode yang digunakan dalam melakukan pengumpulan data yaitu dengan cara sensus dan sampling. Sensus yaitu pengumpulan data dari semua atau seluruh elemen dari objek penelitian. Sedangkan sampling adalah metode yang digunakan dalam pengumpulan data dengan mengaplikasi setengah atau sebagian dari objek penelitian. Pengambilan sampel dilakukan dengan dua cara, yaitu secara acak (random) dan tidak acak (non random).

Pengambilan sampel secara acak yaitu memilih elemen dari populasi yang dijadikan objek dengan kesempatan yang sama untuk terpilih. Cara pemilihan dapat dilakukan dengan undian dan tabel angka, teknik ini disebut dengan probability sampling. Pengambilan sampel secara tidak acak merupakan suatu pemilihan elemen dari populasi di mana setiap elemen tidak memiliki kesempatan yang sama untuk terpilih sebagai sampel, teknik ini disebut dengan non-probability sampling.

Variabel adalah segala sesuatu yang akan dijadikan objek pengamatan dalam penelitian. Variabel juga diartikan sebagai faktor tidak tetap atau gejala yang dapat diubah – ubah. Misalnya, nilai – nilai mata kuliah mahasiswa. Variabel pada dasarnya bersifat kualitatif namun ditandai dengan angka. Misalnya, nilai matematika adalah gejala kualitatif, tetapi dilambangkan dengan angka seperti 70, 85, 53, 64 dan sebagainya. Umur juga merupakan gejala kualitatif yang dilambangkan dengan angka seperti 10 tahun, 20 tahun, dan sebagainya.

Beberapa cara yang digunakan untuk melakukan pengukuran terhadap variabel adalah:

1. Skala nominal

Skala pengukuran yang diberikan kepada objek tidak menunjukkan tingkatan atau arti tertentu, tetapi berupa label. Contohnya pembagian enam warna pada permen mentos. Pada pembagian ini tidak ada aturan khusus yang berlaku dalam mengurutkannya, semua warna memiliki kesempatan menjadi yang pertama atau terakhir.

2. Skala ordinal

Skala pengukuran di mana angka yang digunakan memiliki tingkatan. Contohnya pengurutan mahasiswa berdasarkan umurnya

3. Skala interval

Skala pengukuran dengan angka pada klasifikasi atau ciri objek yang sama dengan skala ordinal dengan menambah jarak atau interval yang sama. Contohnya pengelompokkan berdasarkan suhu badan

4. Skala rasio

Skala pengukuran yang bersifat membedakan, memiliki tingkatan yang jaraknya pasti, dan setiap nilai kategori diukur dari titik yang sama. Contohnya berat badan, tinggi badan, dan lain sebagainya

Berdasarkan konteks hubungannya, variabel memiliki jumlah lebih dari satu di setiap penelitian. Hal ini membuat setiap variabel saling berhubungan, jika dilihat dari konteksnya dapat dibedakan menjadi:

1. Variabel bebas atau independent variables
Variabel yang nilainya akan memengaruhi variabel lainnya, yaitu variabel terikat
2. Variabel terikat atau dependent variables
Variabel yang nilainya tergantung pada nilai variabel lainnya
3. Variabel moderator dan variable intervening
Variabel yang berpengaruh terhadap variabel terikat, tetapi pengaruhnya tidak diutamakan
4. Variabel perancu atau confounding variable
Variabel yang memiliki hubungan dengan variabel bebas dan variabel terikat, tetapi bukan variabel antara
5. Variabel kendali
Variabel yang memengaruhi variabel terikat, tetapi perannya sebagai variabel yang netral dalam penelitian
6. Variabel rambang
Variabel yang ikut memengaruhi variabel terikat, namun pengaruhnya tidak begitu berarti sehingga variabel ini sering diabaikan dalam penelitian

Ada sejumlah variabel yang memungkinkan seorang peneliti untuk melakukan intervensi maupun tidak melakukan intervensi. Atas dasar ini, variabel dibedakan menjadi:

1. Variabel dinamis
Peneliti dapat melakukan manipulasi atau intervensi terhadap variabel. Contohnya metode mengajar, teknik pelatihan, strategi pembiasaan, dan lain sebagainya
2. Variabel statis
Peneliti tidak dapat melakukan manipulasi atau intervensi terhadap variabel. Contohnya jenis kelamin, umur, status perkawinan, dan lain sebagainya

Dalam mengoperasionalkan variabel dilakukan dengan beberapa cara di antaranya yaitu:

1. Menjabarkan variabel teoritis ke dalam konsep empiris dan konsep analitis dalam bentuk indikator – indikator yang terukur
2. Menjelaskan variabel kedalam bentuk aspek atau sub variabel. Aspek yang dimaksud merupakan fokus atau sudut pandang dari peneliti mengenai alasan menggunakan variabel tersebut
3. Menjelaskan aspek atau sub variabel tersebut dalam bentuk indikator – indikator yang terukur berupa skala, misalnya skala nominal, skala ordinal, skala interval serta skala rasio

Definisi operasional dari variabel dapat dijabarkan dengan narasi atau matrik, yang di dalamnya terkandung nama variabel, definisi variabel, cara ukur variabel, alat ukur variabel, hasil ukur variabel, dan jenis skala variabel.

9.2 Proses Analisis Dan Penyajian Data

Data merupakan serangkaian informasi berupa angka atau pernyataan yang dapat dimanfaatkan dalam penggambaran suatu kejadian atau peristiwa, pengambilan keputusan atau membuat kebijakan. Data – data yang dimaksud dapat berupa data yang masih mentah (belum melalui proses pengolahan) maupun data yang sudah menjadi olahan atau data olahan.

Dalam melakukan analisis data, ada beberapa tahapan yang harus dilalui agar analisis penelitian menghasilkan informasi sesuai dengan kebutuhan yang diinginkan.

Tahapan – tahapan tersebut antara lain:

1. Editing
Tahapan paling awal dalam melakukan analisis data, hal ini dilakukan untuk memeriksa apakah isian formulir atau kuesioner sudah lengkap, jelas, relevan, dan konsisten
2. Koding
Proses dalam merubah data berbentuk huruf menjadi angka atau bilangan. Misalnya variabel untuk pendidikan dilakukan koding 1=

SD, 2= SMP, 3= SMA, 4= PT dan sebagainya. Manfaat dari pemberian koding adalah mempercepat saat dilakukan entry data dan memudahkan dalam analisis data

3. Processing

Setelah dilakukan pengkodean, data tersebut dapat dilakukan analisa

4. Cleaning

Hal ini dilakukan untuk memeriksa kembali data yang sudah dimasukkan apakah ada kekeliruan atau tidak

Analisis data memiliki alat analisis yang membedakan data menjadi dua yaitu data kategorik dan data numerik. Data kategorik atau kualitatif yaitu data dari hasil pengelompokan, seperti jenis kelamin, jenis pekerjaan, tingkat pendidikan, dan lain – lain. Sedangkan data numerik atau kuantitatif yaitu variabel data dari hasil perhitungan dan pengukuran. Contohnya jumlah anak, jumlah pengunjung rumah sakit, tekanan darah, dan lain – lain. Variabel kategorik secara umum berisi variabel berskala nominal dan ordinal, variabel numerik berisi variabel berskala interval dan rasio.

Secara garis besar analisis data dilakukan untuk memperoleh penjelasan dari masing – masing variabel, membandingkan dan menguji teori atau konsep dengan informasi yang telah ditemukan, menemukan rancangan baru dari data yang sudah dikumpulkan, mendapatkan konsep baru yang diuji berlaku untuk umum atau hanya dapat berlaku pada kondisi tertentu.

Analisis data didahului dengan mengulas seluruh data yang diperoleh dari berbagai sumber, seperti wawancara, pengamatan, catatan yang ditulis. Dalam analisis data terdapat dua model pendekatan, yaitu analisis kuantitatif secara deskriptif dan analisis kuantitatif secara inferensial. Keduanya memiliki karakteristik yang berbeda dari segi teknik analisis maupun tujuan yang akan dihasilkan dari analisisnya.

Pada dasarnya semua data saling berhubungan dengan berbagai peristiwa atau kejadian yang ada di dunia ini. Salah satu metode statistik yang digunakan untuk mencari hubungan antara dua variabel. Analisis korelasi adalah salah satu metode yang digunakan untuk menilai kedekatan hubungan (korelasi) antara dua variabel. Koefisien korelasi bernilai antara -1 sampai dengan +1. Lambang dari koefisien ini adalah r .

Hubungan A dan B dapat bernilai positif dan negatif, contohnya apabila kenaikan (penurunan) A diikuti kenaikan (penurunan) B bernilai positif, sedangkan jika kenaikan penggunaan pupuk (A) diikuti penurunan (kenaikan) B. Untuk melihat hubungan di antara kedua variabel digunakan dengan tampilan dalam bentuk diagram pencar. Diagram pencar yang teratur menunjukkan adanya hubungan antara variabel A dan B. Hubungan positif pada diagram pencar ditandai dengan gerakan dari kiri ke bawah atas kanan. Hubungan negatif ditunjukkan dengan gerakan pencar dari kiri atas ke kanan bawah.

Analisis regresi terjadi apabila hubungan variabel A dan B untuk perkiraan ramalan variabel B (sebagai variabel terikat) karena perubahan variabel A (sebagai variabel bebas). Pembentukan garis persamaan yang lurus atau linear disebut dengan persamaan regresi. Garis regresi dapat diletakkan pada data diagram sehingga selisih penyimpangan positif dan negatif dari titik – titik ini berada di atas dan di bawah garis lurus yaitu nol.

Data statistik tidak hanya dianalisis tetapi juga disajikan dalam bentuk tabel, grafik maupun tulisan, sesuai dengan kebutuhan dan dapat dimengerti oleh peneliti maupun pembaca. Pengertian tabel yaitu sekumpulan angka – angka yang tersusun menurut kategori – kategori (misalnya jumlah pegawai menurut pendidikan dan masa kerja, jumlah penjualan menurut jenis barang dan daerah penjualan, jumlah produksi menurut jenis barang dan kantor cabang, dan lain sebagainya) sehingga dapat dengan mudah dalam membuat analisis data.

Pada umumnya sebuah tabel terbagi menjadi master table (tabel induk) dan text table (tabel rincian). Tabel induk menyajikan kumpulan data yang masih dalam bentuk mentah dan disajikan dalam lampiran suatu laporan pengumpulan data. Tabel rincian yaitu penjelasan dari data yang bersumber dari tabel induk, contohnya distribusi frekuensi, distribusi relative, distribusi kumulatif, serta tabel silang (kontingensi table = cross tabulasi).

Dalam penyajiannya, tabel mencakup beberapa bagian antara lain:

1. Nama (judul) dan identifikasi

Penempatan judul yaitu berada di atas tabel dengan bahasa yang jelas dan singkat, atau dengan sederhana. Urutan redaksi secara umum menerangkan tentang ciri – ciri data yang terdapat dalam tabel, jika judul disusun dalam beberapa baris maka susunannya berupa piramida yang terbalik atau dalam baris – baris yang rapi dan teratur.

Jika terdapat beberapa tabel dalam satu tulisan, maka perlu diberikan penomoran secara urut atau sistematis

2. Catatan pendahuluan dan catatan di bawah tabel

Catatan pendahuluan dan catatan di bawah tabel adalah bagian dari integral tabel. Catatan pendahuluan umumnya ditempatkan langsung di bawah nama tabel dalam bentuk yang agak kurang menonjol jika dibandingkan dengan namanya sendiri

3. Sumber data

Sumber data diletakkan langsung di bawah tabel setelah catata, dengan keterangan harus lengkap (penulis, nama buku, jilid, halaman buku, penerbit, dan lain – lain)

4. Presentase

Perincian pos – pos keterangan dalam kompartimen berupa angka – angka yang digunakan dalam tabel harus ditulis dengan jelas

5. Jumlah

Jika jumlah angka pada tabel merupakan sesuatu yang penting dan perlu diperlihatkan maka diletakkan pada sisi atas atau sisi kiri dalam nama kolom. Jumlah juga ditulis dalam bentuk huruf tebal atau huruf besar

6. Unit

Unit pengukuran angka – angka yang tertera pada tabel umumnya sudah jelas dan tidak diragukan

7. Bentuk tabel

Terdapat beberapa bentuk tabel yaitu tabel satu arah (memuat keterangan mengenai suatu hal atau satu karakteristik), tabel dua arah (memuat hubungan dua hal atau dua karakteristik), tabel tiga arah (memuat tiga hal atau tiga karakteristik, misalnya jumlah kendaraan perusahaan menurut merk, jenis, dan umur kendaraan)

Seperti halnya tabel, grafik juga dibagi menjadi beberapa macam di antaranya histogram frekuensi, poligon frekuensi, ogif, diagram garis, diagram batang, diagram pinca (pie), diagram tebar, box wishker plot, stem and leaf plot. Histogram frekuensi adalah grafik yang berbentuk batang untuk menyajikan

data kontinyu. Biasanya histogram digambarkan dengan mudah apabila distribusi frekuensinya memiliki interval kelas yang sama bagi setiap kelas.

Penyajian data berbentuk poligon frekuensi dimanfaatkan untuk data kontinyu seperti pada histogram. Grafik poligon frekuensi dibuat berdasarkan hubungan puncak – puncak dari suatu balok – balok histogram. Kegunaan dari grafik adalah dapat melakukan perbandingan penyebaran beberapa masalah yang digambar dalam satu gambar.

Grafik kontinyu dalam bentuk kumulatif adalah ogiv. Dari perpotongan ogiv kurang dari (less than) dan besar dari (more than), akan didapatkan nilai yang tepat untuk letak dan besarnya nilai modus. Dengan kata lain, distribusi data pada kurva ogiv dibentuk dari mencari nilai frekuensi kumulatifnya. Diagram batang diaplikasikan dalam penyajian data diskrit atau data dengan skala nominal maupun ordinal. Perbedaan dari histogram dan diagram batang yaitu, balok pada histogram menyambung karena mewakili data kontinyu sedangkan diagram batang terpisah.

Ketika ingin menyajikan data dari waktu ke waktu atau perubahan dari suatu tempat ke tempat lain, diagram yang dapat digunakan yaitu diagram garis. Penggunaan diagram garis yaitu untuk data diskrit atau data skala nominal. Luas lingkaran adalah 360 derajat, dibagi berdasarkan data yang diperoleh merupakan penggambaran dari diagram pinca (pie). Diagram ini berbentuk lingkaran dengan sajian data menggunakan bentuk derajat. Diagram pinca diperlukan untuk menyajikan data diskrit atau data dengan skala nominal dan ordinal yang disebut dengan data kategorik.

Selain dalam bentuk tabel dan grafik, penyajian data juga berbentuk tulisan. Hampir semua bentuk laporan dari pengumpulan data diserahkan dalam bentuk tertulis, mulai dari proses pengambilan sampel, pelaksanaan pengumpulan data hingga hasil analisis data tersebut.

Bab 10

Uji Chi-Square

10.1 Karakteristik Uji Chi-Square

Dalam analisis bivariat, terdapat beberapa uji yang dapat digunakan, salah satunya adalah uji Chi-square. Pada tabel 10.1 disajikan perbedaan dari beberapa uji statistik yang digunakan dalam analisis bivariat.

Uji chi-square (χ^2) adalah uji nonparametrik, yaitu pengujian yang tidak melibatkan penggunaan parameter populasi apa pun, digunakan untuk membandingkan informasi kategorik dengan apa yang kita harapkan berdasarkan pengetahuan sebelumnya (Jackson, 2015). Uji chi-square (χ^2) merupakan uji hipotesis tentang perbedaan proporsi antara beberapa kelompok data (Polit and Beck, 2018). Peneliti sering melakukan beberapa uji chi-square dalam satu sampel.

Tabel 10.1: Guide to Major Bivariate Statistical Tests (Polit and Beck, 2018)

Name	Test Statistic	Purpose	Measurement Level	
			Independent Variable	Dependent Variable
t-test for independent groups	t	To test the difference between the means of two independent groups (e.g., experimental vs. control, men vs. women)	Nominal	Continuous*
t-test for paired groups	t	To test the difference between the means of a paired group (e.g., pretest vs. posttest for the same people)	Nominal	Continuous*
Analysis of variance (ANOVA)	F	To test the difference among means of 3+ independent groups	Nominal	Continuous*
Repeated measures ANOVA	F	To test the difference among means of 3+ related groups, e.g., the same group over time, or to compare 2+ groups over time	Nominal	Continuous*
Pearson's correlation coefficient	r	To test the existence and strength of a relationship between two variables	Continuous*	Continuous*
Chi-squared test	χ^2	To test the difference in proportions in 2+ independent groups	Nominal (or ordinal, few categories)	Nominal (or ordinal, few categories)

*Continuous measures are on an interval- or ratio-level scale.

10.1.1 Tujuan Uji Chi-Square

Dalam penelitian kesehatan uji chi-square digunakan dalam analisis hubungan antara variabel independen berupa data kategorik (nominal atau ordinal) dengan variabel dependen berupa data kategorik (nominal atau ordinal). Sebagai contoh, sebuah penelitian menentukan apakah tingkat pengetahuan responden berhubungan dengan kejadian diare. Uji chi-square digunakan dalam menguji hubungan variabel tingkat pengetahuan (kategori dengan klasifikasi rendah dan tinggi) dengan variabel kejadian diare (kategori dengan klasifikasi ya dan tidak).

Uji chi-square juga digunakan untuk membandingkan dua (atau lebih) proporsi untuk menentukan apakah keduanya berbeda secara signifikan satu sama lain (Hulley et al., 2013). Sebagai contoh, sebuah penelitian menentukan apakah resiko terkena Hipertensi serupa di antara orang yang rutin berolahraga setidaknya dua kali seminggu dibandingkan dengan mereka yang tidak atau jarang berolahraga. Perbandingan terjadinya resiko hipertensi diantara 2 kategori responden tersebut secara statistik menggunakan uji chi-square.

10.1.2 Prinsip Dasar Uji Chi-Square

Prinsip dasar uji chi-square adalah membandingkan frekuensi yang terjadi (observasi) dengan frekuensi harapan (ekspektasi). Bila nilai frekuensi observasi dengan nilai frekuensi harapan sama, maka dikatakan tidak ada perbedaan yang bermakna (signifikan), sebaliknya, bila nilai frekuensi observasi dan nilai frekuensi harapan berbeda, maka dikatakan ada perbedaan

yang bermakna (signifikan) (Hastono, 2017). Nilai data kedua variabel disajikan dalam bentuk tabel silang dalam mempermudah analisis uji chi-square. Uji chi-square selalu dua sisi (two-sided).

Chi-square meringkas perbedaan antara frekuensi yang diamati (observed frequencies) dan frekuensi yang diharapkan (expected frequencies) dalam tabel bivariat. Pembuktian dengan uji chi-square menggunakan formula sebagai berikut:

$$X^2 = \sum \frac{(O - E)^2}{E}$$

$$df = (k-1)(n-1)$$

keterangan:

O = nilai observasi

E = nilai ekspektasi (harapan)

k = jumlah kolom

b = jumlah baris

Hastono (2017) menjelaskan bahwa jika frekuensi sangat kecil, penggunaan uji Chi-Square mungkin kurang tepat. Keterbatasan uji Chi-Square yaitu:

1. Tidak boleh ada sel yang mempunyai nilai harapan (nilai E) kurang dari 1.
2. Tidak boleh ada sel yang mempunyai nilai harapan (nilai E) kurang dari 5, lebih dari 20% dari jumlah sel.
3. Bila keterbatasan tersebut terjadi pada tabel 2 x 2 (ini berarti tidak bisa menggabung kategorik-kategoriknya lagi), maka dianjurkan menggunakan uji Fisher's Exact.

10.1.3 Sifat Uji Chi-Square

Beberapa sifat Uji Chi-square (X^2) dijelaskan oleh Chan (2015), sebagai berikut:

1. Ketika ada kesesuaian yang sempurna (perfect agreement) antara frekuensi yang diamati dan yang diharapkan, nilai Chi-square nol, dan H_0 harus diterima.
2. Ketika ada kesepakatan yang erat (close agreement) antara frekuensi yang diamati dan yang diharapkan, nilai Chi-square kecil (dan nilai-p besar).
3. Ketika ada kesepakatan yang buruk (poor agreement) antara frekuensi yang diamati dan yang diharapkan, nilai Chi-square besar (dan nilai-p kecil), dan H_0 harus ditolak.

10.2 Analisis Uji Chi-Square

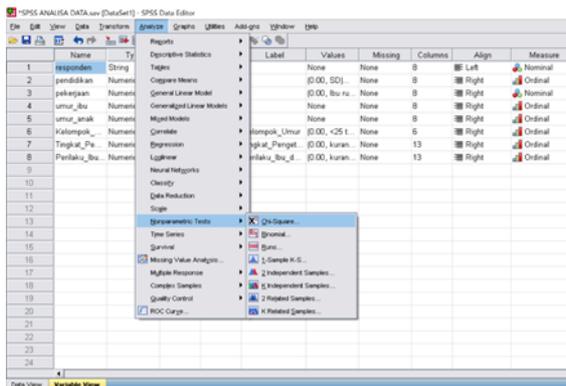
Ada dua pilihan menu dalam *Statistic Program for Social Science* (SPSS) yang digunakan dalam uji chi-square, yaitu: pengujian non parametrik dan tabulasi silang (cross-tabulation) (Wagner III, 2019).

10.2.1 Pengujian Non Parametrik

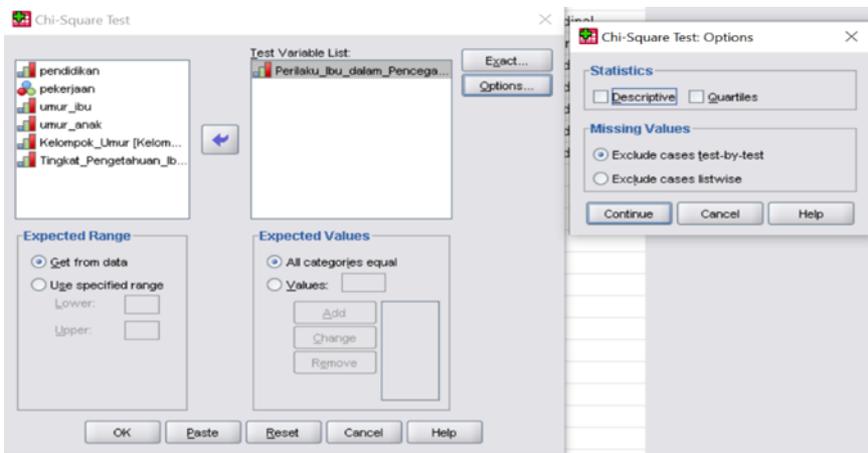
Bila distribusi data populasi yang akan diuji berbentuk normal, maka proses pengujian dapat digunakan dengan pendekatan uji statistik parametrik, namun bila distribusi data tidak normal atau tidak diketahui distribusinya maka dapat digunakan pendekatan uji statistik non parametrik. Bila jenis variabelnya kategorik (kualitatif), maka bentuk distribusinya tidak normal, sehingga uji non parametrik dapat digunakan (Hastono, 2017).

Analisis uji Chi-square dalam pengujian non parametrik, dengan menggunakan pilihan opsi menu berikut:

Analisis → Tes Nonparametrik → Legacy Dialogs → Chi-Square



Sekarang, di kotak dialog "Uji Chi-square", pilih variabel yang Anda ingin uji. Dalam kasus ini, variabel yang dipilih adalah perilaku ibu.



Bila Anda klik tombol "Option", Anda akan diberikan kotak dialog yang memungkinkan Anda meminta informasi tambahan dan menentukan metode kasus yang seharusnya. Klik "Continue", lalu klik "OK" di Kotak dialog "Uji Chi-square". Output akan diberikan SPSS seperti berikut ini:

→ Chi-Square

Frequencies

Perilaku Ibu dalam Pencegahan ISPA			
	Observed N	Expected N	Residual
cukup	13	20.0	-7.0
baik	27	20.0	7.0
Total	40		

Test Statistics

	Perilaku Ibu dalam Pencegahan ISPA
Chi-Square	4.900 ^a
df	1
Asymp. Sig.	.027

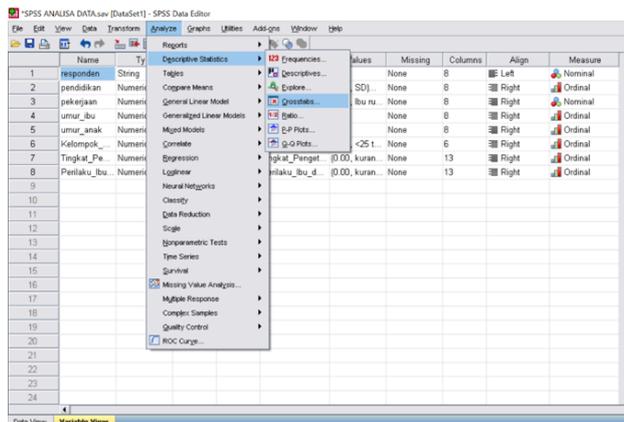
a. 0 cells (.0%) have expected frequencies less than 5. The minimum expected cell frequency is 20.0.

Pada output SPSS di atas, Tabel pertama memberikan jumlah yang diamati dan diharapkan, dan residual atau perbedaan. Tabel kedua menghasilkan nilai chi-square, derajat kebebasan dan mengungkapkan apakah nilai tersebut signifikan secara statistik. Dalam output analisis ini, didapatkan nilai chi-square = 4,9 dengan 1 derajat kebebasan (df) dan signifikan secara statistik ($p = 0,027$).

10.2.2 Cross-Tabulation

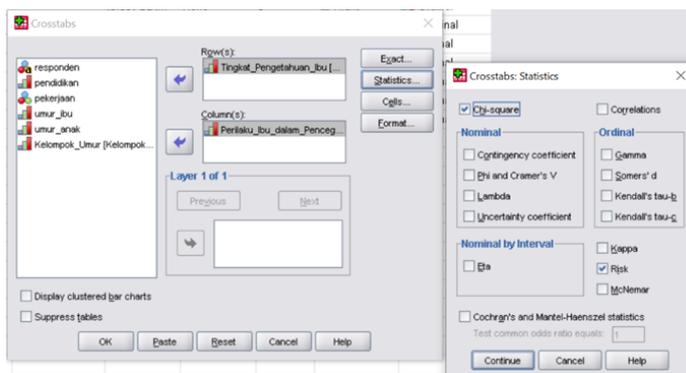
Analisis chi-square yang kedua adalah dengan menggunakan cross-tabulation, dengan menggunakan pilihan opsi menu berikut:

Analyze → Descriptive Statistics → Crosstabs

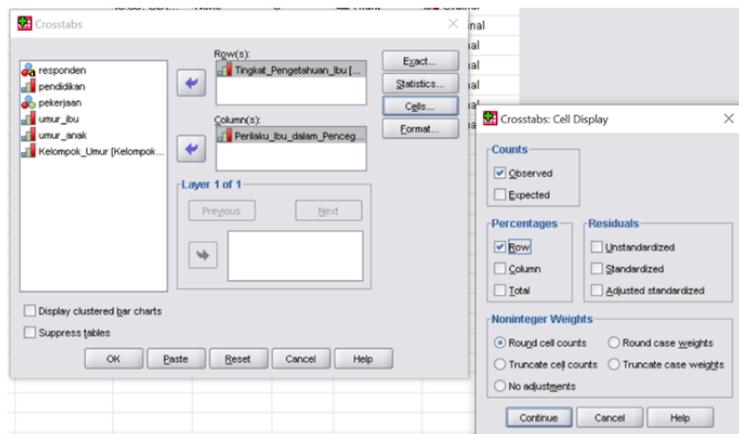


Dari menu crosstab, ada dua kotak yang harus diisi, pada kotak “Row(s)” diisi variabel independen (variabel bebas), dalam contoh ini variabel tingkat pengetahuan masuk ke kotak “Row(s)”. Pada kotak “Column(s)” diisi variabel dependen (terikat), dalam contoh ini variabel perilaku ibu dalam pencegahan Infeksi Saluran Pernafasan Akut (ISPA) masuk ke kotak “Column(s)”.

Selanjutnya, klik option “Statistics”, klik pilihan “Chi Square” dan klik pilihan “Risk”



Lalu, klik option “Cells”, maka akan muncul tampilan berikut, dan klik “Row”. Lalu klik Continue dan klik ok di kotak dialog ”Uji Chi-Square”.



Output akan diberikan oleh SPSS seperti berikut ini:

Case Processing Summary

	Cases					
	Valid		Missing		Total	
	N	Percent	N	Percent	N	Percent
Pengetahuanibu * Perilaku_Ibu_dalam_Pencegahan_ISPA	40	100.0%	0	.0%	40	100.0%

Pengetahuanibu * Perilaku_Ibu_dalam_Pencegahan_ISPA Crosstabulation

		Perilaku_Ibu_dalam_Pencegahan_ISPA		Total	
		cukup	baik		
Pengetahuanibu	kurang	Count	11	18	29
		% within Pengetahuanibu	37.9%	62.1%	100.0%
	baik	Count	2	9	11
		% within Pengetahuanibu	18.2%	81.8%	100.0%
Total		Count	13	27	40
		% within Pengetahuanibu	32.5%	67.5%	100.0%

Chi-Square Tests

	Value	df	Asymp. Sig. (2-sided)	Exact Sig. (2-sided)	Exact Sig. (1-sided)
Pearson Chi-Square	1.418 ^a	1	.023		
Continuity Correction ^b	.661	1	.042		
Likelihood Ratio	1.519	1	.022		
Fisher's Exact Test				.028	.021
Linear-by-Linear Association	1.382	1	.024		
N of Valid Cases ^a	40				

a. 1 cells (25,0%) have expected count less than 5. The minimum expected count is 3,58.

b. Computed only for a 2x2 table

Risk Estimate

	Value	95% Confidence Interval	
		Lower	Upper
Odds Ratio for Pengetahuanibu (kurang / baik)	2.750	.499	15.143
For cohort Perilaku_Ibu_dalam_Pencegahan_ISPA = cukup	2.086	.548	7.946
For cohort Perilaku_Ibu_dalam_Pencegahan_ISPA = baik	.759	.509	1.130
N of Valid Cases	40		

10.3 Interpretasi Uji Chi-Square

Hastono, (2017) menjelaskan beberapa aturan yang berlaku pada analisis dan interpretasi uji Chi Square adalah sebagai berikut:

1. Bila pada 2 x 2 dijumpai nilai Expected (harapan) kurang dari 5, maka yang digunakan adalah “Fisher’s Exact Test”
2. Bila tabel 2 x 2, dan tidak ada nilai $E < 5$, maka uji yang dipakai sebaiknya “Continuity Correction”
3. Bila tabelnya lebih dari 2 x 2, misalnya 3 x 2, 3 x 3 dsb, maka digunakan uji “Pearson Chi Square”

4. Uji “Likelihood Ratio” dan “Linear-by-Linear Association”, biasanya digunakan untuk keperluan lebih spesifik, misalnya analisis stratifikasi pada bidang epidemiologi dan juga untuk mengetahui hubungan linier dua variable kategorik, sehingga kedua jenis ini jarang digunakan.
5. Untuk mengetahui adanya nilai E kurang dari 5, dapat dilihat pada footnote dibawah kotak Chi-Square Test.
6. Uji Chi-square hanya dapat digunakan untuk mengetahui ada/tidaknya hubungan dua variabel, namun tidak dapat mengetahui derajat/kekuatan hubungan dua variable. Pada bidang Kesehatan, untuk mengetahui besar/kekuatan hubungan digunakan nilai OR atau RR. Nilai OR digunakan untuk jenis penelitian Cross Sectional dan Case Control, sedangkan nilai RR digunakan bila jenis penelitiannya Kohort.

Pada output analisis SPSS yang digunakan pada analisis (10.2.1) diatas, terlihat tabel silang antara variabel pengetahuan ibu dengan perilaku ibu dalam pencegahan ISPA, dengan masing-masing angka di tiap selnya, baik jumlah kasus dan persentasenya. Persentase disini adalah persentase baris bukan persentase jumlah dari tiap kategorinya. Dalam menyajikan hasil keseluruhan, yang digunakan adalah jumlah persentase dari tiap kategori yang ada pada tiap variabel terhadap seluruh jumlah responden. Namun pada bagian ini akan dijelaskan menggunakan hasil tabel silang yang disajikan dari analisis SPSS.

Dari hasil analisis data diatas, diketahui terdapat 18 (62,1%) dari ibu dengan pengetahuan kurang yang memiliki perilaku pencegahan ISPA yang baik, dan 9 (81,8%) ibu dengan pengetahuan baik yang memiliki perilaku pencegahan ISPA yang baik. Pada footnote di bawah kotak Chi-Square Test tertulis nilainya 1 cell (25,0 %), dimana terdapat expected (harapan) kurang <5 , dengan demikian kita menggunakan uji Chi Square yang sudah dilakukan koreksi yaitu Fisher's Exact Test, dengan p value 0,028. Maka dapat disimpulkan bahwa ada perbedaan perilaku Ibu dalam pencegahan ISPA antara ibu yang memiliki pengetahuan kurang dengan ibu yang memiliki pengetahuan baik, atau dengan kata lain ada hubungan antara pengetahuan ibu dengan perilaku ibu dalam pencegahan ISPA.

Pada hasil di atas nilai OR terdapat pada baris Odds ratio yaitu 2,75 (95% CI: 0,499 – 15,143). Sedangkan nilai RR terlihat dari baris For Cohort yaitu

besarnya 2,086 (95% CI: 0,548 – 7,946). Data yang digunakan ini berasal dari penelitian Cross Sectional, maka kita dapat menggunakan nilai $OR = 2,75$. Sehingga dapat disimpulkan kembali bahwa Ibu yang memiliki pengetahuan baik mempunyai peluang 2,75 kali untuk memiliki perilaku pencegahan ISPA dibandingkan ibu yang memiliki pengetahuan kurang.

Bab 11

Korelasi dan Regresi Linier

11.1 Karakteristik Uji Chi-Square

Teknik yang paling umum digunakan untuk menyelidiki hubungan antara dua variabel kuantitatif adalah korelasi dan regresi linier. Korelasi (Sarwono, 2008) yaitu mengukur kekuatan hubungan linier antara sepasang variabel, sedangkan regresi mengungkapkan hubungan dalam bentuk persamaan, misalnya pada pasien yang menghadiri unit kecelakaan dan gawat darurat, kita bisa menggunakan korelasi dan regresi untuk menentukan apakah ada hubungan antara usia dan tingkat urea. Untuk menghitung interval kepercayaan dan uji hipotesis, diasumsikan bahwa kesalahan pada independen dan distribusi yang normal dengan menggunakan mean nol dan varians (Sugiyono, 2006).

Metode kuadrat terkecil memperkirakan β_0 dan β_1 serta berbagai kuantitas lain yang menggambarkan ketepatan perkiraan dan kesesuaian garis lurus atau linier dengan data, karena garis perkiraan jarang akan cocok dengan data secara tepat, istilah untuk perbedaan antara yang sebenarnya dan yang sesuai nilai data harus ditambahkan. Persamaan tersebut kemudian menjadi rumus sebagai berikut;

$$y_j = b_0 + b_1 x_j + e_j$$

$$= y_j + e_j$$

J adalah jumlah observasi (baris), b_0 estimasi β_0 , b_1 estimasi β_1 dan e_j adalah selisih antara nilai data aktual y_j dan nilai konstan yang diberikan oleh persamaan regresi, yang sering disebut sebagai y_j . Perbedaan ini biasanya disebut sebagai sisa. Perhatikan bahwa persamaan regresi linier adalah model matematika yang menggambarkan hubungan antara X dan Y. Pada banyak kasus, model tidak hanya untuk mendefinisikan hubungan yang tepat antara dua variabel namun digunakan sebagai perkiraan hubungan yang tepat. Pada bagian analisis juga akan menentukan seberapa dekat variabel tersebut.

Perhatikan juga bahwa persamaan memprediksi Y dari X. Nilai Y bergantung pada nilai X. Pengaruh variabel lainnya terhadap nilai Y disatukan menjadi residual. Regresi linier mengacu pada sekelompok teknik untuk menyesuaikan dan mempelajari hubungan garis lurus Antara dua variabel. Regresi linier memperkirakan koefisien regresi β_0 dan β_1 dalam persamaan di mana X adalah variabel bebas, Y adalah variabel terikat, β_0 adalah perpotongan Y, β_1 adalah gradient, dan ϵ adalah kesalahannya

11.2 Analisis Korelasi

Salah satu yang paling sering digunakan dari indeks reliabilitas ini adalah koefisien korelasi. Koefisien korelasi atau sering disebut korelasi adalah indeks yang berkisar dari -1 hingga 1 (Tanzeh, 2009). Jika nilainya mendekati nol maka bernilai tidak ada hubungan linier. Ketika korelasi semakin mendekati plus atau minus satu, bernilai hubungan yang semakin kuat. Sebuah nilai satu (atau negatif satu) menunjukkan hubungan linear sempurna antara dua variable (Sudarmanto, 2005). Korelasi adalah parameter bivariat berdistribusi normal. Distribusi ini digunakan untuk mendeskripsikan asosiasi antara dua variabel. Asosiasi ini tidak termasuk pernyataan sebab dan akibat, artinya variabel tidak diberi label sebagai independen dan dependen, karena satu variabel tidak bergantung pada variabel yang lain (Ghazali, 2005). Setiap variabel dianggap sebagai dua variabel acak yang tampaknya bervariasi. Hal yang perlu diperhatikan dalam regresi linier adalah Y diasumsikan sebagai variabel acak

dan X diasumsikan sebagai variabel tetap. Dalam analisis korelasi, Y dan X diasumsikan sebagai variabel acak.

Analisis korelasi mengacu pada derajat hubungan antar variabel. Tapi itu tidak menjelaskan tentang variabel mana yang menyebabkan dan mana yang merupakan akibat. Studi korelasi antara dua variabel tersebut disebut sederhana dan antara lebih dari dua variabel parsial atau ganda (Montgomery, 2012). Korelasi dapat dipelajari dengan dua metode, metode diagram dan metode matematika, setelah titik potong dan kemiringan diperkirakan menggunakan kuadrat terkecil, berbagai indeks dipelajari untuk menentukan kemampuannya dalam hal perkiraan (Hasan, 2004).

11.3 Koefisien Korelasi

Analisis korelasi sering disebut sebagai koefisien korelasi. Koefisien korelasi merupakan nilai yang memperlihatkan kuat atau tidaknya hubungan linier antara dua variabel. Koefisien korelasi sering dilambangkan dengan huruf r di mana nilai r mempunyai bermacam makna dari -1 sampai $+1$, penjabaran karakteristik koefisien korelasi (Uyanto, 2009) sebagai berikut;

1. Rentang r adalah antara -1 dan 1 , inklusif
2. Jika $r = 1$, pengamatan dilakukan pada garis lurus dengan kemiringan positif
3. Jika $r = -1$, pengamatan dilakukan pada garis lurus dengan kemiringan negatif.
4. Jika $r = 0$, maka tidak ada hubungan linier antara kedua variabel
5. R adalah ukuran hubungan linier (garis lurus) antara dua variabel.
6. Nilai r tidak berubah jika X atau Y dikalikan dengan konstanta atau jika konstanta ditambahkan.
7. Arti fisik dari r secara matematis abstrak dan mungkin tidak terlalu membantu. Korelasinya adalah kosinus sudut yang dibentuk oleh perpotongan dua vektor diruang dimensi- N . Komponen vektor pertama adalah nilai X sedangkan komponen vektor kedua adalah nilai-nilai yang sesuai dari Y . Komponen-komponen ini disusun

sedemikian rupa sehingga yang pertama dimensi kedua sesuai dengan observasi kedua, dan sebagainya.

Karakteristik-karakteristik yang dimiliki koefisien korelasi ini akan membantu peneliti untuk melakukan analisis. Koefisien korelasi menurut (Rubiyanto, 2009) sendiri mempunyai beberapa penafsiran yang didapat dari beberapa cara. Berikut beberapa penafsirannya, yaitu :

1. Jika Y dan X distandarisasi dengan mengurangkan rata-rata dan membaginya dengan deviasi standarnya, maka korelasi adalah kemiringan regresi dari Y standar pada X standar.
2. Korelasi adalah kovariansi standar antara Y dan X.
3. Korelasi adalah rata-rata geometri lereng regresi Y atas X dan X pada Y.
4. Korelasinya adalah akar kuadrat dari R-kuadrat, menggunakan tanda dari kemiringan regresi Y atas X.

Rumus yang sesuai untuk penghitungan koefisien korelasi adalah sebagai berikut;

$$\begin{aligned}
 r &= \frac{\sum_{j=1}^N w_j (X_j - \bar{X})(Y_j - \bar{Y})}{\sqrt{\left[\sum_{j=1}^N w_j (X_j - \bar{X})^2 \right] \left[\sum_{j=1}^N w_j (Y_j - \bar{Y})^2 \right]}} \\
 &= \frac{s_{XY}}{\sqrt{s_{XX} s_{YY}}} \\
 &= \pm \sqrt{b_{YX} b_{XY}} \\
 &= \text{sign}(b_{YX}) \sqrt{R^2}
 \end{aligned}$$

s_{XY} adalah kovariansi antara X dan Y, b_{XY} adalah kemiringan dari regresi X atas Y serta b_{YX} adalah kemiringan dari regresi Y atas X. s_{XY} dihitung menggunakan rumus, sebagai berikut;

$$s_{XY} = \frac{\sum_{j=1}^N w_j (X_j - \bar{X})(Y_j - \bar{Y})}{N - 1}$$

Koefisien korelasi populasi, ρ , didefinisikan untuk dua variabel acak, U dan W, sebagai berikut;

$$\begin{aligned}\rho &= \frac{\sigma_{UW}}{\sqrt{\sigma_U^2 \sigma_W^2}} \\ &= \frac{E[(U - \mu_U)(W - \mu_W)]}{\sqrt{\text{Var}(U)\text{Var}(W)}}\end{aligned}$$

Perhatikan bahwa definisi ini tidak merujuk pada satu variabel sebagai dependen dan variabel lainnya sebagai independen. Sebaliknya, itu sederhana mengacu pada dua variabel acak.

11.4 Uji Hipotesis Korelasi

Pengujian hipotesis tentang koefisien korelasi populasi, seperti $\rho = \rho_0$, Ketika $\rho_0 = 0$, pengujian tersebut identik dengan uji-t yang digunakan untuk menguji hipotesis dengan slope sebesar nol. Uji statistik dihitung menggunakan rumus berikut ini;

$$t_{N-2} = \frac{r}{\sqrt{\frac{1-r^2}{N-2}}}$$

Pengujian dengan $\rho_0 \neq 0$ akan berbeda dengan pengujian terkait yang kemiringannya sudah ditentukan, bukan nol dan bernilai. Terdapat dua metode untuk melakukan pengujian hipotesis korelasi.

Metode 1. Metode ini menggunakan distribusi koefisien korelasi, menggunakan hipotesis nol yang $\rho = \rho_0$ dan menggunakan distribusi koefisien korelasi sampel, kemungkinan diperolehnya korelasi sampel koefisien, r , dapat dihitung. Kemungkinan ini adalah signifikansi statistik dari pengujian tersebut. Metode ini membutuhkan asumsi bahwa kedua variabel terdistribusi normal bivariat.

Metode 2. Metode ini menggunakan fakta bahwa transformasi Fisher's z , yang diberikan oleh rumus berikut ini;

$$F(r) = \frac{1}{2} \ln\left(\frac{1+r}{1-r}\right)$$

Jika mendekati distribusi normal dengan mean, maka akan menjadi sebagai berikut;

$$\frac{1}{2} \ln \left(\frac{1+\rho}{1-\rho} \right)$$

Dan variansi

Untuk menguji hipotesis dengan $\frac{1}{N-3}$, dapat menghitung z dengan menggunakan rumus berikut;

$$\begin{aligned} z &= \frac{F(r) - F(\rho_0)}{\sqrt{\frac{1}{N-3}}} \\ &= \frac{\ln \left(\frac{1+r}{1-r} \right) - \ln \left(\frac{1+\rho_0}{1-\rho_0} \right)}{2\sqrt{\frac{1}{N-3}}} \end{aligned}$$

Fakta yang digunakan yakni bahwa z kira-kira terdistribusi sebagai distribusi normal standar dengan mean sama dengan nol dan variansi sama dengan satu. Metode ini membutuhkan dua asumsi. Pertama, kedua variabel mengikuti bivariat distribusi normal. Kedua, distribusi z didekati dengan distribusi normal standar. Metode ini menjadi populer karena menggunakan distribusi normal yang tersedia secara umum daripada distribusi korelasi yang tidak jelas. Namun, karena membuat asumsi tambahan, itu tidak seakurat metode 1.

11.5 Interval Kepercayaan Korelasi

Interval kepercayaan $100(1 - \alpha)\%$ untuk ρ dapat dibangun menggunakan salah satu dari dua metode hipotesis dijelaskan di atas. Interval kepercayaan dihitung dengan mencari, baik secara langsung menggunakan Metode 2 atau dengan pencarian menggunakan Metode 1, semua nilai ρ_0 yang uji hipotesisnya tidak ditolak. Kumpulan nilai ini menjadi interval kepercayaan (Montgomery, 2012). Berhati-hatilah untuk tidak membuat kesalahan umum dengan mengasumsikan bahwa selang kepercayaan ini terkait dengan transformasi interval kepercayaan pada lereng β_1 . Dua interval kepercayaan itu tidak sederhana transformasi satu sama lain.

11.6 Koefisien Korelasi Rank Spearman

Koefisien korelasi Rank Spearman adalah analog nonparametrik populer dari koefisien korelasi biasa. Statistik ini dihitung dengan mengganti nilai data dengan peringkatnya dan menghitung koefisien korelasi dari jajaran. Koefisien ini sebenarnya adalah ukuran asosiasi dari korelasi, karena peringkat tidak berubah oleh transformasi monotonik dari aslinya data. Ketika N lebih besar dari 10, distribusi koefisien korelasi peringkat Spearman dapat menggunakan dengan distribusi koefisien korelasi reguler.

11.7 Analisis Regresi Linier

Pada penelitian ataupun percobaan, model regresi sering digunakan untuk mengetahui atau memprediksi pengaruh variabel X terhadap variabel Y yang diteliti. Pembahasan ini akan dibatasi pada regresi linier sederhana yakni tentang hubungan sebab akibat antara dua variabel yang dijabarkan dalam suatu garis lurus atau linier (Sudarmanto, 2005). Analisis regresi difungsikan untuk mengetahui hubungan sebab akibat antara variabel X dan Y berdasarkan teori yang ada.

Penjelasan alasan penggunaan analisis regresi linier, yaitu :

1. Deskripsi

Analisis ini digunakan untuk menemukan persamaan yang mendeskripsikan hubungan antara dua variabel. Tujuan ini membuat asumsi yang paling sedikit.

2. Estimasi Koefisien

Ini adalah alasan terbanyak untuk melakukan analisis regresi. Analisis mungkin memiliki hubungan teoritis dan analisis regresi akan mengkonfirmasi. Kemungkinan besar, ada minat khusus pada besaran dan tanda koefisien, seringkali tujuan regresi ini tumpang tindih dengan yang lain.

3. Prediksi

Perhatian utama disini adalah untuk memprediksi variabel respon, seperti penjualan, waktu pengiriman, efisiensi, tingkat hunian di

rumah sakit, reaksi menghasilkan beberapa proses kimiawi, atau kekuatan beberapa logam. Prediksi ini mungkin sangat tepat dalam perencanaan, pemantauan, atau evaluasi beberapa proses atau sistem. Ada banyak asumsi dan kualifikasi yang harus dibuat dalam kasus, misalnya tidak boleh melakukan ekstrapolasi di luar rentang data serta perkiraan interval membutuhkan asumsi normalitas sebagai patokan.

4. Kontrol

Model regresi dapat digunakan untuk memantau dan mengendalikan sistem, misalnya Anda mungkin ingin mengkalibrasi sistem pengukuran atau mempertahankan variabel respon dalam pedoman tertentu. Saat model regresi digunakan untuk mengendalikan tujuan, maka variabel independen harus dikaitkan dengan variabel dependen secara kausal. Hubungan fungsional ini harus berlanjut seiring waktu, jika tidak maka harus dilakukan modifikasi model.

Bagian ini akan menyajikan rincian teknis dari analisis regresi menggunakan campuran penjumlahan dan notasi matriks, rumus akan menyertakan bobot w_j . Berikut penjabaran rumus dan penjelasannya;

Least Squares

Dengan menggunakan notasi ini, perkiraan kuadrat terkecil ditemukan menggunakan persamaan di bawah ini;

$$\mathbf{b} = (\mathbf{X}'\mathbf{W}\mathbf{X})^{-1} \mathbf{X}'\mathbf{W}\mathbf{Y}$$

Perhatikan bahwa jika bobot tidak digunakan, ini akan berkurang menjadi seperti berikut;

$$\mathbf{b} = (\mathbf{X}'\mathbf{X})^{-1} \mathbf{X}'\mathbf{Y}$$

Nilai prediksi dari variabel dependen didapatkan dari rumus berikut ini;

$$\hat{\mathbf{Y}} = \mathbf{b}'\mathbf{X}$$

Residu dihitung menggunakan rumus berikut ini;

$$\mathbf{e} = \mathbf{Y} - \hat{\mathbf{Y}}$$

Estimated Variances

Perkiraan varian residual dihitung dengan menggunakan persamaan berikut ini;

$$s^2 = \frac{\mathbf{e}'\mathbf{W}\mathbf{e}}{N - 2}$$

Perkiraan varian dari koefisien regresi dihitung dengan menggunakan rumus berikut ini;

$$\begin{aligned} V\begin{pmatrix} b_0 \\ b_1 \end{pmatrix} &= \begin{pmatrix} s_{b_0}^2 & s_{b_0 b_1} \\ s_{b_0 b_1} & s_{b_1}^2 \end{pmatrix} \\ &= s^2 (\mathbf{X}'\mathbf{W}\mathbf{X})^{-1} \end{aligned}$$

Uji hipotesis intercept dan slope

Uji hipotesis intercept dan slope ini penggunaannya dengan asumsi residual berdistribusi normal, uji hipotesis dapat dilakukan dengan menggunakan distribusi t siswa dengan N-2 derajat kebebasan menggunakan rumus sebagai berikut;

$$t_{b_0} = \frac{b_0 - B_0}{s_{b_0}} \qquad t_{b_1} = \frac{b_1 - B_1}{s_{b_1}}$$

Biasanya, nilai hipotesis B_0 dan B_1 adalah nol, namun tidak harus demikian.

Interval Kepercayaan intercept dan slope

Interval kepercayaan 100(1- α)% untuk intercept, β_0 , dapat dilihat pada rumus berikut;

$$b_0 \pm t_{1-\alpha/2, N-2} s_{b_0}$$

Interval kepercayaan 100(1- α)% untuk slope, β_1 , dapat dilihat pada rumus berikut;

$$b_1 \pm t_{1-\alpha/2, N-2} S_{b_1}$$

Interval Kepercayaan Y yang diberikan untuk X

Interval kepercayaan untuk rata-rata Y pada nilai X tertentu, katakanlah X_0 , menjadikan rumus berikut;

$$b_0 + b_1 X_0 \pm t_{1-\alpha/2, N-2} S_{Y_m | X_0}$$

Perhatikan bahwa interval kepercayaan ini mengasumsikan bahwa ukuran sampel di X adalah N. Interval prediksi 100(1- α)% pada nilai Y untuk individu pada nilai X tertentu, katakanlah X_0 , dapat dirumuskan sebagai berikut;

$$b_0 + b_1 X_0 \pm t_{1-\alpha/2, N-2} S_{Y_f | X_0}$$

Interval kepercayaan X yang diberikan untuk Y

Jenis analisis ini disebut prediksi terbalik atau kalibrasi. Interval kepercayaan 100 (1- α)% untuk mean nilai X pada nilai Y tertentu dihitung sebagai berikut. Pertama, hitung X dari Y menggunakan rumus berikut;

$$\hat{X} = \frac{Y - b_0}{b_1}$$

Persamaan untuk menghitung interval menggunakan rumus berikut;

$$\frac{(\hat{X} - g\bar{X}) \pm A \sqrt{\frac{(1-g)}{N} + \frac{(\hat{X} - \bar{X})^2}{\sum_{j=1}^N w_j (X_j - \bar{X})^2}}}{1-g}$$

Interval kepercayaan 100(1- α)% untuk nilai individual X pada nilai Y, yaitu;

$$\frac{(\hat{X} - g\bar{X}) \pm A \sqrt{\frac{(N+1)(1-g)}{N} + \frac{(\hat{X} - \bar{X})^2}{\sum_{j=1}^N w_j (X_j - \bar{X})^2}}}{1-g}$$

R-squared

Beberapa ukuran goodness-of-fit dari model regresi untuk data yang diusulkan, namun sejauh ini yang sering digunakan adalah R^2 . R^2 adalah kuadrat dari koefisien korelasi. Ini adalah proporsi variasi Y dihitung melalui variasi X (Riduwan and Akdon, 2007). R^2 bervariasi antara nol (tidak ada hubungan linier) dan satu (linier sempurna). R^2 secara resmi dikenal sebagai koefisien determinasi.

Rumus untuk R^2 , sebagai berikut;

$$R^2 = 1 - \left(\frac{e'W e}{Y'W Y - \frac{(1'W Y)^2}{1'W 1}} \right)$$

$$= \frac{SS_{Model}}{SS_{Total}}$$

R^2 mungkin adalah ukuran paling populer penggunaannya pada model regresi dengan data peneliti. R^2 juga dapat didefinisikan sebagai rasio atau persentase, jika yang digunakan bentuk rasio maka nilainya berkisar dari nol hingga satu. Nilai R^2 mendekati nol menunjukkan tidak ada hubungan linier, sedangkan nilai yang mendekati satu menunjukkan kesesuaian linier yang sempurna, walaupun sering digunakan, R^2 tidak dapat digunakan secara sembarangan atau diinterpretasikan tanpa dukungan plot pencar.

Berikut adalah beberapa kualifikasi penafsirannya, yaitu

1. Variabel independen tambahan. Ini dimungkinkan untuk meningkatkan R^2 dengan menambahkan lebih banyak variabel independen, tetapi variabel bebas tambahan ini dimungkinkan akan menyebabkan peningkatan kesalahan kuadrat rata-rata, ini merupakan situasi yang tidak menguntungkan. Ini biasanya terjadi jika ukuran sampelnya kecil.
2. Rentang variabel independen. R^2 dipengaruhi oleh rentang variabel independen. R^2 meningkat saat rentang X meningkat dan menurun saat rentang X menurun.
3. Besaran lereng. R^2 tidak mengukur besarnya lereng
4. Linieritas. R^2 tidak mengukur kesesuaian model linier. Uji ini mengukur kekuatan komponen linier dari model. Misalkan hubungan

- antara X dan Y adalah lingkaran sempurna, meskipun jika ada hubungan yang sempurna antar variabel, nilai R^2 akan menjadi nol.
5. Prediktabilitas. R^2 besar tidak selalu berarti prediktabilitas tinggi, begitu pula jika R^2 rendah tidak berarti prediktabilitas yang buruk.
 6. No-intercept model. Definisi dari R^2 diasumsikan bahwa ada intersep dalam model regresi. Ketika intersep tidak digunakan dalam model maka akan mengubah definisi R^2 .
 7. Ukuran sampel. R^2 sangat sensitif terhadap jumlah observasi. Semakin kecil ukuran sampel, maka semakin besar nilainya

11.8 Asumsi Regresi Linier

Analisis Regresi Linier merupakan analisis yang sering digunakan untuk mengukur pengaruh antar dua variabel atau lebih (Wahyono, 2010), namun dalam penggunaannya ada beberapa asumsi yang harus diperhatikan oleh peneliti sebelum menggunakan analisis ini untuk penelitiannya.

Asumsi berikut yang harus dipertimbangkan ketika menggunakan analisis regresi linier.

1. Linieritas

Model regresi linier berhubungan dengan garis lurus Antara Y dan X . Hubungan lengkung apapun diabaikan. Asumsi ini paling mudah dievaluasi dengan menggunakan plot pencar. Ini harus dilakukan diawal analisis. Pola nonlinier juga dapat muncul di plot sisa. Kurangnya uji kesesuaian juga disediakan.

2. Varians Konstan

Varians residual diasumsikan konstan untuk semua nilai X . Asumsi ini dapat dideteksi dengan merencanakan residual melawan variabel independen. Jika plot sisa ini menunjukkan bentuk persegi panjang, kita bias mengasumsikan bahwa varians konstan, namun jika plot sisa menunjukkan irisan yang naik turun maka varians diasumsikan tidak konstan (heteroskedastisitas) dan harus diperbaiki. Tindakan korektif untuk varian tidak konstan adalah dengan menggunakan regresi linier berbobot atau untuk mengubah Y atau X masuk sedemikian rupa

sehingga varians lebih mendekati konstan. Transformasi menstabilkan varian dikenal dengan logaritma Y.

3. Penyebab khusus

Penyebab khusus menyebabkan varian tidak konstan, non normalitas atau masalah lain dengan model regresi.

4. Normalitas

Normalitas terjadi saat uji hipotesis dan batas kepercayaan akan digunakan, residunya diasumsikan mengikuti normal distribusi.

5. Independence

Residu diasumsikan tidak berkorelasi satu sama lain atau berarti bahwa Y juga tidak berkorelasi. Asumsi ini dapat dilanggar dengan dua cara, yakni Kesalahan spesifikasi model atau data urutan waktu. Berikut penjabarannya :

a. Kesalahan model spesifikasi

Jika variable independen penting dihilangkan atau jika bentuk fungsional salah digunakan, residu mungkin tidak independen. Solusi untuk masalah ini adalah menemukan bentuk fungsional yang tepat atau memasukkan variable independen yang tepat dan menggunakan regresi ganda.

b. Data urutan waktu

Setiap kali analisis regresi dilakukan pada data yang diambil dari waktu ke waktu, residual mungkin akan berkorelasi. Korelasi antar residual ini disebut korelasi serial. Korelasi serial positif artinya residual pada periode waktu j cenderung memiliki tanda yang sama dengan residual periode waktu (j-k), di mana k adalah jeda dalam periode waktu. Korelasi serial negative berarti residual dalam periode waktu j cenderung bertanda berlawanan dengan sisa periode waktu (j-k).

Bab 12

Analisis Multivariat

12.1 Pendahuluan

Tahapan analisis data dalam suatu penelitian berperan penting dalam menggali informasi dari penelitian yang telah dilaksanakan. Penelitian yang menggunakan banyak variabel sering menggunakan teknik analisis multivariat (Minto, 2007).

Ada beberapa istilah yang sering digunakan dalam statistika diantaranya (Priatna, 2007):

1. Analisis univariat ialah analisis yang hanya memakai 1 (satu) variabel
2. Analisis bivariat adalah analisis yang memakai 2 (dua) variabel
3. Elemen adalah sesuatu yang menjadi objek penelitian
4. Variabel merupakan objek yang berubah-ubah nilainya berdasarkan waktu atau tempat
5. Sampel merupakan sebagian dari populasi
6. Sampling merupakan cara pengumpulan data yang hasilnya merupakan perkiraan atau estimasi

7. Sampling eror merupakan kesalahan yang ada pada data riset, diakibatkan oleh riset dilaksanakan berdasarkan pengambilan sampel. Sampling eror bermanfaat mengukur seberapa teliti data hasil riset.

Analisis multivariat adalah pengembangan dari analisis univariat dan analisis bivariat. Dari kata multi yang berarti banyak serta variate yang berarti variabel, jadi analisis multivariat merupakan analisis pada banyak variabel. Analisis multivariat bisa diilustrasikan terdapatnya suatu permasalahan ataupun gap yang diakibatkan oleh tidak terdapatnya kesamaan antara harapan dan realitas. Setiap permasalahan tentu terdapat faktor-faktor yang menjadi penyebabnya, umumnya penyebabnya lebih dari satu. Bila permasalahan disebut variabel dependen (Y) serta faktor penyebab disebut variabel independen (X) sehingga masalah (Y) adalah fungsi dari $X_1, X_2, X_3, \dots, X_n$. Fenomena ini dikenal sebagai fenomena multivariat. Dapat disimpulkan, analisis multivariat adalah metode statistik yang bertujuan menganalisis variabel yang banyak secara bersamaan (Pranoto, 2018).

Data multivariat merupakan data yang terkumpul dari dua ataupun lebih observasi dengan mengukur observasi yang ada dengan banyak karakteristik. Analisis data multivariat terbagi menjadi tiga kategori, yaitu teknik dependensi, interdependensi, dan teknik persamaan struktural. Teknik dependensi mempunyai dua tipe variabel yaitu variabel independen dan dependen, terdiri dari berbagai macam jenis analisis yakni regresi berganda, korelasi kanonikal, analisis diskriminan, dan manova. Sedangkan teknik interdependensi mempunyai satu tipe variabel yaitu variabel dependen, yang terdiri dari analisis kluster, analisis faktor, dan multidimensional scaling (Wijaya and Budiman, 2016).

Teknik analisis multivariat menganalisis 2 (dua) atau lebih variabel independen dengan 1 (satu) variabel dependen pada saat yang sama. Banyaknya sampel pada analisis multivariat merupakan hal yang penting untuk diberikan perhatian. Aturan yang berlaku adalah dibutuhkan minimal 10 responden pada setiap variabel, sehingga jika pada penelitian terdiri dari 10 variabel, maka banyaknya sampel yang dibutuhkan adalah minimal 10 variabel dikalikan 10 responden = 100 responden.

Dari analisis multivariat dapat diketahui tentang (Hastono, 2006):

1. Variabel independen memiliki pengaruh terbesar kepada variabel dependen

2. Diketahui tentang adakah variabel lain yang memengaruhi hubungan variabel independen terhadap variabel dependen atau tidak
3. Jenis hubungan variabel-variabel independen dengan variabel dependen diketahui memiliki hubungan atau pengaruh langsung atau tidak langsung

12.2 Konfounding dan Interaksi

Sebelum menjalankan analisis multivariat sebaiknya kita perlu mengetahui lebih dulu tentang konsep konfounding dan juga konsep interaksi. Seorang peneliti harus menjaga dengan baik seluruh rangkaian penelitian agar pengaruh atau hubungan yang akan dibuktikan benar adanya, atau dengan kata lain variabel dependen memang dipengaruhi oleh variabel independen, bukan dipengaruhi variabel lain. Konfounding dan interaksi secara konseptual adalah berbeda, tetapi konfounding dan interaksi sama-sama melibatkan adanya kaitan antara dua variabel ataupun lebih, yang kemudian semua variabel tambahan yang bisa mengganggu hubungan tersebut bisa diperhitungkan, variabel yang diperhitungkan tersebut dikenal dengan sebutan kovariat (Mawarni, 2008).

12.2.1 Konfounding

Konfounding berasal dari bahasa latin yang memiliki arti “menuang bersama”. Sehingga bisa didefinisikan sebagai bias atau kesalahan pada estimasi faktor risiko/ pajanan yang menyebabkan terjadinya penyakit yang akan diteliti yang disebabkan oleh tidak sebandingnya antara populasi tidak terpajan serta populasi terpajan (Putra and Sutarga, 2018).

Konfounding ataupun yang dikenal dengan variabel pengganggu atau perancu merupakan variabel yang merencanakan hubungan antara variabel pajanan (exposure) dan variabel terpajan (outcome). Variabel ini dikatakan pengganggu karena kehadirannya selaku penyebab terbentuknya exposure, tetapi disaat yang sama, kehadirannya juga selaku penyebab terbentuknya outcome (Putri, 2020).

Konfounding adalah keadaan kesalahan pada saat estimasi dampak pajanan yang menyebabkan terjadinya masalah kesehatan/ penyakit, yang disebabkan

oleh tidak balance/ seimbang nya perbandingan antara kelompok non expose dan kelompok expose. Permasalahan tersebut disebabkan pada dasarnya telah terdapat ketidaksamaan risiko kejadian penyakit di kelompok non expose dan kelompok expose. Maksudnya risiko terjadinya penyakit di kedua kelompok tidak sama walaupun pun pada masing-masing kelompok expose telah dihilangkan (Hastono, 2006).

Variabel independen pada sebuah riset umumnya memengaruhi variabel dependen. Sebagai contoh, bila suatu riset ingin mengetahui “apakah kurang berolahraga menimbulkan peningkatan berat badan?”, sehingga peningkatan berat badan merupakan variabel dependen sedangkan kurang berolahraga merupakan variabel independen. Variabel perancu merupakan variabel lain yang juga mempunyai pengaruh pada variabel dependen. Variabel perancu semacam variabel independen ekstra yang memiliki pengaruh tersembunyi terhadap variabel dependen. Variabel perancu bisa mengakibatkan dua permasalahan utama diantaranya meningkatkan varian serta memunculkan bias. Sebagai contoh, jika dilakukan penelitian kepada 200 orang (100 laki-laki dan 100 perempuan), diketahui jika kurang berolahraga mengakibatkan peningkatan berat badan. permasalahan pada riset ini adalah minimnya variabel kontrol. Sebagai contoh, pemakaian placebo, ataupun pembagian kelompok secara acak. Jadi peneliti dengan benar tidak dapat menjamin apakah kurang berolahragalah yang mengakibatkan peningkatan berat badan (Statistics How To, 2017).

Sebagian variabel perancu adalah pola konsumsi. Bisa jadi pula laki-laki makannya lebih banyak daripada perempuan; ini pula dapat menjadikan variabel jenis kelamin sebagai variabel perancu. Variabel lain yang dapat jadi perancu adalah variabel pekerjaan, berat badan, dan juga usia. Desain penelitian yang kurang baik semacam ini bisa menimbulkan bias. Misalnya, bila semua perempuan dalam riset ini berumur paruh baya, dan semua laki-laki berumur 16 tahun, umur menjadi berpengaruh langsung pada peningkatan berat badan. Hal tersebut menjadikan umur sebagai variabel perancu (Statistics How To, 2017).

Memperhatikan variabel perancu dibutuhkan pada saat mendesain studi serta pada saat analisis data. Ketidakberhasilan mengenali dan mengendalikan variabel perancu bisa menyebabkan kesalahan pada interpretasi hasil (Andrade, 2007).

Beberapa sifat confounder atau variabel perancu adalah (Putra and Sutarga, 2018):

1. Agar disebut variabel perancu, suatu covariat mesti adalah faktor risiko terjadinya penyakit tersebut pada populasi dasar (unexposed base population) yang tidak mengalami pajanan exposure. Dengan kata lain: perancu mesti adalah faktor risiko terjadinya penyakit tersebut.
2. Covariat berkaitan dengan exposure pada populasi dasar tetapi tidak mesti menjadi faktor risiko dari exposure.
3. Covariat disebut bukan perancu bila kaitannya dengan exposure di populasi dasar berupa exposure \square covariat. Dalam perihal ini covariat adalah variabel intermediate yang menjadi perantara hubungan exposure dan covariat

Variabel konfounding sesungguhnya dapat diabaikan saja, namun pengabaian ini malah akan membuat kualitas riset yang dilaksanakan berkurang (Putri, 2020).

Terdapat beberapa cara untuk menghindari atau mengendalikan variabel *confounding* antara lain ialah (Andrade, 2007):

1. *Retriction* atau disebut dengan “batasan”; seumpama, bila pemakaian terapi penggantian hormon membuat kacau risiko penyakit Alzheimer pada studi hipotetis mengenai kemanjuran vaksin pada kondisi tersebut, sampel pada riset tersebut bisa dibatasi untuk perempuan yang belum memakai terapi penggantian hormon. Pembatasan tersebut menjadikan lebih susah untuk memperoleh sampel dengan jumlah yang mencukupi serta membatasi generalisasi (validitas eksternal) sebuah penelitian.
2. *Matching* atau disebut dengan “mencocokkan”; pada contoh di poin 1 di atas, setiap perempuan yang sudah memakai terapi penggantian hormon bisa disandingkan dengan perempuan yang belum pernah memakai terapi penggantian hormon. Proses matching bisa jadi tidak gampang. Misalnya, umur, riwayat keluarga, jenis kelamin, pendidikan, pemakaian penggantian hormon, diet, pemakaian obat

anti-inflamasi jangka panjang, serta faktor-faktor lain seluruhnya ialah variabel perancu dalam setiap penelitian tentang risiko penyakit Alzheimer dan akan sangat susah untuk mencocokkan subjek satu persatu pada seluruh variabel tersebut. Permasalahan lain dengan matching ialah bahwa, kalau subjek penelitian dicocokkan pada variabel tertentu, pengaruh variabel itu (pada hasil penelitian) tidak bisa diukur secara statistik.

3. *Stratification*; pada contoh di atas, perempuan dengan atau tanpa riwayat terapi penggantian hormon bisa diacak secara independen untuk menerima plasebo atau vaksin atau, risiko penyakit Alzheimer dengan plasebo versus vaksin bisa dipelajari secara independen pada perempuan dengan atau tanpa riwayat terapi penggantian hormon. Namun, stratifikasi dapat menjadi masalah jika terdapat beberapa variabel perancu karena beberapa lapisan stratifikasi akan menjadi penting.
4. Uji statistik; analisis kovarians digunakan digunakan untuk mengendalikan variabel perancu tunggal serta salah satu dari beberapa jenis analisis regresi bisa digunakan untuk secara bersamaan mengendalikan beberapa variabel perancu. Tetapi, terdapat beberapa asumsi yang mesti dipenuhi untuk analisis tersebut serta batasan tertentu pada hasil yang diperoleh.

12.2.2 Interaksi

Suatu variabel bisa disebut sebagai variabel konfounding atau pengganggu, yang kemudian diterapkan analisis kovarians, butuh rasanya dilihat apakah terdapat interaksi pada tiap perlakuan terhadap variabel yang diduga sebagai konfounding. Interaksi adalah suatu keadaan di mana hubungan yang diteliti berbeda pada perbedaan tingkatan variabel atau nilai. Pengamatan terdapatnya interaksi tidak berdasarkan dugaan (Mawarni, 2008).

Dampak modifikasi atau interaksi adalah heterogenitas atau adanya keberagaman dampak dari suatu expose pada tingkat expose yang lain. Sehingga dampak expose pada terjadinya masalah kesehatan/ penyakit tidak sama pada kelompok expose yang lain. Tidak terdapatnya interaksi, artinya dampak expose seragam atau homogen. Modifikasi dampak adalah sesuatu

yang penting pada analisis sebab saat menganalisis kita mesti tentukan apakah akan menyajikan dampak bersama (yang terkendali konfounder) atau dampak yang terpisah untuk setiap strata (Hastono, 2006).

Cara mendeteksi ada atau tidaknya interaksi ataupun faktor pengganggu atau keduanya, mendapatkan nilai perkiraan odds rasio (OR) untuk setiap strata, dan memperkirakan kehomogenan atau keseragaman OR pada setiap strata tidak hanya bisa dilakukan dengan Analisis Mantel-Haenszel yang memiliki tabel kontingensi 2x2 yang bisa digunakan untuk analisis berstrata, namun juga dapat diperoleh dengan memakai model regresi logistik. walaupun pemakaian model regresi logistik dapat memberi cara yang relatif lebih cepat dan efektif untuk mendapatkan nilai perkiraan OR pada setiap strata atau tingkatan, melihat ada atau tidaknya interaksi serta faktor pengganggu, dan menaksir dugaan kehomogenan OR antar setiap strata, tetapi pada kalangan para peneliti lapangan dan statistisi yang memiliki latar belakang pendidikan epidemiologi lebih kearah tetap memakai analisis berstrata selaku pra analisis Mantel-Haenszel, khususnya pada riset ilmu kesehatan (Khrisna, 2017).

Bila didapatkan interaksi antar variabel expose dengan variabel yang lain pada analisis multivariat, maka angka koefisien, seperti nilai OR, mesti ditampilkan secara terpisah berdasarkan strata atau tingkatan dari variabel tersebut. Nilai OR yang tercantum pada variabel berubah jadi tidak berlaku serta nilai OR pada setiap strata mesti dihitung (Hastono, 2006).

Terdapat 2 (dua) macam interaksi yaitu (Departemen Biostatistika FKM UI, 2010):

1. Interaksi Aditif
2. Nilai asosiasi berbentuk risk/ rate difference (AR) digunakan menilai kehadiran Interaksi aditif. Interaksi aditif berguna bagi kepentingan program kesehatan masyarakat maupun intervensi pencegahan terjadinya penyakit
3. Interaksi Multiplikatif
4. Nilai asosiasi berbentuk risk/rate ratio (RR/OR) digunakan menilai Interaksi multiplikatif. Interaksi multiplikatif berarti untuk menerangkan hubungan kausal.

12.3 Teknik Analisis Multivariat

Umumnya permasalahan di bidang kesehatan/ medis sifatnya kompleks dan tidak sering terjadinya satu gangguan kesehatan/ penyakit diakibatkan faktor tunggal. Contoh, faktor perilaku, lingkungan, fisik, dan akses terhadap pelayanan adalah faktor risiko penyakit jantung koroner (PJK). Analisis bivariat tidak dapat digunakan untuk mempelajari hubungan suatu masalah kesehatan dengan berbagai faktor risikonya. Teknik analisis multivariat adalah teknik analisis yang menganalisis hubungan 2 (dua) atau lebih variabel independen dengan 1 (satu) variabel dependen. Misalnya pada riset faktor determinan kematian bayi, faktor risiko PJK, dan lainnya (Widarsa, Putra and Astuti, 2016).

Analisis multivariat merupakan bagian dari statistik inferensial, yang menganalisis data sampel dari populasi. Karakteristik populasi yang diteliti adalah karakteristik sampel yang diperoleh secara acak dari populasi. Statistik inferensial berupaya menyimpulkan fenomena ataupun hipotesis yang diuji pada sebuah penelitian. Adanya kesalahan sampling (sampling error) membuat peneliti yang memakai statistik inferensial mesti menentukan tingkat kesalahan yang dapat diterima. Alpha (α) adalah tingkat kesalahan tipe I yaitu pendekatan yang tersering dipakai untuk menentukan besarnya tingkat kesalahan yang dapat diterima (Pranoto, 2018).

Analisis multivariat sebagai alat yang umum digunakan pada analisis statistika, terdiri dari beberapa jenis. Teknik analisis multivariat dikelompokkan menjadi 3 (tiga) kelompok yaitu: (1) teknik dependen, (2) teknik interdependen, dan (3) teknik persamaan struktural. Teknik dependen adalah ketika variabel independen memengaruhi variabel dependen. Sementara teknik interdependen adalah teknik yang seluruh variabel berpengaruh satu sama lain atau seluruh variabel adalah variabel independen. Sedang teknik persamaan struktural adalah teknik yang variabel independen dan dependen dianalisis dengan simultan (Widarjono, 2010).

12.3.1 Teknik Dependen

Analisis multivariat yang memakai teknik dependen mempunyai tujuan untuk mencari tahu pengaruh atau memprediksi nilai variabel terikat atau variabel dependen berdasarkan pada beberapa variabel independen yang memengaruhi (Priatna, 2007). Terdapat lebih dari satu jenis teknik dependen pada analisis multivariat. Pembagian kelompoknya didasarkan pada 2 (dua) hal yaitu: (1)

tipe pengukuran data terhadap variabel independen maupun variabel dependen dan (2) jumlah variabel dependen. Menurut jumlah variabel dependennya, analisis multivariat dibagi menjadi apakah mempunyai satu, dua atau lebih variabel dependen. Selanjutnya jika telah diketahui jumlah variabel dependen maka pembagian kelompoknya berdasarkan jenis data variabel independen maupun variabel dependennya (Widarjono, 2010).

Apabila jumlah variabel dependennya hanya satu, ada 5 (lima) jenis analisis multivariat yang bisa digunakan di antaranya yaitu (Priatna, 2007):

1. Anova (Analysis of variance)

Memiliki tujuan untuk melihat pengaruh lebih dari satu variabel independen yang memiliki skala nominal/ ordinal (data kategorik) terhadap variabel dependen yang memiliki skala interval/ rasio (data numerik).

2. Ancova (Analysis of covariance)

Memiliki tujuan untuk melihat beda rerata dari variabel dependen sehubungan dengan pengaruh variabel independen terkendali. Kovariat adalah sebutan untuk variabel independen yang metrik (interval dan rasio) sedangkan faktor adalah sebutan untuk variabel independen yang kategorik (nonmetrik: nominal dan ordinal). Ancova ialah analisis yang sangat pas untuk faktor/ variabel independennya berupa data kategorik.

3. Regresi Berganda

Merupakan tata cara yang cocok digunakan untuk permasalahan riser yang menggunakan 1 (satu) variabel dependen (Y) yang memiliki data numerik yang berhubungan dengan lebih dari 1 (satu) variabel independen (X) yang memiliki skala pengukuran interval/ rasio (kuantitatif) maupun nominal/ ordinal (kualitatif). Tujuannya untuk memprediksi nilai Y, bila nilai dari seluruh variabel independennya diketahui. Metode kuadrat terkecil dipakai untuk membuat persamaan regresi linier berganda. Tidak hanya itu, persamaan regresi linier berganda juga digunakan untuk melihat nilai pengaruh dari semua variabel independen yang ada pada persamaan.

4. Analisis Diskriminan

Analisis diskriminan umumnya dipergunakan untuk mengevaluasi klasifikasi objek dan memisahkan objek sesuai dengan klasifikasinya (Suhartono, 2008). Bertujuan untuk mengetahui perbedaan kelompok variabel dependen yang memiliki data berskala nominal/ ordinal dan memprediksikan peluang atau probabilitas bahwa sebuah objek penelitian akan menjadi anggota kelompok tertentu yang berdasar pada lebih dari satu variabel independen yang memiliki data dengan skala interval/ rasio. Analisis diskriminan baik digunakan bila variabel dependen berbentuk kelompok.

5. Analisis Konjoin

Menggambarkan sebuah ukuran kuantitatif tentang kepentingan relatif (relatif importance) berbagai atribut suatu produk.

Apabila jumlah variabel dependennya lebih dari satu, maka ada 2 (dua) jenis analisis multivariat yang bisa digunakan diantaranya yaitu (Widarjono, 2010):

1. Analisis kolonikal

Sebuah metode statistik yang dimanfaatkan untuk mengetahui strata asosiasi linier pada 2 (dua) perangkat variabel, di mana setiap perangkat terdiri dari lebih dari satu variabel. Manfaat utama dari analisis kolonikal adalah untuk melihat hubungan linieritas antara beberapa variabel dependen dengan variabel-variabel independen yang berperan selaku prediktor (Wijaya and Budiman, 2016).

2. Analisis Manova

Tujuan analisis manova pada dasarnya sama dengan Anova, yaitu ingin melihat apakah terdapat beda yang nyata pada semua variabel dependen yang ada diantara grup (variabel independen). Biasanya jumlah sampel untuk analisis Manova yang baik adalah ada 20 kasus untuk setiap grup (Santoso, 2002). Analisis Manova dipakai untuk penelitian yang memiliki variabel independen bersifat kualitatif, sedangkan variabel dependennya bersifat kuantitatif lebih dari 1 (satu). Analisis manova memiliki tujuan untuk mengetahui apakah terdapat beda yang signifikan atau tidak antara rata-rata kelompok (centroid). Analisis manova juga berguna untuk melihat perbedaan

antar kelompok yang ada dipengaruhi oleh variabel independen yang mana saja (Widarjono, 2010).

12.3.2 Teknik Interdependen

Analisis multivariat yang memakai metode interdependen/ saling ketergantungan, digunakan guna mengetahui faktor yang menjadi penyebab munculnya masalah atau membantu mengetahui informasi yang diinginkan. Pada banyak kasus, sering dialami kesusahan dalam memisahkan variabel independen dan dependen. Teknik interdependen dipakai untuk mengetahui struktur data, tanpa membedakan antara variabel independen dan dependen. Tujuan dari analisis teknik interdependen adalah menguji bagaimana dan mengapa variabel-variabel yang diteliti berhubungan satu dengan yang lainnya. Pembagian teknik interdependen berdasarkan pada variabel yang ada yaitu apakah bersifat non metrik/ kualitatif atau metrik/ kuantitatif (Widarjono, 2010).

Apabila jenis variabel ialah metrik maka terdapat 3 (tiga) jenis analisis, yakni (Widarjono, 2010):

1. Analisis Faktor

Analisis faktor adalah salah satu teknik interdependensi pada analisis multivariat yang pada umumnya digunakan untuk mengeksplorasi struktur hubungan yang terjalin dalam suatu kelompok variabel (Suhartono, 2008). Analisis faktor adalah analisis yang digunakan untuk memilih variabel baru atau faktor yang memiliki jumlah lebih sedikit jika dibandingkan dengan banyaknya variabel asli di mana semua faktor tersebut tidak memiliki korelasi satu dengan yang lainnya (multikolinearitas). Variabel baru itu mesti memuat sebanyak mungkin informasi yang terdapat dalam variabel asli. Informasi yang hilang pada proses reduksi banyaknya variabel harus seminimal mungkin (Priatna, 2007). Analisis faktor berguna memperkecil data yang jumlahnya banyak guna menentukan faktor yang bisa menjelaskan varians yang sementara diteliti dengan lebih jelas pada kelompok variabel yang memiliki jumlah lebih besar (Wijaya and Budiman, 2016).

2. Analisis Kluster

Analisis kluster adalah analisis guna menyatukan unsur serupa sebagai objek penelitian menjadi kelompok (cluster) yang berbeda dan saling asing (mutually exclusive). Kelompok (cluster) dibentuk dengan kriteria khusus yang melihat data yang ditunjukkan oleh nilai banyak variabel (Priatna, 2007). Analisis kluster membagi ke dalam kelompok obyek yang ada sehingga setiap obyek memiliki kemiripan dengan yang lain pada sebuah kluster. Hasil kluster suatu obyek mesti mempunyai eksternal (between cluster) heterogenitas yang tinggi dan mempunyai internal (within cluster) homogenitas yang tinggi. Jika pengelompokan sukses, maka obyek pada satu kluster akan saling dekat antara satu dengan yang lain bila diplot secara geometri dan kluster yang berbeda akan saling menjauh satu dengan yang lain (Wijaya and Budiman, 2016).

3. Skala Multidimensi

Memiliki tujuan untuk membuat penilaian pelanggan tentang kesamaan (similarity) atau preferensi (perasaan lebih suka) ke dalam jarak (distances) yang diwakili dalam ruang multidimensional. Bila objek X dan Y mendapat penilaian konsumen selaku pasangan objek termirip jika dibandingkan dengan pasangan lain, objek X dan Y pada metode penskalaan multidimensional akan diposisikan sebaik mungkin sehingga bila dibandingkan dengan jarak pasangan objek yang lainnya, jarak antara objek dalam ruang multidimensional akan lebih pendek/ kecil (Priatna, 2007).

12.3.3 Teknik Persamaan Struktural

Hubungan antar variabel pada sebuah riset kerap kali tidak sesederhana seperti yang dikemukakan pada teknik dependen ataupun teknik interdependen. Saat ini berkembang dengan pesat metode statistika multivariat yang berupaya menganalisis hubungan secara simultan variabel independen dan dependen. Metode ini adalah pemodelan persamaan struktural (Structural Equation Modeling = SEM) (Widarjono, 2010).

SEM adalah pengertian umum yang diberikan pada metode analisis multivariat yang mempunyai karakteristik yaitu:

1. Membuat perkiraan hubungan-hubungan dependen yang berkaitan satu sama lain.
2. Mempunyai kemampuan untuk mempresentasikan konsep-konsep yang tidak teramati secara langsung.

SEM adalah kumpulan alat atau metode-metode statistika yang memungkinkan memperoleh model hubungan dan pengujian suatu rangkaian hubungan yang cukup rumit secara simultan. Hubungan yang rumit tersebut bisa dibandingkan antara 1 (satu) ataupun lebih dari 1 (satu) variabel independen dengan 1 (satu) ataupun lebih dari 1 (satu) variabel dependen. Setiap variabel independen dan dependen bisa berupa faktor (yang terdiri dari lebih dari satu variabel indikator). Variabel-variabel tersebut bisa berupa suatu variabel tunggal yang dinilai atau diukur secara langsung pada suatu proses penelitian. SEM merupakan perpaduan antara analisis regresi berganda dan faktor (Nugroho, 2008).

Metode-metode yang dapat digunakan untuk membuat model persamaan struktural diantaranya adalah: analisis faktor konfirmatif, analisis kovarians, analisis regresi majemuk, dan LISREL. Asumsi dasar mesti dipenuhi dan pastinya disesuaikan dengan metode analisis yang dipakai (Minto, 2007).

12.4 Analisis Multivariat dalam Penelitian Kesehatan

Terdapat 2 (dua) analisis multivariat yang paling sering dipakai dalam riset kesehatan, yakni analisis regresi linier berganda (multiple regression linier) dan analisis regresi logistik. Skala pengukuran variabel dependen pada penelitian menjadi penentu pemilihan jenis analisis regresi yang digunakan. Regresi logistik dipakai jika variabel dependennya adalah kategorik. Sedangkan regresi linier berganda dipakai jika variabel dependennya adalah numerik (Dahlan, 2013).

Analisis regresi linier berganda adalah perkembangan dari regresi linier sederhana (simple regression linier). Pada analisis regresi linier berganda jenis

hubungannya adalah beberapa variabel independen terhadap 1 (satu) variabel dependen. Sebagai contoh jika ingin tahu tentang beberapa faktor yang berhubungan dengan tekanan darah sistolik (variabel Y) analisis dilaksanakan dengan menyertakan kadar kolesterol darah (variabel X1), berat badan (variabel X2), dan kadar glukosa darah (variabel X3) (Yusuf, 2003).

Tidak sama dengan regresi linier sederhana, regresi linier berganda dipakai guna menduga pengaruh beberapa variabel independen yang memiliki skala kuantitatif terhadap satu variabel dependen. Supaya proses pengujian regresi linier berganda bisa berjalan sempurna, peneliti sangat dianjurkan untuk memakai skala interval. Hal ini diperlukan supaya data penelitian yang didapatkan terdistribusi normal. Data yang terdistribusi dengan normal adalah persyaratan untuk melaksanakan pengujian regresi linier berganda (Iman, 2012).

Tata cara analisis regresi linier berganda adalah sebagai berikut (Dahlan, 2013):

1. Menyeleksi variabel yang memiliki nilai $p < 0,25$ pada analisis bivariat untuk disertakan pada analisis multivariat.
2. Melaksanakan analisis multivariat. Analisis multivariat regresi linier berganda terbagi jadi 3 teknik, yaitu enter, forward, dan backward.
3. Melaksanakan interpretasi hasil. Dengan analisis multivariat regresi linier berganda bisa didapatkan beberapa hal yakni: (1) variabel yang memengaruhi variabel dependen didapatkan dari nilai p setiap variabel, (2) semua variabel yang berpengaruh terhadap variabel dependen berurut berdasarkan kekuatan hubungan. Urutan kekuatan hubungannya dilihat dari besar nilai r nya (koefisien korelasi), (3) rumus atau model untuk memperkirakan variabel dependen untuk regresi linier berganda yakni: $y = \text{konstanta} + a_1x_1 + a_2x_2 + \dots + a_nx_n$, diketahui y adalah nilai variabel dependen, x adalah nilai dari variabel independen, dan a adalah nilai koefisien setiap variabel,.
4. Memperhitungkan nilai kualitas dari rumus diketahui melalui nilai diskriminan dengan mengetahui nilai R^2 serta kalibrasinya melalui hasil pengujian Anova. Rumus disebut memiliki diskriminasi baik bila nilai R^2 semakin dekat dengan nilai 1. Rumus disebut memiliki nilai kalibrasi baik jika nilai “ p ” dari pengujian Anova $< 0,05$.

5. Memperhitungkan nilai syarat atau dugaan. Regresi linier memiliki dugaan liniaritas, independensi, homogenitas, normalitas, dan juga eksistensi

HEILGauss merupakan singkatan dari asumsi yang dipakai pada regresi linier ganda, yakni (Sudirman, 2017):

1. Homocedasticity: varian nilai variabel Y sama untuk seluruh nilai X. Homocedasticity bisa didapatkan dengan membuat pembuatan plot residual. Varian homogen pada nilai X apabila titik sebaran tidak memiliki pola tertentu serta tersebar merata pada sekitar garis titik nol residual.
2. Eksistensi (variabel random): pada semua nilai variabel X, variabel Y merupakan variabel random dengan mean dan varian tertentu. Asumsi ini berhubungan dengan metode pengambilan sampel. Asumsi eksistensi diketahui dari nilai variabel residual model. Asumsi terpenuhi apabila residual memperlihatkan adanya nilai mean dan sebaran.
3. Independensi: suatu kondisi di mana setiap nilai Y bebas antara satu dengan yang lain. Sehingga nilai dari setiap individu saling berdiri sendiri. Asumsi independensi terpenuhi apabila nilai Durbin Watson antara -2 hingga +2.
4. Liniaritas: nilai mean variabel Y untuk suatu kombinasi dari X_1, X_2, \dots, X_k berada pada garis linier yang terbentuk dari persamaan regresi. Model berbentuk linier apabila hasil uji Anovanya signifikan.
5. Gauss (normalitas): variabel Y memiliki distribusi normal untuk semua pengamatan variabel X. Dapat diketahui dari Normal P-P Plot residual. Asumsi normalitas terpenuhi apabila data tersebar pada sekitar garis diagonal serta mengikuti arah garis diagonalnya.

Bab 13

Regresi Logistik

13.1 Pendahuluan

Analisis regresi adalah metode analisis yang sangat populer di kalangan peneliti. Menurut Hosmer dan Lemeshow (2000), regresi logistik adalah suatu metode analisis statistika untuk mendeskripsikan hubungan antara variabel terikat yang memiliki dua kategori atau lebih dengan satu atau lebih peubah bebas berskala kategori atau kontinu. Regresi logistik merupakan suatu metode analisis regresi dengan variabel respon merupakan variabel biner atau kategorik, untuk variabel responnya bersifat biner atau dikotomis yang terdiri dari dua kategori yaitu 0 dan 1 (Hosmer, Lemeshow dan Sturdivant, 2013). Analisis regresi logistik dapat digunakan untuk menelaah faktor-faktor yang memengaruhi Indeks Prestasi Mahasiswa (IPK) (Tampil, Komaliq, & Langi, 2017).

Regresi logistik merupakan jenis regresi yang mempunyai ciri khusus, yaitu variabel dependennya berbentuk variabel kategorik (terutama yang dikotomis, artinya yang terdiri dari dua kelompok, misalnya sehat/sakit, baik/kurang baik). Perbedaan antara regresi linier dengan regresi logistik terletak pada jenis variabel dependennya. Regresi linier digunakan apabila variabel dependennya numerik, sedangkan regresi logistik digunakan pada data yang dependennya berbentuk kategorik dikotom. Tujuan dari analisis regresi logistik adalah untuk mendapatkan model yang paling baik (fit) dan sederhana (parsimony) yang

dapat menggambarkan hubungan variabel independen dengan variabel dependen. Regresi logistik berfungsi apabila ingin mengestimasi suatu probabilitas kejadian pada dependen yang dikotom (Riyanto, 2012). Keunggulan analisis regresi adalah kemampuannya dalam meramalkan atau memprediksi nilai variabel terikatnya (Yudiatmaja, 2013).

Regresi logistik terbagi menjadi dua yaitu regresi logistik sederhana dan regresi logistik ganda. Regresi logistik sederhana digunakan bila ingin mempelajari hubungan antara satu variabel independen dengan satu variabel dependen yang bersifat dikotomus. Regresi logistik ganda digunakan bila ingin mempelajari hubungan antara beberapa variabel independen dengan satu variabel dependen yang bersifat dikotomus (Riyanto, 2012).

Sedangkan menurut Sepang, Komalig, Hatidja (2012), regresi logistik dapat dibagi menjadi regresi logistik biner, regresi logistik multinomial dan regresi logistik ordinal. Model regresi logistik biner digunakan untuk menganalisis hubungan antara satu variabel respon dan beberapa variabel prediktor, dengan variabel responnya berupa data kualitatif dikotomi yaitu bernilai 1 untuk menyatakan keberadaan sebuah karakteristik dan bernilai 0 untuk menyatakan ketidakberadaan sebuah karakteristik

Regresi Logistik (Logit) merupakan suatu metode analisis statistika yang mendeskripsikan hubungan antara peubah respon (dependent variable) yang bersifat kualitatif memiliki dua kategori atau lebih dengan satu atau lebih peubah penjelas (independent variable) berskala kategori atau interval. Faktor kunci untuk mendapatkan hasil duga regresi logistik yang baik, adalah besaran jumlah responden yang representatif dengan keragaman relatif tinggi. Oleh karena itu validasi data menjadi faktor penentu dan krusial dilakukan sebelum analisis data (Hendayana, 2015).

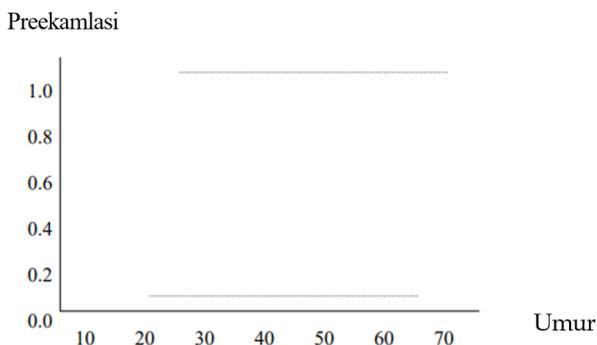
13.2 Analisis Regresi Logistik

Pada regresi logistik, variabel independennya boleh campuran antara variabel kategorik dan numerik. Namun sebaiknya variabel independennya berupa kategorik karena dalam menginterpretasikan hasil analisis akan lebih mudah. Perbedaan antara regresi linier dengan regresi logistik terletak pada jenis variabel dependennya. Regresi linier digunakan apabila variabel dependennya numerik sedangkan regresi logistik digunakan pada data yang dependennya

berbentuk kategori yang dikotom. Untuk memahami lebih jelas tentang regresi logistik coba kita lihat contoh analisis penelitian yang mempelajari hubungan antara variabel umur dengan kejadian preeklamsi pada ibu hamil. Pengamatan dilakukan pada 100 orang sampel didapatkan hasil sebagai berikut :

Nomor	: 1 2 3 4 5 6 7 8 9 10 11	100
Umur	: 20 22 23 24 25 27 28 29 30 32 33	70
Preeklamsi	: 0 0 1 0 0 1 0 1 1 0 0	1

Nomor merupakan nomor urut responden dan preeklamsi merupakan variabel kejadian preeklamsi pada ibu hamil. Variabel preeklamsi diberi kode 1 bila responden menderita preeklamsi dan diberi kode 0 bila mereka tidak menderita preeklamsi.



Gambar 13.1: Grafik hubungan umur dengan kejadian preeklamsi

Bila data tersebut kita perlakukan analisisnya dengan menggunakan regresi linier misalnya dibuat penyajian dalam bentuk diagram tebar (scatter plot) seperti terlihat pada gambar 13.1, maka pola hubungannya tidak jelas terlihat. Tebaran data pada scatter plot membentuk dua garis yang sejajar. Diagram tebar menunjukkan adanya kecenderungan kejadian penyakit preeklamsi yang lebih sedikit pada responden yang berusia muda. Walaupun grafik tersebut telah dapat menggambarkan/ menjelaskan variabel dependen (kejadian preeklamsi) yang cukup jelas, namun grafik tersebut tidak mampu menggambarkan dengan lebih tajam/ jelas hubungan antara umur dengan kejadian preeklamsi.

Untuk mempertajam analisis kita, sekarang dicoba untuk mengelompokan variabel independen (variabel umur) dan menghitung nilai tengah (dalam hal

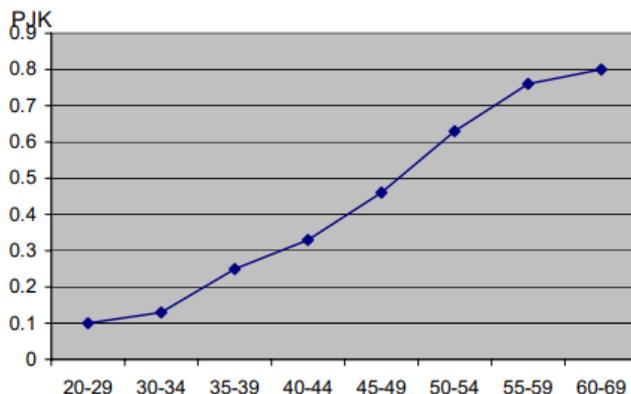
ini menghitung proporsi) variabel dependen (variable preeklamsi) untuk setiap kelompok variabel umur. Hasil pengelompokan variabel umur dan kejadian preeklamsi dapat dilihat pada tabel berikut :

Tabel 13.1 Proporsi kejadian preeklamsi dengan umur

Umur	Jumlah	Preeklamsi		Proporsi kejadian
		E Preeklamsi	Preeklamsi	Preeklamsi
20-29	10	9	1	0.10
30-34	15	13	2	0.13
35-39	12	9	3	0.25
40-44	15	10	5	0.33
45-49	13	7	6	0.46
50-54	8	3	5	0.63
55-59	17	4	13	0.76
60-69	10	2	8	0.80
Total	100	57	43	0.43

Pada tabel terlihat bahwa ada peningkatan proporsi kejadian preeklamsi pada kelompok umur yang semakin tua/ lanjut. Kemudian kita coba sajikan data tersebut dengan grafik dan hasilnya dapat dilihat pada gambar.2 berikut :

Preeklamsi



Gambar 13.2: Peningkatan proporsi kejadian preeklamsi pada kelompok umur

Pada grafik terlihat jelas tentang adanya peningkatan yang tidak linear antara proporsi kejadian preeklamsi dengan peningkatan umur. Diawali peningkatan yang landai, kemudian meningkat tajam dan kemudian landai kembali, garis tersebut menyerupai bentuk huruf S. Kalau kita cermati, pembuatan diagram terbar merupakan cara untuk mendeteksi/ mengetahui hubungan pada analisis regresi linier namun ada sedikit perbedaan dalam hal meringkas variabel dependennya. Seperti kita ketahui bahwa pada regresi linier kita ingin mengestimasi nilai mean variabel dependen berdasarkan setiap nilai variabel independen. Nilai tersebut disebut mean kondisional yang dinyatakan dengan $E(Y/x)$, dengan Y sebagai dependen dan x sebagai independen. $E(Y/x)$ adalah nilai Y yang diharapkan berdasarkan nilai x . Misal Y variabel tekanan darah dan x variabel umur, maka untuk mengetahui estimasi tekanan darah berdasarkan umur, dihitung rata-rata (mean) tekanan darah pada masing-masing nilai umur. Pada regresi linier nilai $E(Y, x)$ akan berkisar antara 0 s/d ∞ ($0 \leq E(Y/x) \leq \infty$).

Pada regresi logistik dapat juga diberlakukan hal tersebut namun ada sedikit perbedaan dalam menghitung rata-rata variabel dependennya (Y). Oleh karena pada regresi logistik variabel dependennya adalah dikotom maka variabel dependen dihitung bukan dengan mean namun dengan menggunakan proporsi. Seperti pada data diatas variabel Y kejadian preeklamsi dan x variabel umur, dihitung mengetahui estimasi kejadian preeklamsi berdasarkan umur, dihitung nilai proporsi kejadian preeklamsi pada tiap kelompok umur. Pada regresi

logistik, nilai $E(Y|x)$ akan selalu berada antara nol dan satu ($0 \leq E(Y|x) \leq 1$) (Jasaputra dan Santosa, 2008)

Kelebihan regresi logistik yaitu:

1. Cocok digunakan pada bidang kesehatan khususnya bidang epidemiologi.
2. Dapat memprediksi risiko terhadap suatu penyakit akibat dari suatu paparan pada desain kohort
3. Menentukan asumsi-asumsi yang lebih sedikit daripada regresi linier
4. Dapat digunakan sebagai model prediksi dan faktor risiko terhadap suatu penyakit
5. Dapat mengetahui pengaruh dari beberapa variabel independen terhadap variabel dependen.

Kekurangan regresi logistik yaitu:

1. Ciri data awal (numerik) hilang karena dijadikan kategorik
2. Prediksi yang didapat tidak seakurat bila dibandingkan regresi linier
3. Pada desain cross sectional dan case control tidak dapat memprediksi faktor risiko yang akan terjadi (Riyanto, 2012)

13.3 Regresi Logistik Sederhana

Model regresi logistik sederhana adalah model regresi logistic dengan 1 prediktor variabel kontinu atau variabel indikator (Harlan, 2018).

Berikut ini adalah asumsi yang digunakan dalam regresi logistik menurut Garson, 2008 dalam Albana dan Alam (2013):

1. Regresi logistik tidak mengasumsikan suatu hubungan yang linear antara variabel respon dengan variabel prediktornya tetapi mengasumsikan hubungan yang linear antara log odds dari variabel responnya dengan variabel prediktornya.

2. Variabel responnya tidak harus berdistribusi normal (tetapi diasumsikan distribusinya berada dalam keluarga distribusi eksponensial, seperti normal, poisson, binomial, gamma).
3. Variabel responnya tidak harus homoskedastis untuk setiap kategori dari variabel prediktornya yaitu tidak ada homogenitas asumsi variansi (variansi tidak harus sama dalam kategori).
4. Galatnya tidak diasumsikan berdistribusi normal.
5. Regresi logistik tidak mengharuskan bahwa semua variabel prediktornya merupakan data interval.
6. Penambahan atau pengurangan alternatif variabel tidak memengaruhi odds yang diasosiasikan.
7. Tidak adanya multikolinearitas.
8. Tidak ada outlier seperti dalam regresi linear.
9. Galat diasumsikan saling bebas.
10. Galat yang rendah dalam variabel bebasnya.
11. Pengkodean berarti (meaningful coding). Koefisien-koefisien logistik akan sulit diinterpretasikan jika kodenya tidak berarti.

13.4 Regresi Logistik Ganda

Model regresi logistik ganda (multiple logistic regression) adalah model regresi logistik dengan lebih daripada 1 prediktor (Harlan, 2018).

Fungsi regresi logistik ganda menurut Riyanto (2012) adalah:

1. Menentukan model matematik yang paling baik untuk menggambarkan hubungan variabel independen dengan variabel dependen
2. Menggambarkan hubungan kuantitatif antara variabel independen (x) dengan variabel dependen (y) setelah dikontrol variabel lain
3. Mengetahui variabel independent (x) mana yang penting (dominan) dalam memprediksi variabel dependen
4. Mengetahui adanya interaksi pada dua atau lebih variabel independen terhadap variabel dependen

13.5 Aplikasi Regresi Logistik

Contoh suatu studi follow-up selama 9 tahun. Dalam studi ini dipelajari mengenai hubungan antara kejadian penyakit jantung koroner (dengan nama variabel PJK) dengan tinggi rendahnya kadar katekolamin dalam darah (nama variabel KAT). Pemberian kode nilai variabel adalah sebagai berikut :

Untuk variabel PJK 1 timbul penyakit jantung koroner

0 tidak ada penyakit jantung koroner

untuk variabel KAT 1 = kadar katekolamin darah tinggi

0 = kadar katekolamin darah rendah

Pertanyaan :

1. Berapa peluang mereka yang kadar katekolaminnya tinggi mempunyai risiko untuk terjadi PJK?
2. Berapa peluang mereka yang kadar katekolaminnya rendah mempunyai risiko untuk terjadi PJK?
3. Bandingkan risiko terjadi jantung koroner antara mereka yang kadar katekolaminnya tinggi dengan yang kadar katekolaminnya rendah.

Jawab :

Dengan model regresi logistik maka pada soal tersebut bentuk modelnya adalah :

$$F(Z) = \frac{1}{1 + e^{-z}}$$

Nilai $f(z)$ dapat diganti dengan $P(X)$, maka rumusnya

$$P(X) = \frac{1}{1 + e^{-z}}$$

Bila $Z = \alpha + \beta_1 KAT$, maka modelnya :

$$P(X) = \frac{1}{1 + e^{-(\alpha + \beta_1 KAT)}}$$

Misalkan didapatkan hasil analisis didapatkan nilai :

$$\alpha = -3,911 \quad \beta = 0,652$$

maka :

$$P(X) = \frac{1}{1 + e^{-(-3,911 + 0,652 \cdot KAT)}}$$

Dari model tersebut coba kita jawab pertanyaan diatas :

- a. Besar risiko terjadinya PJK pada mereka yang kadar katekolaminnya tinggi. Oleh karena kadar katekolaminnya tinggi diberi angka 1, maka masukan nilai KAT=1 pada model diatas. Hasilnya :

$$P(X) = \frac{1}{1 + e^{-(-3,911 + 0,652 \cdot 1)}} = 0,037 \text{ atau sekitar } 4\%$$

Jadi mereka dengan kadar katekolamin tinggi dalam darah mempunyai risiko untuk terjadinya PJK sebesar 4 %.

- b. Oleh karena kadar katekolamin rendah diberi angka 0, maka masukan nilai KAT=0 pada model diatas. Hasilnya :

$$P(X) = \frac{1}{1 + e^{-(-3,911 + 0,652 \cdot 0)}} = 0,019 \text{ atau sekitar } 2\%$$

Jadi mereka dengan kadar katekolamin rendah dalam darah mempunyai risiko untuk terjadinya PJK sebesar 2%.

- c. Besar resiko kedua kelompok tersebut adalah :

$$\frac{P1(X)}{P0(X)} = \frac{0,037}{0,019} = 1,947 = 2,0$$

Angka tersebut diatas sebenarnya adalah risiko relatif (RR) yang diperoleh secara direk. Arti dari angka diatas adalah, mereka yang kadar katekolaminnya tinggi mempunyai risiko terjadi PJK dua (2) kali lebih tinggi dibandingkan mereka yang pada katekolaminnya rendah. Model regresi logistik digunakan pada data yang dikumpulkan melalui rancangan terakhir, parameternya dicari melalui estimasi rasio odds (OR) yang merupakan perhitungan eksponensial β dari persamaan garis regresi logistik. Jadi nilai OR dapat dihitung dari nilai risk (RR) dengan cara indirek. Rasio Odds : $e\beta$ (Jasaputra dan Santosa, 2008)

Daftar Pustaka

- Albana, M., & Alam, F. M. D. P. (2013). Aplikasi Regresi Logistik Ordinal untuk Menganalisa Tingkat Kepuasan Pengguna Jasa Terhadap Pelayanan di Stasiun Jakarta Kota. Universitas Pakuan, Bogor.
- Alkaf, Halid Nuraida. (2009). Metodologi Penelitian Penelitian. Ciputat: Islamic Research publishing.
- Andrade, C. (2007) 'Confounding', *Indian journal of psychiatry*, 49(2), pp. 129–131. doi: 10.4103/0019-5545.33263.
- Anwar, R. (2005). PROSEDUR PEMILIHAN UJI HIPOTESIS. http://pustaka.unpad.ac.id/wp-content/uploads/2010/05/prosedur_pemilihan_uji_hipotesis.pdf
- Arifin, M. H. (2014) 'Konsep-konsep Dasar Statistika', Pengantar Statistik Sosial, pp. 1–45. Available at: <http://repository.ut.ac.id/4315/1/ISIP4215-M1.pdf>.
- Arikunto, S. (2006). *Prosedur Penelitian: Suatu Pengantar Praktik*. Jakarta: Rineka Cipta.
- Armitage, P., Berry, G. and Matthews, J. N. S. (2002) *Statistical Methods in Medicals Research*. Fourth Edi. USA: Blackwell Publishing Company.
- Artaya, P. (2019) 'Sign Test [Uji Tanda]', *Qualitative Research Analysis Method*, (January), pp. 4–6. doi: 10.13140/RG.2.2.33051.92968.
- Astutik, D. (2011) *Pengujian Hipotesis Dua Sampel Independen Berdasarkan Uji Mann-Whitney dan Uji Kolmogorov Smirnov Dua Sampel Serta Simulasinya Dengan Program SPSS*. Available at: <https://lib.unnes.ac.id/6214/>.
- Bajwa, Z. J. (2017) *EDUCATIONAL STATISTICS*. 1 st. Islamabad: Allama Iqbal Open University.

- Banerjee, A., Jadhav, S. L., & Bhawalkar, J. S. (2009). Probability, clinical decision making and hypothesis testing (pp. 64–69). <https://doi.org/10.4103/0972-6748.57864>
- Benson McClave and Sincich. (2011). *Statistik Untuk Bisnis Dan Ekonomi*. 12th Edn. Erlangga.
- Bridges, R. T., Weiss, N. P. and Hassett, M. J. (1992) *Introductory Statistics, The Mathematical Gazette*. doi: 10.2307/3619161.
- Bruce, P. and Bruce, A. (2017) *Practical Statistics for Data Scientists*. Sebastopol: O'Reilly Media, Inc.
- Chan, B. K. (2015) *Biostatistics for epidemiology and public health using R*. Springer Publishing Company.
- Dahlan, M. S. (2008) 'Statistik untuk kedokteran dan kesehatan by M. Sopiudin Dahlan (z-lib.org).pdf'.
- Dahlan, M. S. (2013) *Statistik Untuk Kedokteran dan Kesehatan*. 5th edn. Jakarta: Penerbit Salemba Medika.
- Dahlan, S. (2008) *Statistik untuk Kedokteran dan Kesehatan*. Jakarta: Salemba Medika.
- Departemen Biostatistika FKM UI (2010) *Konfounding dan Interaksi*. Available at: <https://staff.blog.ui.ac.id/besral/files/2010/02/3-confounding-interaksi.pdf> (Accessed: 10 March 2021).
- Douglas Lind, Marchal, William G. & Wathen, S. A. (2013). *Teknik-Teknik Statistik Dalam Bisnis Dan Ekonomi*. Edisi 12. Jakarta: Salemba Empat.
- Evans James R. (2007). *Statistics, Data Analysis, & Decision Modelling*. Third Edit. New Jersey: Prentice Hall.
- Furchan, A. (2004.) *Pengantar Penelitian dalam Pendidikan*. Yogyakarta: Pustaka Pelajar.
- Ghazali, I. (2005) 'Aplikasi Analisis Multivariate dengan Program SPSS: Edisi Ketiga', Semarang: BP. Universitas Diponegoro.
- Hadi, Amirul. (1998). *Metodologi Penelitian Pendidikan II*. Bandung: Pustaka Setia.
- Harlan, J. (2018). *Analisis Regresi Logistik*. Depok: Penerbit Gunadarma

- Haryono. (1998). *Metode penelitian pendidikan II*. Bandung: PT. Remaja Rosdakarya.
- Hasan, I. (2004) 'Analisis data penelitian dengan statistik'. Jakarta: Bumi Aksara.
- Hastono, S. P. (2006) *Analisis Multivariat*. Departemen Biostatistik FKM UI. Available at: <https://pdfcoffee.com/sutanto-analisis-multivariat-pdf-free.html>.
- Hastono, S. P. (2016) *Analisa Data Bidang Kesehatan*. Depok: Rajagrafindo Persada.
- Hastono, S. P. (2017) *Analisis data pada bidang kesehatan*. Cetakan Ke. Depok: Rajawali Press.
- Hastono, Sutanto Priyo. (2011). *Analisis Data Kesehatan*. Fakultas Kesehatan Masyarakat Universitas Indonesia.
- Hastono, Sutanto Priyo. (2013). *Statistik Kesehatan*. Jakarta: Rajawali Pers
- Hendayana, R. (2015). Penerapan metode regresi logistik dalam menganalisis adopsi teknologi pertanian. *Informatika Pertanian*, 22(1), 1-9.
- Heryana, A. (2020) 'Uji McNemar Dan Uji Wilcoxon (Uji Hipotesa Non-Parametrik Dua Sampel' , (May). doi: 10.13140/RG.2.2.17682.48325.
- Hosmer, D.W., dan S. Lemeshow. (2000). *Applied Logistic Regression*. Edisi ke-2. John Wiley and Sons Inc, Canada.
- Hosmer, D.W., S. Lemeshow dan R. X. Sturdivant. (2013). *Applied Logistic Regression*. Edisi ke-3. John Wiley and Sons Inc, Canada.
- Hulley, S. B. et al. (2013) *Designing clinical research*. Fourth Edi. Lippincott Williams & Wilkins.
- Iman, M. (2012) *Pemanfaatan SPSS dalam Penelitian Bidang Kesehatan dan Umum*. Citapustaka Media Perintis: Hoboken, NJ.
- Jackson, S. L. (2015) *Research methods and statistics: A critical thinking approach*. Cengage Learning.
- Jasaputra, D.K., dan S. Santosa. (2008). *Metodologi Penelitian Biomedis Edisi 2*. Bandung: Danasamartha Sejahtera Utam (DSU)

- Jusmiana, A. and Herianto (2020) *Suplemen Materi Statistik Terapan dalam Ilmu Kesehatan*. Makassar: UPRI. doi: 10.31219/osf.io/cr3ug.
- Kariadinata, R. and Abdurrahman, M. (2012) 'Dasar-Dasar Statistik Pendidikan', p. 13.
- Kariadinata, R. and Abdurrahman, M. (2012). *Dasar-Dasar Statistik Pendidikan*. Bandung: Pustaka Setia.
- Khrisna (2017) 'Uji Mantel Haenszel'. Available at: <http://datariset.com/artikel/detail/uji-mantel-haenszel>.
- Lies Maria Hamzah, Imam Awaludin, E. M. (2016). *Pengantar Statistika Ekonomi*. Bandar Lampung: Anugrah Utama Raharja (AURA).
- Liese, F. and Miescke, K. J. (2008) *Springer Series in Statistics Springer Series in Statistics*. Springer. doi: 10.1007/978-0-387-73194-0.
- Lotka, A. J. and Odell, C. W. (1926) 'Educational Statistics.', *Journal of the American Statistical Association*, 21(155), p. 374. doi: 10.2307/2277079.
- Maiti and Bidinger (1981) 'Statistik Deskriptif', *Journal of Chemical Information and Modeling*, 53(9), pp. 1689–1699.
- Mamik. (2015) *Metodologi Kualitatif*. Taman Sidoarjo: Zifatama Publisher.
- Manglapy, Y. (2014). *Pengujian Hipotesis*. http://eprints.dinus.ac.id/6260/1/TM_4.pdf
- Margono, S. (2004). *Metodologi Penelitian Pendidikan*. Jakarta: PT. Rineka Cipta.
- Mawarni, A. (2008) 'Pengendalian Variabel Pengganggu / Confounding dengan Analisis Kovarians', in *Analisis Kovarians*, pp. 1–18. Available at: http://eprints.undip.ac.id/762/1/ANALISIS_COVARIANS.pdf.
- Minto, W. (2007) 'Teknik Analisis Data Multivariat dengan Structural, Equation Modelling', *Jurnal Teknik Industri FTI UPNV Jatim*, 2(2), pp. 124–139.
- Montgomery, D. C. (2012) *Statistical quality control*. Wiley Global Education.
- Nazir. (2005). *Metode Penelitian*. Jakarta: Ghalia Indonesia.
- Nugroho, S. (2008) *Buku Referensi 'Statistika Multivariat Terapan'*. 1st edn. Bengkulu: UNIB Press.

- Paiva, A. (2010). What is hypothesis testing ?
http://www.sci.utah.edu/~arpaiva/classes/UT_ece3530/hypothesis_testing.pdf
- Peacock, J. and Peacock, P. (2010) Oxford Handbook of Medical Statistics, Oxford Handbook of Medical Statistics. doi: 10.1093/med/9780199551286.001.0001.
- Plichta, Stacey B; Garzon, L. S. (2009) Statistics for Nursing and Allied Health. United States of America: Wolters Kluwer Health | Lippincott Williams & Wilkins.
- Polit, D. F. and Beck, C. T. (2018) Essentials of nursing research: Appraising evidence for nursing practice. Ninth Edit. Lippincott Williams & Wilkins.
- Pranoto, B. (2018) Metode SEM untuk Penelitian Manajemen dengan AMOS 22.00, LISREL 8.80 dan SMART-PLS 3.0. Available at: <https://docplayer.info/52351203-Analisis-multivariat-bab-ii.html>.
- Prasetyo, S. and Ariawan, I. (2008) Biostatistik dasar untuk rumah sakit, Fkm Ui. Depok: FKM UI.
- Priatna, B. A. (2007) 'Teknik-Teknik Analisis Multivariat Terkini Yang Sering Digunakan Dalam Penelitian', Jurusan Pendidikan Matematika FPMIPA UPI, pp. 1–5.
- Purwanto. S. K Suharyadi. (2010). Statistika Untuk Ekonomi Dan Keuangan Modern. Edisi 2. Jakarta: Salemba Empat.
- Putra, I. W. G. A. E. and Sutarga, I. M. (2018) 'Tutorial Kesalahan Sitematik (Bias) dan Cara Penanggulangannya', pp. 1–13. Available at: https://simdos.unud.ac.id/uploads/file_pendidikan_1_dir/a0a7a25ece94487a2b953ea4bc9fc332.pdf.
- Putri, B. D. (2020) "'Confounding Variable_s" Haruskah Dibuang Begitu Saja?' Available at: <http://news.unair.ac.id/2020/07/26/confounding-variable-s-haruskah-dibuang-begitu-saja/>.
- Riduwan, A. and Akdon, A. (2007) 'Rumus dan data dalam analisis statistika', Bandung: Alfabeta.
- Riffenburgh, R. H. and Gillen, D. L. (2020) Statistics in Medicine. Fourth Edi. London: Academic Press.

- Riffenburgh, R. H. and Gillen, D. L. (2020) *Statistics in Medicine*. Fourth Edi. Edited by R. E. Teixeira. UK: Stacy Masucci.
- Riyanto, A. (2012). *Penerapan Analisis Multivariat dalam Penelitian Kesehatan*. Yogyakarta: Nuha Medika
- Riyanto, Agus. (2013). *Statistik Inferensial untuk Kesehatan*. Yogyakarta: Nuha Medika
- Rubiyanto, R. (2009) 'Metode Penelitian Pendidikan Surakarta'. PGSD FKIP UMS.
- Rusydi, A. and Fadhli, M. (2018) *STATISTIKA PENDIDIKAN: Teori dan Praktik Dalam Pendidikan*. Edited by S. Saleh. Medan: CV. Widya Puspita.
- Sabri, Luknis; Hastono, S. P. (2018) *Statistik Kesehatan*. Edisi 1, Ce edn. Depok: PT RajaGrafindo Persada.
- Sadik, K. (2015). *Pengujian Hipotesis*. https://www.stat.ipb.ac.id/en/uploads/STK211/STK211_09.pdf
- Santoso, S. (2002) *Buku Latihan SPSS Statistik Multivariat*. Jakarta: PT. Elex Media Komputindo.
- Sarwono, J. (2008) 'Penelitian Pendekatan Kuantitatif dan Kualitatif', Bandung: Lembaga Penelitian, Universitas Komputer Indonesia.
- Sataloff, R. T., Johns, M. M. and Kost, K. M. (2013) 'Pengantar Statistika Ekonomi'.
- Sawyer, S. F. (2009) 'Analysis of Variance: The Fundamental Concepts', *Journal of Manual & Manipulative Therapy*, 17(2), pp. 27E-38E. doi: 10.1179/jmt.2009.17.2.27E.
- Selvamuthu, D. and Das, D. (2018) *Introduction to statistical methods, design of experiments and statistical quality control*, *Introduction to Statistical Methods, Design of Experiments and Statistical Quality Control*. doi: 10.1007/978-981-13-1736-1.
- Sepang, F., H. Komalig, D. Hatidja. (2012). *Penerapan Regresi Logistik untuk Menentukan*
- Setiawan, D. A. (2015) 'Pengantar Teori Probabilitas', (June 2015), p. 215.

- Setiawan, K. (2019) Buku Ajar Metodologi Penelitian (Anova Satu Arah). Bandar Lampung: Universitas Lampung.
- Setyawan, F. E. B. (2017). Pengantar Metodologi Penelitian (Statistika Praktis). Sidoarjo: Zifatama Jawara.
- Shreffler J, H. M. (2020). Hypothesis Testing, P Values, Confidence Intervals, and Significance - StatPearls - NCBI Bookshelf. <https://www.ncbi.nlm.nih.gov/books/NBK557421/>
- Solution, S. (2013). Hypothesis Testing - Statistics Solutions. <https://www.statisticssolutions.com/hypothesis-testing/>
- Sony Faisal Rinaldi, B. M. (2017). "Metodelogi Penelitian dan Statistik". in Bahan Ajar Teknologi Laboratorium Medis (LTM). Jakarta: Kementerian Kesehatan Republik Indonesia.
- Statistics How To (2017) 'Confounding Variable: Simple Definition and Example'. Available at: <http://www.statisticshowto.com/experimental-design/confounding-variable/>.
- Stewart, A. (2018) Basic Statistics and Epidemiology, Basic Statistics and Epidemiology. doi: 10.1201/9781315383286.
- Stommel, M. and Katherine, J. D. (2014) Statistics for Advanced Practice Nurses and Health Professionals, Statistics for Advanced Practice Nurses and Health Professionals. doi: 10.1891/9780826198259.
- Sudarmanto, R. G. (2005) 'Analisis regresi linear ganda dengan SPSS'.
- Sudirman, E. (2017) Analisis Multivariat. Available at: <https://adoc.pub/analisis-multivariat.html> (Accessed: 13 March 2021).
- Sugiyono, P. D. (2006) 'Statistika untuk penelitian', Bandung: CV. Alfabeta.
- Sugiyono. (2001). Statistika untuk Penelitian. Bandung: Alfabeta.
- Suhartono (2008) Analisis Data Statistik dengan R. Lab. Statistik Komputasi ITS Surabaya.
- Sukmadinata, N. S. (2011). Metode Penelitian Pendidikan. Bandung: Remaja Rosadakarya.
- Sunjoyo et al. (2013) Statistika Parametrik & Non Parametrik_Aplikasi SPSS Untuk Smart Riset.pdf. Bandung: Alfabet.

- Supranto J. (2008). *Statistik Teori dan Aplikasi Jilid 1*. Edisi 7. Jakarta: Erlangga.
- Sutanto Priyo Hastono. (2013). *Statistik Kesehatan*. Jakarta: Rajawali Pers.
- Tabachnick, B. and Fidell, L. (2007) *Experimental Designs Using ANOVA*. California.
- Tampil, Y., Komaliq, H., & Langi, Y. (2017). Analisis Regresi Logistik Untuk Menentukan Faktor-Faktor Yang Mempengaruhi Indeks Prestasi Kumulatif (IPK) Mahasiswa FMIPA Universitas Sam Ratulangi Manado. *d'CARTESIAN*, 6(2), 56-62.
- Tanzeh, A. (2009) 'Pengantar metode penelitian'. Yogyakarta: teras.
- Tyastirin, E. and Hidayati, I. (2017) *Statistik parametrik untuk penelitian kesehatan*. I. Edited by E. Pribadi Teguh. Surabaya: Program Studi Arsitektur UIN Sunan Ampel.
- Uyanto, S. S. (2009) 'Pedoman analisis data dengan SPSS', Yogyakarta: Graha Ilmu, 282.
- Wagner III, W. E. (2019) *Using IBM® SPSS® statistics for research methods and social science statistics*. 5th editio. Sage Publications.
- Wahyono, T. (2010) 'Analisis Regresi dengan MS Excel 2007 dan SPSS 17', Jakarta: Elex Media Komputindo. E-Book.
- Widarjono, A. (2010) *Analisis Statistika Multivariat Terapan*. Yogyakarta: Unit Penerbit dan Percetakan Sekolah Tinggi Ilmu Manajemen YKPN.
- Widarsa, T., Putra, I. W. G. A. E. and Astuti, P. A. S. (2016) *Modul Analisis Data Untuk Variabel Outcome Berskala Nominal Dua Kategori (Binary Outcome)*. Program Studi Kesehatan Masyarakat Fakultas Kedokteran Universitas Udayana.
- Wijaya, T. and Budiman, S. (2016) *Analisis Multivariat Untuk Penelitian Manajemen*. Yogyakarta: Penerbit Pohon Cahaya. Available at: 21-2-2018.
- Yudiatmaja, F. (2013). *Analisis Regresi Dengan Menggunakan Aplikasi Komputer Statistik*. Gramedia Pustaka Utama.
- Yusuf Susi Febriani. (2015). *Metodelogi Kesehatan*. Padang: Darmais Press.
- Yusuf, M. A. N. (2003) *Modul Terapan 'Analisis Data Multivariat' Konsep dan Aplikasi Regresi Linier Ganda*. Depok.

Biodata Penulis



I Made Sudarma Adiputra, Lahir di Tabanan Bali pada tanggal 14 November 1983. Menyelesaikan pendidikan Sarjana Keperawatan dan Profesi di Wira Husada Yogyakarta pada tahun 2009, mulai tahun 2009 mengabdikan diri sebagai staf pengajar di STIKES Wira Medika Bali pada Program Studi Ilmu Keperawatan. Pada tahun 2011 diberi kesempatan untuk melanjutkan pendidikan pascasarjana pada Program Studi Magister Ilmu Kesehatan Masyarakat (Konsentrasi Epidemiologi) di Universitas Udayana dan saat ini penulis tercatat sebagai mahasiswa aktif di program studi S3 Ilmu Kedokteran Universitas Udayana konsentrasi ilmu kesehatan masyarakat. Selain aktif menjadi staf pengajar di program studi ilmu keperawatan, mulai tahun 2017 penulis juga mengampu mata kuliah biostatistik dan metodologi penelitian pada Program Studi Rekam Medis dan Informasi Kesehatan, penulis juga aktif dalam bidang penelitian dan pengabdian kepada masyarakat.

Pengalaman organisasi, sejak tahun 2016 menjadi sekretaris Program Studi Profesi Ners sampai tahun 2017, dari tahun 2017-sekarang diberi kesempatan untuk mengelola Program Studi Rekam Medis dan Informasi Kesehatan. Selain aktif dalam organisasi internal kampus, penulis juga aktif pada Asosisasi Perguruan Tinggi Rekam Medis dan Informasi Kesehatan (APTIRMIKI) sebagai ketua kordinator wilayah 8 (Bali-Nusra).



Ns. Deborah Siregar, S.Kep., M.K.M Lahir di Tarutung, 19 Juli 1987, menyelesaikan pendidikan S1 dan Profesi Ners Fakultas Keperawatan Universitas Padjadjaran (2005-2010). Kemudian melanjutkan pendidikan Magister Kesehatan Masyarakat Universitas Indonesia (2014-2016). Saat ini bekerja sebagai dosen di Fakultas Keperawatan Universitas Pelita Harapan sejak tahun 2016 dengan mata kuliah yang diampu adalah Keperawatan Komunitas Keluarga, Health Promotion Behavior, dan Introduction Nursing Process and Theory. Aktif menjadi editor dalam Nursing Current Jurnal

Keperawatan UPH. Buku kolaborasi yang telah dihasilkan bersama penulis lainnya adalah Anatomi dan Fisiologi untuk Mahasiswa Kebidanan (2020), Konsep Dasar Keperawatan (2020), Gizi Kesehatan dan Penyakit (2020), Surveilans Kesehatan dalam Kondisi Bencana (2020), Surveilans Kesehatan Masyarakat (2020), Asuhan Keperawatan Kebutuhan Dasar Manusia (2020), Ilmu Kesehatan Anak (2020), Keperawatan Keluarga (2020), Keperawatan Komunitas (2020), Penyakit Berbasis Lingkungan (2021), Keperawatan Bencana (2021), dan Promosi Kesehatan dan Perilaku Kesehatan (2021), Etika Keperawatan (2021), dan Pengantar Proses Keperawatan (2021).



Dina Dewi Anggraini, S.S.T.Keb., M.Kes., lahir di Kediri pada tanggal 12 Agustus 1990. Merupakan anak kedua dari Drs. H. Jumari Anang Siswoko dan Hj. Siti Marindun, S.Pd. Menyelesaikan kuliah di Universitas Kediri dan mendapat gelar Ahli Madya Kebidanan pada tahun 2012, Sarjana Sains Terapan Kebidanan pada tahun 2013 dan Magister Kesehatan peminatan Kesehatan Ibu dan Anak di Universitas Airlangga pada tahun 2016. Pada tahun 2017 diangkat menjadi Dosen Kebidanan di Perguruan Tinggi Swasta yaitu Universitas Kediri. Kemudian pada

tahun 2019 diangkat menjadi Dosen Kebidanan di Perguruan Tinggi Negeri di Poltekkes Kemenkes Semarang sampai sekarang.



Ahmad Irfandi lahir di Medan, pada 22 April 1992. Ia tercatat sebagai lulusan S1 Universitas Sumatera Utara dan S2 Universitas Indonesia. Pria yang kerap disapa Fandi ini adalah anak dari pasangan alm. Ismet (ayah) dan Yusnidar Lubis (Ibu). Saat ini Ahmad Irfandi merupakan seorang dosen di Program Studi Kesehatan Masyarakat Universitas Esa Unggul.



Ns. Ni Wayan Trisnadewi, S.Kep.,M.Kes Lahir di Gianyar Bali pada tanggal 18 Agustus 1984 dari pasangan I Wayan Baktiaksa dan Ni Wayan Bukti. Mulai tahun 2009 sebagai pendidik di STIKES Wira Medika Bali. Menyelesaikan pendidikan S1 Keperawatan dan Profesi di Universitas Pembangunan Nasional Veteran Jakarta pada tahun 2009, menyelesaikan Magister Kesehatan Masyarakat di Universitas Udayana pada tahun 2013. Berperan aktif dalam meningkatkan Kesehatan melalui peran serta masyarakat dan aktif dalam kegiatan sosial dan menggemari bidang pemberdayaan masyarakat.



Marlynda Happy Nurmalita Sari, S.ST, MKM, lahir di Sragen pada tanggal 29 Maret 1989. Saya menyelesaikan kuliah dan mendapat gelar sarjana Kebidanan di Universitas Sebelas Maret tahun 2011 dan magister Ilmu Kesehatan Masyarakat peminatan Kesehatan Reproduksi di Universitas Indonesia tahun 2015. Pada tahun 2011 diangkat menjadi Dosen di Perguruan Tinggi Swasta Akademi Kebidanan Pelita Ilmu Depok. Kemudian pada tahun 2019 diangkat menjadi Dosen Perguruan Tinggi Negeri di Poltekkes Kemenkes Semarang sampai sekarang.



Ns. Ni Putu Wiwik Oktaviani, S.Kep.,M.Kep

Lahir di Denpasar Bali pada tanggal 1 Oktober 1986 dari pasangan I Wayan Wirta dan Ni Wayan Widiantari. Mulai tahun 2011 sebagai pendidik di STIKES Wira Medika Bali. Menyelesaikan pendidikan S1 Keperawatan dan Profesi di Universitas Pembangunan Nasional Veteran Jakarta pada tahun 2011, menyelesaikan Magister Keperawatan di Universitas Indonesia pada tahun 2015. Berperan aktif dalam meningkatkan kesehatan

melalui peran serta masyarakat dan aktif dalam kegiatan sosial dan menggemari bidang pemberdayaan masyarakat.

Felicya Angelista lahir di Jakarta, pada 2 November 1994. Ia tercatat sebagai lulusan Institut Bisnis Nusantara.. Wanita yang kerap disapa Feli ini adalah anak dari pasangan Michael Ekel (ayah) dan Novita Ratumbusang (ibu). Felicya Angelista bukanlah orang baru di dunia hiburan Tanah Air. Ia kerap wara-wiri di layar kaca dan layar lebar Indonesia. Pada 2017 lalu, Feli berhasil meraih penghargaan emas dalam ajang Silet Awards dalam kategori artis tersilet.



Puji Laksmi, lahir di Sragen pada tanggal 17 Juni 1984, menyelesaikan Sarjana dan Profesi Ners pada Program Studi Ilmu Keperawatan, Fakultas Kedokteran, Universitas Gadjah Mada pada tahun 2008. Tahun 2015 penulis menyelesaikan program pendidikan Magister Kesehatan Masyarakat dari Fakultas Kesehatan Masyarakat, Universitas Indonesia. Penulis memiliki pengalaman sebagai dosen di Sekolah Tinggi Ilmu Kesehatan Jayakarta, Jakarta dari tahun 2008-2019 dan saat ini penulis

merupakan dosen/staff pengajar pada Program Studi Kesehatan Masyarakat, Fakultas Ilmu Kesehatan, Universitas Siliwangi. Penulis aktif dalam melakukan penelitian dan pengabdian kepada masyarakat serta mempublikasikan hasil kegiatan tersebut di berbagai jurnal.



Dr. Agus Supinganto, S.Kep., Ns., M.Kes Lahir di Lombok Barat pada tanggal 7 Agustus 1971 dari pasangan M.Mulyoto (Alm) dan Hj. Supingani. Mulai 1997 sebagai pendidik di STIKES YARSI Mataram NTB. Selanjutnya menyelesaikan pendidikan D4 Perawat pendidik di PSIK Universitas Airlangga, S1 Keperawatan dan Profesi di STIKES Muhammadiyah Gombang Jawa Tengah dan S2 AAK di FKM Universitas Airlangga. Tahun 2019 menyelesaikan pendidikan Doktorat S3 Ilmu Kedokteran Biomedik peminatan Ilmu Kesehatan

Masyarakat di Universitas Udayana Denpasar Bali. Memiliki visi berperan aktif meningkatkan peran serta masyarakat dalam meningkatkan derajat kesehatan dan aktif dalam kegiatan sosial dan menggemari bidang pemberdayaan masyarakat.



Ns. Martina Pakpahan, S. Kep., M.K.M lahir di Jakarta, 26 Januari 1986. Penulis menamatkan pendidikan Sarjana Keperawatan pada tahun 2008 dan pendidikan Profesi Ners pada tahun 2009 dari Fakultas Ilmu Keperawatan Universitas Indonesia (FIK-UI). Pada tahun 2016, penulis menamatkan pendidikan magister kesehatan masyarakat dari Fakultas Kesehatan Masyarakat Universitas Indonesia (FKM-UI). Penulis pernah bekerja sebagai

perawat di Rs. Jantung dan Pembuluh darah Harapan Kita (2009-2017). Saat ini penulis bekerja sebagai dosen di Fakultas Keperawatan Universitas Pelita Harapan (UPH) dengan mata kuliah yang diampu yaitu; Keperawatan Komunitas, Keperawatan Keluarga, Health Promotion and Behaviour, Riset dan Statistik, Introduction Nursing Process and Theory dan Communication in Healthcare. Penulis juga aktif di Fakultas sebagai sekretaris Research, Community Service and Training Committee (RCTC) dan Editor Nursing Current Jurnal Keperawatan UPH. Beberapa buku yang telah dihasilkan penulis, berkolaborasi bersama penulis lainnya yaitu; Belajar dari Covid-19: Perspektif Ekonomi & Kesehatan (2020), Ilmu Obstetri dan Ginekologi untuk Kebidanan (2020), Konsep Dasar Keperawatan (2020), Gizi Kesehatan dan Penyakit (2020), Dasar-Dasar Komunikasi Kesehatan (2020), Asuhan Keperawatan

Pada Kebutuhan Dasar Manusia (2020), Ilmu Kesehatan Anak (2020), Surveilans Kesehatan Masyarakat (2020), Keperawatan Keluarga (2020), Keperawatan Komunitas (2020), Penyakit Berbasis Lingkungan (2021), Promosi Kesehatan dan Perilaku Kesehatan (2021), dan Pengantar Proses Keperawatan: Konsep, teori dan aplikasi (2021).



Yana Listyawardhani, lahir di Kendal pada tanggal 23 Agustus 1992, menyelesaikan kuliah dan mendapat gelar Sarjana Sains Terapan pada tahun 2015. Penulis merupakan alumnus Jurusan Bidan Pendidik Universitas Sebelas Maret Surakarta. Pada tahun 2016 mengikuti Program Magister Ilmu Kesehatan Masyarakat dengan peminatan epidemiologi dan biostatistika dan lulus pada tahun 2018 dari Universitas Sebelas Maret Surakarta. Pada tahun 2021 diangkat menjadi Dosen di Universitas Siliwangi Tasikmalaya.



Fahrul Islam, SKM, MKM Lahir di Hila-hila (Bulukumba), pada 22 Maret 1987. Menyelesaikan pendidikan S1 di Fakultas Kesehatan Masyarakat Universitas Hasanuddin (2005-2009). Kemudian melanjutkan pendidikan S2 di Fakultas Kesehatan Masyarakat Universitas Indonesia (2014-2016). Saat ini bekerja sebagai dosen di Jurusan Kesehatan Lingkungan Poltekkes Kemenkes Mamuju sejak tahun 2017 dengan mata kuliah yang diampu adalah Teknik Pengambilan Sampel, Fisika Lingkungan, Penyehatan Udara, Toksikologi Lingkungan dan, Penyehatan Makanan dan Minuman .



Murti Ani, SST, M.Kes, lahir di Pati pada tanggal 31 Januari 1989. Menyelesaikan kuliah di Universitas Sebelas Maret Surakarta dan mendapat gelar Ahli Madya Kebidanan pada tahun 2012, Sarjana Sains Terapan Kebidanan pada tahun 2012 dan Magister Kesehatan pada tahun 2014. Pada tahun 2013 diangkat menjadi Dosen Kebidanan di STIKes Madani Yogyakarta. Pada tahun 2015 menjadi dosen di Prodi DIII Kebidanan Blora Poltekkes Kemenkes Semarang sampai sekarang. Beberapa karya buku yang sudah pernah dibuat yaitu Buku Ajar Promosi

Kesehatan Untuk Mahasiswa Kebidanan, Buku Ajar Asuhan Kebidanan Persalinan & Bayi Baru Lahir Untuk Mahasiswa Kebidanan, Konsep Dasar Keperawatan Maternitas, Dasar-Dasar Komunikasi Kesehatan, Pengantar Kebidanan.

STATISTIK KESEHATAN

TEORI & APLIKASI

Dalam kehidupan sehari-hari, kita acapkali disuguhkan data statistik, statistik digunakan sebagai alat untuk mengubah data menjadi Informasi yang mudah dipahami. Statistik mempunyai peranan yang sangat penting dalam berbagai bidang, tak terkecuali dalam dunia Pendidikan dan penelitian. Penggunaan statistik dalam penelitian khususnya dalam penelitian kesehatan sangat berguna dalam melakukan finalisasi data sehingga dapat menguji hipotesis suatu penelitian.

Pada buku ini akan lebih detail membahas tentang:

1. Pengantar Statistik Kesehatan
2. Probabilitas
3. Populasi dan Sampling
4. Uji Hipotesis
5. Statistik Deskriptif
6. Uji Beda Berpasangan (Kategorik dan Numerik)
7. Uji Beda Dua Kelompok (Kategorik dan Numerik)
8. Uji Beda Lebih Dua Kelompok (Kategorik dan Numerik)
9. Analisis Data Proporsi
10. Uji Chi-Square
11. Korelasi dan Regresi Linier
12. Analisis Multivariat
13. Regresi Logistik



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

ISBN 978-623-342-042-6

