



Buku Ajar **BIOSTATISTIK**

Tim Penulis :

Ishak S. Beno, S.Si., M.Sc

Ni Putu Ayu Wulan N, S.KM., M.Kes

Ni Putu Eka Febianingsih, S.KM., MPH

Mariene Wiwin Dolang, S.KM., M.Kes

Titik Suparwati, S.Si., M.Si

Henrietta Imelda Tondong, SKM., MPH

Dr. Rita Raya, S.Si., M.Si

Putu Ratna Suprema Dewi, S.KM., MPH

Felix Reba, S.Si., M.Sc

SONPEDIA.COM

PT. Sonpedia Publishing Indonesia

BUKU AJAR BIOSTATISTIK

Penulis :

Ishak S. Beno, S.Si., M.Sc
Ni Putu Ayu Wulan N, S.KM., M.Kes
Ni Putu Eka Febianingsih, S.KM., MPH
Mariene Wiwin Dolang, S.KM., M.Kes
Titik Suparwati, S.Si., M.Si
Henrietta Imelda Tondong, SKM., MPH
Dr. Rita Raya, S.Si., M.Si
Putu Ratna Suprima Dewi, S.KM., MPH
Felix Reba, S.Si., M.Sc

Penerbit

SONPEDIA.COM
PT. Sonpedia Publishing Indonesia

BUKU AJAR BIOSTATISTIK

Tim Penulis :

Ishak S. Beno, S.Si., M.Sc
Ni Putu Ayu Wulan N, S.KM., M.Kes
Ni Putu Eka Febianingsih, S.KM., MPH
Mariene Wiwin Dolang, S.KM., M.Kes
Titik Suparwati, S.Si., M.Si
Henrietta Imelda Tondong, SKM., MPH
Dr. Rita Raya, S.Si., M.Si
Putu Ratna Suprima Dewi, S.KM., MPH
Felix Reba, S.Si., M.Sc

ISBN : 978-623-8598-17-5

Editor :

Putu Intan Daryaswanti

Penyunting :

Efitra

Desain sampul dan Tata Letak :

Yayan Agusdi

Penerbit :

PT. Sonpedia Publishing Indonesia

Redaksi :

Jl. Kenali Jaya No 166 Kota Jambi 36129 Telp. +6282177858344

Email : sonpediapublishing@gmail.com

Website : www.buku.sonpedia.com

Anggota IKAPI : 006/JBI/2023

Cetakan Pertama, Maret 2024

Hak cipta dilindungi undang-undang

Dilarang memperbanyak karya tulis ini dalam bentuk dan dengan cara Apapun tanpa ijin dari penerbit

KATA PENGANTAR

Puji syukur kepada Tuhan yang Maha Esa, telah memberikan rahmat dan hidayah-Nya sehingga penulis dapat menyelesaikan buku ini dengan baik. Buku ini berjudul “***Buku Ajar Biostatistik***”. Tidak lupa kami ucapkan terima kasih bagi semua pihak yang telah mendukung dan kontribusi dalam penulisan dan penerbitan buku ini. Buku ini dirancang untuk memberikan pemahaman yang komprehensif tentang Biostatistik bagi mahasiswa kesehatan. Buku Ajar Biostatistik ini merupakan sebuah buku yang dapat digunakan oleh dosen dalam melaksanakan kegiatan pembelajaran khususnya Program Studi Kesehatan. Buku ini umum dapat digunakan sebagai panduan dan referensi mengajar menyesuaikan Rencana Pembelajaran Semester di lingkungan kampus dan lainnya.

Secara garis besar, buku ajar ini membahas tentang pemahaman dasar tentang konsep utama dalam Biostatistik. Buku ini disusun secara sistematis mencakup beberapa topik penting, mulai dari statistik deskriptif, statistik, data dan variabel, statistik inferensial, penyajian data kuantitatif, penyajian data kualitatif, tendensi sentral, probabilitas, uji validitas dan reliabilitas dan uji korelasi.

Ditulis dengan bahasa yang jelas dan mudah dipahami, buku ajar ini dirancang untuk digunakan oleh dosen dalam kegiatan pembelajaran mahasiswa. Kami berharap buku ini memberikan wawasan berharga dan menjadi panduan yang berguna dalam memahami dan mempelajari biostatistik. Buku ini mungkin masih terdapat kekurangan dan kelemahan. Oleh karena itu, saran dan kritik para pemerhati sungguh penulis harapkan. Semoga buku ini memberikan manfaat dan menambah khasanah ilmu pengetahuan.

Jayapura, Maret 2024

Tim Penulis

DAFTAR ISI

KATA PENGANTAR	ii
DAFTAR ISI.....	iii
KEGIATAN BELAJAR 1 STATISTIK DESKRIPTIF	1
DESKRIPTIF, KOMPETENSI DAN PETA KONSEP PEMBELAJARAN	
A. PENGANTAR.....	2
B. JENIS VARIABEL.....	4
C. UKURAN TENDENSI SENTRAL	6
D. UKURAN VARIABILITAS/ SEBARAN	8
E. RANGKUMAN	11
F. TES FORMATIF	12
G. LATIHAN.....	13
KEGIATAN BELAJAR 2 STATISTIK, DATA DAN VARIABEL	15
DESKRIPTIF, KOMPETENSI DAN PETA KONSEP PEMBELAJARAN	
A. KONSEP DASAR STATISTIK	16
B. KATA-KATA DASAR DALAM STATISTIK.....	17
C. RUANG LINGKUP STATISTIK.....	18
D. DATA DAN SUMBER DATA.....	20
E. VARIABEL	22
F. RANGKUMAN	23
G. TES FORMATIF	24
H. LATIHAN.....	25
KEGIATAN BELAJAR 3 BIOSTATISTIK INFERENSIAL	26
DESKRIPTIF, KOMPETENSI DAN PETA KONSEP PEMBELAJARAN	
A. KONSEP DASAR BIOSTATISTIK INFERENSIAL.....	27
B. ESTIMASI.....	29
C. UJI HIPOTESIS	30
D. RANGKUMAN	36
E. TES FORMATIF	36
F. LATIHAN.....	37
KEGIATAN BELAJAR 4 PENYAJIAN DATA KUANTITATIF.....	38

DESKRIPTIF, KOMPETENSI DAN PETA KONSEP PEMBELAJARAN	
A. PENGERTIAN PENYAJIAN DATA KUANTITATIF	39
B. CARA PENYAJIAN DATA KUANTITATIF	39
C. RANGKUMAN	47
D. TES FORMATIF	48
E. LATIHAN.....	48
KEGIATAN BELAJAR 5 PENYAJIAN DATA KUALITATIF	49
DESKRIPTIF, KOMPETENSI DAN PETA KONSEP PEMBELAJARAN	
A. PENGANTAR PENYAJIAN DATA KUALITATIF	50
B. JENIS-JENIS PENYAJIAN DATA KUALITATIF	51
C. RANGKUMAN	60
D. TES FORMATIF	61
E. LATIHAN.....	61
KEGIATAN BELAJAR 6 TENDENSI SENTRAL.....	62
DESKRIPTIF, KOMPETENSI DAN PETA KONSEP PEMBELAJARAN	
A. PENGERTIAN TENDENSI SENTRAL	63
B. SYARAT-SYARAT TENDENSI SENTRAL	63
C. MEAN.....	63
D. MEDIAN	66
E. MODUS	70
F. RANGKUMAN	72
G. TES FORMATIF	73
H. LATIHAN.....	73
KEGIATAN BELAJAR 7 PROBABILITAS	75
DESKRIPTIF, KOMPETENSI DAN PETA KONSEP PEMBELAJARAN	
A. KONSEP-KONSEP DASAR PROBABILITAS	76
B. DEFINISI PROBABILITAS	78
C. MENGHITUNG PROBABILITAS SUATU PERISTIWA	81
D. PROBABILITAS BERSYARAT	83
E. RANGKUMAN	87
F. TES FORMATIF	88
G. LATIHAN.....	89
KEGIATAN BELAJAR 8 UJI VALIDITAS DAN RELIABILITAS	90

DESKRIPTIF, KOMPETENSI DAN PETA KONSEP PEMBELAJARAN

A. UJI VALIDITAS	91
B. UJI RELIABILITAS	91
C. TAHAPAN DAN INTERPRETASI UJI VALIDITAS.....	93
D. TAHAPAN DAN INTERPRETASI UJI RELIABILITAS	96
E. RANGKUMAN	98
F. TES FORMATIF	98
G. LATIHAN.....	99
KEGIATAN BELAJAR 9 UJI KORELASI	100
DESKRIPTIF, KOMPETENSI DAN PETA KONSEP PEMBELAJARAN	
A. PENGANTAR UJI KORELASI	101
B. DATA DENGAN TIES DAN TANPA TIES	102
C. KORELASI PEARSON.....	102
D. KORELASI SPEARMAN.....	105
E. KORELASI KENDALL.....	109
F. RANGKUMAN	114
G. TES FORMATIF	114
H. LATIHAN.....	115
DAFTAR PUSTAKA	116
TENTANG PENULIS	121

KEGIATAN BELAJAR 1

STATISTIK DESKRIPTIF

DESKRIPSI PEMBELAJARAN

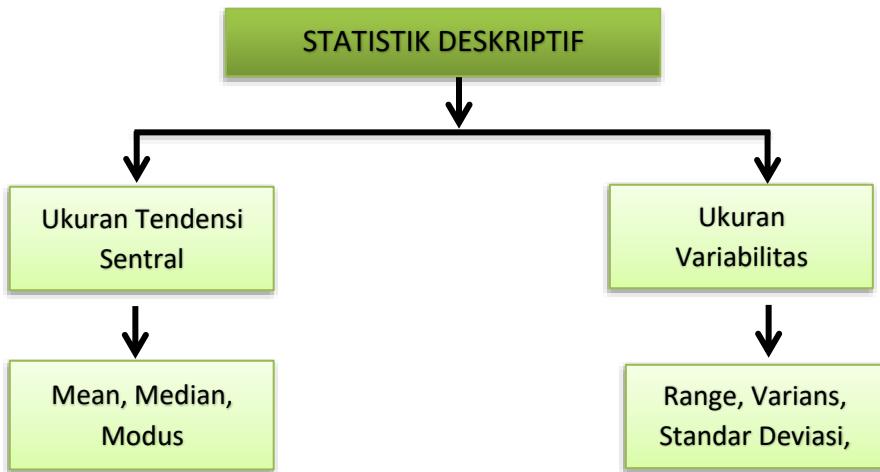
Statistik deskriptif adalah cabang statistik yang berfokus pada pengorganisasian, peringkasan, dan interpretasi data. Dalam bab ini, kita akan mempelajari konsep dasar statistik deskriptif, sebuah cabang penting dari analisis statistik yang membantu kita memahami dan meringkas karakteristik dari sebuah set data. Sebagai alat fundamental yang berguna menganalisis dan memahami data di berbagai bidang, statistik deskriptif menyediakan kita alat untuk mengatur, menyajikan, dan mendeskripsikan data dengan cara yang bermakna, sehingga memudahkan interpretasi dan analisis.

KOMPETENSI PEMBELAJARAN

Setelah mengikuti perkuliahan ini diharapkan mahasiswa dan mahasiswi memiliki kemampuan dan pemahaman tentang:

1. Teknik dasar yang digunakan untuk mendeskripsikan kumpulan data, termasuk ukuran tendensi sentral (seperti rata-rata, median, dan modus) dan ukuran dispersi (seperti rentang, varians, dan deviasi standar). Mereka akan mengeksplorasi bagaimana ukuran-ukuran ini menawarkan wawasan tentang distribusi dan variabilitas titik data dalam kumpulan data.
2. Keterampilan (*skills*) yang diperlukan untuk meringkas dan menginterpretasikan data secara efektif, meletakkan dasar untuk teknik analisis statistik lanjutan dalam studi selanjutnya.

PETA KONSEP PEMBELAJARAN



A. PENGANTAR

Statistik adalah cabang matematika yang berurusan dengan pengumpulan, analisis, interpretasi, dan penyajian data numerik dalam jumlah besar. Pada dasarnya, statistik adalah kumpulan data kuantitatif. Dua kategori besar dalam merangkum dan menganalisis data disebut sebagai statistik deskriptif dan inferensial. Pada bab ini kita akan membahas tentang statistik deskriptif. Lebih lanjut, statistik inferensial melibatkan penggunaan uji statistik seperti *uji-t*, *analisis varians (ANOVA)*, dan lain sebagainya akan dibahas di bab 4.

Ketika meringkas data, statistik deskriptif dan grafik digunakan untuk menampilkan data secara ringkas. Menampilkan data berguna untuk mengungkap pola dalam data secara efektif dengan tujuan menyajikan hasil dari sebuah studi. Oleh karena itu, fokus kita akan lebih pada deskripsi satu variabel pada satu waktu, atau yang disebut data univariat.

Ketika menggambarkan hubungan antara dua atau lebih variabel, tampilan data bivariat atau multivariat harus dipertimbangkan. Namun, seperti yang akan ditunjukkan di sini, langkah pertama dalam menggambarkan data bivariat adalah dengan menggunakan statistik deskriptif sederhana. Statistik deskriptif, atau singkatnya statistik, sering digunakan pada sampel untuk memperkirakan karakteristik populasi. Karakteristik, atau ciri-ciri, yang kita ukur pada individu atau sumber lain sering disebut variabel, karena mereka bervariasi dari satu individu ke individu lainnya. Pengukuran yang diperoleh berbeda karena berbagai alasan, termasuk variabilitas karena kesalahan pengukuran, lingkungan genotipe, dan sejenisnya.

Dalam penelitian kuantitatif, sebelum memutuskan apakah data sebuah hipotesis dapat dikonfirmasi atau ditolak, mengonfirmasi atau menyangkal hipotesis Anda dan apakah data tersebut dapat digeneralisasi ke populasi yang lebih besar., setelah mengumpulkan data, langkah pertama analisis statistik adalah mendeskripsikan karakteristik respons, seperti rata-rata satu variabel (misalnya, usia), atau hubungan antara dua variabel (misalnya, usia dan kreativitas). Berperan sebagai fondasi yang akan menjadi dasar bagi sebagian besar eksplorasi statistik lanjutan, pada bab ini kita akan mengeksplorasi konsep dan teknik penting yang membentuk landasan statistik deskriptif. Anda akan belajar cara mengungkap cerita yang tersembunyi di dalam kumpulan data, mengidentifikasi tren, dan memahami variabilitas data. Keterampilan ini tidak hanya sangat berharga untuk penelitian akademis, tetapi juga penting untuk membuat keputusan yang tepat di berbagai bidang profesional.

Dalam pembahasan selanjutnya akan dipelajari tentang ukuran tendensi sentral, seperti mean, median, dan modus, yang membantu memahami di mana letak nilai "khas" dalam kumpulan data. Selain itu, Anda juga akan mempelajari ukuran dispersi, seperti varians dan standar deviasi, yang menjelaskan penyebaran titik data di sekitar nilai pusat ini.

Namun, statistik deskriptif bukan hanya tentang angka-angka. Dalam penerapan aplikatif, statistik deskriptif juga tentang visualisasi. Dengan demikian, pada bab ini juga akan diajarkan cara membuat dan menafsirkan representasi grafis, seperti histogram, plot kotak, dan plot sebar, yang menawarkan wawasan intuitif ke dalam struktur dan pola data. Sehingga setelah menyelesaikan bab ini, Anda tidak hanya akan memiliki pemahaman yang kuat tentang statistik deskriptif, tetapi juga keterampilan untuk menganalisis dan menginterpretasikan data dengan percaya diri dalam upaya akademis dan profesional Anda di masa depan. Jadi, marilah menelusuri lebih lanjut dan buka kekuatan statistik deskriptif, secara spesifik yang berhubungan dengan dua ukuran yang sering ditampilkan pertama kali dari data. Namun, alangkah lebih baik kita membahas secara singkat tentang jenis-jenis variable (lengkapannya akan dibahas di bab 2).

B. JENIS VARIABEL

Terdapat banyak variabel yang dapat ditemui dalam bidang biologi molekuler dan medis. Namun, semua variable yang dijumpai ini umumnya terbagi menjadi dua kategori utama, yakni variabel kualitatif dan kuantitatif. Variabel kualitatif, disebut juga variabel kategori, mewakili karakteristik atau atribut yang bersifat nonnumerik dan tidak dapat diukur dalam skala numerik. Sebaliknya, data yang ada dikelompokkan ke dalam kelompok atau kelas yang berbeda, misalnya golongan darah dan mutasi genetik. Golongan darah adalah contoh klasik dari variabel kualitatif dalam biologi dan kedokteran. Ini mengkategorikan individu ke dalam golongan darah yang berbeda seperti A, B, AB, atau O. Klasifikasi ini analog dengan menyortir objek ke dalam wadah berbeda berdasarkan warnanya – ini adalah perbedaan kategori tanpa nilai numerik.

Contoh lainnya, dalam genetika dan biologi molekuler, mutasi genetik adalah variabel kualitatif yang menggambarkan variasi urutan DNA. Mutasi dapat dikategorikan berdasarkan jenisnya (misalnya *substitusi*, *penyisipan*, *penghapusan*) atau dampaknya terhadap fungsi gen (misalnya *diam*, *salah arti*, *tidak bermakna*). Klasifikasi ini mirip dengan menyortir potongan *puzzle* ke dalam berbagai bentuk – hal ini membantu peneliti mengkategorikan dan memahami beragam perubahan genetik yang mendasari penyakit dan sifat.

Variabel kuantitatif, di sisi lain, adalah jenis variabel dalam statistik yang mewakili besaran numerik. Variabel-variabel ini diukur pada skala numerik, memungkinkan dilakukannya operasi matematika seperti penjumlahan, pengurangan, perkalian, dan pembagian. Jenis variabel ini dalam biostatistik, seperti halnya bidang lainnya, sangat penting untuk mengukur fenomena biologis, menganalisis hasil eksperimen, dan menarik kesimpulan yang berarti tentang kesehatan, penyakit, dan hasil pengobatan. Baik itu pembacaan tekanan darah, jumlah sel, tingkat ekspresi gen, atau dosis obat, variabel numerik ini berfungsi sebagai dasar analisis statistik dan kesimpulan ilmiah dalam biostatistik. Sebagai contoh, tekanan darah merupakan variabel kuantitatif yang biasa dipelajari. Tekanan darah diukur dalam milimeter air raksa (mmHg) dan mewakili kekuatan yang diberikan oleh sirkulasi darah pada dinding pembuluh darah. Sama seperti alat pengukur tekanan pada pipa air, pembacaan tekanan darah memberikan informasi numerik yang dapat dibandingkan dan dianalisis.

Pemahaman singkat tentang jenis-jenis variabel statistik ini selanjutnya dapat menjadi landasan dalam pembahasan lebih lanjut. Dalam mendeskripsikan *dataset*, ada dua jenis ukuran yang biasa digunakan, yaitu ukuran tendensi sentral dan ukuran variabilitas atau sebaran.

C. UKURAN TENDENSI SENTRAL

Ukuran tendensi sentral adalah ukuran statistik yang bertujuan untuk mendeskripsikan pusat atau nilai tipikal dari suatu kumpulan data (dataset). Ukuran-ukuran ini memberikan nilai tunggal yang mewakili "rata-rata" atau "tengah" dari data. Ukuran tendensi sentral yang umum mencakup *mean*, *median*, dan *modus*.

a. Mean

Mean atau Rata-rata adalah kecenderungan sentral suatu data, yaitu suatu bilangan yang tersebar di seluruh data. Mean, atau *M*, adalah metode yang paling umum digunakan untuk mencari rata-rata.

Untuk mencari *mean*, cukup jumlahkan semua nilai respons dan bagi hasilnya dengan jumlah total respons. Jumlah seluruh tanggapan atau pengamatan disebut *N*. Secara matematis dapat dirumuskan sebagai:

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

b. Median

Median adalah nilai yang membagi data menjadi 2 bagian sama besar, yaitu banyaknya suku di sebelah kanannya sama dengan banyaknya suku di sebelah kirinya bila data diurutkan menaik atau menurun.

Untuk mencari median, urutkan setiap nilai jawaban dari yang terkecil hingga yang terbesar. Maka median adalah angka yang berada di tengah. Jika ada dua angka di tengahnya, tentukan meannya.

c. Modus

Modusnya adalah nilai respons yang paling populer atau paling sering. Suatu kumpulan data tidak boleh mempunyai modus, satu modus, atau lebih dari satu modus.

Untuk menemukan modusnya, urutkan kumpulan data Anda dari terendah ke tertinggi dan temukan respons yang paling sering muncul.

Dalam biostatistik, ukuran-ukuran ini sangat penting untuk merangkum dan menafsirkan data yang berkaitan dengan fenomena biologis dan kesehatan. Mereka memberikan wawasan berharga mengenai distribusi variabel biologis seperti usia pasien, tingkat tekanan darah, atau hasil pengobatan.

Mari kita pertimbangkan kumpulan data hipotetis yang mewakili usia pasien dalam uji klinis di bawah ini:

ID Pasien	Usia
1	45
2	52
3	35
4	60
5	28
6	45
7	50
8	38
9	42
10	55

Usia rata-rata pasien dihitung dengan menjumlahkan seluruh usia lalu membaginya dengan jumlah total pasien.

$$\begin{aligned}
 \text{Mean} &= \frac{(45 + 52 + 35 + 60 + 28 + 45 + 50 + 38 + 42 + 55)}{10} \\
 &= \frac{450}{10} \\
 \text{Mean} &= 45
 \end{aligned}$$

Usia median adalah nilai tengah dari kumpulan data jika disusun dalam urutan menaik. Jika jumlah observasinya ganjil, maka mediannya adalah nilai tengahnya. Jika jumlah observasinya genap, maka mediannya adalah rata-rata dari dua nilai tengah. Pertama, kita perlu mengurutkan usia:

28, 35, 38, 42, 45, 45, 50, 52, 55, 60.

Karena ada 10 observasi, maka mediannya adalah rata-rata dari nilai ke-5 dan ke-6.

$$\begin{aligned}\text{Median} &= \frac{(45 + 45)}{2} \\ \text{Median} &= 45\end{aligned}$$

Nilai yang paling sering muncul (modus) dalam adalah umur yang terjadi dengan frekuensi tertinggi, jadi Modus = 45.

D. UKURAN VARIABILITAS/ SEBARAN

Ada terdapat empat ukuran variabilitas dalam statistics, yaitu *range*, *varians (variance)*, *deviasi standar (standard deviation)* dan *Interquartile range*. Untuk lebih memahami perbedaan ukuran-ukuran ini, perhatikanlah penjelasan berikut:

- a. Rentang (*range*) adalah ukuran variabilitas yang paling sederhana dan mewakili perbedaan antara nilai terbesar dan terkecil dalam suatu kumpulan data. Ini memberikan gambaran singkat tentang seberapa tersebar data. Secara matematis dirumuskan sebagai:

$$\text{range} = \text{nilai maksimum} - \text{nilai minimum}$$

- b. *Varians (variance)* mengukur deviasi kuadrat rata-rata setiap titik data dari mean. Ini mengukur penyebaran titik data di sekitar mean. *Varians* dapat dituliskan dalam rumus sebagai berikut:

$$\sigma^2 = \frac{\sum_{i=1}^n (x_i - \mu)^2}{N}$$

di mana x_i merepresentasikan setiap titik data, μ adalah mean populasi, dan N adalah ukuran populasi.

- c. Deviasi standar (*standard deviation*) adalah akar kuadrat dari varians. Ini mengukur jarak rata-rata titik data dari mean dan dinyatakan dalam satuan yang sama dengan data asli. Secara matematis dapat dituliskan ke dalam persamaan berikut:

$$\sigma = \sqrt{\frac{\sum_{i=1}^n (x_i - \mu)^2}{N}}$$

- d. Rentang Interkuartil (*Interquartile range - IQR*) merupakan selisih antara kuartil ketiga (Q_3) dan kuartil pertama (Q_1) suatu kumpulan data. Ini memberikan ukuran penyebaran 50% tengah data dan tidak terlalu terpengaruh oleh outlier dibandingkan rentangnya. IQR dapat dituliskan sebagai

$$IQR = Q_3 - Q_1,$$

dengan Q_1 adalah kuartil pertama (persentil ke-25) dan Q_3 adalah kuartil ketiga (persentil ke-75) dari kumpulan data.

Mari kita pertimbangkan kumpulan data hipotetis yang mewakili pembacaan tekanan darah (sistolik) sepuluh pasien:

ID Pasien	Tekanan Darah (mmHg)
1	120
2	130
3	125
4	140
5	118

ID Pasien	Tekanan Darah (mmHg)
6	135
7	128
8	132
9	125
10	138

Penyelesaian:

- Range = nilai maksimum – nilai minimum
Dari data kita dapat mengetahui nilai minumnya adalah 118 dan nilai maksimumnya adalah 140. Dengan demikian range data ini adalah 22 mmHg
- Untuk menentukan variance, pertama kita perlu menghitung nilai rata-rata menggunakan rumus *mean*, sehingga memberikan hasil *mean* adalah 126.10. Selanjutnya kita dapat mencari nilai varians dengan menghitung total selisih setiap nilai data dengan mean yang diperoleh:

$$\begin{aligned}\sigma^2 &= \frac{(120 - 126.1)^2 + (130 - 126.1)^2 + \dots + (138 - 126.1)^2}{10} \\ \sigma^2 &= \frac{574.49}{10} \\ \sigma^2 &= 57.449\end{aligned}$$

- Selanjutnya untuk menghitung deviasi standar, kita hanya perlu menghitung nilai akar dari varians yang telah diperoleh, sehingga menjadi:

$$\begin{aligned}\sigma &= \sqrt{57.44957} \\ &\approx 7.58 \text{ mmHg}\end{aligned}$$

- Akhirnya, kita dapat menentukan IQR data ini, yaitu dengan cara pertama-tama kita perlu mencari kuartil pertama (*Q1*) dan kuartil ketiga (*Q3*) dari data tekanan darah. Kemudian, kita

akan menghitung IQR menggunakan rumus $IQR = Q3 - Q1$. Untuk mencari $Q1$ dan $Q3$, pertama-tama kita perlu menyusun datanya dalam urutan kecil-besar, 118, 120, 125, 125, 128, 130, 132, 135, 138, 140. Ada 10 titik data, jadi median ($Q2$) adalah rata-rata dari nilai ke-5 dan ke-6:

$$\begin{aligned} Q2 &= \frac{(128 + 130)}{2} \\ &= 129 \text{ mmHg} \end{aligned}$$

Sekarang, untuk mencari $Q1$, kita perlu mencari median dari bagian bawah data (nilai di bawah $Q2$):

$$\begin{aligned} Q1 &= \text{median} \{118, 120, 125, 125, 128\} \\ &= 125 \text{ mmHg} \end{aligned}$$

dan untuk mencari $Q3$, kita perlu mencari median dari separuh atas data (nilai di atas $Q2$):

$$\begin{aligned} Q3 &= \text{median} \{132, 135, 138, 140\} \\ &= 136.5 \text{ mmHg} \end{aligned}$$

Sekarang, kita dapat menghitung IQR :

$$\begin{aligned} IQR &= Q3 - Q1 \\ &= 136.5 - 125 \\ &= 11.5 \text{ mmHg} \end{aligned}$$

E. RANGKUMAN

Sebagai kesimpulan, statistik deskriptif dalam biostatistik memainkan peranan penting dalam pengukuran kuantitatif dan kualitatif dalam menganalisis data biologi/ medis. Melalui eksplorasi ukuran tendensi sentral, seperti mean, median, dan mode, serta ukuran dispersi, termasuk varians dan deviasi standar, kita memperoleh wawasan berharga mengenai distribusi dan variabilitas fenomena biologis. Alat statistik ini memungkinkan karakterisasi variabel biologi/ medis, mulai dari tekanan darah dan

jumlah sel hingga mutasi genetik dan konsentrasi obat, sehingga memfasilitasi pemahaman yang lebih mendalam tentang kesehatan, penyakit, dan hasil pengobatan. Salah satu hal penting yang dapat diambil dari bab ini adalah pentingnya merangkum dan mendeskripsikan data dengan cara yang bermakna dan dapat ditafsirkan. Dengan menggunakan statistik deskriptif, seseorang (peneliti) dapat menyusun kumpulan data yang kompleks menjadi rangkuman yang ringkas, memfasilitasi komunikasi yang lebih jelas dan pengambilan keputusan yang tepat dalam penelitian biologi dan praktik klinis.

Secara keseluruhan, manfaat statistik deskriptif dalam biostatistik bermacam-macam. Teknik statistik ini memberi peneliti alat untuk mengukur, menganalisis, dan menafsirkan variabel biologis, yang mengarah pada kemajuan dalam ilmu kedokteran, pengobatan yang dipersonalisasi, dan intervensi kesehatan masyarakat. Dengan memanfaatkan kekuatan statistik deskriptif, kita dapat mengungkap misteri dunia biologis, mengungkap wawasan baru mengenai kesehatan dan penyakit manusia, dan pada akhirnya meningkatkan kualitas hidup individu dan populasi.

F. TES FORMATIF

1. Manakah dari ukuran berikut yang paling dipengaruhi oleh nilai ekstrem atau outlier dalam kumpulan data?
 - a. Berarti
 - b. median
 - c. Modus
 - d. Jangkauan
2. Apa yang diukur dengan standar deviasi?
 - a. Sebaran titik data di sekitar median
 - b. Perbedaan rata-rata antara titik data dan rata-rata
 - c. Perbedaan antara nilai maksimum dan minimum dalam suatu dataset

- d. 50% bagian tengah dari data.
3. Ukuran statistik manakah yang secara langsung mewakili penyebaran 50% tengah data dalam kumpulan data?
- Mean
 - Median
 - Rentang Interkuartil (IQR)
 - Standar Deviasi
4. Apa yang diwakili oleh mode dalam kumpulan data?
- Nilai tengah dari dataset
 - Nilai yang paling sering muncul
 - Selisih antara nilai maksimum dan minimum
 - Nilai rata-rata kumpulan data
5. Jika kuartil pertama (Q1) suatu kumpulan data adalah 20 dan kuartil ketiga (Q3) adalah 30, berapakah rentang antarkuartil (IQR)?
- 10
 - 20
 - 5
 - 15

G. LATIHAN

Diberikan data usia 20 pasien yang mengalami kondisi medis gagal jantung,

ID Pasien	Usia
1	55
2	62
3	48
4	70

ID Pasien	Usia
5	52
6	60
7	57
8	65
9	45
10	58
11	50
12	72
13	56
14	63
15	59
16	47
17	54
18	61
19	49
20	64

Hitunglah *mean*, *median*, *range*, *standar deviasi*, dan *interquartile range (IQR)* usia pasien yang mengalami gagal jantung.

KEGIATAN BELAJAR 2

STATISTIK, DATA DAN VARIABEL

DESKRIPSI PEMBELAJARAN

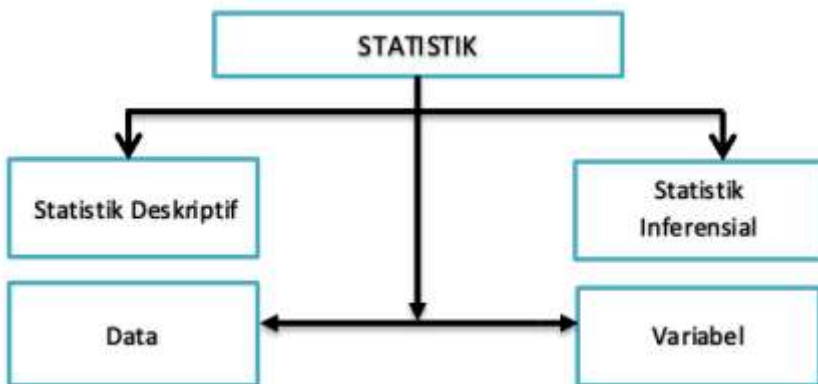
Pada bab ini mahasiswa mempelajari gambaran umum mengenai konsep dasar statistik. Bab ini memberikan landasan yang diperlukan untuk memahami istilah dan konsep dasar statistik. Mahasiswa juga mempelajari secara umum mengenai data dan variabel dalam statistik.

KOMPETENSI PEMBELAJARAN

Setelah mengikuti perkuliahan ini diharapkan mahasiswa dan mahasiswi memiliki pengetahuan dan kemampuan:

1. Mampu menguraikan konsep dasar statistik
2. Mampu menjelaskan mengenai statistic deskriptif
3. Mampu menjelaskan mengenai statistic inferensial
4. Mampu membedakan data dan variabel

PETA KONSEP PEMBELAJARAN



A. KONSEP DASAR STATISTIK

Statistika merupakan cabang ilmu yang menyediakan metode untuk membantu kita memahami data. Namun, kami meyakini bahwa pemanfaatan metode statistik dengan bijak dapat memberikan serangkaian alat yang efektif untuk mendapatkan wawasan tentang dunia sekitar kita. Metode statistik diterapkan secara luas dalam berbagai bidang seperti bisnis, kedokteran, pertanian, ilmu sosial, ilmu alam, dan disiplin ilmu terapan seperti teknik. Dengan peningkatan penggunaan analisis statistik di berbagai sektor, kesadaran akan pentingnya literasi statistik atau pengetahuan tentang tujuan dan metode statistik telah meningkat, sehingga literasi statistik dianggap sebagai komponen pokok dalam program pendidikan yang komprehensif. Bidang statistika memberikan pengajaran tentang cara membuat penilaian yang cerdas dan mengambil keputusan yang tepat saat dihadapkan dengan ketidakpastian dan variasi.

Statistik merupakan bidang studi yang terkait dengan dua aspek utama: (1) pengumpulan, pengorganisasian, peringkasan, dan analisis data, dan (2) penarikan kesimpulan mengenai suatu set data ketika hanya sebagian dari data tersebut yang diamati. Individu yang terlibat dalam kegiatan statistik harus dapat menafsirkan dan mengkomunikasikan hasilnya sesuai dengan kebutuhan situasi. Secara sederhana, kita dapat menyatakan bahwa data adalah angka yang mengandung informasi, dan tujuan statistik adalah untuk menyelidiki serta mengevaluasi sifat dan makna informasi yang terkandung dalam angka tersebut.

Statistik adalah alat yang diterapkan di berbagai sektor, seperti bisnis, pendidikan, psikologi, pertanian, dan ekonomi, dan di banyak bidang lainnya. Ketika berurusan dengan data yang bersumber dari ilmu biologi dan kedokteran, istilah yang digunakan untuk merinci penggunaan alat dan konsep statistik adalah biostatistik.

Fungsi statistik pada dasarnya terbagi menjadi enam bagian, yaitu:

1. Deskripsi (Penggambaran): Memberikan gambaran data, seperti statistik kelahiran, angka kesakitan, angka kematian ibu saat melahirkan, dan sebagainya.
2. Komparasi: Membandingkan data satu dengan yang lainnya untuk menemukan perbedaan atau kesamaan.
3. Korelasi: Menganalisis kekuatan hubungan antara data dari satu variabel dengan data dari variabel lainnya.
4. Regresi (Kecenderungan): Memprediksi pengaruh data dari satu variabel terhadap data dari variabel lainnya atau melihat kecenderungan masa depan.
5. Mengambil Keputusan: Berdasarkan hasil perhitungan statistik atau fakta dari analisis data, dapat diambil keputusan yang informasional.
6. Informasi: Menyediakan data sekunder sebagai bahan informasi untuk orang lain yang membutuhkannya.

B. KATA-KATA DASAR DALAM STATISTIK

Terdapat beberapa kata dasar dalam statistic. Kita harus mempelajari terlebih dahulu arti dari beberapa kata berikut.

1. Populasi

Populasi adalah sekumpulan orang yang mempunyai karakteristik tertentu. Populasi dalam penelitian dibagi menjadi dua yaitu populasi target dan populasi terjangkau. Populasi target dalam suatu penelitian adalah kelompok individu yang menjadi fokus penelitian atau kelompok individu yang menjadi subjek generalisasi hasil penelitian. Karakteristik yang menjadi ciri dari populasi target biasanya meliputi informasi demografis seperti jenis kelamin dan usia, serta karakteristik klinis seperti orang yang mengalami penyakit diabetes, stunting dan lainnya. Contoh populasi target adalah balita (karakteristik demografi) stunting (karakteristik klinis). Populasi terjangkau adalah bagian dari populasi target yang dapat diakses oleh peneliti

dengan mudah dalam rangka menjalankan penelitian. Untuk memastikan bahwa populasi dapat dijangkau, maka perlu menambahkan ciri administratif seperti lokasi/tempat dan waktu kedalam karakteristik populasi. Contoh populasi terjangkau adalah Balita stunting di Kabupaten XX pada tahun 2023.

2. Sampel

Sampel merupakan bagian dari populasi. Kriteria eligibilitas yaitu pertimbangan dalam pemilihan sampel, apakah seseorang memenuhi syarat untuk menjadi partisipan dalam suatu penelitian atau tidak. Terdapat kriteria eligibilitas sampel yaitu kriteria inklusi dan eksklusi. Kriteria inklusi adalah karakteristik utama dari populasi yang diinginkan peneliti. Calon partisipan penelitian harus memenuhi semua kriteria inklusi untuk dapat berpartisipasi dalam penelitian. Kriteria eksklusi adalah karakteristik yang dapat mengganggu pengumpulan data, *follow up*, dan keselamatan peserta penelitian. Jika calon partisipan memenuhi salah satu kriteria eksklusi maka mereka dikeluarkan dari partisipasi penelitian.

3. Data

Data merupakan bahan yang diperlukan statistic. Dalam statistic data didefinisikan sebagai angka.

4. Variabel

Pengamatan terhadap suatu karakteristik, dan ditemukan nilai yang berbeda pada setiap orang, tempat atau benda. Contohnya berat badan dan tinggi badan tiap anak akan menghasilkan nilai yang berbeda. Maka berat badan dan tinggi badan merupakan variable.

C. RUANG LINGKUP STATISTIK

Statistik berdasarkan tujuan analisisnya, statistik dibedakan menjadi statistik deskriptif dan statistik inferensial.

1. Statistik Deskriptif

Statistik deskriptif merupakan cabang statistik yang bertujuan untuk memberikan deskripsi atau gambaran obyek yang sedang diteliti melalui data sampel, tanpa melakukan analisis atau membuat kesimpulan yang berlaku secara umum. Statistik deskriptif menggambarkan data melalui berbagai metode, seperti penyajian dalam tabel, diagram, pengukuran tendensi sentral seperti rata-rata hitung, rata-rata ukur, dan rata-rata harmonik. Selain itu, statistik deskriptif juga menggunakan pengukuran penempatan, seperti median, kuartil, desil, dan persentil, serta pengukuran penyimpangan, termasuk range, rentang antar kuartil, rentang semi antar kuartil, simpangan rata-rata, simpangan baku, varians, koefisien varians, dan angka baku.

2. Statistik Inferensial

Statistik inferensial adalah suatu prosedur di mana kita membuat kesimpulan atau generalisasi tentang suatu populasi berdasarkan informasi yang diperoleh dari sampel yang diambil dari populasi tersebut. Pada umumnya, agar kesimpulan yang ditarik mengenai suatu populasi memiliki validitas, diperlukan sampel yang diambil dengan metode ilmiah dan mewakili populasi secara adekuat. Terdapat beragam tipe sampel yang dapat diambil dari suatu populasi, namun tidak semuanya dapat menjadi dasar yang valid untuk membuat kesimpulan mengenai populasi tersebut. Terdapat beberapa metode pengambilan sampel yaitu *probability sampling* dan *non-probability sampling*. *Probability sampling* adalah metode *sampling* di mana setiap item yang dipilih, dipilih dengan mempertimbangkan peluang setiap item. Melalui pengambilan sampel probabilitas, peluang yang lebih tinggi terciptanya sampel yang mewakili populasi dapat ditingkatkan. *Probability sampling* terdiri dari *simple random sampling*, *systematic random sampling*, *stratified random sampling*, dan *cluster random sampling*. *Non-probability sampling* merupakan metode pemilihan *sampling* dimana tidak setiap individu memiliki

kesempatan terpilih sebagai sample. Pertimbangan menggunakan metode ini adalah respon rate rendah dan tidak semua individu pada populasi tepat dijadikan sampel. *Non-probability sampling* terdiri dari *purposive sampling*, *convenience sampling*, *quota sampling* dan *snowball sampling*.

D. DATA DAN SUMBER DATA

Data merupakan materi yang diperlukan dalam statistika. Dalam konteks statistika, data didefinisikan sebagai nilai numerik. Nilai numerik terbagi menjadi dua kategori, yakni nilai yang berasal dari pengukuran dan nilai yang dihasilkan melalui proses perhitungan. Sebagai contoh, ketika seorang kader Posyandu menimbang seorang balita, pengukuran tersebut menghasilkan nilai numerik, misalnya 10 kg. Sementara itu, nilai numerik lainnya bisa didapatkan ketika seorang administrator rumah sakit menghitung jumlah pasien, yang mungkin menghasilkan angka, seperti 20 pasien. Setiap angka tersebut bersifat unik dan menyimpan informasi yang beragam, dan ketika angka-angka tersebut digabungkan, maka kita mendapatkan data.

Kegiatan statistika dilaksanakan sebagai respons terhadap kebutuhan untuk menjawab pertanyaan tertentu. Pertanyaan tersebut dapat berasal dari berbagai pihak, seperti seorang dokter yang ingin mengetahui manfaat relatif dari suatu prosedur perawatan, atau seorang administrator yang ingin mendapatkan jawaban mengenai area tertentu seperti kualitas layanan atau pemanfaatan fasilitas yang diminati. Untuk merespons pertanyaan tersebut dengan tepat, pendekatan statistika diperlukan, sehingga membutuhkan pengumpulan data yang sesuai untuk diolah. Data ini biasanya dapat diperoleh dari satu atau lebih sumber berikut:

1. Catatan Rutin

Catatan harian merupakan bagian rutin dari setiap organisasi salah satu contohnya adalah rumah sakit, di mana kegiatan dan aktivitas terdokumentasi secara teratur. Rekam medis rumah

sakit termasuk dalam catatan rutin yang menyimpan beragam informasi mengenai pasien. Selain itu, catatan akuntansi rumah sakit juga mencakup data yang signifikan tentang keuangan rumah sakit. Jika data yang dibutuhkan dapat ditemukan dalam catatan rutin tersebut, maka sumber data dapat berasal dari catatan rutin tersebut.

2. Survey

Jika data yang diperlukan tidak dapat ditemukan dalam catatan rutin, sumber data yang memungkinkan adalah melalui survei. Sebagai contoh, seorang administrator layanan kesehatan mungkin ingin mengumpulkan informasi mengenai moda transportasi yang digunakan pasien saat mengunjungi klinik atau mendapatkan informasi tentang jarak tempuh pasien dari rumah ke layanan kesehatan. Jika pertanyaan-pertanyaan ini tidak tercakup dalam formulir administrasi, maka survei dapat dilakukan untuk mengumpulkan informasi yang dibutuhkan.

3. Eksperimen

Beberapa data dapat diperoleh melalui percobaan atau eksperimen. Sebagai contoh, seorang apoteker yang ingin mengetahui strategi terbaik untuk meningkatkan kepatuhan pasien dalam mengonsumsi obat mungkin akan melakukan serangkaian eksperimen. Melalui percobaan ini, apoteker dapat mengidentifikasi dan menilai berbagai strategi, sehingga dapat menemukan pendekatan yang paling efektif untuk meningkatkan kepatuhan pasien.

4. Sumber Eksternal

Data juga dapat diperoleh dari sumber eksternal yang telah dipublikasikan, seperti bank data yang tersedia secara komersial. Dengan kata lain, ada kemungkinan bahwa pertanyaan yang sama telah diajukan oleh orang lain, dan jawaban yang mereka peroleh dapat diterapkan pada situasi yang kita hadapi saat ini.

E. VARIABEL

Pengamatan terhadap suatu ciri, yang kemudian menunjukkan variasi nilai pada individu, lokasi, atau objek yang berbeda, disebut sebagai variabel. Beberapa contoh variabel melibatkan tekanan darah diastolik, detak jantung, tinggi badan pria dewasa, berat badan anak prasekolah, dan usia pasien yang datang ke klinik gigi.

Variabel kuantitatif adalah variabel yang dapat diukur. Misalnya, kita dapat memperoleh pengukuran tinggi badan, berat badan, dan usia pasien yang datang ke poli mata. Ini adalah contoh variabel kuantitatif. Pengukuran yang dilakukan terhadap variabel kuantitatif menyampaikan informasi mengenai jumlah.

Variabel kualitatif menyampaikan konsep bahwa terdapat beberapa karakteristik yang tidak dapat diukur seperti tinggi badan, berat badan, atau umur. Sebaliknya, karakteristik ini hanya dapat dikategorikan atau dikelompokkan ke dalam kategori tertentu. Contohnya, ketika seseorang didiagnosis menderita suatu penyakit, atau ketika seseorang diidentifikasi sebagai bagian dari suatu kelompok etnis, atau bahkan ketika kita menyatakan bahwa suatu objek memiliki atau tidak memiliki suatu karakteristik tertentu. Pengukuran pada variabel kualitatif dilakukan melalui proses pengkategorian atau pemberian label pada suatu karakteristik. Informasi yang dihasilkan dari pengukuran variabel kualitatif memberikan wawasan tentang atribut atau karakteristik yang dimiliki oleh individu, tempat, atau objek dalam kategori-kategori tersebut.

Variabel diskrit adalah hitungan atau jumlah dari suatu hal. Contoh variabel diskrit yaitu jumlah penerimaan harian di rumah sakit umum adalah variabel acak diskrit karena jumlah penerimaan setiap hari harus direpresentasikan oleh angka bulat, seperti 0, 1, 2, atau 3. Jumlah gigi berlubang, jumlah anak sekolah dasar, jumlah anak yang ada dalam keluarga adalah contoh lain dari variabel diskrit.

Variabel kontinu dapat mengambil nilai apa pun dalam interval nilai yang relevan yang ditetapkan oleh variabel tersebut, hanya terbatas oleh ketelitian proses pengukuran. Contoh dari variabel kontinu melibatkan berbagai pengukuran yang dapat diambil pada individu, seperti tinggi badan (162,4 cm), berat badan (15,6 kg), dan lingkaran lengan atas (23,5 cm).

F. RANGKUMAN

Statistika adalah cabang ilmu yang memberikan metode untuk memahami data. Statistik diterapkan luas dalam berbagai bidang, dan literasi statistik dianggap penting dalam pendidikan. Statistik berkaitan dengan pengumpulan, pengorganisasian, analisis data, serta penarikan kesimpulan dari data yang hanya sebagian diobservasi. Tujuan statistik adalah menyelidiki dan mengevaluasi informasi dalam angka. Statistik dibagi menjadi statistik deskriptif (penggambaran) dan statistik inferensial (penarikan kesimpulan). Kata-kata dasar dalam statistik yang perlu kita pahami adalah populasi, sampel, data dan variabel. Statistik terbagi menjadi deskriptif (memberikan gambaran) dan inferensial (penarikan kesimpulan). Statistik deskriptif menggunakan metode seperti tabel, diagram, dan pengukuran tendensi sentral. Statistik inferensial membuat kesimpulan atau generalisasi dari sampel terhadap populasi dengan metode ilmiah. Terdapat metode *sampling* yaitu *probability sampling* (random) dan *non-probability sampling* (tidak random). Data didefinisikan sebagai nilai numerik. Pengumpulan data melibatkan catatan rutin, survei, eksperimen, atau sumber eksternal. Variabel adalah pengamatan terhadap suatu ciri. Kuantitatif dapat diukur (tinggi, berat), sementara kualitatif hanya dapat dikategorikan. Diskrit adalah hitungan (jumlah pasien), kontinu dapat mengambil nilai dalam interval yang ditetapkan. Contoh variabel melibatkan pengukuran seperti tinggi, berat, dan jumlah gigi berlubang.

G. TES FORMATIF

1. Bagian dari populasi yang dipilih untuk dianalisis disebut....
 - a. Sampel
 - b. Populasi
 - c. Data
 - d. Variabel
 - e. Statistik

2. Metode statistik yang melibatkan pengumpulan, penyajian dan karakterisasi sekumpulan data untuk mendeskripsikan dengan tepat berbagai fitur yang ada dari kumpulan data disebut....
 - a. Statistik inferensial
 - b. Statistik deskriptif
 - c. Sampling
 - d. Metode Ilmiah
 - e. Variabel

3. Statistik inferensial digunakan ketika....
 - a. Memberi gambaran objek yang diteliti
 - b. Melakukan sensus lengkap pada suatu populasi
 - c. Menyajikan data dalam bentuk grafik
 - d. Mengambil hasil dari sebuah sampel dan membuat kesimpulan tentang suatu populasi
 - e. Menyajikan data dalam bentuk tabel

4. Jumlah balita yang terdapat dalam keluarga merupakan contoh dari....
 - a. Variabel kualitatif
 - b. Variabel diskrit
 - c. Variabel kontinu
 - d. Variabel kategorikal
 - e. Variabel nominal

5. Statistik inferensial digunakan ketika....

- a. Memberi gambaran objek yang diteliti
- b. Melakukan sensus lengkap pada suatu populasi
- c. Menyajikan data dalam bentuk grafik
- d. Mengambil hasil dari sebuah sampel dan membuat kesimpulan tentang suatu populasi
- e. Menyajikan data dalam bentuk tabel

H. LATIHAN

- 1. Jelaskan perbedaan antara statistik deskriptif dan inferensial!
- 2. Penelitian yang dilakukan Yayasan XXX di Kabupaten Gianyar pada tahun 2020 dilakukan pada 886 balita dengan tujuan untuk mengetahui prevalensi balita yang mengalami stunting. Jelaskan siapa yang menjadi populasi target, populasi terjangkau dan populasi dari penelitian tersebut!

KEGIATAN BELAJAR 3

BIOSTATISTIK INFERENSIAL

DESKRIPSI PEMBELAJARAN

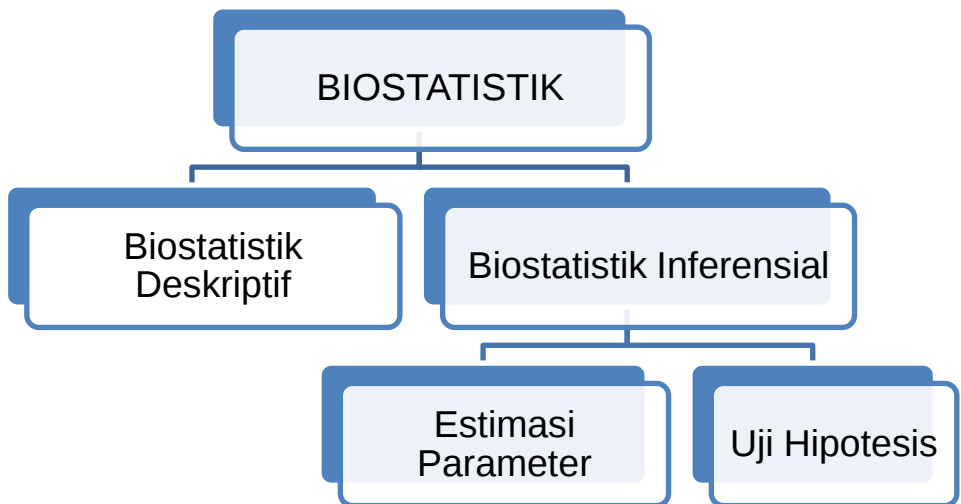
Pada bab ini mahasiswa mempelajari konsep dasar biostatistik inferensial. Diharapkan mahasiswa mampu memahami konsep biostatistik inferensial dan ruang lingkup pengujiannya.

KOMPETENSI PEMBELAJARAN

Setelah mengikuti perkuliahan ini diharapkan mahasiswa dan mahasiswi memiliki pengetahuan dan kemampuan untuk:

1. Menjelaskan konsep dasar biostatistik inferensial.
2. Mampu memahami dan menjelaskan konsep estimasi parameter
3. Mampu memahami dan menjelaskan konsep uji hipotesis

PETA KONSEP PEMBELAJARAN



A. KONSEP DASAR BIOSTATISTIK INFERENSIAL

Biostatistik merupakan ilmu terapan dari ilmu statistik. Statistik berasal dari bahasa Latin yaitu “Status” yang dalam Bahasa Inggris berarti “State” dan berarti “Negara” dalam Bahasa Indonesia. Statistik diartikan segala sesuatu yang menyebutkan keterangan atau angka-angka yang dibutuhkan oleh negara. Statistik kemudian lebih lanjut didefinisikan sebagai kumpulan fakta yang berbentuk angka-angka yang disusun dalam bentuk tabel, grafik, atau gambar. Kemudian ilmu pengetahuan yang mempelajari penciptaan, pengembangan, dan penerapan teknik atau metode sehingga ketidakpastian inferensial dapat dievaluasi, disebut dengan Statistika (Nuryadi, Astuti, Utami, & Budiantara, 2017). Ilmu statistika dapat diterapkan dalam berbagai disiplin ilmu.

Statistik adalah ilmu yang digunakan untuk membuat kesimpulan tentang fenomena tertentu berdasarkan. Sampel yang relatif terbatas. Bidang statistik mempunyai dua bidang utama, yaitu Statistik Matematika dan Statistika Terapan. Statistik matematika berkaitan dengan pengembangan metode baru tentang inferensi statistik dan membutuhkan pengetahuan yang rinci tentang matematika untuk penerapannya. Sedangkan Statistik Terapan melibatkan penerapan statistik matematika pada bidang ilmu tertentu, seperti ekonomi, psikologi, dan kesehatan masyarakat (Rosner, 2011)

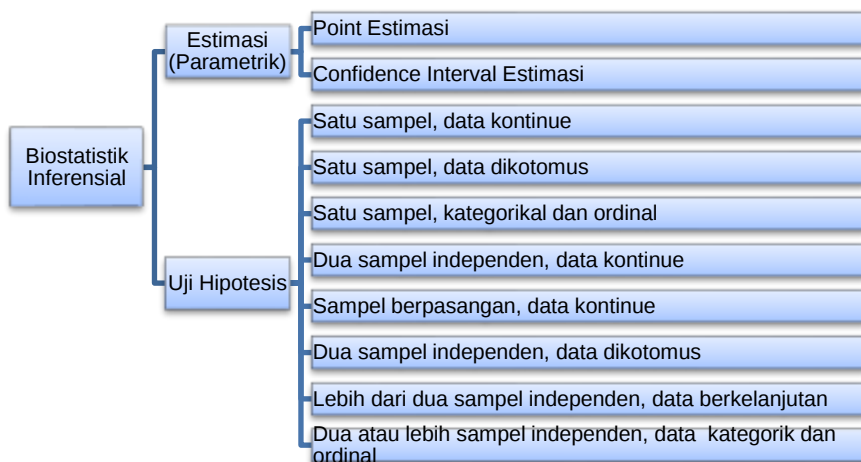
Statistik dapat dibagi menjadi 3 garis besar berdasarkan kriteria tertentu (Hasan, 2016). Ketiga kriteria tersebut antara lain:

1. Statistik berdasarkan cara pengolahan data, yaitu:
 - a. Statistik Deskriptif
 - b. Statistik Inferensial
2. Statistik berdasarkan bentuk parameter
 - a. Statistik Parametrik
 - b. Statistik Non Parametrik

3. Statistik berdasarkan ruang lingkup penggunaan disiplin ilmu, yaitu Statistik Sosial, statistik Pendidikan, Statistik Ekonomi, Statistik Pertanian, Statistik Kesehatan, dan lain sebagainya

Penerapan ilmu statistika pada ilmu kesehatan disebut dengan Biostatistika. Biostatistika adalah ilmu terapan berupa pengembangan dan penerapan penalaran metode statistika dalam menangani; menganalisis; dan memecahkan masalah di bidang ilmu biologi; kesehatan masyarakat; kesehatan umum; serta dibidang penelitian biomedis, klinis dan berbasis populasi. Prinsip biostatistik didasarkan pada matematika terapan dan mencakup alat dan teknik untuk mengumpulkan data kemudian merangkum, menganalisis, dan menafsirkan hasil tersebut. Prinsip tersebut mencakup pembuatan dan penarikan kesimpulan yang memperhitungkan ketidakpastian secara tepat (Sullivan, 2018).

Biostatistik Inferensial atau induktif adalah bagian dari ilmu statistik yang mempelajari penafsiran dan penarikan kesimpulan dari data yang telah tersedia. Biostatistik inferensial berhubungan dengan pendugaan populasi dan pengujian hipotesis dari data atau fenomena pada populasi. Dapat dikatakan bahwa biostatistik inferensial merupakan kelanjutan dari biostatistik deskriptif (Hasan, 2016). Biostatistik Inferensial mencakup dua bisang utama, yaitu pengukuran Estimasi dan Pengujian Hipotesis. Estimasi berkaitan dengan memperkirakan nilai parameter populasi tertentu; pengujian hipotesis berkaitan dengan pengujian apakah nilai parameter populasi sama dengan nilai tertentu (Rosner, 2011). Secara Garis besar dapat digambarkan sebagai berikut:



Gambar 3.1 Bagan Lingkup Pengukuran Biostatistik Inferensial
 Sumber: (Rosner, 2011; Sahu, Pal, & Das, 2015; Sullivan, 2018)

B. ESTIMASI

Estimasi merupakan penggunaan statistik sampel untuk menghasilkan perkiraan tentang parameter populasi yang tidak diketahui. Estimasi adalah proses menentukan kemungkinan nilai parameter populasi (misalnya rata-rata populasi sebenarnya atau proporsi populasi) berdasarkan sampel acak. Sampel harus bersifat representatif atau mewakili populasi (Rosner, 2011; Sahu et al., 2015; Sullivan, 2018). Terdapat dua jenis estimasi yang dapat digunakan sebagai pendugaan populasi, yaitu *point estimate* (estimasi titik) dan *Confidence Interval Estimate* (estimasi interval kepercayaan). Estimasi titik untuk parameter populasi adalah estimasi bernilai tunggal dari parameter tersebut dan digunakan untuk memberi gambaran spesifik tentang perkiraan parameter populasi.

Estimasi interval kepercayaan adalah kisaran nilai untuk parameter populasi dengan tingkat kepercayaan yang melekat (misalnya, kepercayaan 95%). Estimasi interval apabila ingin menentukan kisaran dimana nilai parameter cenderung berada. Apabila rentang interval kepercayaannya sempit maka dapat dikatakan estimasi titik sangat tepat. Namun jika luas, maka lebih merujuk pada penggunaan estimasi interval untuk menjelaskan parameter populasi (Rosner, 2011; Sullivan, 2018). Berikut adalah contoh Kelurahan Kawan memiliki populasi 450 bayi. Kemudian dilakukan sebuah survei terhadap sampel 150 bayi di Kelurahan Kawan, didapatkan bahwa rerata berat badan bayi adalah 2100 gr (95% CI: 2000 – 2300). Jika 95% CI dalam rentang sempit maka gambaran berat badan bayi di Kelurahan Kawan dapat digambarkan dengan nilai rerata (2100gr). Namun jika rentang 95% CI luas maka sebaiknya menggunakan estimasi interval untuk menjelaskan gambaran berat badan bayi di Kelurahan Kawan.

C. UJI HIPOTESIS

Bagian kedua dari inferensi statistik adalah uji hipotesis. Pada pengujian hipotesis, pernyataan atau hipotesis dihasilkan dari parameter populasi, dan sampel di digunakan untuk menilai kemungkinan bahwa hipotesis tersebut adalah benar. Hipotesis didasarkan atas informasi, teori, hasil penelitian sebelumnya yang tersedia dan keyakinan peneliti tentang parameter tersebut. Proses pengujian hipotesis melalui pembuatan dua hipotesis yang saling bersaing. Satu hipotesis mencerminkan tidak ada perbedaan, tidak ada hubungan, atau tidak ada pengaruh. Hipotesis ini disebut dengan hipotesis nol). Kemudian hipotesis lainnya merupakan hipotesis alternatif atau hipotesis penelitian. Hipotesis alternatif atau hipotesis penelitian mencerminkan keyakinan peneliti berupa terdapat perbedaan, terdapat hubungan, atau terdapat pengaruh. Pengujian hipotesis bergantung pada klasifikasi yang tepat dari

variabel penelitian dan kelompok pembanding yang digunakan (Sullivan, 2018).

Pengujian hipotesis dilakukan melalui lima langkah, yaitu:

1. Langkah 1: Siapkan hipotesis dan tentukan tingkat signifikansinya
2. Langkah 2: Pilih uji statistik yang sesuai
3. Langkah 3: Siapkan aturan keputusan. Aturan keputusan adalah pernyataan yang menunjukkan hal yang mendasari keadaan untuk menolak hipotesis nol. Aturan pengambilan untuk pengujian tertentu tergantung pada tiga faktor utama, yaitu: hipotesis penelitian, uji statistik yang digunakan, dan tingkat signifikansi
4. Langkah 4: Hitung statistik pengujian. Langkah ini dilakukan melalui substitusi data sampel yang diamati kedalam uji statistik yang sesuai dan tepat berdasarkan langkah 2.
5. Langkah 5: Kesimpulan. Kesimpulan akhir dibuat dengan membandingkan uji statistik (yang merupakan ringkasan informasi yang diamati pada sampel). Kesimpulan akhir adalah menolak hipotesis nol atau tidak menolak hipotesis nol.

Pengujian hipotesis yang dapat dilakukan bergantung pada masing-masing variabel dan skala data yang akan diuji. Diantaranya adalah:

1. Uji Hipotesis satu sampel dengan data kontinue
Tujuan pengujian hipotesis satu sampel dengan hasil data kontinue adalah untuk membandingkan rata-rata suatu populasi (μ) dengan rata-rata yang diketahui (μ_0). Hal penting dalam penyusunan hipotesis uji satu sampel adalah bahwa rata-rata yang ditentukan pada hipotesis nol adalah pembanding yang rasional dan sesuai. Nilai rata-rata yang digunakan sebagai pembanding umumnya diperoleh dari hasil penelitian sebelumnya atau dari laporan lain yang relevan.
Contoh hipotesis:
 $H_0: \mu = 205$
 $H_1: \mu < 205$

Dengan $\alpha = 0.05$

2. Uji Hipotesis satu sampel dengan data dikotomus

Tujuan uji hipotesis satu sampel dengan hasil data dikotomus adalah untuk membandingkan proporsi keberhasilan dalam satu populasi dengan proporsi yang diketahui (p_0). Proporsi yang diketahui tersebut umumnya diperoleh dari penelitian sebelumnya atau laporan lain (terkadang disebut sebagai kendali historis). Hal penting dalam menyusun hipotesis dalam uji satu sampel bahwa proporsi yang ditentukan dalam hipotesis nol adalah pembandingan yang adil dan masuk akal. Contoh hipotesis:

$$H_0: p = 0.211$$

$$H_1: p < 0.211$$

Dengan $\alpha = 0.05$

3. Uji Hipotesis satu sampel dengan data kategorikal dan ordinal
Variabel kategorikal dan ordinal adalah variabel yang mempunyai lebih dari dua tanggapan atau kategori berbeda, dan tanggapan dapat diurutkan dan bertingkat (ordinal) atau tanggapan yang tidak diurutkan (kategorikal). Tujuan prosedur uji hipotesis satu sampel dengan hasil kategorikal dan atau ordinal adalah:

- Untuk membandingkan distribusi tanggapan dengan distribusi yang diketahui
- Untuk membandingkan proporsi peserta dalam setiap kategori tanggapan dengan distribusi yang telah diketahui

Distribusi yang diketahui dapat berasal dari hasil penelitian sebelumnya atau laporan terkait. Contoh hipotesis:

$$H_0: \text{Distribusi respon adalah } 0.60; 0.25; 0.15$$

$$H_1: H_0 \text{ adalah salah}$$

Dengan $\alpha = 0.05$

Hal penting yang perlu diperhatikan adalah bahwa hipotesis penelitian ditulis dalam bentuk kata-kata atau narasi, bukan simbol. Hipotesis penelitian sebagaimana dinyatakan mencakup

perbedaan distribusi tanggapan dari yang ditentukan dalam hipotesis nol

4. Uji Hipotesis dua sampel independen dengan data kontinue

Uji hipotesis dua sampel independen dengan hasil kontinue bertujuan untuk mengetahui adanya perbedaan rata-rata antara dua kelompok yang berbeda atau dua kelompok yang terpisah. Kedua kelompok tersebut dapat ditentukan berdasarkan atribut tertentu, misalkan antara laki-laki dan perempuan, antara kelompok sehat dan kelompok sakit atau dapat ditentukan oleh peneliti (antara kelompok perlakuan dan kelompok kontrol). Uji hipotesis ini tepat untuk dua kelompok pembandingan independen dan hasil yang diinginkan bersifat berkelanjutan (misalnya: tekanan darah, kolesterol total, kadar Hemoglobin, penurunan berat badan dan variabel dengan data kontinue lainnya). Pada uji hipotesis ini diperlukan perhitungan statistik deskriptif pada masing-masing kelompok sampel, yang kemudian dibandingkan. Contoh Hipotesis:

$$H_0: \mu_1 = \mu_2$$

$$H_1: \mu_1 \neq \mu_2$$

$$\text{Dengan } \alpha = 0.05$$

5. Uji Hipotesis sampel berpasangan dengan data kontinue

Uji hipotesis sampel berpasangan dengan hasil kontinue bertujuan untuk membandingkan skor antara dua kelompok sampel yang sama atau berpasangan. Hal ini berarti dua kelompok data yang dibandingkan tersebut berasal dari satu kelompok sampel atau partisipan yang diukur dua kali. Data yang dianalisis dapat berupa data pengukuran sebelum dan sesudah suatu perlakuan. Ketika sampel bersifat dependen atau berpasangan, maka fokus uji adalah pada skor perbedaan disetiap sampel atau diantara pasangan, dan uji hipotesis didasarkan pada perbedaan rerata atau *mean difference* atau dengan simbol μ_d . Hipotesis nol menyatakan tidak ada perbedaan rerata. Contoh hipotesis:

$$H_0: \mu_d = 0$$

$$H_1: \mu_d \neq 0$$

Dengan $\alpha = 0.05$

6. Uji Hipotesis dua sampel independen dengan data dikotomis
- Uji Hipotesis dua sampel independen dengan hasil dikotomis mempertimbangkan situasi dimana terdapat dua kelompok pembanding independen dan hasil yang diinginkan bersifat dikotomis (misalnya keberhasilan/kegagalan; sakit/sehat; anemia/normal). Tujuan dari uji hipotesis ini adalah untuk membandingkan proporsi keberhasilan antara kedua kelompok.

Data sampel yang relevan adalah ukuran sampel di setiap kelompok pembanding (n_1 dan n_2) dan proporsi sampel ($\bar{p}_1$ dan $\bar{p}_2$), yang dihitung dengan mengambil rasio jumlah keberhasilan dengan ukuran sampel di setiap kelompok yang dihitung dengan mengambil rasio jumlah keberhasilan dengan ukuran sampel di setiap kelompok, misalnya $\bar{p}_1 = x_1/n_1$ dan $\bar{p}_2 = x_2/n_2$. Beberapa contoh penulisan pada uji hipotesis dua sampel independen dengan hasil dikotomis antara lain:

- a. Perbedaan proporsi yang signifikan antara 2 kelompok independen

$$H_0: \bar{p}_1 = \bar{p}_2$$

$$H_1: \bar{p}_1 \neq \bar{p}_2$$

- b. Perbedaan proporsi menggunakan *Risk Difference* (RD), *Relative Risk* (RR) dan *Odds Rasio* (OR) sebagai ukuran dampak, maka:

$$H_0: RD = 0$$

$$H_1: RD \neq 0$$

atau

$$H_0: RR = 1$$

$$H_1: RR \neq 1$$

atau

$$H_0: OR = 1$$

$$H_1: OR \neq 1$$

7. Uji Hipotesis lebih dari dua sampel independen dengan data kontinue

Teknik untuk menguji perbedaan lebih dari dua mean independen merupakan perpanjangan dari prosedur dua sampel independen dan disebut analisis varians (ANOVA). Teknik ANOVA berlaku jika terdapat lebih dari dua perbandingan independen. Salah satu contohnya adalah uji klinis yang dirancang untuk mengevaluasi efektivitas suatu obat untuk pengurangan suatu penyakit. Peneliti kemudian membandingkan tiga jenis obat terstandar melalui penelitian eksperimental. Contoh penelitian observasional misalnya penelitian untuk membandingkan rerata kadar kolesterol pada orang yang memiliki indeks masa tubuh (IMT) kurang, normal, dan *overweight*. Contoh hipotesis yang dapat dibangun adalah:

$$H_0: \mu_1 = \mu_2 = \mu_3$$

$$H_1: \mu_1 \neq \mu_2 \neq \mu_3$$

8. Uji Hipotesis untuk dua atau lebih sampel independen dengan hasil kategorik dan ordinal

Uji Hipotesis untuk dua atau lebih sampel independen dengan hasil kategorik dan ordinal bertujuan untuk membandingkan distribusi respon terhadap variabel hasil kategorikal atau ordinal di antara beberapa kelompok pembanding independen. Uji tersebut disebut uji independensi χ^2 dan hipotesis nolnya adalah tidak ada perbedaan dalam distribusi respons terhadap hasil antar kelompok pembanding. Salah satu contoh adalah, sebuah uji coba secara acak dirancang untuk mengevaluasi efektivitas metode pereda nyeri yang baru dikembangkan untuk mengatasi nyeri haid atau dismenore primer pada remaja putri. Uji coba tersebut melibatkan tiga kelompok yaitu kelompok perlakuan metode 1, kelompok perlakuan metode 2, dan kelompok ketiga adalah kelompok kontrol. Ketiga kelompok independen tersebut merupakan variabel kategorik. Uji hipotesis dilakukan untuk mengetahui apakah terdapat perbedaan proporsi remaja yang melaporkan adanya penurunan skala nyeri diantara ketiga metode perlakuan.

Contoh hipotesis yang dapat dibangun adalah:

H_0 = Perlakuan dan adanya penurunan nyeri bersifat independen

H_1 = H_0 Salah

D. RANGKUMAN

Biostatistik Inferensial adalah bagian dari ilmu statistik yang mempelajari penafsiran dan penarikan kesimpulan dari data yang telah tersedia. Biostatistik inferensial berhubungan dengan pendugaan populasi dan pengujian hipotesis dari data atau fenomena pada populasi. Dua bidang utama biostatistik inferensial antara lain adalah pengukuran Estimasi dan Pengujian Hipotesis. Estimasi berkaitan dengan memperkirakan nilai parameter populasi tertentu. Terdapat dua jenis estimasi yang dapat digunakan sebagai pendugaan populasi, yaitu *point estimate* (estimasi titik) dan *Confidence Interval Estimate* (estimasi interval kepercayaan). Pengujian hipotesis berkaitan dengan pengujian apakah nilai parameter populasi sama dengan nilai tertentu.

E. TES FORMATIF

1. Biostatistik yang diterapkan untuk menafsirkan dan menarik dan membuat kesimpulan dari data yang telah tersedia disebut dengan?
 - a. Biostatistik deskriptif
 - b. Biostatistik inferensial
 - c. Biostatistik deduktif
 - d. Biostatistik uji hipotesis
 - e. Biostatistik uji instrumen
2. Pendugaan parameter populasi apabila rentang interval kepercayaan sangat sempit maka dapat menggunakanuntuk menjelaskan parameter populasi

- a. Hipotesis untuk dua atau lebih sampel
- b. Hipotesis untuk satu sampel
- c. Hipotesis untuk sampel
- d. Estimasi titik
- e. Estimasi interval

F. LATIHAN

Berikan contoh uji hipotesis untuk tiga kelompok sampel independen dengan data kontinue, lengkapi dengan tahapan atau langkah penyusunan hipotesisnya!

KEGIATAN BELAJAR 4

PENYAJIAN DATA KUANTITATIF

DESKRIPSI PEMBELAJARAN

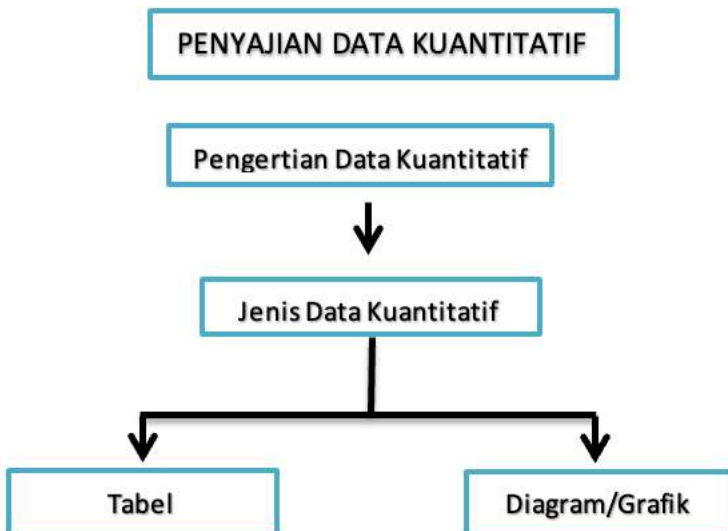
Pada bab ini mahasiswa mempelajari tentang konsep penyajian data dan cara menyajikan data kuantitatif. Diharapkan mahasiswa memiliki pemahaman dan mampu untuk menyajikan data kuantitatif.

KOMPETENSI PEMBELAJARAN

Setelah mengikuti perkuliahan ini diharapkan mahasiswa dan mahasiswi memiliki pengetahuan dan kemampuan:

1. Mampu menguraikan definisi penyajian data
2. Mampu menjelaskan kegunaan penyajian data
3. Mampu menjelaskan bentuk penyajian data

PETA KONSEP PEMBELAJARAN



A. PENGERTIAN PENYAJIAN DATA KUANTITATIF

Penyajian data merupakan salah satu bagian yang sangat penting dalam pembuatan sebuah laporan atau penelitian. Penyajian data berfungsi untuk memberikan gambaran awal dari hasil pengumpulan data, informasi data lebih cepat dimengerti, dan memudahkan proses analisis data. Tahapan penyajian data dimulai dengan editing, koding, kemudian membuat penyajian data tersebut. Editing data merupakan proses memeriksa data mentah yang dikumpulkan untuk mendeteksi kesalahan yang selanjutnya dapat dilakukan koreksi segera. Koding merupakan proses untuk menempatkan angka atau simbol lain pada setiap jawaban sehingga data dapat dimasukkan ke dalam sejumlah kategori atau kelas untuk mempermudah penyajian data.

Penyajian data kuantitatif merupakan penyajian data yang berbentuk angka. Jika data yang berbentuk angka tersebut bervariasi, maka harus dikelompokkan kedalam kelompok angka atau yang biasa disebut dengan kelas, Dimana prosesnya lebih dikenal dengan penyederhanaan data. Data dapat disajikan dalam bentuk tabel dan grafik/diagram. Tabel menyajikan data ke dalam bentuk baris atau kolom sedemikian rupa sehingga memberikan informasi lebih kepada peneliti. Sedangkan grafik menyajikan data dari tabel tersebut menjadi bentuk visual yang lebih informatif lagi. Penyajian data ini tidak hanya sangat membantu peneliti untuk mengetahui gambaran data awal, namun digunakan juga pada analisis inti penelitian atau pelaporan.

B. CARA PENYAJIAN DATA KUANTITATIF

1. Penyajian Data Dalam Bentuk Tabel

Penyajian dalam bentuk tabel dapat merangkum data ke dalam bentuk baris atau kolom. Adapun baris dan kolom tersebut dapat berupa kategori-kategori dan angka frekuensi. Beberapa hal yang diperlukan untuk membuat suatu tabel antara lain:

- a. Judul (*Title*)
- b. Judul Baris (*Stub*)
- c. Judul Kolom (*Box Heading*)
- d. Badan Tabel (*Body*)
- e. Bagian kiri bawah tabel umumnya terdapat bagian untuk menuliskan catatan yang diberikan (jika perlu) atau dapat juga “sumber” yang menjelaskan pengutipan data yang digunakan.
- f. Jika terdapat data tentang waktu, hendaknya disusun secara berurutan.
- g. Jika ada data tentang kategori, hendaknya disusun menurut kebiasaan. Sebagai contoh, laki-laki terlebih dahulu, kemudian wanita; untung terlebih dahulu, kemudian rugi; besar terlebih dahulu, kemudian kecil; dan bagus terlebih dahulu, kemudian jelek atau rusak.

Terdapat 2 jenis tabel, yaitu :

- a. Tabel satu arah (*one-way table*)
Tabel satu arah merupakan tabel yang menunjukkan keterangan mengenai satu variabel atau karakteristik saja. Berikut adalah contoh tabel satu arah.

Tabel 4.1 Jumlah Pasien Berdasarkan Jenis Kelamin yang Memeriksa Kesehatan Di Puskesmas Mawar Pada Bulan Januari Tahun 2024

Jenis Kelamin	Jumlah Responden
Laki-Laki	68
Perempuan	92
Jumlah	160

Sumber : Data Primer, 2024

b. Tabulasi silang (*lebih dari satu arah / two-way table*)

Merupakan metode tabulasi untuk merangkum data dengan dua atau lebih variabel secara bersamaan. Tabel ini bisa dijadikan untuk perbandingan suatu hal. Pada bentuk tabel, sisi (kolom) sebelah kiri dan baris atas menyatakan kelas untuk kedua variabel yang digunakan. Metode tabulasi silang ini dapat digunakan jika :

- 1) Salah Satu Variabel Bersifat Kualitatif Dan Lainnya Kuantitatif
- 2) Kedua Variabel Berupa Variabel Kualitatif
- 3) Kedua Variabel Berupa Variabel Kuantitatif

Tabel 5.2 Jumlah Pasien Berdasarkan Jenis Kelamin dan Usia yang Memeriksa Kesehatan Di Puskesmas Mawar Pada Bulan Januari Tahun 2024

Usia	Jumlah Responden	
	Laki-Laki	Perempuan
< 21 Tahun	30	35
21-30 Tahun	15	16
31-40 Tahun	16	12
40-49 Tahun	34	23
> 49 Tahun	65	24
Jumlah	160	110

Sumber : Data Primer, 2024

2. Penyajian Data Dalam Bentuk Grafik/Diagram

Penyajian data ke dalam tabel yang telah dibahas sebelumnya hanya menunjukkan data berupa angka-angka. Sebagai kelanjutannya, kita dapat membuat penyajian data dari tabel tersebut menjadi grafik-grafik yang lebih menarik dan informatif. Grafik adalah bentuk penyajian data statistik berupa gambar-gambar visual. Oleh karena itu penyajian data menggunakan grafik akan mempermudah dalam penyampaian sebuah data. Pengguna data akan lebih mudah memahami keadaan data

yang ditampilkan melalui grafik daripada tabel karena manusia cenderung lebih menyukai sesuatu yang disampaikan secara visual daripada melalui kata-kata. Penyajian data menggunakan grafik atau diagram merupakan pengembangan dari penyajian data menggunakan tabel. Penyajian data menggunakan grafik haruslah mudah dimengerti oleh pembaca, sebab terkadang pemberian informasi kepada umum hanya disajikan dalam bentuk grafik tanpa adanya tabel sebagai informasi yang mendukung.

a. Diagram Garis (*Line Chart*)

Diagram garis adalah diagram berisi garis atau plot yang menghubungkan titik-titik data dan menunjukkan kuantitas dari data-data tersebut. Diagram garis merupakan suatu grafik yang berupa garis lurus yang menghubungkan titik tengah suatu data dengan data yang lainnya. Dalam diagram garis waktu biasanya ditempatkan pada sumbu datar (sumbu X) dan nilai data pada sumbu tegak (sumbu Y) sehingga diperoleh titik-titik koordinat. Jika titik-titik yang berturut-turut dihubungkan oleh garis lurus, maka akan diperoleh diagram garis.

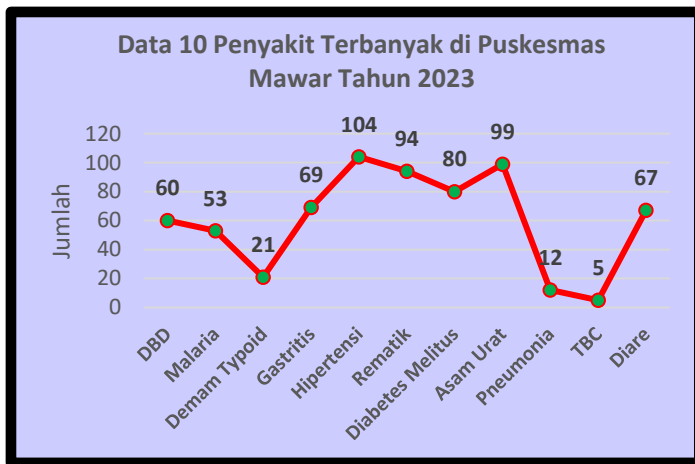
Kelebihan: Diagram garis biasanya digunakan untuk menaksir atau memperkirakan data berdasarkan pola-pola yang telah diperoleh. Diagram garis ada yang tunggal dan majemuk, diagram garis majemuk yaitu dalam satu gambar terdapat lebih dari satu garis. Diagram garis majemuk biasanya digunakan untuk membandingkan dua keadaan atau lebih yang mempunyai hubungan.

Kekurangan: Hanya digunakan untuk data yang berkala, tidak bisa data yang lainnya. Pembaca juga harus teliti dalam membaca diagram ini.

Adapun Langkah-langkah pembuatan diagram garis, adalah

- 1) Menentukan data yang akan diplot ke diagram
- 2) Menuliskan judul diagram garis

- 3) Membuat garis horizontal dan garis vertikal dengan nama variabel dan skala yang sesuai dengan data
- 4) Memplot atau memasukkan data secara satu persatu dengan membuat garis horizontal atau memberikan koordinat titik data
- 5) Menghubungkan semua titik data dengan garis lurus



Gambar 4.3 Diagram Garis

b. Diagram Batang (*Bar Chart*)

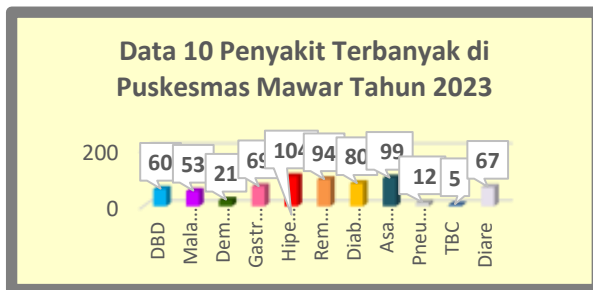
Diagram batang adalah bentuk diagram yang menyajikan data statistik atau informasi dalam bentuk batang persegi panjang. Diagram batang merupakan grafik yang mewakili data dalam bentuk persegi panjang vertikal atau horizontal. Data pada diagram batang bisa digambarkan secara tegak maupun mendatar. Diagram batang memiliki fungsi memudahkan dalam menunjukkan perbandingan antara beberapa kumpulan data yang berbeda. Diagram batang lebih mudah untuk dibaca, dimana pembaca bisa langsung melihat puncak batang, sehingga bisa dengan mudah mengetahui kategori mana yang terbesar dan yang terkecil.

Kelebihan: Paling sederhana dan paling umum.

Kekurangan: diagram batang hanya disajikan data yang telah dikelompokkan berdasarkan kategori. Diagram batang tidak dapat menampilkan data umum dari tiap orang atau benda yang dicatat.

Adapun Langkah-langkah untuk membuat diagram batang pada suatu data ialah:

- 1) Atur sumbu secara horizontal (datar) dan vertikal (tegak).
- 2) Untuk grafik batang vertikal, tulis nilai atau nama data pada sumbu horizontal dan frekuensi setiap nilai atau nama tersebut pada sumbu vertikal. Sedangkan untuk grafik batang horizontal, tulis nilai atau nama data pada sumbu vertikal dan tulis frekuensi masing – masing nilai atau nama tersebut pada sumbu horizontal dan buatlah label pada kedua sumbu.
- 3) Buat persegi panjang yang cocok dengan nilai atau nama setiap data dengan frekuensinya.
- 4) Beri nama diagram balok tersebut berdasarkan data yang disediakan.



Gambar 4.4 Diagram Batang

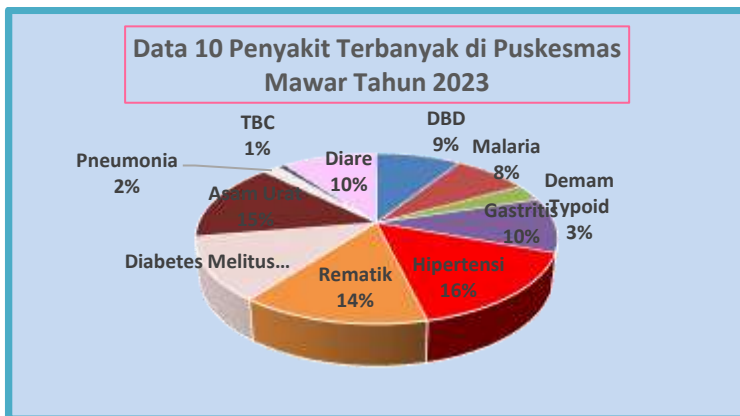
c. Diagram Lingkaran (*Pie Chart*)

Grafik lingkaran merupakan sebuah grafik yang berbentuk lingkaran yang dibagi-bagi menjadi beberapa irisan.

Umumnya digunakan untuk menyajiakan data dalam bentuk persentase. Secara umum, fungsi diagram lingkaran adalah menampilkan persentase data secara keseluruhan berdasarkan periode waktu tertentu. Jika dibandingkan dengan jenis diagram lainnya, fungsi diagram lingkaran bukan untuk menampilkan perubahan data dari waktu ke waktu, melainkan hanya pembagiannya saja pada periode yang telah ditentukan. Tiap bagian menunjukkan kategori data yang terlebih dahulu diubah kedalam persentase.

Kelebihan: Tempat untuk membuat diagram lingkaran tidak terlalu besar. Diagram lingkaran sangat berguna untuk menunjukkan dan membandingkan proporsi dari data.

Kekurangan: Diagram lingkaran tidak dapat menunjukkan frekuensinya karena sudah dirubah kedalam persentase.



Gambar 4.5 Diagram Batang

d. *Piktogram*

Diagram Piktogram adalah penyajian data statistik dengan menggunakan lambang-lambang. Meskipun penyajian data dengan piktogram sederhana, akan tetapi pemakaiannya

sangat terbatas. Biasanya piktogram dipakai untuk menyajikan data yang nilainya cukup besar dengan nilai-nilai data yang telah dibulatkan. Diagram ini sering dipakai untuk mendapatkan gambaran kasar suatu hal dan sebagai alat visual bagi orang awam. Setiap satuan jumlah tertentu dibuat sebuah simbol sesuai dengan macam datanya. Gambar-gambar atau lambang-lambang yang digunakan dibuat semenarik mungkin, sehingga lebih jelas dan mampu mewakili jumlah tertentu untuk satu gambar dan lambang tersebut.

Kelebihan: Lebih mudah dalam menjumlahkan, karena setiap gambar mewakili jumlah tertentu dan tampilan yang lebih menarik.

Kekurangan: Kelemahan dari diagram ini adalah kurang efisien tempat, serta sulit dalam penggambaran untuk nilai yang tidak penuh.



Gambar 4.6 Diagram Piktogram

e. *Kartogram (Peta Statistik)*

Kartogram atau peta statistik digunakan untuk melihat atau menunjukkan lokasi suatu wilayah beserta atribut atau

karakteristiknya. Grafik ini sangat penting digunakan pada analisis data spasial. Kartogram juga digunakan untuk peta yang menunjukkan kepadatan penduduk, curah hujan, hasil pertanian, hasil pertambangan, dsb. Diagram peta atau kartogram merupakan diagram yang digunakan dalam pembuatan peta yang menggunakan suatu perbandingan. Dalam pembuatannya digunakan peta geografis tempat data terjadi. Dengan demikian, kartogram ini melukiskan keadaan dihubungkan dengan tempat kejadiannya.



Gambar 4.7 Diagram Kartogram

C. RANGKUMAN

Berdasarkan uraian di atas, penyajian data kuantitatif merupakan penyajian data yang berbentuk angka yang dapat disajikan dalam bentuk tabel dan grafik/diagram. Tabel terdiri dari tabel satu arah dan tabel dua arah. Grafik/diagram terdiri dari diagram garis, batang, lingkaran, piktogram, dan kartogram. Dari beberapa penjelasan dan contoh penyajian data yang telah dibahas sebelumnya, maka dapat kita simpulkan beberapa manfaat tabel dan grafik, di antaranya: a. Meringkas data. b. Data kuantitatif berupa distribusi frekuensi, persen distribusi frekuensi, diagram. c.

Dapat digunakan untuk melakukan eksplorasi data. d. Membuat tabulasi silang dan diagram sebaran data.

D. TES FORMATIF

1. Berikut ini yang termasuk grafik, kecuali:
 - a. Batang
 - b. Garis
 - c. Lingkaran
 - d. Kartograf
 - e. Dua arah
2. Metode tabulasi untuk merangkum data dengan dua atau lebih variabel secara bersamaan adalah
 - a. Tabel dua arah
 - b. Tabel satu arah
 - c. Diagram batang
 - d. Diagram garis
 - e. Kartograf

E. LATIHAN

Pada bulan Januari tahun 2024 terdapat 100 pasien yang di rawat Inap di Rumah Sakit Melati, berikut adalah rincian jenis penyakit yang diderita: DBD (5 Pasien), DM (25 Pasien), Diare (6 Pasien), Kecelakaan Lalu Lintas (12 Pasien), Kanker (3 Pasien), Stroke (17 Pasien), Hipertensi (32 Pasien). Buatlah data tersebut dalam tabel, diagram batang, dan diagram lingkaran.

KEGIATAN BELAJAR 5

PENYAJIAN DATA KUALITATIF

DESKRIPSI PEMBELAJARAN

Pada bab ini mahasiswa mempelajari pengertian dan konsep dasar, metode dan langkah-langkah dalam penyajian data kualitatif. Diharapkan mahasiswa memiliki wawasan dan kemampuan yang mendalam untuk menyajikan data kualitatif dalam penelitian. Lebih lanjut mahasiswa diharapkan dapat menyajikan data dalam bentuk laporan penelitian yang sesuai dengan sistematika penelitian yang berlaku umum. Selain itu, secara khusus mahasiswa diharapkan mampu untuk menguraikan t dalam suatu penelitian eknik penyajian data dan menyusunnya sesuai sistematika penulisan.

KOMPETENSI PEMBELAJARAN

Setelah mengikuti perkuliahan ini, mahasiswa diharapkan memiliki pengetahuan dan mampu:

1. Menjelaskan pengertian data kualitatif
2. Menjelaskan tujuan dan fungsi penyajian data.
3. Menjelaskan cara dan langkah-langkah penyajian data.

PETA KONSEP PEMBELAJARAN



A. PENGANTAR PENYAJIAN DATA KUALITATIF

Penyajian data merupakan suatu kegiatan dalam penyusunan laporan hasil penelitian sehingga hasil penelitian dapat dipahami dan dianalisis sesuai dengan tujuan yang akan dicapai. Data yang disajikan haruslah sederhana dan jelas agar mudah dibaca dan dipahami. Adapun tujuan penyajian data adalah sebagai berikut:

1. Memberikan gambaran yang sistematis tentang kejadian yang merupakan hasil penelitian atau observasi.
2. Data lebih cepat dipahami dan dimengerti.
3. Memudahkan dalam membuat dan menyusun analisis data.
4. Membuat proses pengambilan keputusan dan kesimpulan lebih tepat, cepat, dan akurat.

Penyajian data dianggap sebagai langkah penting selama analisis data kualitatif atau tahapan tahap penulisan (Burke et al., 2005; Coffey & Atkindo, 1996; Dey, 1993; Eisner, 1997; Grbich, 2007; Lofland, Snow, Anderson, & Lofland, 2006; Miles & Huberman, 1994; Radnofsky, 1996. Slone, 2009; Yin, 2011). Tampilan data dalam format grafik yang merupakan cara untuk menggambarkan informasi secara ringkas dan efisien, mengilustrasikan rincian yang diberikan dalam informasi tekstual yang lebih panjang. Adapun tampilan visual menyediakan ruang multidimensi untuk mengatur data dan menunjukkan hubungan antara berbagai bagian data yang relevan (Dey, 1993). Tujuan utama diagram adalah untuk menyediakan akses siap pakai terhadap informasi dan menyampaikan pesan, penemuan, atau perspektif tertentu mengenai data atau topik tertentu (Iliinsky, 2010; Lengler & Eppler, 2007). Representasi grafis memungkinkan pembaca memperoleh wawasan, mengembangkan pemahaman yang rumit, atau mengapresiasi pengetahuan baru.

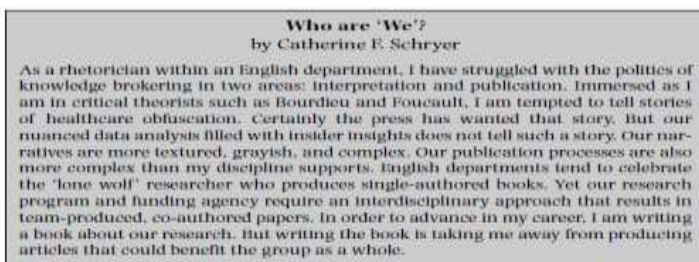
B. JENIS-JENIS PENYAJIAN DATA KUALITATIF

Berikut ini dijelaskan metode penyajian data yang sesuai dengan hasil penelitian dengan pendekatan kualitatif:

1. Tampilan Kotak

Secara harfiah, ini adalah teks yang dibingkai dalam sebuah kotak. Peneliti menggunakan jenis tampilan ini untuk menyorot narasi tertentu yang dianggap cukup penting untuk diekstraksi dari teks dan bingkai dalam sebuah kotak. Dia menekankan kepentingan atau poin relevansi penulis. Pajangan kotak menonjol dan sejauh ini merupakan bentuk pajangan paling sederhana. Saat melakukan tinjauan ini, kami menemukan bahwa sebagian besar tampilan kotak mempertahankan warna latar belakang putih pada artikel, namun beberapa latar belakang diberi bayangan, seperti pada kasus yang dipilih (lihat Gambar 5.1).

Dalam artikel yang mengambil contoh berikut, Lingard, Schryer, Spafford, dan Campbell (2007) menggunakan empat kotak tampilan untuk melaporkan catatan pribadi dan refleksi mereka mengenai penelitian yang mereka lakukan. Hasilnya, jenis tampilan ini membantu mereka memisahkan konseptualisasi teoritis yang lebih luas yang direpresentasikan dalam format teks dari narasi pribadi mereka yang dibingkai dalam tampilan kotak.

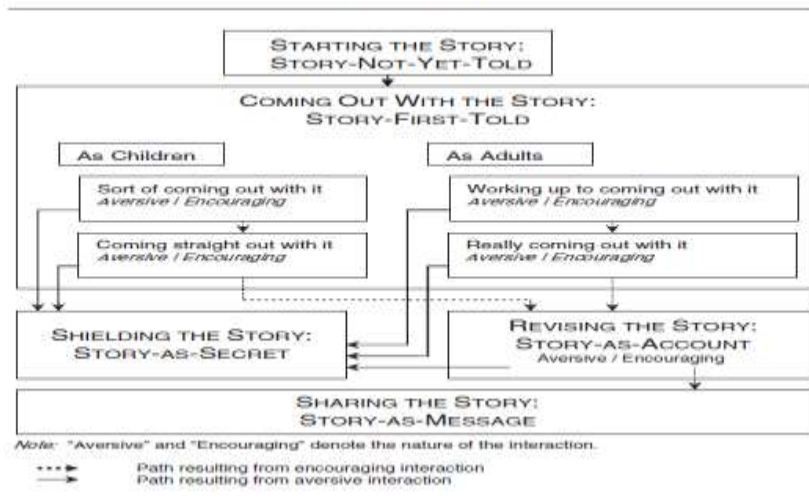


Gambar 5.1 Contoh tampilan kotak
(Sumber: Verdinelli, S., & Scagnoli, N. I, 2013)

digunakan untuk menghubungkan dan/atau menunjukkan arah aliran. Diagram alur dapat memiliki bentuk yang berbeda-beda, namun karakteristik utamanya adalah diagram tersebut mewakili jalur, lintasan, dan proses.

Dalam contoh diagram alur berikut, Draucker dan Martsof (2008) menggambarkan proses inti psikososial dalam menceritakan pelecehan seksual pada masa kanak-kanak (lihat Gambar 3). Persegi panjang mewakili lima proses, hubungan dan arahnya. Para peserta penelitian ini mengklasifikasikan penyampaian cerita sebagai tindakan yang tidak menyenangkan atau memberi semangat, yang mana hal ini sangat penting dalam penceritaan lebih lanjut. Panah mewakili jalur umum untuk memberitahukan lintasan.

The Theoretical Framework "Storying Childhood Sexual Abuse"

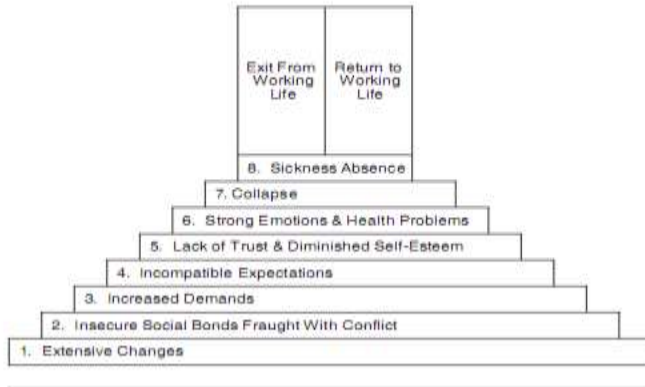


Gambar 5.3. Contoh diagram alur.
 (Sumber: Verdinelli, S., & Scagnoli, N. I, 2013)

4. Tangga atau proses langkah demi langkah

Diagram ini mewakili dimensi perkembangan fenomena tertentu sepanjang waktu. Tangga tersebut menunjukkan fase-fase atau langkah-langkah berbeda yang dialami peserta dalam kondisi tertentu. Representasi langkah-langkah tersebut menunjukkan suatu lompatan atau perubahan dalam proses memperoleh sesuatu. Mirip dengan diagram alur, tangga menunjukkan kemajuan suatu proses, tetapi tangga juga menunjukkan tingkat perubahan. Setiap langkah dalam sebuah tangga mewakili dasar untuk langkah berikutnya. Diagram alur atau jaringan menunjukkan putaran atau siklus proses atau pengalaman, sedangkan tangga bersifat searah. Tidak ada definisi formal yang diberikan untuk jenis diagram ini dalam literatur. Dalam bukunya, Miles dan Huberman (1994) menyajikan diagram yang mereka sebut sebagai “tangga abstraksi”, yang menggambarkan kemajuan analitis langkah demi langkah dari data ke abstraksi. Dalam contoh berikut, Eriksson, Starrin, dan Janson (2008) mempelajari penyebab Para penulis menyajikan temuan mereka sebagai sebuah tangga dengan delapan langkah yang menunjukkan proses kehilangan emosi yang dialami partisipan sebelum kelelahan. Selain sebagai pajangan tangga, contoh ini juga dapat dianggap sebagai pajangan visual metaforis. Inti dari tampilan jenis ini adalah menggunakan simbol umum atau dikenal untuk menyorot secara visual apa yang diwakili. Dalam hal ini, penggunaan tangga secara visual memberi informasi kepada pembaca bahwa ada jalur menanjak atau maju.

The Burnout Stairs: A Step-by-Step Process to Sickness Absence Due to Burnout



Gambar 5.4 Contoh tampilan tangga.
(Sumber: Verdinelli, S., & Scagnoli, N. I, 2013)

5. Matriks

Matriks adalah tabel yang didasarkan pada “klasifikasi silang dua dimensi atau lebih, variabel, ataupun suatu konsep yang relevan dengan topik atau topik yang sedang diminati” (Lofland et al., 2006, hal. 214). Baik matriks maupun tabel, banyak digunakan dengan berbagai maksud dan tujuan mulai dari pendaftaran demografis informasi hingga ilustrasi hasil yang kompleks. Dalam contoh berikut, LeGreco dan Tracy (2009) menggunakan matriks dengan dua kolom dan empat baris untuk menampilkan empat fase penelusuran wacana. Tugas unik dari setiap fase adalah poin-poin lancip. Matriks dioperasikan sebagai tampilan organisasi dan deskriptif dari topik yang diteliti. (lihat Gambar 5.5).

Table 1
Discourse Tracing Methodological Overview

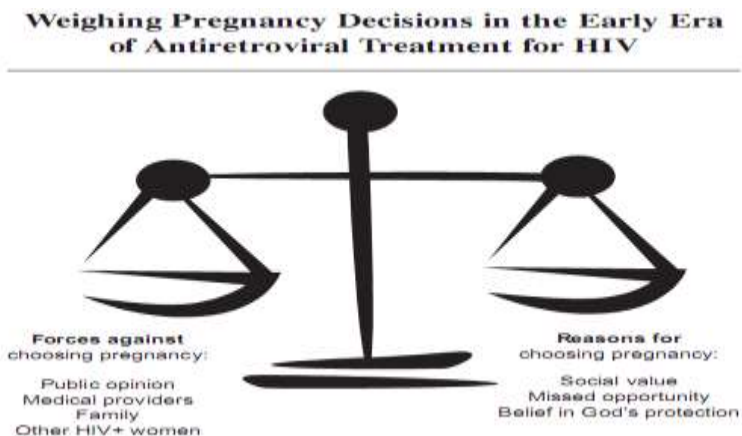
Discourse Tracing Overview	
Phase	Tasks
Phase 1: Research design	• Define the case using a rupture or turning point • Review the literature to outline potential research directions
Phase 2: Data management	• Gather data from a variety of sources that span micro, meso, and macro levels of discourse • Order data chronologically
Phase 3: Data analysis	• Read over this data for emergent themes and issues • Create structured questions based upon literature and emergent themes and apply those questions to the data • Write the case study based upon answers to structured questions and with a focus upon the formation, interpretation, and appropriation of discursive practices
Phase 4: Evaluation	• Address theoretical conclusions of the case • Develop practical implications and recommendations that may apply to other cases

Gambar 5.5 Contoh tampilan matriks.
(Sumber: Verdinelli, S., & Scagnoli, N. I, 2013)

6. Tampilan Visual Metaforis

Jenis tampilan ini menggambarkan sifat topik atau tema yang ditemukan secara metaforis dan terkadang puitis. Tampilan visual metaforis mewakili temuan dan hasil melalui tanda atau simbol umum atau konsensus yang tertanam dalam budaya. Menurut Lengler dan Eppler (2007), visualisasi metafora menyampaikan wawasan yang kompleks menggunakan templat yang sederhana dan umum. Tampilan ini memiliki kekhasan dalam melayani dua fungsi atau menggunakan dua lapisan representasi. Pada satu tingkat, mereka menyajikan informasi secara grafis dengan cara yang terstruktur dan terorganisir, namun dalam. Selain itu, ciri-ciri utama metafora digunakan untuk menyampaikan representasi tingkat kedua, yang memberikan wawasan atau pengetahuan implisit tentang informasi yang diwakili. Mirip dengan jaringan, tampilan visual metaforis menggambarkan model, teori, atau kerangka kerja dan mewakili hubungan antara tema dan subtema.

Pada contoh berikut, Barnes dan Murphy (2009) memilih skala untuk mewakili bagaimana perempuan dengan HIV menimbang atau menilai keputusan mereka untuk hamil (lihat Gambar 6). Jadi, kata-kata yang digunakan dalam tampilan ini menunjukkan satu tingkat representasi dan skala itu sendiri menunjukkan tingkat kedua. Skala tersebut antara lain melambangkan gagasan tentang keadilan, keseimbangan, stabilitas, atau ketidakstabilan.



Gambar 5.6 Contoh tampilan visual metaforis.
(Sumber: Verdinelli, S., & Scagnoli, N. I, 2013)

7. Diagram Venn yang dimodifikasi

John Venn membuat diagram jenis ini sekitar tahun 1880 untuk menunjukkan semua kemungkinan hubungan antar kelompok benda. Ini terdiri dari lingkaran yang tumpang tindih dengan persimpangan yang mewakili area umum di antara mereka ((Baron, 1969). Artikel penelitian kualitatif biasanya menggunakan versi diagram Venn yang dimodifikasi untuk menampilkan aspek bersama dari suatu konsep, kategori, atau proses. Diagram Venn juga telah banyak digunakan untuk menggambarkan model atau kerangka konseptual.

Dalam contoh berikut, Carr (2008) mengeksplorasi pentingnya asuhan keperawatan spiritual (lihat Gambar 7). Penulis menggunakan diagram Venn yang dimodifikasi untuk menggambarkan proses sentral, subproses, dan kualitas yang menggambarkan makna asuhan keperawatan spiritual. Penulis menyatakan bahwa ia memilih lingkaran karena mewakili keutuhan, dan lingkaran yang tumpang tindih dan terputus menggambarkan pengaruh timbal balik dari proses tersebut terhadap satu sama lain serta sifat dinamisnya. Garis putus-putus dalam lingkaran mewakili keterbukaan terhadap pengaruh lingkungan eksternal dari semua proses dan subproses yang terkait dengan asuhan keperawatan spiritual.

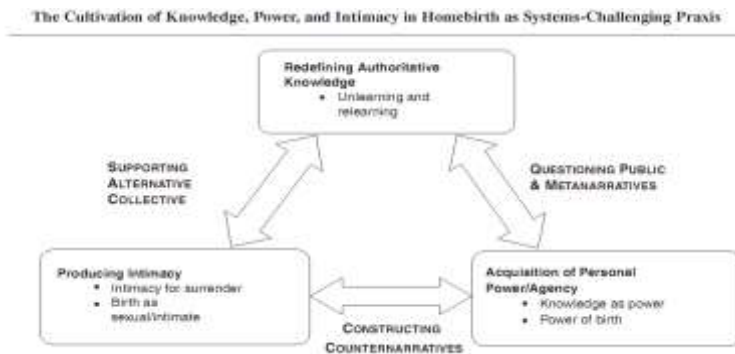


Gambar 5.7 Contoh diagram Venn yang dimodifikasi.
(Sumber: Verdinelli, S., & Scagnoli, N. I, 2013)

8. Jaringan

Jaringan didefinisikan sebagai “kumpulan ‘node’ atau titik-titik yang dihubungkan oleh garis (‘link’)” (Miles & Huberman, 1994,p.94. Jaringan memungkinkan fokus berada pada banyak variabel sekaligus. Di dalam studi kualitatif, jaringan biasanya digunakan untuk menunjukkan kerangka kerja, model ataupun teori yang menunjukkan temuan tentang fenomena yang diteliti. Tampilan jaringan menunjukkan hasil atau analisis konseptual, dan mereka membantu menggambarkan hubungan atau

hubungan antara aspek teoritis yang diteliti. Selain itu, jenis tampilan ini menggambarkan hubungan antara tema dan kategori serta subkategorinya. Pada contoh berikut, Cheyney (2008) menggunakan jaringan untuk merepresentasikan hubungan yang ada di antara tiga tema utama dan dominan yang muncul dari data dan hubungan di antara ketiga tema tersebut (lihat Gambar 5.8).

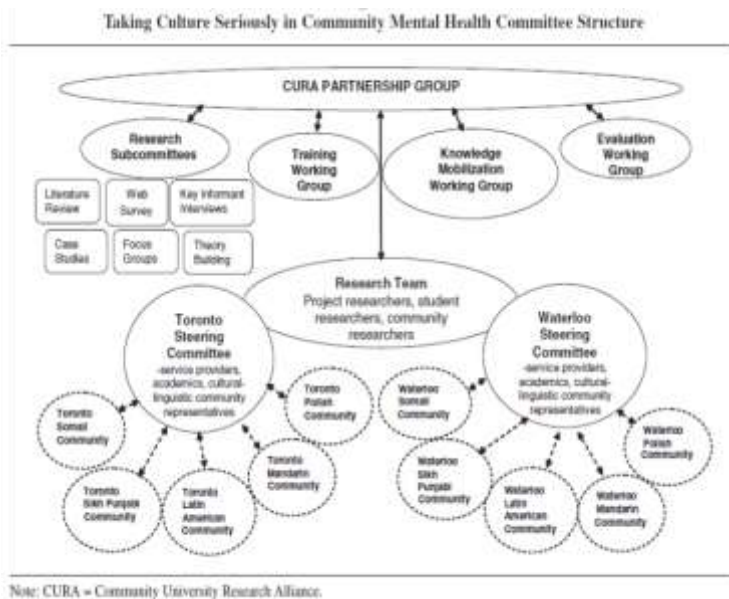


Gambar 5.8 Contoh tampilan jaringan.
(Sumber: Verdinelli, S., & Scagnoli, N. I, 2013)

9. Taksonomi

Taksonomi adalah klasifikasi yang disusun dalam struktur hierarki, diurutkan berdasarkan hubungan supertipe-subtipe. Subtipe mempertahankan semua properti supertipe, dengan properti tambahan atau kendala. Misalnya, bunga mawar adalah sejenis bunga; jadi, setiap mawar adalah bunga, tetapi tidak setiap bunga adalah mawar. Spradley (1980) mengembangkan analisis taksonomi yang menunjukkan bagaimana suatu domain diorganisasikan dan bagaimana aktor, objek, aktivitas—komponen—dari suatu sistem saling berhubungan satu sama lain. Dalam tinjauan kami, ditemukan bahwa taksonomi mewakili berbagai jenis klasifikasi dan digunakan untuk mengklasifikasikan atau mengatur informasi. Penulis menggunakan taksonomi untuk menggambarkan

bagan organisasi tempat pengumpulan data, pendekatan bola salju dalam menghubungkan dengan peserta, dan struktur hierarki organisasi yang sedang diteliti. Dalam contoh berikut, Westhues dkk. (2008) menggunakan tampilan untuk menyebutkan berbagai organisasi dan komite pengarah yang mendukung proyek penelitian mereka (lihat Gambar 5.9).



Gambar 5.9 Contoh sebuah taksonomi.
(Sumber: Verdinelli, S., & Scagnoli, N. I, 2013)

C. RANGKUMAN

Penyajian data merupakan kegiatan dalam pembuatan laporan hasil penelitian yang dapat dipahami dan dianalisis sesuai dengan tujuan yang akan dicapai. Adapun tujuan penyajian data adalah: a. memberikan gambaran yang sistematis tentang kejadian yang merupakan hasil penelitian atau observasi. b. data lebih cepat dipahami dan dimengerti. c. memudahkan dalam membuat suatu analisis data. d. membuat proses pengambilan keputusan dan

kesimpulan lebih tepat, cepat dan akurat. Sedangkan Metode penyajian data yang umum digunakan adakalah metode Tampilan kotak, Pemodelan pohon keputusan, Flowchart, Tangga atau proses langkah demi langkah, Matriks, Tampilan Visual Metaforis, Diagram Venn yang dimodifikasi, Jaringan, atau Taksonomi.

D. TES FORMATIF

1. Penyajian data merupakan proses dalam suatu penelitian berupa....
 - a. Memberikan visualisasi dari hasil penelitian.
 - b. Mengumpulkan data menggunakan metode pengumpulan data
 - c. Mengkaji referensi ilmiah yang tepat terkait topik penelitian
 - d. Mengubah jawaban hasil kuesioner menjadi kode-kode.
2. Berikut ini yang bukan merupakan tujuan penyajian data adalah...
 - a. Memudahkan dalam membuat analisis data menyulitkan
 - b. Data lebih cepat dimengerti.
 - c. Memberikan gambaran yang sistematis tentang kejadian yang merupakan hasil penelitian atau observasi.
 - d. Menyulitkan proses pengambilan keputusan dan kesimpulan lebih tepat, cepat dan akurat.

E. LATIHAN

Dalam penyajian data, ada beberapa metode yang digunakan yaitu menggunakan Tampilan kotak, Pemodelan pohon keputusan, Flowchart, Tangga atau proses langkah demi langkah, Matriks, Tampilan Visual Metaforis, Diagram Venn yang dimodifikasi, Jaringan, atau Taksonomi. Jelaskan masing-masing metode yang digunakan tersebut dan berilah contoh kasusnya.

KEGIATAN BELAJAR 6

TENDENSI SENTRAL

DESKRIPSI PEMBELAJARAN

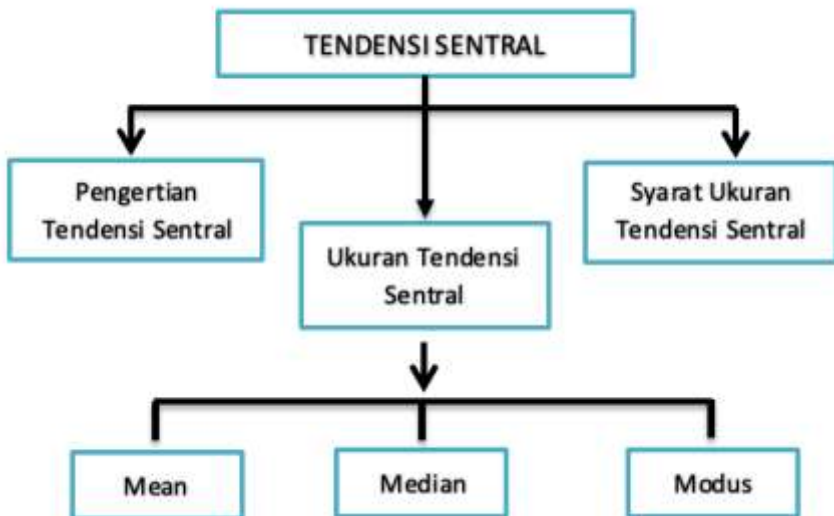
Pada bab ini mahasiswa mempelajari ukuran Tendensi Sentral. Diharapkan mahasiswa memiliki pemahaman dan kemampuan untuk menghitung nilai-nilai Tendensi Sentral.

KOMPETENSI PEMBELAJARAN

Setelah mengikuti perkuliahan ini diharapkan mahasiswa dan mahasiswi memiliki pengetahuan dan kemampuan:

1. Mampu menguraikan definisi ukuran Tendensi Sentral
2. Mampu menjelaskan syarat ukuran Tendensi Sentral
3. Mampu melakukan penghitungan Mean, Median, Modus pada data tunggal dan data kelompok

PETA KONSEP PEMBELAJARAN



A. PENGERTIAN TENDENSI SENTRAL

Ukuran pemusatan merupakan salah satu bentuk analisis statistik Deskriptif untuk Data atau Variabel yang berskala Numerik (Interval atau Rasio). Tendensi Sentral adalah pengukuran statistik untuk menentukan skor tunggal yang menetapkan pusat dari distribusi. Data yang disajikan dengan ukuran-ukuran Tendensi Sentral akan lebih mudah dipahami dibandingkan dengan data dalam keadaan tersebar. Pengukuran Tendensi Sentral meliputi mean, median dan modus.

B. SYARAT-SYARAT TENDENSI SENTRAL

1. Harus dapat mewakili rangkaian data
2. Perhitungannya harus didasarkan pada seluruh data
3. Perhitungannya harus objektif
4. Perhitungannya harus mudah
5. Dalam suatu rangkaian hanya ada 1 nilai sentral.

C. MEAN

Mean biasa diterjemahkan dengan rata-rata atau rerata. Istilah Mean akan tetap dipakai disini oleh karena sudah lazim digunakan dalam statistik. Istilah mean saja merupakan metode yang paling banyak digunakan untuk menggambarkan ukuran Tendensi Sentral. Untuk menentukan atau menghitung Mean dibedakan menjadi 2 macam, yaitu:

1. Mean Data Tunggal

Penghitungan nilai Mean pada data tunggal ini dapat dibedakan lagi menjadi 2 macam, yaitu data tunggal dengan frekuensi =1 (satu) dan data tunggal dengan frekuensi >1 (lebih dari satu).

- a. Mean pada data tunggal dengan frekuensi =1

Untuk menghitung nilai rata-rata atau Mean pada data tunggal dengan frekuensi yang sama dengan 1 (=1) dapat digunakan rumus:

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

Keterangan:

$\sum X$: jumlah nilai
 N : banyaknya frekuensi
 $\bar{X}$: rata-rata

Contoh:

Data berat badan 7 orang ibu hamil dalam tabel distribusi frekuensi sebagai berikut:

Berat Badan Ibu Hamil (X)	Frekuensi (f)
45	1
50	1
55	1
60	1
65	1
$\sum X = 275$	$N = 5$

Berdasarkan data pada tabel, maka nilai rata-rata atau Mean data tersebut adalah $275/5 = 55$

- b. Mean pada data tunggal dengan frekuensi >1

Untuk menghitung nilai rata-rata atau Mean pada data tunggal dengan frekuensi yang lebih dari 1 (>1) dapat digunakan rumus:

$$\bar{X} = \frac{\sum fX}{\sum f}$$

Keterangan:

$\sum f$: jumlah total frekuensi

$\sum fX$: jumlah frekuensi x nilai

$\bar{X}$: rata-rata

Contoh:

Data distribusi berat badan 50 orang ibu hamil dalam tabel distribusi frekuensi sebagai berikut:

Berat Badan Ibu Hamil (X)	Frekuensi (f)	f(X)
45	8	360
50	11	550
55	9	495
60	5	300
65	7	455
	$\sum f = 40$	$\sum fX = 2160$

Berdasarkan data pada tabel, maka nilai rata-rata atau Mean data tersebut adalah $2160/40 = 54$.

2. Mean Data Kelompok

Menghitung nilai rata-rata atau Mean pada data kelompok (interval) dimulai dengan terlebih dahulu menghitung atau mengetahui nilai atau titik tengah (midpoint) data tiap kelas. Untuk menghitung atau menentukan midpoint pada tiap kelas adalah dengan menggunakan $\frac{1}{2}$ dari nilai batas kelas atas + batas kelas bawah. Untuk menghitung nilai rata-rata atau Mean pada data kelompok dapat digunakan rumus:

$$\bar{X} = \frac{\sum fX}{\sum f}$$

Keterangan:

$\sum f$: jumlah total frekuensi

X : nilai midpoint tiap kelas

$$X = \frac{\text{nilai batas kelas atas} + \text{nilai batas kelas bawah}}{2}$$

$\sum fX$: jumlah frekuensi x nilai midpoint tiap kelas

$\bar{X}$: rata-rata

Contoh:

Data distribusi berat badan 100 orang ibu hamil dalam tabel distribusi frekuensi sebagai berikut:

Kelas	Frekuensi	Midpoint (X)	fX
50-54	11	52	572
55-59	12	57	684
60-64	23	62	1426
65-69	24	67	1608
70-74	21	72	1512
75-79	9	77	693
	$\sum f = 100$		$\sum fX = 6495$

Berdasarkan data pada tabel, maka nilai rata-rata atau Mean data tersebut adalah $6495/100 = 64,9$.

D. MEDIAN

Median (nilai tengah), adalah suatu nilai yang membatasi 50% dari frekuensi distribusi sebelah atas dan 50% frekuensi distribusi sebelah bawah atau merupakan nilai tengah dari rangkaian data yang telah tersusun secara teratur. Atau sebagai ukuran letak, karena median membagi distribusi menjadi 2 bagian yang sama. Median menentukan letak data setelah data itu disusun menurut urutan nilainya. Untuk menentukan atau menghitung Median dibedakan menjadi 2 macam, yaitu:

1. Median Data Tunggal

Langkah-langkahnya antara lain:

- Urutkan data dari terkecil ke terbesar atau dari terbesar ke terkecil.
- Tentukan letak median dengan formulasi $\frac{(N+1)}{2}$

- c. Untuk kasus jumlah data ganjil, nilai tengah dari observasi yang sudah diurutkan merupakan nilai median sementara untuk kasus jumlah data genap, nilai median merupakan rata-rata dari dua data yang berada pada letak median untuk data yang sudah diurutkan.

Rumus posisi median data jumlah ganjil:

$$\frac{N + 1}{2}$$

Keterangan:

N : jumlah data (jumlah frekuensi)

Contoh:

Jumlah data ganjil

Berat Badan Ibu Hamil (X)	Frekuensi (f)
45	1
50	1
55	1
60	1
65	1
Total	$\Sigma f = 5$

Berdasarkan data pada tabel, maka nilai Median pada data tunggal dengan jumlah data ganjil tersebut adalah $(5+1)/2 = 3$. Berdasarkan hasil perhitungan posisi median, maka nilai Median data tersebut adalah nilai yang terletak pada urutan ke 3, yaitu 55.

Rumus posisi Median data jumlah genap:

Keterangan:

N : jumlah data (jumlah frekuensi)

Jumlah data genap

Berat Badan Ibu Hamil (X)	Frekuensi (f)
45	1
50	1

55	1
60	1
70	1
75	1
Total	$\Sigma f = 6$

Berdasarkan data pada tabel, maka nilai Median pada data tunggal dengan jumlah data genap tersebut adalah:

$$\begin{aligned}
 & (X_{6/2} + X_{(6+2)/2})/2 \\
 & = (X_3 + X_4)/2 \\
 & = (55 + 60)/2 \\
 & = 57,5
 \end{aligned}$$

Berdasarkan hasil perhitungan posisi median, maka nilai Median data tersebut adalah 57,5.

2. Median Data Kelompok

Untuk menentukan Median pada data kelompok dapat dilakukan dengan rumus berikut:

$$Mdn = B + p \left[\frac{\frac{1}{2}n - F}{f} \right]$$

Keterangan:

B : batas bawah kelas median (dimana median akan terletak)

p : panjang kelas median

n : ukuran sampel (banyaknya data)

f : frekuensi kelas median (diambil berdasarkan frekuensi terbanyak)

F: jumlah semua frekuensi dengan tanda kelas lebih kecil dari kelas Median

Langkah-langkah yang harus dilakukan untuk menentukan Median pada data kelompok adalah:

- Menentukan frekuensi kelas Median (f) dengan cara: melihat jumlah frekuensi yang terbanyak/terbesar.

- b. Menentukan kelompok atau kelas atau interval dimana median akan terletak.
- c. Menentukan panjang kelas median (p) yaitu menghitung jarak antara batas bawah sampai batas atas kelas median.
- d. Menentukan batas bawah kelas median (B), karena ini adalah data kelompok, maka cara menentukan batas bawah kelas median adalah dengan menjumlah batas atas kelas sebelum kelas median dengan batas bawah kelas median dibagi dua.
- e. Menentukan jumlah semua frekuensi dari kelas atau interval atau kelompok yang lebih kecil dari kelas median (F) yaitu dengan menjumlahkan semua frekuensi (f) pada kelompok atau kelas atau interval yang lebih kecil atau sebelum kelas median.
- f. Mengidentifikasi jumlah data atau banyaknya data atau sampel (n).

Contoh:

Data distribusi nilai mata kuliah Biostatistik 30 orang mahasiswa dalam tabel distribusi frekuensi sebagai berikut:

Kelas	Frekuensi
39-44	4
45-50	4
51-56	7
57-62	6
63-68	4
69-74	5
Total	$\sum f = 30$

Untuk dapat menentukan Median pada data tersebut, maka harus dilakukan langkah-langkah sebagai berikut:

- a. Menentukan frekuensi kelas Median (f) = 7
- b. Menentukan kelompok atau kelas atau interval dimana Median akan terletak = interval 51-56
- c. Menentukan panjang kelas Median (p) = 6

- d. Menentukan batas bawah kelas Median $(B) = (50+51)/2 = 50,5$
- e. Menentukan jumlah semua frekuensi dari Kelas atau Interval atau Kelompok yang lebih kecil dari kelas Median $(F) = (4+4) = 8$
- f. Mengidentifikasi jumlah data atau banyaknya data atau sampel $(n) = 30$

Selanjutnya angka-angka tersebut dimasukkan ke dalam rumus:

$$\begin{aligned}
 \text{Mdn} &= 50,5 + 6 \{(1/2 (30)-8)\}/7 \\
 &= 50,5 + 6 \{(7)/7\} \\
 &= 50,5 + 6 \\
 &= 56,5
 \end{aligned}$$

Jadi Median dari data tersebut adalah 56,5.

E. MODUS

Modus, merupakan nilai data yang memiliki frekuensi terbesar atau dengan kata lain, nilai data yang paling sering terjadi. Ukuran ini juga dalam keadaan tidak disadari sering dipakai untuk menentukan rata-rata data kualitatif. Modus juga dibedakan menjadi 2 macam, yaitu Modus pada data tunggal dan Modus pada data kelompok.

1. Modus Data Tunggal

Modus pada data tunggal dapat ditentukan berdasarkan data yang jumlah frekuensinya paling banyak.

Contoh:

Data berat badan 7 orang ibu hamil dalam tabel distribusi frekuensi sebagai berikut:

Berat Badan Ibu Hamil (X)	Frekuensi (f)
45	5
50	3
55	3
60	9

65	7
70	2
75	4
Total	$\sum f = 33$

Berdasarkan data yang mempunyai frekuensi terbanyak, yaitu 9 adalah nilai 60. Jadi nilai $M_o = 60$.

2. Modus Data Kelompok

Modus pada Data Kelompok dapat ditentukan dengan cara menggunakan Rumus berikut ini:

$$M_o = b + p \left[\frac{b_1}{b_1 + b_2} \right]$$

Keterangan:

Frekuensi kelas modus ditentukan berdasarkan jumlah frekuensi data yang terbanyak.

b: batas bawah kelas modus (yaitu: kelas dimana modus akan terletak)

p: panjang kelas modus.

b₁ : frekuensi kelas modus dikurangi frekuensi kelas interval yang mempunyai tanda kelas lebih kecil (sebelum) kelas modus.

b₂ : frekuensi kelas modus dikurangi frekuensi kelas interval yang mempunyai tanda kelas lebih besar (sesudah) kelas modus.

Contoh:

Data distribusi nilai mata kuliah Biostatistik 30 orang mahasiswa dalam tabel distribusi frekuensi sebagai berikut:

Kelas	Frekuensi
39-44	4

45-50	4
51-56	7
57-62	6
63-68	4
69-74	5
Total	$\sum f = 30$

Langkah-langkah untuk dapat menentukan Modus pada data tersebut, adalah sebagai berikut:

- Menentukan frekuensi kelas Modus (f) = 7
- Menentukan kelompok atau kelas atau interval dimana modus akan terletak = interval 51-56
- Menentukan panjang kelas modus (p) = 6
- Menentukan batas bawah kelas modus (b) = $(50+51)/2 = 50,5$
- Menentukan b_1 , yaitu frekuensi kelas modus dikurangi frekuensi kelas interval yang mempunyai tanda kelas lebih kecil (sebelum) kelas modus = $7-4 = 3$
- Menentukan b_2 , yaitu : frekuensi kelas modus dikurangi frekuensi kelas interval yang mempunyai tanda kelas lebih besar (sesudah) kelas modus = $7-6 = 1$

Selanjutnya angka-angka tersebut dimasukkan ke dalam rumus:

$$mo = 50,5 + 6 (3/4)$$

$$= 50,50 + 4,5$$

$$= 55$$

Jadi modus atau Mo dari data tersebut di atas adalah 55.

F. RANGKUMAN

Nilai sentral atau Tendensi Sentral adalah nilai dalam rangkaian data yang mewakili rangkaian data tersebut. Disebut juga sebagai

ukuran letak/lokasi karena menunjukkan letak dari pusat atau sekumpulan data. Ukuran Pemusatan atau Tendensi Sentral terdiri atas Mean, Median dan Modus. Masing-masing ukuran tersebut memiliki sifat dan kekhususan sendiri, sehingga penggunaannya harus disesuaikan dengan tujuan analisa data. Penghitungan atau penentuan Mean, Median dan Modus dapat dilakukan baik pada data tunggal maupun data kelompok, yang masing-masing mempunyai rumus dan ketentuan yang berbeda satu sama lainnya.

G. TES FORMATIF

1. Penghitungan ukuran Tendensi Sentral dapat dilakukan pada data
 - a. Ganjil
 - b. Tunggal
 - c. Kelompok
 - d. C dan D benar
2. Yang termasuk ukuran Tendensi Sentral adalah....
 - a. Mean
 - b. Median
 - c. Modus.
 - d. Benar semua

H. LATIHAN

Tentukan Mean, Median dan Modus dari data tentang distribusi data berikut:

Kelas	Frekuensi
25-29	23
30-34	20
35-39	27
40-44	16

45-49	10
50-54	31
55-59	22
60-64	23

KEGIATAN BELAJAR 7

PROBABILITAS

DESKRIPSI PEMBELAJARAN

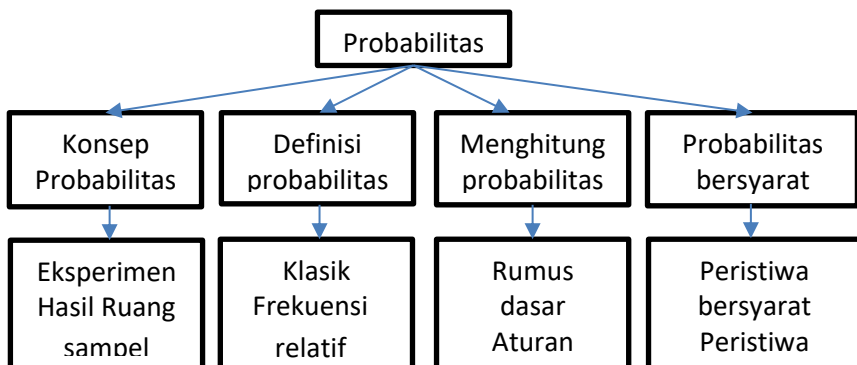
Dalam bab ini mahasiswa akan mempelajari tentang konsep-konsep dasar probabilitas; beberapa definisi probabilitas seperti definisi klasik, definisi frekuensi relatif dan definisi aksiomatis; cara menghitung probabilitas suatu peristiwa menggunakan rumus-rumus dasar dan aturan penjumlahan probabilitas; probabilitas bersyarat dan peristiwa-peristiwa bebas.

KOMPETENSI PEMBELAJARAN

Setelah mempelajari materi ini diharapkan para mahasiswa memiliki pengetahuan dan kemampuan:

1. Menjelaskan konsep-konsep dasar probabilitas
2. Menyebutkan definisi probabilitas dan mencari besarnya probabilitas suatu peristiwa
3. Menyebutkan definisi dan menggunakan teori probabilitas bersyarat
4. Menentukan adanya peristiwa-peristiwa yang bebas atau tidak bebas

PETA KONSEP PEMBELAJARAN



A. KONSEP-KONSEP DASAR PROBABILITAS

Kata probabilitas atau kemungkinan atau peluang sering ditemui dalam kehidupan sehari-hari. Hal ini terlihat saat seseorang menyelesaikan masalah yang memuat unsur ketidakpastian, dimana berbagai hasil dapat terjadi tetapi hasil yang pasti tak dapat ditentukan sebelumnya dengan tepat. Misalnya: apakah hari ini cuaca akan cerah? berapa besar peluang tim sepak bola Indonesia akan menang melawan tim Brunei? apakah saya bisa lulus tes penerimaan CPNS? Biasanya digunakan kata kemungkinan untuk merepresentasikan secara subjektif muncul atau terjadinya suatu peristiwa. Derajat keyakinan terhadap kemungkinan muncul atau terjadinya suatu peristiwa sangat ditentukan oleh ada tidaknya pengetahuan tentang peristiwa tersebut. Ketika derajat keyakinan tersebut dibutuhkan dalam bentuk kuantitas, maka dibutuhkan suatu ukuran. Seluruh konsep mengenai kemungkinan muncul atau terjadinya suatu peristiwa dikembangkan menjadi teori peluang atau teori probabilitas.

Ada beberapa istilah yang sering digunakan terkait dengan probabilitas, antara lain:

1. Eksperimen adalah proses yang menghasilkan hasil pengamatan, perhitungan, atau pengukuran.

Contoh:

- Melempar dua keping uang logam.
- Mengukur tinggi seseorang
- Mengamati penambahan berat badan bayi setiap bulan di suatu puskesmas.

2. Hasil dari suatu eksperimen adalah hasil yang muncul atau terjadi jika eksperimen dilakukan.

Contoh:

- Misalkan eksperimen adalah melempar 2 keping uang logam sekaligus. Hasil yang mungkin terjadi adalah MM (dapat Muka pada keping logam1 dan Muka pada keping logam2), atau MB (dapat Muka pada keping logam1 dan

Belakang pada keping logam2), atau BM (dapat Belakang pada keping logam1 dan Muka pada keping logam2) atau BB (dapat Belakang pada keping logam1 dan Belakang pada keping logam2).

- Pada eksperimen mengukur tinggi badan seorang mahasiswa, hasil eksperimennya adalah ukuran antara 150 cm dan 180 cm, misalkan 160,5 cm.
3. Ruang Sampel (ditulis S) suatu eksperimen adalah himpunan semua hasil yang muncul atau terjadi dari suatu eksperimen yang diberikan.

Contoh:

- Misalkan eksperimen melempar dua keping uang logam sekaligus, maka

$$S = \{MM, MB, BM, BB\}$$

- Misalkan eksperimen mengukur tinggi badan seorang mahasiswa, maka

$$S = \{x/150 < x < 180\}$$

Menurut banyaknya hasil yang mungkin terjadi, ruang sample S dibedakan atas dua macam, yaitu ruang sampel diskrit (contoh 1) dan ruang sampel kontinu (contoh 2).

4. Peristiwa adalah suatu himpunan hasil atau himpunan bagian dari S .

Contoh:

Pada eksperimen melempar sebuah dadu satu kali dimana

A = peristiwa memperoleh mata genap

B = peristiwa memperoleh mata ganjil

C = peristiwa memperoleh mata yang < 6

D = peristiwa memperoleh mata yang ≥ 3

maka diperoleh $A = \{2, 4, 6\}$

$$B = \{1, 3, 5\}$$

$$C = \{1, 2, 3, 4, 5\}$$

$$D = \{3, 4, 5, 6\}$$

B. DEFINISI PROBABILITAS

Berdasarkan perbedaan pendekatan yang digunakan oleh para ahli, ada beberapa definisi probabilitas, yaitu:

1. Definisi Klasik

Misalkan suatu eksperimen menghasilkan N hasil yang tidak mungkin terjadi bersama-sama dan masing-masing mempunyai probabilitas yang sama. Misalkan pula H adalah suatu peristiwa (himpunan bagian dari S) yang mempunyai hasil sebanyak $n(H)$. Maka probabilitas terjadinya peristiwa H didefinisikan sebagai:

$$P(H) = \frac{n(H)}{N} \dots\dots\dots (1)$$

Contoh:

- a. Pada eksperimen melempar sekeping uang logam 3 kali, dengan K adalah peristiwa mendapat sisi muka (M) pada lemparan kedua dan L adalah peristiwa mendapat hasil ketiga lemparan sama, diperoleh:

$S = \{MMM, MBM, MMB, BMM, BMB, MBB, BBM, BBB\}$

$K = \{MMM, MMB, BMM, BMB\} \rightarrow n(K) = 4$

$L = \{MMM, BBB\} \rightarrow n(L) = 2$

sehingga

$P(K) = \frac{n(K)}{N} = \frac{4}{8} = \frac{1}{2}$

$P(L) = \frac{n(L)}{N} = \frac{2}{8} = \frac{1}{4}$

- b. Sebungkus permen berisi 9 rasa coklat, 4 rasa mangga, dan 7 rasa melon. Bila diambil satu permen secara acak, hitunglah probabilitasnya memperoleh:

- satu rasa coklat
- satu rasa mangga atau melon

Jawab:

Misalkan A , B , dan C menyatakan peristiwa yang terambil adalah masing-masing rasa coklat, mangga, dan melon. Jumlah permen (N) = 20, $n(A)$ = 9, $n(B)$ = 4, $n(C)$ = 7.

Semuanya mempunyai probabilitas yang sama untuk terambil.

- Karena ada 9 rasa coklat dari 20 permen, maka probabilitas terambil rasa coklat adalah: $P(A) = \frac{n(A)}{N} = \frac{9}{20}$
- Karena ada 11 rasa mangga atau melon dari 20 permen, maka probabilitas terambil rasa mangga atau melon adalah:

$$P(B \cup C) = \frac{11}{20}$$

2. Definisi Frekuensi Relatif

Misalkan bahwa suatu eksperimen dengan ruang sampel S diulang-ulang dibawah kondisi yang sama. Bila n bertambah besar sampai tak hingga ($n \rightarrow \infty$), untuk setiap peristiwa H yang merupakan himpunan bagian dari S , probabilitas peristiwa H adalah nilai limit dari frekuensi relatif peristiwa H . Dengan demikian, jika peristiwa H terjadi sebanyak $n(H)$ kali dari keseluruhan eksperimen sebanyak n , dimana n mendekati tak hingga, maka probabilitas peristiwa H , didefinisikan sebagai:

$$P(H) = \lim_{n \rightarrow \infty} \frac{n(H)}{n} \dots\dots\dots (2)$$

Contoh:

Suatu eksperimen melempar sebuah dadu sebanyak 1000 kali. Frekuensi munculnya mata dadu disajikan pada tabel berikut:

Tabel. 7.1 Eksperimen Melempar Sebuah Dadu

Mata Dadu	1	2	3	4	5	6
Frekuensi	155	166	185	166	160	168

Bila H menyatakan peristiwa munculnya mata-mata dadu tersebut, maka probabilitas peristiwa H untuk masing-masing mata dadu adalah:

$$P(H) = P(1) = \frac{155}{1000}$$

$$P(H) = P(2) = \frac{166}{1000}$$

$$P(H) = P(3) = \frac{185}{1000}$$

$$P(H) = P(4) = \frac{166}{1000}$$

$$P(H) = P(5) = \frac{160}{1000}$$

$$P(H) = P(6) = \frac{168}{1000}$$

3. Definisi Aksiomatis

Misalkan suatu eksperimen dengan ruang sampel S . $\mathcal{A}$ himpunan semua peristiwa dan P suatu fungsi bernilai nyata pada $\mathcal{A}$ dengan daerah hasil $[0,1]$. Dikatakan bahwa $P(H)$ adalah probabilitas terjadinya peristiwa H , jika dipenuhi:

- $0 \leq P(H) \leq 1$, untuk setiap $H \in \mathcal{A}$
- $P(S) = 1$
- Untuk sembarang peristiwa-peristiwa $H_1, H_2, \dots$, saling asing ($H_i \cap H_j = \emptyset$, untuk $i \neq j$), maka

$$P(\cup_{i=1}^{\infty} H_i) = \sum_{i=1}^{\infty} P(H_i) \dots\dots\dots (3)$$

Contoh:

Misalkan suatu eksperimen menghasilkan empat hasil yang mungkin (A, B, C, dan D) dan saling asing dengan $P(A) = 0,21$, $P(B) = 0,54$, $P(C) = 0,10$, $P(D) = 0,25$. Apakah P merupakan fungsi peluang?

Jawab:

Walaupun sifat (1) dan (3) dipenuhi tetapi sifat (2) tidak dipenuhi karena $P(S) = P(A \cup B \cup C \cup D)$

$$= 0,21 + 0,54 + 0,10 + 0,25 = 1,10 > 1$$

maka P bukan fungsi peluang.

C. MENGHITUNG PROBABILITAS SUATU PERISTIWA

Berikut ini akan disajikan rumus-rumus probabilitas yang banyak digunakan untuk melakukan perhitungan-perhitungan probabilitas suatu peristiwa.

Rumus-Rumus Dasar

1. Probabilitas peristiwa mustahil adalah nol, ditulis:

$$P(\emptyset) = 0$$

Hal ini dapat dibuktikan dengan memisalkan A sembarang peristiwa dalam S . Maka $A \cup \emptyset = A$

$$P(A \cup \emptyset) = P(A)$$

$$P(A) + P(\emptyset) = P(A)$$

Dengan mengurangi kedua ruas dengan $P(A)$, maka diperoleh $P(\emptyset) = 0$

2. Jika peristiwa A menghindarkan terjadinya peristiwa A^c , maka

$$P(A^c) = 1 - P(A)$$

Peristiwa ini disebut juga sebagai peristiwa komplementer.

3. Jika A dan B sembarang peristiwa, dan $A \subset B$, maka

$$P(A) \leq P(B)$$

Karena $B = B \cap S = B \cap (A \cup A^c)$, maka dengan hukum distributif diperoleh

$$B = (B \cap A) \cup (B \cap A^c) \text{ dengan } (B \cap A) \cap (B \cap A^c) = \emptyset$$

$$P(B) = P(B \cap A) + P(B \cap A^c)$$

Karena $A \subset B$, maka $B \cap A = A$, $P(B \cap A) = P(A)$

Sehingga diperoleh $P(B) = P(A) + P(B \cap A^c)$, yang berarti $P(A) \leq P(B)$

4. Jika A dan B sembarang peristiwa, maka

$$P(A) = P(A \cap B) + P(A \cap B^c),$$

Karena $A = A \cap S = A \cap (B \cup B^c)$, maka dengan hukum distributif diperoleh $A = (A \cap B) \cup (A \cap B^c)$. Berdasarkan sifat (3) definisi aksiomatis diperoleh:

$$P(A) = P(A \cap B) + P(A \cap B^c)$$

Aturan Penjumlahan Probabilitas

Jika A dan B dua peristiwa sembarang yang saling beririsan, maka probabilitas terjadinya kedua peristiwa A atau B secara matematis dirumuskan sebagai berikut:

$$P(A \cup B) = P(A) + P(B) - P(A \cap B) \quad \dots\dots\dots (4)$$

Karena $A \cup B$ dapat dituliskan sebagai gabungan dua peristiwa A dan $(A^c \cap B)$ yang saling asing,

$$A \cup B = A \cup (A^c \cap B)$$

sehingga berdasarkan sifat (3) definisi aksiomatis diperoleh

$$P(A \cup B) = P(A) + P(A^c \cap B)$$

Telah dibuktikan bahwa $P(A) = P(A \cap B) + P(A \cap B^c)$ dan

$$P(A^c \cap B) = P(B) - P(A \cap B)$$

dengan demikian

$$P(A \cup B) = P(A) + P(B) - P(A \cap B)$$

Akibatnya: Jika A dan B dua peristiwa yang saling asing (kedua peristiwa A dan B tidak memiliki irisan), maka probabilitas terjadinya kedua peristiwa A atau B adalah jumlah dari probabilitas A dan probabilitas B. Secara matematis ditulis sebagai berikut:

$$P(A \cup B) = P(A) + P(B) \quad \dots\dots\dots (5)$$

Contoh:

1. Probabilitas seorang mahasiswa lulus statistika $\frac{1}{3}$ dan probabilitasnya lulus biologi $\frac{3}{4}$. Bila probabilitasnya lulus kedua mata kuliah $\frac{1}{4}$, berapakah probabilitasnya lulus paling sedikit satu mata kuliah?

Jawab: Misalkan M adalah peristiwa lulus statistika dan B menyatakan peristiwa lulus biologi,

$$P(M) = \frac{1}{3}, P(B) = \frac{3}{4}, \text{ dan } P(M \cap B) = \frac{1}{4},$$

$$\text{Diperoleh, } P(M \cup B) = P(M) + P(B) - P(M \cap B) = \frac{4+9-3}{12} = \frac{5}{6}$$

2. Berapakah peluangnya memperoleh jumlah 7 atau 10 jika dua dadu dilemparkan?

Jawab: Misalkan: A = peristiwa muncul jumlah 7

B = peristiwa muncul jumlah 10

$S = \{(1,1), (1,2), (1,3), (1,4), (1,5), \dots, (5,6), (6,6)\};$

$N = 36$

$A = \{(1,6), (2,5), (3, 4), (4,3), (5,2), (6,1)\}; P(A) = \frac{6}{36}$

$B = \{(4,6), (5,5), (6,4)\}; P(B) = \frac{3}{36}$

A dan B merupakan kejadian yang saling asing karena jumlah 7 dan 10 tidak mungkin terjadi pada lemparan yang sama

$(A \cap B = \{ \quad \})$, sehingga,

$$P(A \cup B) = P(A) + P(B) = \frac{6}{36} + \frac{3}{36} = \frac{9}{36} = \frac{1}{4}$$

D. PROBABILITAS BERSYARAT

1. Peristiwa Bersyarat

Pada perhitungan probabilitas sebelumnya, perhatian kita adalah pada probabilitas terjadinya suatu peristiwa jika suatu eksperimen dilakukan. Namun dalam praktek, masalah yang sering muncul adalah menghitung probabilitas terjadinya suatu peristiwa apabila peristiwa yang lain sudah terjadi. Misalnya dalam eksperimen melempar 2 buah dadu yang setimbang, dan misalkan dadu pertama muncul mata 4. Maka dengan diketahuinya informasi ini, berapa probabilitas bahwa jumlah kedua mata dadu sama dengan 9? Jika K menyatakan peristiwa diperolehnya jumlah mata dadu 9 sedangkan L menyatakan peristiwa bahwa dadu pertama muncul mata 4, maka probabilitas yang baru diperoleh tersebut dinamakan probabilitas bersyarat terjadinya K bila diketahui L terjadi.

Misalkan K dan L sembarang peristiwa dalam ruang sampel S. Probabilitas bersyarat dari peristiwa K bila diketahui L terjadi, didefinisikan sebagai berikut:

$$P(K/L) = \frac{P(K \cap L)}{P(L)}, \text{ dengan } P(L) > 0 \quad \dots\dots\dots (6)$$

Contoh:

2. Pada eksperimen melempar 2 buah dadu satu kali. Berapakah probabilitas memperoleh jumlah mata pada kedua dadu sama dengan 7 jika hasil lemparan dadu pertama muncul mata 3?

Jawab:

Misalkan Q peristiwa muncul jumlah 7 pada kedua dadu, sedangkan R peristiwa bahwa muncul mata 3 pada lemparan dadu pertama.

Maka:

$$S = \{(1,1), (1,2), (1,3), (1,4), (1,5), \dots, (5,6), (6,6)\};$$

$$N = 36$$

$$Q = \{(1,6), (2,5), (3,4), (4,3), (5,2), (6,1)\};$$

$$R = \{(3,1), (3,2), (3,3), (3,4), (3,5), (3,6)\};$$

$$n(R) = 6; \quad P(R) = \frac{6}{36}$$

$$Q \cap R = \{(3,4)\};$$

$$n(Q \cap R) = 1; \quad P(Q \cap R) = \frac{1}{36}$$

Sehingga probabilitas mendapat jumlah mata pada kedua dadu sama dengan 7 jika hasil lemparan dadu pertama muncul mata dadu 3 adalah:

$$P(Q/R) = \frac{1/36}{6/36} = \frac{1}{6}$$

3. Sebuah rumah sakit swasta membuka lowongan pekerjaan bagi para perawat untuk mengisi kekurangan tenaga perawat di rumah sakit tersebut. Tercatat pelamar sebanyak 100 orang terdiri atas orang yang berpengalaman kurang dari tiga tahun dan lebih dari tiga tahun, berstatus menikah dan tidak menikah. Seluruh pelamar dianggap mempunyai kesempatan yang sama untuk diterima sebagai pegawai/perawat. Ringkasan para pelamar disajikan dalam tabel berikut:

Tabel 7.2 Ringkasan Para Pelamar

Pengalaman	Nikah	Tidak Nikah	Total
Pengalaman < 3 tahun	25	40	65
Pengalaman > 3 tahun	11	24	35
Jumlah	36	64	100

Bila diketahui orang yang dipilih adalah yang mempunyai pengalaman di atas tiga tahun, berapakah probabilitas orang tersebut tidak menikah?

Jawab:

Misalkan diberikan notasi sebagai berikut

L : peristiwa pelamar yang dipilih berpengalaman lebih dari tiga tahun

M : peristiwa pelamar yang dipilih berstatus tidak menikah

Berdasarkan data pada tabel diperoleh $S = 100$

Jumlah pelamar yang berpengalaman lebih dari tiga tahun,

$$n(L) = 35 \text{ dan } P(L) = \frac{35}{100}$$

Jumlah pelamar yang dipilih berstatus tidak menikah dan berpengalaman lebih dari tiga tahun, $n(M \cap L) = 24$

$$P(M \cap L) = \frac{24}{100}, \text{ Maka diperoleh } P(M/L) = \frac{P(M \cap L)}{P(L)} = \frac{24/100}{35/100} =$$

$$\frac{24}{35}$$

Jadi probabilitas orang yang terpilih tidak menikah bila orang tersebut memiliki pengalaman lebih dari tiga tahun adalah

$$\frac{24}{35}.$$

Persamaan (6) dapat dinyatakan dengan

$$P(K \cap L) = P(L).P(K/L) ; P(L) > 0$$

atau

$$P(K \cap L) = P(K).P(L/K) ; P(K) > 0$$

yang dalam teori probabilitas dikenal sebagai aturan perkalian.

2. Peristiwa Bebas

Dalam perhitungan $P(K/L)$ sebelumnya, jelas bahwa probabilitas terjadinya K bergantung pada terjadinya L. Dengan kata lain, probabilitas terjadinya K harganya ditentukan relatif terhadap terjadinya L. Dalam hal ini K dikatakan tidak bebas dari L. Suatu peristiwa K dikatakan bebas dari peristiwa L jika probabilitas terjadinya K tidak terpengaruh oleh terjadi tidaknya peristiwa L.

Jika dua peristiwa K dan L saling bebas, maka berlaku

$$P(K \cap L) = P(K) \cdot P(L) \quad \dots\dots\dots (7)$$

Bagaimana dengan kebebasan n buah peristiwa?

Untuk n buah peristiwa $(K_1, K_2, \dots, K_n)$ yang saling bebas, maka persamaan (7) dapat diperluas menjadi:

$$P(K_1 \cap K_2 \cap \dots \cap K_n) = P(K_1) \cdot P(K_2) \dots P(K_n) \dots\dots\dots (8)$$

Contoh:

Misalkan sebuah mata uang dilempar 3 kali dengan

Q = peristiwa muncul sisi M pada lemparan I

R = peristiwa muncul sisi M pada lemparan II

T = peristiwa muncul sisi M pada lemparan I dan II

Maka:

$$S = \{MMM, MBM, MMB, BMM, MBB, BBM, BMB, BBB\}$$

$$Q = \{MMM, MMB, MBM, MBB\};$$

$$P(Q) = \frac{4}{8}$$

$$R = \{MMM, MMB, BMB, BMM\};$$

$$P(R) = \frac{4}{8}$$

$$T = \{MMM, MMB\};$$

$$P(T) = \frac{2}{8}$$

- $Q \cap R = \{MMM, MMB\};$

$$P(Q \cap R) = \frac{2}{8}$$

$$P(Q) \cdot P(R) = \frac{1}{2} \cdot \frac{1}{2} = \frac{1}{4}$$

Karena $P(Q \cap R) = P(Q) \cdot P(R) = \frac{1}{4}$, maka Q dan R adalah dua peristiwa yang bebas.

- $Q \cap T = \{\text{MMM, MMB}\};$

$$P(Q \cap T) = \frac{2}{8}$$

$$P(Q) \cdot P(T) = \frac{1}{2} \cdot \frac{1}{4} = \frac{1}{8}$$

Karena $P(Q \cap T) \neq P(Q) \cdot P(T)$, maka Q dan T adalah dua peristiwa yang tidak bebas.

E. RANGKUMAN

1. Eksperimen adalah proses yang memberikan hasil pengamatan, pengukuran atau perhitungan.
Ruang sampel (S) suatu eksperimen adalah himpunan semua hasil yang muncul atau terjadi dari suatu eksperimen yang diberikan.
Peristiwa adalah himpunan hasil atau himpunan bagian dari ruang sampel S.
2. Misalkan S menghasilkan N hasil dan H peristiwa dalam S yang mempunyai hasil sebanyak $n(H)$, maka probabilitas terjadinya peristiwa H didefinisikan sebagai: $P(H) = \frac{n(H)}{N}$
3. Definisi aksiomatis: andaikan suatu eksperimen dengan ruang sampel S. Dikatakan bahwa $P(H)$ adalah probabilitas terjadinya peristiwa H, jika dipenuhi:
 - a. $0 \leq P(H) \leq 1$
 - b. $P(S) = 1$
 - c. Untuk sembarang peristiwa-peristiwa $H_1, H_2, \dots$, saling asing, ($H_i \cap H_j = \emptyset$, untuk $i \neq j$),
maka $P(\bigcup_{i=1}^{\infty} H_i) = \sum_{i=1}^{\infty} P(H_i)$

4. Misalkan K dan L sembarang peristiwa dalam ruang sampel S. Probabilitas bersyarat dari peristiwa K bila diketahui L terjadi, didefinisikan sebagai berikut:

$$P(K/L) = \frac{P(K \cap L)}{P(L)}, \text{ dengan } P(L) > 0$$

5. Suatu peristiwa K dikatakan bebas dari peristiwa L jika probabilitas terjadinya K tidak terpengaruh oleh terjadi tidaknya peristiwa L atau jika

$$P(K \cap L) = P(K) \cdot P(L)$$

6. Jika $P(K \cap L) \neq P(K) \cdot P(L)$, maka K dan L dikatakan saling bergantung.

F. TES FORMATIF

1. Misalkan G dan H peristiwa-peristiwa dengan $P(G \cup H) = \frac{3}{4}$, $P(G^c) = \frac{2}{3}$, dan $P(G \cap H) = \frac{1}{4}$. Maka $P(H)$
- $\frac{1}{2}$
 - $\frac{2}{3}$
 - $\frac{3}{4}$
 - $\frac{1}{3}$
2. Dua buah dadu dilempar bersama-sama satu kali. Misalkan C peristiwa bahwa jumlah kedua mata dadu lebih dari 10. D peristiwa mata dadu pertama muncul mata prima. E peristiwa bahwa kedua dadu muncul mata sama. Maka peristiwa:
- C dan D bebas
 - D dan E tidak bebas
 - C dan D tidak bebas
 - C dan E saling asing

G. LATIHAN

1. Misalkan D dan E peristiwa-peristiwa dengan $P(D) = \frac{1}{28}$, $P(E) = \frac{3}{8}$, dan $P(D \cap E) = \frac{1}{4}$. Hitunglah:
 - a. $P(D^c)$
 - b. $P(D \cup E)$
 - c. $P(E \cap D^c)$
2. Dua buah dadu dilemparkan satu kali dan yang diperhatikan adalah jumlah mata dadu yang muncul. Carilah probabilitas diperoleh jumlah mata kedua dadu lebih dari atau sama dengan 8, jika:
 - a. muncul mata 4 pada dadu pertama
 - b. muncul mata 4 pada paling sedikit satu dadu
3. Misalkan sebuah mata uang dilempar 3 kali dengan
 - a = peristiwa muncul sisi M pada lemparan II
 - b = peristiwa muncul sisi M pada lemparan I dan IIApakah a dan b merupakan dua peristiwa bebas?

KEGIATAN BELAJAR 8

UJI VALIDITAS DAN RELIABILITAS

DESKRIPSI PEMBELAJARAN

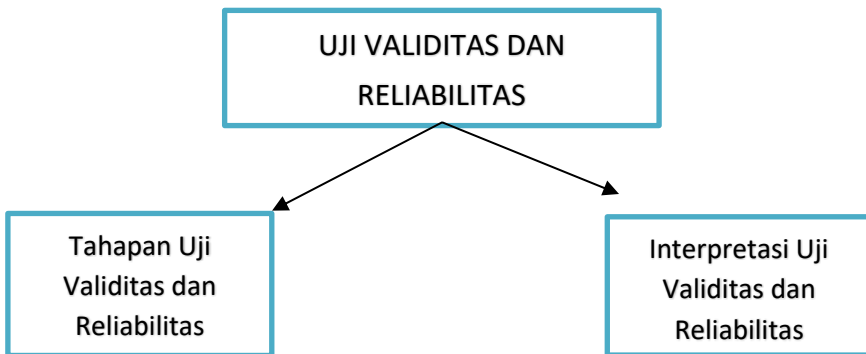
Pada bab ini mahasiswa memahami tahapan dan interpretasi uji validitas dan reliabilitas pada instrument penelitian.

KOMPETENSI PEMBELAJARAN

Setelah mengikuti perkuliahan ini mahasiswa diharapkan:

1. Mampu memahami tahapan uji validitas dan reliabilitas.
2. Mampu memahami interpretasi uji validitas dan reliabilitas

PETA KONSEP PEMBELAJARAN



A. UJI VALIDITAS

Validitas berasal dari kata *validity* yang mempunyai arti sejauh mana ketepatan dan kecermatan suatu alat ukur dalam menunjukkan hasil pengukurannya (Azwar, 1986). Uji validitas digunakan untuk mengukur sah atau valid tidaknya suatu instrumen penelitian. Suatu instrumen penelitian dalam hal ini kuesioner dikatakan valid jika pertanyaan pada kuesioner mampu untuk mengungkapkan sesuatu yang akan diukur oleh kuesioner tersebut (Ghozali, 2009).

Suatu kuesioner dikatakan valid jika pertanyaan tersebut pada kuesioner dapat mengungkapkan sesuatu yang diukur oleh kuesioner. Misalnya, kita ingin mengukur kepuasan pasien. Untuk melihat tingkat kepuasan pasien, pasien tersebut diberi lima pertanyaan, maka lima pertanyaan tersebut harus tepat mengungkapkan bagaimana kepuasan pasien. Dalam uji pengukuran validitas terdapat dua macam yaitu Pertama, mengkorelasikan antar skor butir pertanyaan (item) dengan total item. Kedua, mengkorelasikan antar masing-masing skor indikator item dengan total skor konstruk.

B. UJI RELIABILITAS

Reliabilitas berasal dari kata *reliability*. Reliabilitas menunjuk pada suatu pengertian bahwa instrumen yang digunakan dalam penelitian untuk memperoleh informasi yang digunakan dapat dipercaya sebagai alat pengumpulan data dan mampu mengungkapkan informasi yang sebenarnya dilapangan (Sitinjak & Sugiharto, 2006).

Reliabilitas adalah alat untuk mengukur suatu kuesioner yang merupakan indikator dari peubah atau konstruk. Suatu kuesioner dikatakan reliabel atau handal jika jawaban seseorang terhadap pernyataan adalah konsisten atau stabil dari waktu ke waktu.

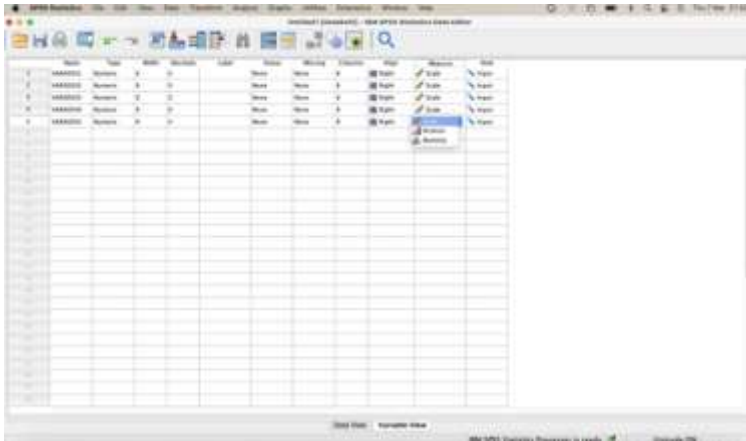
Reliabilitas suatu test merujuk pada derajat stabilitas, konsistensi, daya prediksi, dan akurasi. Pengukuran yang memiliki reliabilitas yang tinggi adalah pengukuran yang dapat menghasilkan data yang reliabel (Ghozali, 2009). Bila suatu alat pengukur dipakai dua kali – untuk mengukur gejala yang sama dan hasil pengukuran yang diperoleh relative konsisten, maka alat pengukur tersebut reliable. Dengan kata lain, realibitas menunjukkan konsistensi suatu alat pengukur di dalam pengukur gejala yang sama.

Reliabilitas, atau keandalan, adalah konsistensi dari serangkaian pengukuran atau serangkaian alat ukur. Hal tersebut bisa berupa pengukuran dari alat ukur yang sama (tes dengan tes ulang) akan memberikan hasil yang sama, atau untuk pengukuran yang lebih subjektif, apakah dua orang penilai memberikan skor yang mirip (reliabilitas antar penilai). Reliabilitas tidak sama dengan validitas. Artinya pengukuran yang dapat diandalkan akan mengukur secara konsisten, tapi belum tentu mengukur apa yang seharusnya diukur. Dalam penelitian, reliabilitas adalah sejauh mana pengukuran dari suatu tes tetap konsisten setelah dilakukan berulang-ulang terhadap subjek dan dalam kondisi yang sama. Kuesioner penelitian dianggap dapat diandalkan bila memberikan hasil yang konsisten untuk pengukuran yang sama. Tidak bisa diandalkan bila pengukuran yang berulang itu memberikan hasil yang berbeda-beda.

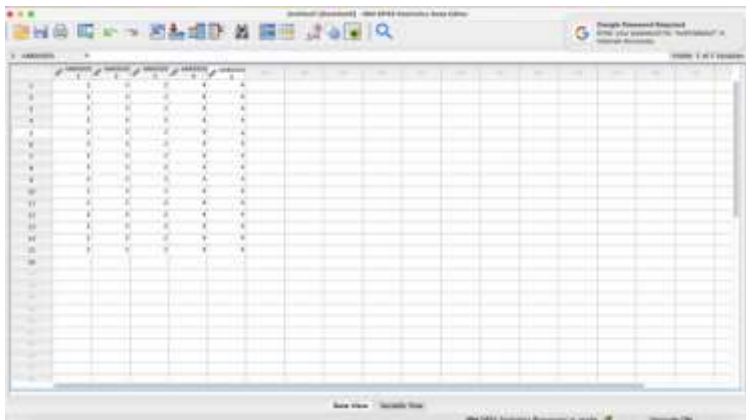
C. TAHAPAN DAN INTERPRETASI UJI VALIDITAS

Adapun Langkah-langkah pengujiannya, yaitu sebagai berikut :

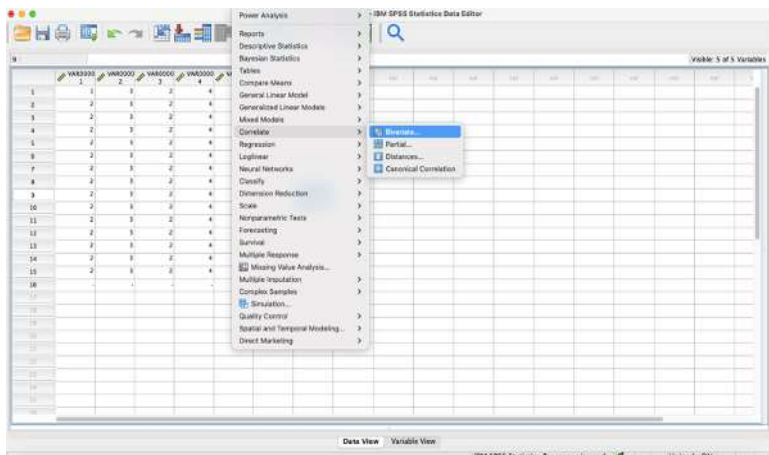
1. Buka aplikasi SPSS. Silahkan atur format yang ada di Variable View. Sesuaikan dengan kriteria data kita. Contohnya seperti ini.



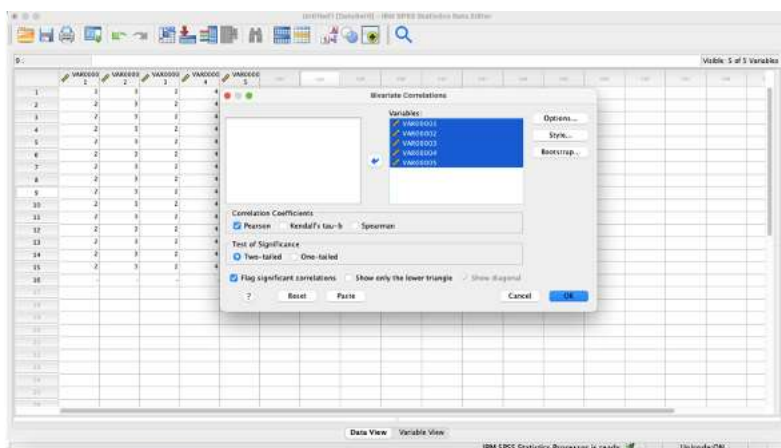
2. Lalu, Input Data Kita di Data View.



3. Selanjutnya, kita akan mencari nilai R statistik atau R hitung. Caranya klik **Analyze> Correlate> Bivariate**. Maka akan muncul kotak dialog.



4. Lalu, pindahkan semua item variabel ke kotak *Variables*. Pada *Correlation Coefficients* beri centang pada *Pearson*. Dibawahnya centang *Two-Tailed* dan juga centang *Flag Significant Correlation*. Lalu, klik OK



5. Maka akan muncul hasilnya. Perhatikan pada kolom 'Correlations'. Nilai yang akan kita uji adalah nilai pada kolom paling bawah bagian X1total yaitu 'Pearson Correlation'. Correlations
6. Apabila tingkat signifikansi item $< 0,05$ maka alat ukur yang digunakan valid.

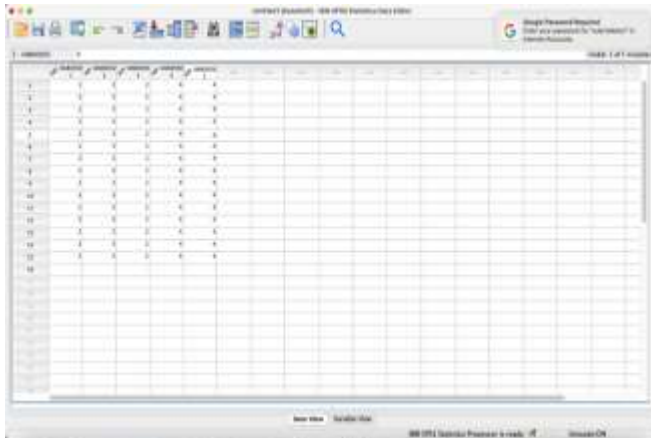
	X1total	X2total	X3total	X4total	X5total	
X1total	Pearson Correlation	1	-.871	-.871	-.888	
	Sig. (2-tailed)		.000	.000	.000	
	Sum of Squares and Cross-products	1.867	-.131	.888	-.131	
	N	30	30	30	30	
X2total	Pearson Correlation	-.871	1	-.871	-.888	
	Sig. (2-tailed)	.000		.000	.000	
	Sum of Squares and Cross-products	-.131	1.867	.888	-.131	
	N	30	30	30	30	
X3total	Pearson Correlation	-.871	-.871	1	-.888	
	Sig. (2-tailed)	.000	.000		.000	
	Sum of Squares and Cross-products	.888	-.131	1.867	-.131	
	N	30	30	30	30	
X4total	Pearson Correlation	-.871	-.871	-.888	1	
	Sig. (2-tailed)	.000	.000	.000		
	Sum of Squares and Cross-products	-.131	-.131	.888	1.867	
	N	30	30	30	30	
X5total	Pearson Correlation	-.888	-.888	-.888	-.888	1
	Sig. (2-tailed)	.000	.000	.000	.000	
	Sum of Squares and Cross-products	-.131	-.131	-.131	.888	1.867
	N	30	30	30	30	30

***. Correlation is significant at the 0.001 level (2-tailed).
 a. Cannot be computed because at least one of the variables is constant.

IBM SPSS Statistics Processor is ready. [OK] [Cancel]

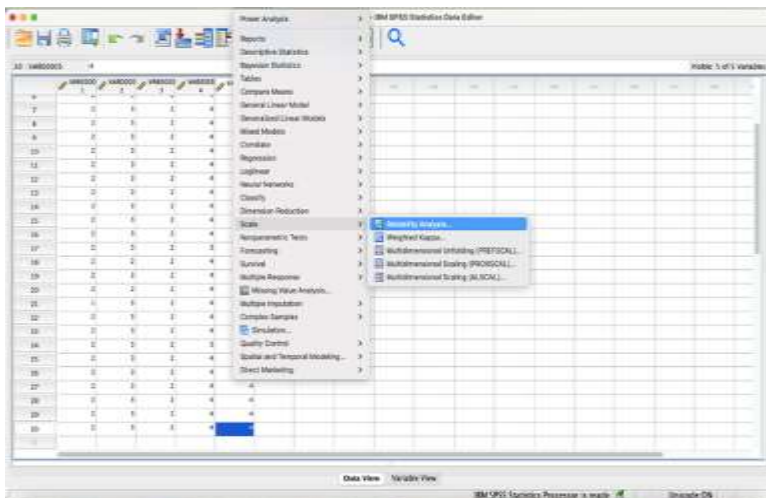
D. TAHAPAN DAN INTERPRETASI UJI RELIABILITAS

Berikut ini adalah contoh data kuesioner dan responden yang akan digunakan pada artikel ini:

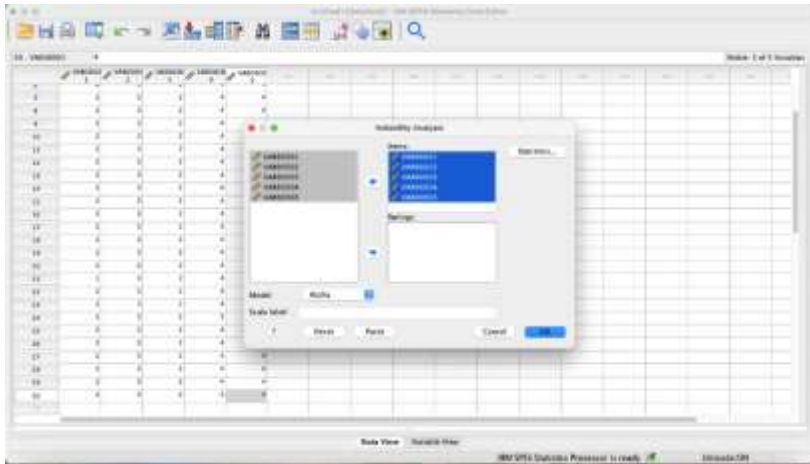


	X1	X2	X3	X4	X5	X6	X7	X8	X9	X10
1	1	2	3	4	5	6	7	8	9	10
2	2	3	4	5	6	7	8	9	10	11
3	3	4	5	6	7	8	9	10	11	12
4	4	5	6	7	8	9	10	11	12	13
5	5	6	7	8	9	10	11	12	13	14
6	6	7	8	9	10	11	12	13	14	15
7	7	8	9	10	11	12	13	14	15	16
8	8	9	10	11	12	13	14	15	16	17
9	9	10	11	12	13	14	15	16	17	18
10	10	11	12	13	14	15	16	17	18	19

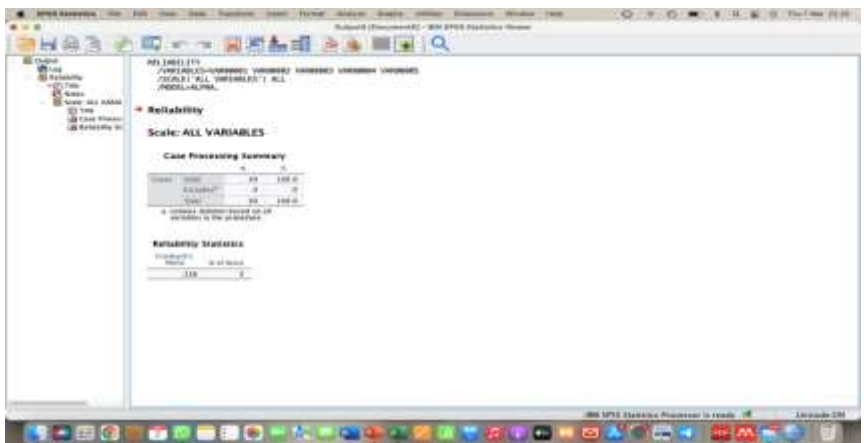
1. Kemudian, klik Analyze > Scale > Reliability Analysis



2. Lalu, akan muncul tampilan Reliability Analysis. Pindahkan seluruh data variabel berupa skala ke kolom Items (ke sebelah kanan) dan pilih model Alpha.



3. Kemudian, klik OK dan lihat hasil perhitungan data pada Output. Interpretasi Hasil Uji Reliabilitas Metode Cronbach's Alpha. Kuesioner dapat dikatakan reliabel apabila nilai Cronbach's Alpha $< 0,05$.



E. RANGKUMAN

Uji validitas digunakan untuk mengukur sah atau valid tidaknya suatu instrumen penelitian. Suatu kuesioner dikatakan valid jika pertanyaan tersebut pada kuesioner dapat mengungkapkan sesuatu yang diukur oleh kuesioner. Dalam uji pengukuran validitas terdapat dua macam yaitu Pertama, mengkorelasikan antar skor butir pertanyaan (item) dengan total item. Kedua, mengkorelasikan antar masing-masing skor indikator item dengan total skor konstruk.

Uji reliabilitas digunakan untuk mengetahui instrumen (kuesioner) yang digunakan dalam penelitian untuk memperoleh informasi yang digunakan dapat dipercaya sebagai alat pengumpulan data dan mampu mengungkapkan informasi yang sebenarnya dilapangan. Kuesioner penelitian dianggap dapat diandalkan bila memberikan hasil yang konsisten untuk pengukuran yang sama. Tidak bisa diandalkan bila pengukuran yang berulang itu memberikan hasil yang berbeda-beda. Interpretasi Hasil Uji Reliabilitas Metode Cronbach's Alpha. Kuesioner dapat dikatakan reliabel apabila nilai Cronbach's Alpha $>0,05$.

F. TES FORMATIF

1. Uji yang digunakan untuk mengetahui konsistensi kuesioner sebagai alat ukur penelitian adalah?
 - a. Chi Square
 - b. Korelasi
 - c. Reliabilitas
 - d. Validitas
 - e. Salah semua
2. Indikator yang bisa digunakan untuk menentukan item pertanyaan dalam kuesioner tersebut valid adalah?
 - a. Nilai signifikansi item pertanyaan $<0,05$
 - b. Nilai Cronbach's Alpha $>0,01$

- c. $R_{hitung} = R_{tabel}$
- d. $R_{hitung} < R_{tabel}$
- e. Semua salah

G. LATIHAN

Sebutkan metode yang bisa digunakan untuk menguji reliabilitas data!

KEGIATAN BELAJAR 9

UJI KORELASI

DESKRIPSI PEMBELAJARAN

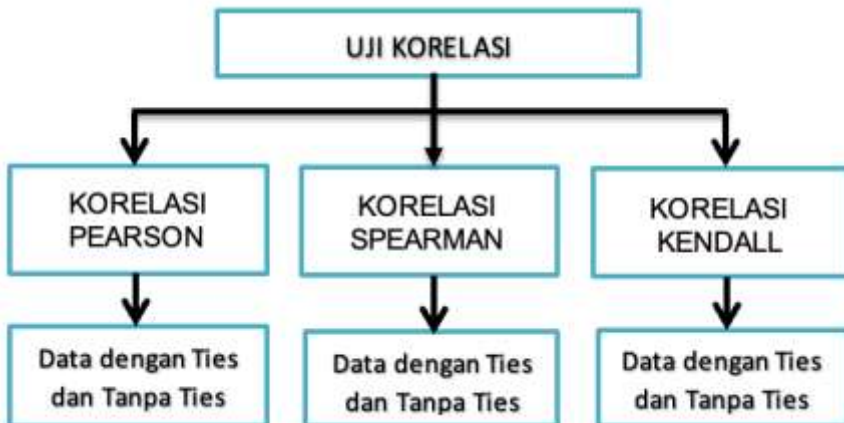
Pada bab ini mahasiswa mempelajari definisi, konsep dasar dan langkah-langkah penyelesaian kasus tiga metode uji korelasi. Diharapkan mahasiswa memiliki wawasan dan kemampuan yang mendalam untuk menyelesaikan kasus-kasus dalam uji korelasi. Juga diharapkan mahasiswa mampu menerapkan masing-masing uji korelasi sesuai data penelitian.

KOMPETENSI PEMBELAJARAN

Setelah mengikuti perkuliahan ini diharapkan mahasiswa dan mahasiswi memiliki pengetahuan dan kemampuan:

5. Mampu menjelaskan tipe data dengan ties dan tanpa ties
6. Mampu menjelaskan korelasi Pearson dengan dan tanpa ties
7. Mampu menjelaskan korelasi Spearman dengan dan tanpa ties
8. Mampu menjelaskan korelasi Kendall dengan dan tanpa ties

PETA KONSEP PEMBELAJARAN



A. PENGANTAR UJI KORELASI

Bagian penting dari siklus pengembangan ilmu data adalah menentukan fitur mana yang harus disertakan dalam sebuah model. Ada sejumlah manfaat menggunakan teknik pemilihan fitur untuk mengurangi data yang masuk ke model Anda (James Lani, 2010).

- Meningkatkan waktu pelatihan. Mengurangi jumlah fitur akan mengurangi waktu dan sumber daya komputasi yang diperlukan untuk melatih model Anda. Fitur yang berlebihan atau tidak informatif memperlambat proses dan tidak memberikan hasil yang maksimal.
- Mengurangi kemungkinan overfitting. Membuat model lebih sedikit fitur akan mengurangi kerumitan. Dimana model yang lebih kompleks justru lebih cenderung tidak mendefinisikan hasil secara baik.
- Peningkatan interpretabilitas model. Dalam banyak kasus, model merupakan faktor penting. Lebih sedikit fitur dapat membuat model kami lebih mudah dijelaskan.

Matriks yang dibahas dalam uji korelasi adalah korelasi Pearson, Kendall, dan Spearman. Semuanya berguna saat menggunakan teknik pemilihan fitur yang disebut sebagai pemilihan fitur berbasis filter. Sesuai namanya, dalam pemilihan fitur berbasis filter, kita menggunakan matriks yang dipilih untuk memfilter kolom yang tidak relevan atau lebih dalam data. Setiap kolom data kita diberi skor menggunakan metode statistik yang sesuai, diberi peringkat berdasarkan skor, dan kemudian disertakan atau dikecualikan dari model akhir. Korelasi Pearson, Kendall, dan Spearman adalah beberapa, tetapi tidak semua, matriks yang dapat digunakan. Perbedaan umum dari tiga metode korelasi ini adalah:

- Pearson, biasanya digunakan untuk kasus dua variabel kontinu kuantitatif yang memiliki hubungan linier.
- Spearman, biasanya digunakan untuk dua variabel kuantitatif jika sebagian variabel linier atau untuk satu variabel ordinal kualitatif dan satu variabel kuantitatif.

- Kendall, sering digunakan untuk dua variabel ordinal kualitatif.

B. DATA DENGAN TIES DAN TANPA TIES

Untuk mengilustrasikan metode dengan dan tanpa ties, maka akan digunakan dua kumpulan data yang berbeda, yang pertama dengan ties dan kedua tanpa ties. Sebagian ilustrasi, di bawah ini adalah contoh kasus data dengan ties dan tanpa ties.

Tabel 9.1. Data dengan Ties

No	x	y
1	-1	-3
2	3	1
3	5	0
4	5	2
5	2	-1

Tabel 9.2. Data Tanpa Ties

No	x	y
1	2	-1
2	3	5
3	5	1

Seperti yang dapat kita lihat pada tabel 1 adalah data dengan ties, karena ada dua pengamatan yang identik dalam variabel x tersebut. Selanjutnya adalah kasus pada tabel 2 yang tanpa ties, karena tidak ada pengamatan yang identik di variable x dan y .

C. KORELASI PEARSON

Korelasi Pearson adalah bentuk korelasi yang paling umum dan biasa dibahas dalam pengantar statistik. Koefisien korelasi Pearson antara dua variabel akan selalu digunakan nilai antara -1 dan 1 . Semakin jauh nilainya dari 0 , semakin kuat hubungannya. Nilai 1 menandakan hubungan linier positif sempurna, -1 mewakili hubungan linier negatif sempurna dan 0 tidak memiliki hubungan. Selanjutnya, karena prosedur untuk menghitung koefisien korelasi Pearson untuk ada dan tanpa ties adalah sama, maka untuk perhitungannya akan digabung (Andale, 2012; Lalinde et al., 2018).

1. Data dengan Ties dan Tanpa Ties

Koefisien korelasi Pearson, dilambangkan dengan r dalam kasus sampel dapat dihitung sebagai berikut:

$$r = \frac{\sum_{i=1}^n x_i y_i - n \bar{x} \bar{y}}{(n-1) s_x s_y} \quad (1)$$

Dimana:

x_i dan y_i : Pengamatan

$\bar{x}$ dan $\bar{y}$: Rata-rata Sampel dari x dan y

s_x dan s_y : Standar Deviasi Sampel x dan y

Berikut akan berikan contoh dan langkah-langkah perhitungan korelasi Pearson secara manual untuk data pada tabel 1:

Langkah 1.

Sebelum menghitung koefisien korelasi, pertama-tama kita perlu menghitung mean dan deviasi standar untuk masing-masing dari kedua variabel tersebut. Untuk menghitung Rata-rata x dan Standar Deviasi x adalah sebagai berikut:

$$\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i \quad \text{dan} \quad s_x = \sqrt{\frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})^2}$$

Rumus yang sama juga digunakan untuk menghitung mean dan standar deviasi dari variabel y . Selanjutnya kita mulai dengan menghitung mean dari kedua variabel tersebut:

$$\bar{x} = \frac{-1 + 3 + 5 + 5 + 2}{5} = 2.8 \quad \text{dan} \quad \bar{y} = \frac{-3 + 1 + 0 + 2 - 1}{5} = -0.2$$

Langkah 2.

Selanjutnya akan dihitung deviasi standar dari setiap variabel. Untuk memudahkan perhitungan, yang terbaik adalah menggunakan tabel. Dimana perhitungan manual yang kemudian disajikan dalam bentuk tabel. Jadi, untuk menghitung nilai deviasi standar, kita membutuhkan jumlah selisih kuadrat antara setiap pengamatan dan mean-nya,

$$\sum_{i=1}^n (x_i - \bar{x})^2 \text{ dan } \sum_{i=1}^n (y_i - \bar{y})^2$$

Kita mulai dengan membuat dua kolom baru dalam tabel, dua kolom tersebut diberi lambang $x_i - \bar{x}$ dan $y_i - \bar{y}$.

Tabel 9.3 Hasil Perhitungan $x_i - \bar{x}$ dan $y_i - \bar{y}$

No	x	y	$x_i - \bar{x}$	$y_i - \bar{y}$
1	-1	-3	-3.8	-2.8
2	3	1	0.2	1.2
3	5	0	2.2	0.2
4	5	2	2.2	2.2
5	2	-1	-0.8	-0.8

Kemudian dihitung $(x_i - \bar{x})^2$ dan $(y_i - \bar{y})^2$ dengan cara mengambil kuadrat dari dua kolom baru yaitu $x_i - \bar{x}$ dan $y_i - \bar{y}$,

Tabel 9.4 Hasil Perhitungan $(x_i - \bar{x})^2$ dan $(y_i - \bar{y})^2$

No	x	y	$x_i - \bar{x}$	$y_i - \bar{y}$	$(x_i - \bar{x})^2$	$(y_i - \bar{y})^2$
1	-1	-3	-3.8	-2.8	14.44	7.84
2	3	1	0.2	1.2	0.04	1.44
3	5	0	2.2	0.2	4.84	0.04
4	5	2	2.2	2.2	4.84	4.84
5	2	-1	-0.8	-0.8	0.64	0.64
Jumlah					24.8	14.8

Selanjutnya dihitung nilai Standar deviasi dan x dan y sebagai berikut:

$$s_x = \sqrt{\frac{24.8}{5-1}} = 2.49 \text{ dan } s_y = \sqrt{\frac{14.8}{5-1}} = 1.92$$

Langkah 3.

Sekarang kita membutuhkan nilai $\sum_{i=1}^n x_i y_i$, jadi perlu ditambahkan kolom baru pada tabel untuk menentukan nilai $x * y$:

Tabel 9.5 Hasil Perhitungan $\sum_{i=1}^n x_i y_i$

No	x	y	$x_i - \bar{x}$	$y_i - \bar{y}$	$(x_i - \bar{x})^2$	$(y_i - \bar{y})^2$	$x * y$
1	-1	-3	-3.8	-2.8	14.44	7.84	3
2	3	1	0.2	1.2	0.04	1.44	3
3	5	0	2.2	0.2	4.84	0.04	0
4	5	2	2.2	2.2	4.84	4.84	10
5	2	-1	-0.8	-0.8	0.64	0.64	-2
Jumlah							14

Langkah 4.

Akhirnya, koefisien korelasi Pearson dapat dihitung dengan memasukkan nilai-nilai yang ditemukan di atas dalam (pers 1):

$$r = \frac{\sum_{i=1}^n x_i y_i - n \bar{x} \bar{y}}{(n-1) s_x s_y} = \frac{14 - (5 \times 2.8 \times -0.2)}{(5-1) \times 2.49 \times 1.92} = 0.88$$

Perhatikan bahwa ada rumus lain untuk menghitung koefisien korelasi Pearson. Misalnya,

$$r = \frac{1}{n-1} \sum_{i=1}^n \left(\frac{x_i - \bar{x}}{s_x} \right) \left(\frac{y_i - \bar{y}}{s_y} \right) = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^n (x_i - \bar{x})^2} \sqrt{\sum_{i=1}^n (y_i - \bar{y})^2}}$$

Semua rumus tentunya akan memberikan hasil yang sama persis. Mengkuadratkan koefisien korelasi Pearson r adalah untuk menentukan koefisien determinasi R^2 dalam konteks regresi linier sederhana.

D. KORELASI SPEARMAN

Korelasi Spearman sangat mirip dengan Kendall. Ini adalah tes non-parametrik yang mengukur hubungan monotonik menggunakan data peringkat. Meskipun sering dapat digunakan secara bergantian dengan Kendall, Kendall lebih kuat dan umumnya merupakan metode yang disukai dari kedua metode korelasi lainnya. Keuntungan dari Spearman adalah lebih mudah untuk menghitungnya, tetapi dalam konteks ilmu data, kecil kemungkinan kita akan melakukan perhitungan secara manual. Pada kasus ini akan disajikan langkah-langkah cara perhitungan manual dengan dua kasus (dengan ties dan tanpa ties) (Statistics Solutions, 2020).

1. Data dengan Ties

Korelasi Spearman (dengan ties) dilambangkan dengan r_s .

$$r_s = \frac{\sum_{i=1}^n R_{x_i} R_{y_i} - n \overline{R_x} \overline{R_y}}{(n-1) S_{R_x} S_{R_y}} \quad (2)$$

Dimana:

R_{x_i} dan R_{y_i} : peringkat untuk dua variabel

$\overline{R_x}$ dan $\overline{R_y}$: rata-rata sampel peringkat untuk dua variabel

S_{R_x} dan S_{R_y} : Standar deviasi sampel peringkat untuk dua variabel

Berikut adalah langkah-langkah cara perhitungan manual uji korelasi Spearman untuk kasus pada tabel 1:

Langkah 1.

Seperti disebutkan sebelumnya, koefisien Spearman didasarkan pada peringkat. Jadi pertama-tama kita perlu menjumlahkan baris observasi (dari terendah ke tertinggi), secara terpisah untuk masing-masing dari dua variabel tersebut. Untuk y , kita melihat bahwa:

- ✓ -3 adalah nilai terkecil, jadi kita tetapkan pada peringkat 1
- ✓ -1 adalah nilai terkecil kedua, jadi kita tetapkan peringkat 2
- ✓ kemudian muncul 0, jadi kita tetapkan pada peringkat 3
- ✓ kemudian datang 1, jadi kita menetapkan peringkat 4
- ✓ terakhir, 2 adalah nilai terbesar, jadi kita tetapkan peringkat 5

Hal yang sama berlaku untuk x , kecuali bahwa di sini kita memiliki dua pengamatan dengan nilai 5 (jadi akan ada ties dalam barisan). Dalam hal ini, kami mengambil peringkat rata-rata:

- ✓ -1 adalah nilai terkecil, jadi kita tetapkan pada peringkat 1
- ✓ 2 adalah nilai terkecil kedua, jadi kita tetapkan pada peringkat 2
- ✓ kemudian 3, jadi kita menetapkan peringkat 3
- ✓ terakhir, dua nilai terbesar adalah milik peringkat 4 dan 5, jadi kita menetapkan keduanya sebagai peringkat 4,5

Berikut disajikan dalam tabel peringkat untuk masing-masing data dari variabel x dan y yang dilambangkan dengan R_x dan R_y yang telah dijabarkan di atas:

Tabel 9.6 Hasil Perangkingan R_x dan R_y

No	x	y	R_x	R_y
1	-1	-3	1	1
2	3	1	3	4
3	5	0	4.5	3
4	5	2	4.5	5
5	2	-1	2	2

Langkah 2.

Untuk menghindari kebingungan dalam langkah-langkah perhitungan, kita menyatukan perhitungan $\bar{R}_x$ dan $\bar{R}_y$ dalam satu tabel seperti dibawah ini:

Tabel 9.7 Hasil Penjumlahan $\bar{R}_x$ dan $\bar{R}_y$

No	x	y	R_x	R_y
1	-1	-3	1	1
2	3	1	3	4
3	5	0	4.5	3
4	5	2	4.5	5
5	2	-1	2	2
Jumlah			3	3

Untuk deviasi standar, kita harus menghitung $\sum_{i=1}^n (Rx_i - \bar{Rx})^2$ dan $\sum_{i=1}^n (Ry_i - \bar{Ry})^2$ menggunakan tabel seperti yang telah kita lakukan pada koefisien Pearson:

Tabel 9.8. Hasil Perhitungan $(Rx_i - \bar{Rx})^2$ dan $(Ry_i - \bar{Ry})^2$

No	x	y	R_x	R_y	$Rx_i - \bar{Rx}$	$Ry_i - \bar{Ry}$	$(Rx_i - \bar{Rx})^2$	$(Ry_i - \bar{Ry})^2$
1	-1	-3	1	1	-2	-2	4	4
2	3	1	3	4	0	1	0	1
3	5	0	4.5	3	1.5	0	2.25	0
4	5	2	4.5	5	1.5	2	2.25	4
5	2	-1	2	2	-1	-1	1	1
Jumlah							9.5	10

Sehingga standar deviasi dengan demikian adalah:

$$s_x = \sqrt{\frac{9.5}{5-1}} = 1.54 \text{ dan } s_y = \sqrt{\frac{10}{5-1}} = 1.58$$

Langkah 3.

Selanjutnya akan dicari nilai $\sum_{i=1}^n Rx_i Ry_i$, jadi kita harus menambahkan kolom baru pada tabel untuk menghitung $Rx_i * Ry_i$,

Tabel 9.9 Hasil Perhitungan $Rx_i * Ry_i$

No	R_x	R_y	$Rx_i - \bar{Rx}$	$Ry_i - \bar{Ry}$	$(Rx_i - \bar{Rx})^2$	$(Ry_i - \bar{Ry})^2$	$Rx_i * Ry_i$
1	1	1	-2	-2	4	4	1
2	3	4	0	1	0	1	12
3	4.5	3	1.5	0	2.25	0	13.5
4	4.5	5	1.5	2	2.25	4	22.5
5	2	2	-1	-1	1	1	4
Jumlah					9.5	10	53

Langkah 4.

Terakhir, koefisien korelasi Spearman dengan ties dapat dihitung dengan memasukkan semua nilai ke dalam (pers 2):

$$r_s = \frac{\sum_{i=1}^n Rx_i Ry_i - n\bar{Rx}\bar{Ry}}{(n-1)s_{Rx}s_{Ry}} = \frac{53 - (5 \times 3 \times 3)}{(5-1) \times 1.54 \times 1.58} = 0.82$$

2. Data Tanpa Ties

Ketika semua nilai awal berbeda dalam setiap variabel, artinya semua peringkat adalah bilangan bulat yang berbeda, tidak ada ties. Dalam kasus khusus tersebut, koefisien Spearman dapat dihitung dengan rumus singkat berikut:

$$r_s = 1 - \frac{6 \sum_{i=1}^n (Rx_i - Ry_i)^2}{n(n^2 - 1)} \quad (3)$$

Pada kasus ini digunakan tabel 2, yaitu sampel data tanpa ties yang telah dijelaskan di awal:

Langkah 1.

Semua pengamatan dalam setiap variabel berbeda, jadi semua peringkat adalah bilangan bulat yang berbeda dan tidak ada ties:

Tabel 9.10 Hasil Perangkingan R_x dan R_y

No	x	y	R_x	R_y
1	3	5	2	3
2	5	1	3	2
3	2	-1	1	1

Langkah 2.

Kita hanya perlu menghitung selisih $R_x - R_y$ dan $(R_x - R_y)^2$ sebagai berikut:

Tabel 9.11. Hasil Perhitungan $R_x - R_y$ dan $(R_x - R_y)^2$

No	x	y	R_x	R_y	$R_x - R_y$	$(R_x - R_y)^2$
1	3	5	2	3	-1	1
2	5	1	3	2	1	1
3	2	-1	1	1	0	0
Jumlah						2

Langkah 3.

Terakhir, kita dapat menghitung koefisien Spearman tanpa ties menggunakan (pers 3) sebagai berikut:

$$r_s = 1 - \frac{6 \sum_{i=1}^n (Rx_i - Ry_i)^2}{n(n^2 - 1)} = 1 - \frac{(6)(2)}{3(3^2 - 1)} = 1 - \frac{12}{24} = 0.5$$

E. KORELASI KENDALL

Kendall dan Spearman biasanya digunakan untuk mengukur hubungan monotonik. Seperti diketahui bahwa hubungan linier berarti dua variabel bergerak bersama dengan kecepatan konstan. Sedangkan hubungan monotonik mengukur seberapa besar kemungkinan dua variabel bergerak ke arah yang sama, tetapi tidak harus dengan kecepatan konstan. Seperti korelasi Pearson, Kendall akan juga menggunakan nilai antara -1 dan 1. Kendall sering digunakan ketika data tidak memenuhi salah satu

persyaratan korelasi Pearson. Kendall bersifat non-parametrik artinya tidak mengharuskan kedua variabel tersebut membentuk kurva lonceng. Karena didasarkan pada nilai peringkat dari setiap variabel, Kendall menggunakan data kontinu, tetapi juga dapat digunakan dengan data ordinal. Pertama-tama kita ilustrasikan kasus tanpa ties dan kemudian ketika ada ties (Statistics Solutions, 2020).

1. Data Tanpa Ties

Ketika tidak ada ties, koefisien Kendall dilambangkan dengan τ_a dan didefinisikan sebagai berikut:

$$\tau_a = \frac{C - D}{C + D} \quad (4)$$

Dimana:

C : jumlah angka pasangan *concordant*

D : jumlah angka pasangan *discordant*

Berikut akan disajikan pasangan data *concordant* dan *discordant* menggunakan data tanpa ties, yaitu data tabel 2:

Tabel 9.12. Sampel Data Tanpa Ties

No	x	y
1	5	1
2	3	5
3	2	-1

Langkah 1.

Kita mulai dengan menghitung pangkat untuk setiap variabel (R_x dan R_y), seperti yang kita lakukan untuk koefisien Spearman:

Tabel 9.13. Hasil Penentuan Pangkat R_x dan R_y

No	x	y	R_x	R_y
1	3	5	2	3
2	5	1	3	2
3	2	-1	1	1

Langkah 2.

Kita kemudian secara acak dipilih variabel referensi diantara R_x dan R_y . Misalkan kita mengambil R_x sebagai variabel referensi, maka kita mengurutkan kumpulan data berdasarkan variabel referensi ini:

Tabel 9.14. Hasil Penentuan Pangkat R_x dan R_y

No	x	y	R_x	R_y
1	2	-1	1	1
2	3	5	2	3
3	5	1	3	2

Lakukan pemeriksaan pada baris R_y satu per satu, apakah baris di bawahnya dalam tabel lebih kecil atau lebih besar. Pada tabel di atas, baris pertama R_y adalah 1. Kita melihat bahwa, 3 tepat berada dibawah 1 dan $3 > 1$. Karena $3 > 1$, ini disebut pasangan yang *concordant*. Baris berikutnya, 2, dimana $2 > 1$. Jadi ini juga merupakan pasangan yang *concordant*. Kita mulai lagi dengan baris kedua R_y , yaitu 3. Sekali lagi, kita melihat baris di bawahnya dan memeriksa apakah lebih besar atau lebih kecil. Di sini, baris di bawahnya adalah 2, dimana $2 < 3$, jadi kita memiliki pasangan yang *discordant*.

Langkah 3.

Kami sekarang menghitung jumlah total pasangan data *concordant* dan *discordant*. Jadi, $C = 2$, dan $D = 1$.

Langkah 4.

Terakhir, kita substitusikan nilai C dan D ke dalam (pers 4) di atas:

$$\tau_b = \frac{C - D}{C + D} = \frac{2 - 1}{2 + 1} = \frac{1}{3} = 0.33$$

2. Data dengan Ties

Koefisien Kendall tau-b, yang membuat penyesuaian dengan ties, didefinisikan sebagai berikut:

$$\tau_b = \frac{C - D}{\sqrt{(C_n^2 - n_x)(C_n^2 - n_y)}} \quad (5)$$

Dimana:

- ✓ C_n^2 : Jumlah Total Pasangan yang Mungkin
- ✓ n_x : Jumlah Kemungkinan Pasangan yang Seri diantara Nilai x
- ✓ n_y : Jumlah Kemungkinan Pasangan yang Seri diantara Nilai y

Mari kita ilustrasikan kasus ini dan rumusnya dengan menggunakan sampel data tabel 1 yang digunakan untuk korelasi Pearson dan korelasi Spearman dengan ties, yaitu:

Tabel 9.15. Sampel Data

No	x	y
1	-1	-3
2	3	1
3	5	0
4	5	2
5	2	-1

Langkah 1.

Sama halnya dengan tanpa ties, proses dengan ties didasarkan pada peringkat. Jadi kita perlu mulai dengan menambahkan peringkat untuk setiap variabel seperti contoh sebelumnya yaitu menentukan nilai R_x dan R_y ,

Tabel 9.16. Hasil Penentuan Pangkat R_x dan R_y

No	x	y	R_x	R_y
1	-1	-3	1	1
2	3	1	3	4
3	5	0	4.5	3
4	5	2	4.5	5
5	2	-1	2	2

Langkah 2.

Salah satu dari dua variabel (x atau y) harus dipilih sebagai variabel referensi. Kali ini, kita tidak memilihnya secara acak, tetapi kita memilih salah satu yang tidak memiliki ties. Dalam kasus ini, variabel x memiliki ties, sedangkan variabel y tidak memiliki ties. Jadi y akan menjadi variabel referensi kita. Hasilnya seperti pada tabel dibawah ini:

Tabel 9.17. Hasil Penentuan Pangkat dengan y Variabel Referensi

No	x	y	R_x	R_y
1	-1	-3	1	1
2	3	1	2	2
3	5	0	3	4.5
4	5	2	4	3
5	2	-1	5	4.5

Langkah 3.

Kita periksa pasangan data yang *concordant* dan *discordant* dengan cara yang sama seperti yang kita lakukan tanpa ties, tetapi kali ini kita tidak menghitung ties. Baris pertama pada kolom R_x adalah 1. Kita membandingkan semua baris di bawahnya dalam tabel dengan nilai tersebut. Proses yang sama diterapkan untuk semua baris. Misalnya, untuk baris kedua kolom R_x , kita memiliki nilai 2. Sekali lagi, semua baris di bawah tabel lebih besar sehingga kami menulis bahwa mereka adalah pasangan yang *concordant* dan seterusnya, sehingga diperoleh pasangan yang *concordant* dan *discordant* yaitu $C = 8$ dan $D = 1$.

Langkah 4

Selanjutnya dihitung nilai kombinasinya. Dalam kasus ini, kita memiliki ukuran sampel adalah 5,

$$C_n^2 = \frac{n!}{2!(n-2)!} = \frac{5!}{2!(5-2)!} = 10$$

Ini berarti jumlah pasangan yang mungkin seri sama adalah:

- ✓ 1 untuk variabel x (satu-satunya pasangan yang mungkin adalah {4.5, 4.5})
- ✓ 0 untuk variabel y (karena semua peringkat berbeda, tidak ada pasangan sama)

Oleh karena itu, $n_x = 1$ dan $n_y = 0$.

Langkah 5

Dengan mensubstitusikan nilai-nilai yang telah diperoleh ke dalam (pers 5) di atas, maka hasilnya sebagai berikut:

$$\tau_b = \frac{C - D}{\sqrt{(C_n^2 - n_x)(C_n^2 - n_y)}} = \frac{8 - 1}{\sqrt{(10 - 1)(10 - 0)}} = \frac{7}{\sqrt{90}} = 0.74$$

F. RANGKUMAN

Pearson, digunakan untuk dua variabel kontinu kuantitatif yang memiliki hubungan linier. Spearman, digunakan untuk dua variabel kuantitatif jika tautannya sebagian linier, atau untuk satu variabel ordinal kualitatif dan satu variabel kuantitatif. Kendall, sering digunakan untuk dua variabel ordinal kualitatif. Ingatlah bahwa koefisien korelasi (tidak peduli apakah itu Pearson, Spearman atau Kendall) berkisar dari -1 hingga 1. Secara garis besar, korelasi positif berarti bahwa nilai tinggi dari satu variabel dikaitkan dengan nilai tinggi dari variabel lainnya (dan sebaliknya). Korelasi negatif berarti bahwa nilai tinggi dari satu variabel dikaitkan dengan nilai rendah dari variabel lainnya. Kesimpulan yang diambil dari koefisien korelasi yang dihitung dalam sampel tidak dapat digeneralisasikan ke populasi tanpa uji statistik yang tepat, yaitu uji korelasi.

G. TES FORMATIF

1. Nilai Koefisien korelasi 1 menunjukkan?
 - a. tidak adanya korelasi apapun.
 - b. korelasi yang sempurna.

- c. tingkat korelasi yang relatif kecil.
 - d. tingkat korelasi yang relatif tinggi.
2. Nilai koefisien korelasi berkisar antara?
 - a. 0 hingga 1.
 - b. -1 hingga 0.
 - c. -1 hingga +1.
 - d. 1 sampai 10.

H. LATIHAN

1. Bagaimana seseorang dapat menentukan uji korelasi yang sesuai dengan data penelitiannya?
2. Apa artinya jika nilai hasil uji korelasi mendekati -1, 1 atau 0?
3. Apa kelebihan dan kekurangan ketiga uji korelasi yang telah dibahas di atas?
4. Gunakan korelasi Pearson, Spearman dan Kendall untuk menyelesaikan dua kasus pada tabel (a dan b) di bawah ini:

Tabel a. Sampel Data

No	x	y
1	2	-1
2	3	5
3	5	2

Tabel b. Sampel Data

No	x	y
1	2	-1
2	3	5
3	5	-1

DAFTAR PUSTAKA

- Agustinus Bandur . (2019). Penelitian Kualitatif Studi Multi-Disiplin Keilmuan Dengan NVivo 12 Plus. Jakarta: Mitra Wacana Media
- Ambrosius, W.T., 2007. *Methods in Molecular Biology: Topics in biostatistics*. 9781588295316
- Amin.I., Aswin.A., Fajar.I., Isnaeni, Iwan.S., Pudjirahaju.A., Sunindya.R.. 2009. *Statistika untuk Praktisi Kesehatan*. Yogyakarta. Graha Ilmu
- Andale. (2012). *Pearson Correlation: Definition and Easy Steps for Use*. WordPress.
- Aunuddin. 2005. *Statistika: Rancangan dan Analisis Data*. Bogor: IPB Press.
- Azwar, Saifudin. 1986. *Validitas dan Reliabilitas*. Jakarta: Rineka Cipta
- Bain, L. J., & Engelhardt, M. (1992). *Introduction to probability and mathematical statistics* (Vol. 4). Belmont, CA: Duxbury Press.
- Cook, A., Netuveli, G., & Sheikh, A. (2004). *Basic Skills in Statistics Basic Skills in Statistics A Guide for Healthcare Professionals*. Class Publishing. www.thepcrj.org
- Daniel, W., & Cross, C. (2013). *Biostatistics: A Foundation for Analysis in The Health Sciences*. United State of America: John Wiley & Sons, Inc.
- Gabrosek, J. (2013). *SPSS Manual for Introductory Applied Statistics: A Variable Approach*. Department of Statistics, Grand Valley State University.

- Gaddis, G.M. and Gaddis, M.L., 1990. Introduction to biostatistics: Part 2, descriptive statistics. *Annals of Emergency Medicine*, 19(3), pp.309-315.
- Gerald J. Miller and Kaifeng Yang. (2008). Handbook of Research Methods in Public Administration, Boca Raton: Taylor & Francis
- Ghozali, Imam. 2019. *Aplikasi Analisis Multivariate Dengan Program IBM SPSS 23*. Semarang: UNDIP
- Hasan, I. (2016). *Pokok-pokok Materi Statistik 1* (2nd ed.). Jakarta: PT Bumi Aksara.
- Hasibuan. A.A. ,Supardi, Syah.D. 2009. *Pengantar Statistik Pendidikan*. Jakarta. Gaung Persada Press.
- Hazra, A. and Gogtay, N., 2016. Biostatistics series module 1: Basics of biostatistics. *Indian journal of dermatology*, 61(1), pp.10-20.
- Hogg, R. V., McKean, J. W., & Craig, A. T. (2013). *Introduction to mathematical statistics*. Pearson Education India.
- Howard Lune & Bruce Berg. (2017). Qualitative Research Methods for the Social Sciences, 9th edition. Boston: Pearson
- IBM. (2021). *SPSS Statistics 28 Brief Guide*. IBM Corporation.
- James Lani. (2010). Correlation (Pearson, Kendall, Spearman). *Statistics Solutions*.
- Joseph A. Maxwell. (2013). Qualitative Research Design. Los Angeles: Sage Publications, 2nd edition
- Kaliyadan, F. and Kulkarni, V., 2019. Types of variables, descriptive statistics, and sample size. *Indian dermatology online journal*, 10(1), pp.82-86.

- Lalinde, J. D. H., Castro, J. F. E., Tarazona, M. E. P., Rodriguez, J. E., Range, J. G. C., Sierra, C. A. T., Torrado, M. K. A., Sierra, S. M. C., & Pirela, V. J. B. (2018). On the proper use of the Pearson correlation coefficient: Definitions, properties and assumptions. *Archivos Venezolanos de Farmacologia y Terapeutica*.
- Lexy J. Moleong. (2000). Metodologi Penelitian Kualitatif. Bandung: Remaja Rosdakarya
- Machfoedz, I., 2016. *Bio Statistika*. Edisi Revisi. Yogyakarta. Fitramaya.
- Majid, U. (2018). Research Fundamentals: Study Design, Population, and Sample Size. *Undergraduate Research in Natural and Clinical Science and Technology (URNCSST) Journal*, 2(1), 1-7.
- Norfai. (2020). *Manajemen Data Menggunakan SPSS* (A. S. Noor, Ed.). Universitas Islam Kalimantan Muhammad Arsyad Al-Banjary.
- Nuryadi, Astuti, T. D., Utami, E. S., & Budiantara, M. (2017). Buku Ajar Dasar-dasar Statistik Penelitian. In *Sibuku Media*.
- Peck, R., Olsen, C., & Devore, J. (2008). *Introduction to Statistics and Data Analysis*. Unites State of America: Introduction to Statistics and Data Analysis
- Priadana, H.M. Sidik dan Sunarsi Denok. 2021. Metode Penelitian Kuantitatif. Tangerang Selatan : Pascal Books
- Riduwan. 2010. *Dasar-dasar Statistika*. Bandung. Alfabeta.
- Riwidikdo, H., 2012. *Statistik Kesehatan*. Yogyakarta. Nuha Medika
- Robert K. Yin. (2011). Qualitative Research from Start to Finish. New York: Guildford Press

- Rosner, B. (2011). *Fundamentals of Biostatistics*. In M. Taylor (Ed.), *Brooks/Cole Cengage Learning* (7th ed.). Boston: Brooks/Cole, Cengage Learning.
- Rumengan, Jemmy. 2017. *Aplikasi statistik dalam penelitian kuantitatif*. Bandung : Citapustaka Media
- Sahu, P. K., Pal, S. R., & Das, A. K. (2015). Estimation and inferential statistics. In *Estimation and Inferential Statistics*. <https://doi.org/10.1007/978-81-322-2514-0>
- Santjaka, A. 2011. *Statistik Untuk Penelitian Kesehatan: Deskriptif, Inferensial, Parametrik dan Non Parametrik*. Yogyakarta. Nuha Medika.
- Schmetterer, L. (2012). *Introduction to mathematical statistics* (Vol. 202). Springer Science & Business Media.
- Shaw, P.A., Johnson, L.L. and Proschan, M.A., 2018. Intermediate topics in biostatistics. In *Principles and practice of clinical research* (pp. 383-409). Academic Press.
- Sinaga, Dameria. 2022. *Metodologi Penelitian (Penelitian Kuantitatif)*. Jakarta : UKI Press
- Sitinjak, Tumpal JR & Sugiarto. 2006. LISREL. Yogyakarta: Graha Ilmu.
- Solikhah, Amyati, 2022. *Biostatistik Sebuah Aplikasi SPSS dalam Bidang Kesehatan dan Kedokteran*. Yogyakarta. Jejak Pustaka.
- Statistics Solutions. (2020). Correlation (Pearson, Kendall, Spearman) - Statistics Solutions. In *Statistics Solutions*.
- Sugiyono. 2015. *Statistika untuk Penelitian*. Bandung. Alfabeta

- Sullivan, L. M. (2018). *Essentials of Biostatistics in Public Health* (Third; A. Clerkin, ed.). Boston: Michael Brown.
- Sumarminingsih, E., & Astutik, S. (2021). *Pengantar Teori Peluang*. Universitas Brawijaya Press.
- Thacker, L. R. (2020). What Is the Big Deal About Populations in Research? *Progress in Transplantation*, 30(1).
- Umami, A., 2021. *Konsep Dasar Biostatistik*. Kediri. CV. Pelita Medika.
- Vehovar, V., Toepoel, V., & Steinmetz, S. (2016). Non-probability Sampling. Dalam *The Sage Handbook of Survey Methodology*. Sage.
- Yuantari, MG.Catur dan Handayani, Sri. 2017. Buku Ajar Biostatistik Deskriptif dan Inferensial. Semarang : Universitas Dian Nuswantoro

TENTANG PENULIS



Ishak S. Beno, S.Si., M.Sc.

Setelah menyelesaikan studi pada program Master of Science in Network Security, dari Departemen Computer Science, Heriot-Watt University, UK, Penulis aktif melaksanakan tugas sebagai Dosen tetap di Program Studi Matematika, Fakultas Matematika dan Ilmu Pengetahuan Alam, Universitas Cenderawasih, Jayapura, Papua. Di samping itu, juga aktif melakukan kegiatan pengabdian dan penelitian sesuai bidang keahlian penulis, yakni, Komputasi Matematika dan Machine Learning. Karya-karya ilmiah yang telah dipublikasikan secara mandiri dan kolaboratif antara lain: Buku-buku Referensi *Big Data Analytics* dan *Visi Komputer*. Artikel-artikel jurnal ilmiah, *Measuring the E-learning Readiness Level in Rural Areas During the Covid-19 Pandemic*; *Analisis Teorema Sisa Cina dalam Dekripsi Data Text Terenkripsi RSA*; dan *Rekayasa Trafik Stokastik Untuk Optimasi Pendapatan Jasa Jaringan*. Selain melaksanakan tanggung jawab tri dharma sebagai dosen, Penulis mendapat kepercayaan mengampu tugas tambahan sebagai Laboratorium Komputasi pada Jurusan Matematika FMIPA Uncen.



Ni Putu Ayu Wulan Noviyanti, S.KM., M.Kes.

Penulis di lahirkan di Pupuan pada tanggal 26 November 1993 Ketertarikan penulis terhadap ilmu kesehatan masyarakat dimulai pada tahun 2011 silam. Hal tersebut membuat penulis memilih untuk melanjutkan pendidikan ke Perguruan Tinggi dan berhasil menyelesaikan studi S1 di prodi Ilmu Kesehatan Masyarakat Universitas Udayana pada tahun 2015. Tahun 2018, penulis menyelesaikan studi S2 di prodi Kesehatan Masyarakat Program Pasca Sarjana Universitas Udayana. Saat ini penulis bekerja sebagai dosen tetap di Program Studi S1 Administrasi Kesehatan Stikes KESDAM IX/Udayana. Penulis saat ini menjabat sebagai Ketua Humas dan Kerjasama di Stikes KESDAM IX/Udayana. Penulis juga aktif dalam kegiatan ilmiah dan organisasi keprofesian. Sehari-harinya bekerja sebagai dosen pengampu mata kuliah Biostatistik, Epidemiologi, Metodologi Penelitian Kesehatan, Dasar Ilmu Gizi, Dasar Ilmu Kesehatan Masyarakat dan Kewirausahaan. Selain itu penulis juga aktif dalam menulis jurnal serta aktif menulis buku ajar dan book chapter.



Ni Putu Eka Febianingsih, S.KM., M.PH.

Kelahiran Badung, Bali pada 10 Pebruari 1987. Lulus Sarjana Kesehatan Masyarakat di Universitas Udayana tahun 2009. Merupakan Alumni Pasca Sarjana *Field Epidemiology Training Programe*, Ilmu Kesehatan Masyarakat, Fakultas Kedokteran Universitas Gadjah Mada (sejak tahun 2016). Sejak tahun 2010 sampai sekarang aktif menjadi dosen di Sekolah Tinggi Ilmu Kesehatan KESDAM IX/Udayana Denpasar. Aktif menulis buku diantaranya Buku Referensi Faktor Risiko



Mariene Wiwin Dolang, SKM., M.Kes.

Seorang penulis dan dosen tetap di STIKes Maluku Husada pada Program Studi Kesehatan Masyarakat. lahir di Ujung Pandang, 19 November 1990. Tahun 2012 lulus Sarjana dari Fakultas Kesehatan Masyarakat Peminatan Biostatistik KKB Universitas Hasanuddin dan pada tahun 2013-2015 menempuh pendidikan Magister Kesehatan Konsentrasi Kesehatan Reproduksi di Universitas Hasanuddin. Buku yang telah ditulis dan terbit berjudul di antaranya: *Gizi dalam kesehatan reproduksi, Buku ajar kecemasan menopause: Kesehatan reproduksi lansia, Dasar Kesehatan reproduksi dan KIA.*



Titik Suparwati, S.Si., M.Si.

Titik Suparwati. Lulus S1 di Program Studi Matematika FMIPA Gadjah Mada tahun 2000. Lulus S2 di Program Pasca Sarjana, Universitas Gadjah Mada tahun 2006. Saat ini adalah dosen tetap Program Studi Matematika, FMIPA, Universitas Cenderawasih. Email:

titiksaa@gmail.com



Henrietta Imelda Tondong, SKM., MPH

Dosen tetap Prodi D-III Kebidanan Palu Jurusan Kebidanan Poltekkes Kemenkes Palu. Lahir di Desa Parigi, 08 Desember 1971 Sulawesi Tengah. Penulis merupakan anak kedua dari dua bersaudara dari pasangan Bapak Marthen Tondong dan Ibu Neltje Kaloh.

Menyelesaikan Pendidikan Sarjana (S1) di Universitas Kristen Indonesia Tomohon (UKIT) dan Pendidikan Pasca Sarjana (S2) di Universitas Gadjah Mada Yogyakarta Prodi.



Dr. Rita Raya, S.Si., M.Si.

Rita Raya lulus S1 di Jurusan Matematika Universitas Hasanuddin tahun 1993. Lulus S2 di Program Studi Statistika, Institut Pertanian Bogor tahun 2001. Lulus S3 di Program Studi Pendidikan Matematika, Universitas Negeri Surabaya tahun 2020. Saat ini adalah dosen tetap Program Studi Statistika, FMIPA, Universitas Cenderawasih. Mengampu mata kuliah Pengantar Teori Peluang, Pengantar Teori Inferensi, Analisis Data, dan Metodologi Penelitian. Penulis juga aktif menulis Jurnal Nasional, Internasional dan Buku. (email: ritaraya68@gmail.com).



Putu Ratna Suprima Dewi, S.KM, SE, MPH

Dosen tetap Stikes Kesdam IX/udayana, kelahiran Klungkung, 24 Oktober 1991. Memiliki latar belakang Pendidikan Sarjana Kesehatan Masyarakat di Program Studi Ilmu Kesehatan Masyarakat, Fakultas Kedokteran, Universitas Udayana (Tahun 2009-2013), dan Sarjana Ekonomi di Program Studi Manajemen, Fakultas Ekonomi dan Bisnis, Undiknas University (Tahun 2010-2014). Kemudian mengambil peminatan Utama Keselamatan dan Kesehatan Kerja, PS. S2 Ilmu Kesehatan Masyarakat, Fakultas Kedokteran, Universitas Gadjah Mada (Tahun 2015-2017).



Felix Reba, S.Si., M.Sc.

Felix Reba. Lulus S1 di Program Studi Matematika FMIPA Universitas Cenderawasih tahun 2008. Lulus S2 di Program Master of Science, Universitas Gadjah Mada tahun 2017. Saat ini adalah dosen tetap Program Studi Matematika, FMIPA, Universitas Cenderawasih. Mengampu mata kuliah Data Mining, Algoritma dan Pemrograman, Komputasi Statistik dan Kecerdasan Buatan. Pernah menjadi pengajar pada Program Digital Talent Scholarship (DTS) untuk Fresh Graduate Academy (FGA) Kemkominfo Jakarta Pusat. Penulis juga aktif menulis Jurnal Nasional, Internasional, Buku dan juga beberapa kali menjadi Narasumber pada Himpunan Matematika Indonesia (IndoMS). (gmail: felix.reba85@gmail.com).

Penerbit :

PT. Sonpedia Publishing Indonesia

Buku Gudang Ilmu, Membaca Solusi
Kebodohan, Menulis Cara Terbaik
Mengikat Ilmu. Everyday New Books

SONPEDIA.COM
PT. Sonpedia Publishing Indonesia

Redaksi :

Jl. Kenali Jaya No 166

Kota Jambi 36129

Tel +6282177858344

Email: sonpediapublishing@gmail.com

Website: www.sonpedia.com