

STATISTIK DASAR

Dr. Ir. Marelianda Al Dianty, S.T., M.Sc., CSRS
Drs. Bambang Suhartawan, M.MT
Dra. Daawia, M.Sc
Dr. Drs. H. Mohzana, S.Pd., M.Pd
Dr. Agus Suprpto, S.P., M.P., IPM
Ni Putu Agustini, S.K.M., M.Si
drg. Putri Erlyn, M.Kes
Febrya Christin H. Buan, S.Kom., M.Stat
Dr. Ir. I Komang Agusjaya Mataram, M.Kes
Achmad, S.E., M.E



STATISTIK DASAR

Penulis:

Dr. Ir. Marelianda Al Dianty, S.T., M.Sc., CSRS

Drs. Bambang Suhartawan, M.MT

Dra. Daawia, M.Sc

Dr. Drs. H. Mohzana, S.Pd., M.Pd

Dr. Agus Suprpto, S.P., M.P., IPM

Ni Putu Agustini, S.K.M., M.Si

drg. Putri Erlyn, M.Kes

Febrya Christin H. Buan, S.Kom., M.Stat

Dr. Ir. I Komang Agusjaya Mataram, M.Kes

Achmad, S.E., M.E



**Penerbit Yayasan
Cendikia Mulia Mandiri**

STATISTIK DASAR

Penulis:

Dr. Ir. Marelianda Al Dianty, S.T., M.Sc., CSRS

Drs. Bambang Suhartawan, M.MT

Dra. Daawia, M.Sc

Dr. Drs. H. Mohzana, S.Pd., M.Pd

Dr. Agus Suprpto, S.P., M.P., IPM

Ni Putu Agustini, S.K.M., M.Si

drg. Putri Erlyn, M.Kes

Febrya Christin H. Buan, S.Kom., M.Stat

Dr. Ir. I Komang Agusjaya Mataram, M.Kes

Achmad, S.E., M.E

Editor:

Paput Tri Cahyono

Penerbit:

Yayasan Cendikia Mulia Mandiri

Redaksi:

Perumahan Cipta No.1

Kota Batam, 29444

Email: cendikiamuliamandiri@gmail.com

ISBN: 978-623-8576-80-7

Terbit: Juli 2024

IKAPI: 011/Kepri/2022

Exp. 31 Maret 2026

Ukuran:

x hal + 149 hal;

14,8cm x 21cm

Cetakan Pertama, 2024.

Hak Cipta Dilindungi Undang-Undang.

Dilarang Keras Memperbanyak Karya Tulis Ini Dalam Bentuk Dan Dengan Cara Apapun
Tanpa Izin Tertulis Dari Penerbit

KATA PENGANTAR

Syukur *alhamdulillah* penulis haturkan kepada Allah Swt. yang senantiasa melimpahkan karunia dan berkah-Nya sehingga penulis mampu merampungkan karya ini tepat pada waktunya, sehingga penulis dapat menghadirkannya dihadapan para pembaca. Kemudian, tak lupa *shalawat* dan salam semoga senantiasa tercurah limpahkan kepada Nabi Muhammad SAW, para sahabat, dan ahli keluarganya yang mulia.

Statistik adalah bahasa yang digunakan untuk menggambarkan dan memahami fenomena kompleks di sekitar kita. Dari ilmu pengetahuan hingga keputusan bisnis, statistik memainkan peran penting dalam membantu kita mengambil keputusan yang lebih baik dan lebih informasional.

Buku ini dirancang untuk memberikan landasan yang kuat dalam konsep dasar statistik, dari pengukuran pemusatan data hingga pengenalan analisis inferensial. Dengan penjelasan yang jelas dan contoh yang relevan, pembaca akan dibimbing melalui berbagai teknik dan aplikasi statistik yang penting.

Penulis menyampaikan terima kasih yang tak terhingga bagi semua pihak yang telah berpartisipasi. Terakhir seperti kata pepatah bahwa” Tiada Gading

Yang Tak Retak” maka penulisan buku ini juga jauh dari kata sempurna, oleh karena itu penulis sangat berterima kasih apabila ada saran dan masukan yang dapat diberikan guna menyempurnakan buku ini di kemudian hari.

2024

Penulis

DAFTAR ISI

KATA PENGANTAR.....	iii
DAFTAR ISI	v
BAB I KONSEP DAN TEKNIK STATISTIKA, DASAR DASAR ANALISIS DAN HUBUNGAN NYA DENGAN PENELITIAN DAN PENGEMBANGAN ILMU PENGETAHUAN	1
1.1. Sejarah.....	1
1.2. Konsep dan Teknik Statistika.....	2
1.2.1. Statistik Deskriptif	3
1.2.2. Statistik Inferensial.....	4
1.3. Data Dalam Statistik.....	7
1.3.1. Sumber Data Statistik	7
1.3.2. Jenis Data Statistik	8
1.3.3. Skala Pengukuran Data.....	8
1.4. Statistik dan Hubungan dengan Pengembangan Penelitian dan Pengembangan Ilmu Pengetahuan	10
BAB II TEKNIK PENYAJIAN DATA STATISTIK	15
2.1. Tabel Frekuensi.....	16
2.2. Diagram Batang.....	18
2.3. Diagram Lingkaran.....	20
2.4. Diagram Garis	22
2.5. Histogram	23
2.6. Poligon Frekuensi	25

**BAB III TEKNIK STATISTIKA UNTUK MENETAPKAN
UKURAN PEMUSATAN DATA, PENYEBARAN DATA,
DAN LOKASI.....27**

- 3.1. Ukuran Pemusatan Data..... 27
- 3.2. Ukuran Penyebaran Data..... 28
- 3.3. Ukuran Lokasi 31
- 3.4. Teknik Statistika Lanjutan 33
- 3.5. Penerapan Teknik Statistika dalam Penelitian
..... 36

**BAB IV TEKNIK STATISTIK UNTUK MENGHITUNG
VARIABILITAS DATA41**

- 4.1. Pengertian Variabilitas Data 41
- 4.2. Penggunaan Teknik Statistik dalam Konteks
Praktis 43
- 4.3. Distribusi Data 45
- 4.4. Grafik dan Diagram..... 47
- 4.5. Koefisien Variasi (Coefficient of Variation) 50
- 4.6. Analisis Lanjutan Mengenai Variabilitas..... 52

**BAB V KONSEP DISTRIBUSI NORMAL DALAM
TERAPAN PENELITIAN KUANTITATIF57**

- 5.1. Mengapa Distribusi Normal Penting dalam
Penelitian Kuantitatif?..... 57
- 5.2. Dasar-Dasar Distribusi Normal 60
- 5.3. Konsep Dasar Distribusi Normal. 60
- 5.4. Properti Distribusi Normal..... 64
- 5.5. Penggunaan Distribusi Normal dalam
Penelitian Kuantitatif..... 65

5.6.	Aplikasi Distribusi Normal dalam Analisis Data.....	67
5.7.	Studi Kasus dan Contoh Penerapan Distribusi Normal.....	69

BAB VI TEKNIK STATISTIK UNTUK MENGUJI HIPOTESIS MENGENAI UJI BEDA LEBIH DARI DUA KELOMPOK..... 73

6.1.	Pendahuluan.....	73
6.2.	Analisis Varian (ANOVA).....	74
6.3.	One-Way ANOVA.....	77
6.4.	Two-Way ANOVA.....	80
6.5.	ANOVA dengan Repeated Measures.....	83

BAB VII TEKNIK STATISTIK UNTUK MENGUJI HIPOTESIS MENGENAI HUBUNGAN DUA VARIABEL 89

7.1.	Hubungan Antar Variabel.....	89
7.1.1.	Definisi dan Jenis Hubungan.....	89
7.1.2.	Korelasi vs. Kausalitas.....	90
7.2.	Korelasi.....	91
7.2.1.	Pengertian Korelasi.....	91
7.2.2.	Hipotesis Korelasi.....	91
7.2.3.	Jenis-Jenis Korelasi.....	92
7.2.4.	Keterbatasan Korelasi.....	93
7.3.	Regresi Linier.....	94
7.3.1.	Pengertian Regresi Linier.....	94
7.3.2.	Hipotesis dalam Regresi Linier.....	94
7.3.3.	Langkah-Langkah Analisis Regresi Linier.....	95

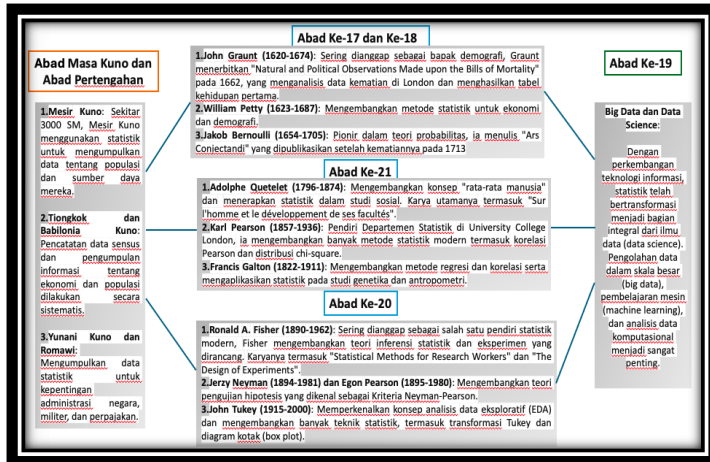
7.3.4.	Keterbatasan Regresi Linier	95
7.4.	Uji Chi-Square untuk Hubungan Dua Variabel Kategorikal.....	96
BAB VIII ANALISIS REGRESI SEDERHANA		101
8.1.	Pendahuluan	101
8.2.	Analisis Regresi Sederhana.....	101
8.3.	Uji Kelinearan dan Keberartian Regresi	107
8.4.	Ketelitian Prediksi.....	115
BAB IX UJI STATISTIK NON PARAMETRIK UJI BEDA DUA KELOMPOK.....		119
9.1.	Uji Wilcoxon Rank-Sum (Uji Mann-Whitney U).....	119
9.2.	Uji Wilcoxon Signed-Rank.....	121
9.3.	Uji Kolmogorov-Smirnov	124
9.4.	Uji Chi-Square untuk Data Berpasangan....	126
BAB X UJI STATISTIK NON PARAMETRIK UJI BEDA LEBIH DARI DUA KELOMPOK.....		131
10.1.	Uji Kruskal-Wallis H	131
10.2.	Uji Friedman	133
10.3.	Uji Cochran	135
10.4.	Uji Jonckheere-Terpstra.....	138
10.5.	Uji Median	140
DAFTAR PUSTAKA.....		143

BAB I

KONSEP DAN TEKNIK STATISTIKA, DASAR DASAR ANALISIS DAN HUBUNGANYA DENGAN PENELITIAN DAN PENGEMBANGAN ILMU PENGETAHUAN

1.1. Sejarah

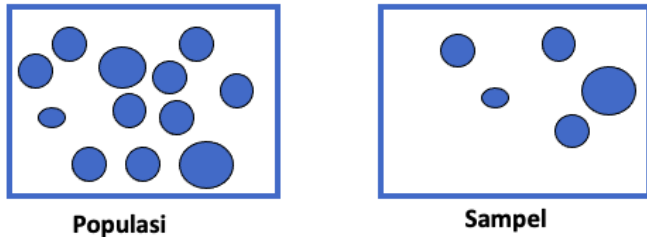
Statistik telah menjadi bagian penting dari kehidupan manusia selama ribuan tahun. Awalnya, statistik digunakan untuk mengumpulkan data tentang populasi, perdagangan, dan sumber daya alam, kemudian berkembang dengan berbagai macam pecahan metode untuk dipergunakan dalam berbagai bidang ilmu. Perkembangan terkini, statistik berkembang lagi ke big data dan data science untuk lebih terintegral lagi sebagaimana pada gambar 1 di bawah.



Gambar 1: Sejarah Statistika

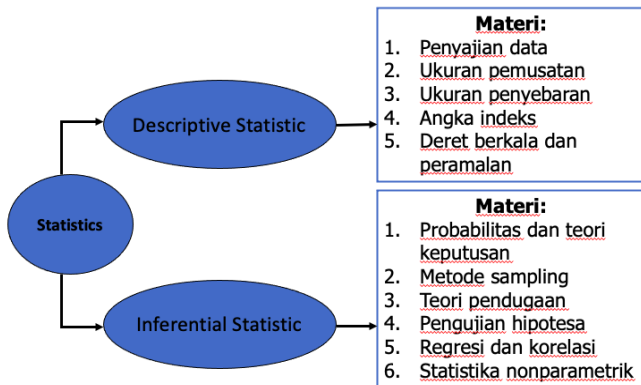
1.2. Konsep dan Teknik Statistika

STATISTIKA adalah pengetahuan yang berkaitan dengan metode, teknik atau cara untuk mengumpulkan, mengolah, menyajikan dan menganalisis data, kemudian menarik kesimpulan dan menginterpretasikan data tersebut. Dalam konsep statistika di kenal dengan adanya populasi dan sampel. Populasi yaitu seluruh kelompok yang menjadi fokus penelitian sedangkan, Sampel adalah Subset dari populasi yang diambil untuk dianalisis (Gambar 2).



Gambar 2. Perbedaan Populasi dan Sampel

Konsep statistika di bagi dua golongan yaitu statistik deskriptif dan statistik inferensial



Gambar 3. Konsep Statistik

1.2.1. Statistik Deskriptif

Statistik deskriptif adalah metode statistik yang digunakan untuk menggambarkan atau meringkas suatu kumpulan data secara kuantitatif. Statistik deskriptif tidak membuat kesimpulan atau

prediksi tentang populasi yang lebih besar, tetapi memberikan gambaran atau ringkasan tentang data yang ada. Beberapa jenis statistic deskriptif yang di gunakan

- a) Pengukuran tendensi sentral: mean, median, dan mode.
- b) Pengukuran dispersi: range, variance, dan standard deviation.
- c) Distribusi data dan bentuk kurva.

1.2.2. Statistik Inferensial

Statistik inferensial adalah cabang statistik yang berfokus pada pengambilan kesimpulan atau generalisasi tentang populasi berdasarkan sampel data. Tidak seperti statistik deskriptif, yang hanya meringkas data, statistik inferensial menggunakan data sampel untuk membuat estimasi, prediksi, atau pengujian hipotesis tentang populasi yang lebih besar. Beberapa konsep utama dalam statistik inferensial antara lain:

- a) **Pengujian Hipotesa:**

Pengujian hipotesis adalah teknik statistik yang digunakan untuk menentukan apakah ada cukup bukti dalam sampel data untuk mendukung suatu pernyataan tentang

populasi. Dalam proses ini, hipotesis nol (H_0) menyatakan bahwa tidak ada efek atau perbedaan, sedangkan hipotesis alternatif (H_1) menyatakan adanya efek atau perbedaan. Pengujian hipotesis melibatkan langkah-langkah seperti menentukan hipotesis, memilih tingkat signifikansi (α), mengumpulkan data, memilih uji statistik yang tepat, menghitung statistik uji, menentukan nilai kritis atau nilai p, dan membuat keputusan. Misalnya, dalam bidang kesehatan, uji t dapat digunakan untuk membandingkan rata-rata tekanan darah antara kelompok yang menerima pengobatan dan kelompok kontrol. Jika nilai $p < 0.05$, hipotesis nol ditolak, menunjukkan bahwa ada perbedaan signifikan antara kelompok. Dengan demikian, pengujian hipotesis membantu dalam membuat keputusan berbasis data dan menentukan hubungan atau perbedaan signifikan dalam penelitian.

b) **Probabilitas**

Probabilitas adalah konsep yang penting dalam mengukur ketidakpastian dan membuat keputusan yang informatif.

Probabilitas suatu kejadian, yang dinyatakan dengan nilai antara 0 dan 1, mencerminkan kemungkinan kejadian tersebut terjadi. Probabilitas teoretis dihitung berdasarkan model matematis, sedangkan probabilitas empiris didasarkan pada hasil percobaan atau data yang sudah ada. Misalnya, dalam bidang kesehatan, probabilitas seseorang memiliki penyakit dapat diperbarui berdasarkan hasil tes medis menggunakan teorema Bayes. Pemahaman yang baik tentang konsep dasar probabilitas dan distribusi probabilitas sangat penting untuk interpretasi yang tepat dan pengambilan keputusan yang berbasis data.

c) **Regresi dan Korelasi**

Analisis regresi adalah teknik statistik yang digunakan untuk memodelkan dan menganalisis hubungan antara variabel independen (X) dan variabel dependen (Y) dengan model linear.

Sedangkan uji korelasi adalah teknik statistik yang digunakan untuk mengukur dan menganalisis hubungan antara dua variabel. Korelasi Pearson mengukur hubungan linier

antara dua variabel kontinu, dengan nilai koefisien berkisar antara -1 sampai +1. Korelasi Spearman, yang mengukur hubungan monotonik, cocok untuk variabel ordinal atau ketika data tidak memenuhi asumsi distribusi normal. Proses uji korelasi melibatkan pengumpulan data, eksplorasi data awal, uji asumsi, perhitungan koefisien korelasi, dan uji signifikansi. Misalnya, dalam bidang kesehatan, uji korelasi Pearson antara tekanan darah dan indeks massa tubuh (IMT) dapat menunjukkan hubungan positif sedang dengan koefisien 0.45 dan nilai $p < 0.01$, menunjukkan signifikansi statistik. Hasil ini menunjukkan bahwa peningkatan IMT cenderung berhubungan dengan peningkatan tekanan darah. Dengan demikian, uji korelasi memberikan wawasan penting tentang hubungan antar variabel yang dapat membantu dalam pengambilan keputusan berbasis data.

1.3. Data Dalam Statistik

1.3.1. Sumber Data Statistik

Sumber data statistik bisa berasal dari data

primer dan data sekunder. Data primer adalah data yang dikumpulkan langsung melalui survei, wawancara, eksperimen, dan observasi lapangan. Sedangkan data sekunder berasal dari sumber yang sudah ada, seperti laporan pemerintah, publikasi ilmiah, database online, dan laporan tahunan perusahaan. Karakteristik sumber data yang baik meliputi reliabilitas, validitas, keterjangkauan, dan relevansi.

1.3.2. Jenis Data Statistik

Data terbagi atas dua jenis data, yang pertama Data Kualitatif yaitu data yang dinyatakan dalam bentuk bukan angka antara lain jenis pekerjaan, status marital, tingkat kepuasan kerja. Serta Data Kuantitatif yaitu Data yang dinyatakan dalam bentuk angka, antara lain lama bekerja, jumlah gaji, usia, hasil ulangan.

1.3.3. Skala Pengukuran Data

a) Data Nominal

Data berskala nominal adalah data yang diperoleh dengan cara kategorisasi atau klasifikasi. Mempunyai **CIRI**: posisi data setara dan tidak bisa dilakukan operasi

matematika (+, -, x,:), contoh nya: jenis kelamin, jenis pekerjaan.

b) Data Ordinal

Data berskala ordinal adalah data yang diperoleh dengan cara kategorisasi atau klasifikasi, tetapi di antara data tersebut terdapat hubungan. Mempunyai **CIRI**: posisi data tidak setara dan tidak bisa dilakukan operasi matematika (+, -, x,:), contoh nya: kepuasan kerja, motivasi.

c) Data Interval

Data berskala interval adalah data yang diperoleh dengan cara pengukuran, di mana jarak antara dua titik skala sudah diketahui. Mempunyai **CIRI**: Tidak ada kategorisasi, bisa dilakukan operasi matematika, contoh: temperatur yang diukur berdasarkan $^{\circ}\text{C}$ dan $^{\circ}\text{F}$, sistem kalender.

d) Data Rasio

Data berskala rasio adalah data yang diperoleh dengan cara pengukuran, di mana jarak antara dua titik skala sudah diketahui dan mempunyai titik 0 absolut. Mempunyai **CIRI**: tidak ada kategorisasi

dan bisa dilakukan operasi matematika, contoh: gaji, skor ujian, jumlah buku.

1.4. Statistik dan Hubungan dengan Pengembangan Penelitian dan Pengembangan Ilmu Pengetahuan

Dalam penelitian ilmiah metode statistika dipergunakan dalam pembuatan rencana penelitian (research design). Statistik memainkan peran penting dalam mengumpulkan, menganalisis, dan menginterpretasi data. Berikut adalah tahapan dan konsep utama dalam statistik dan penelitian:

a) Identifikasi Masalah

Identifikasi masalah merupakan langkah krusial dalam proses penelitian, karena menentukan arah dan fokus penelitian. Langkah-langkah yang umum dilakukan dalam identifikasi masalah meliputi observasi awal, kajian literatur, diskusi dengan ahli, pengumpulan data pendahuluan, dan analisis data awal. Masalah penelitian yang baik harus relevan, spesifik, terukur, dapat dipecahkan, dan signifikan. Metode kualitatif dan kuantitatif dapat digunakan dalam proses ini, dengan alat bantu seperti kuesioner dan software analisis data.

Identifikasi masalah yang tepat memastikan penelitian yang efektif dan memberikan kontribusi yang berarti dalam bidang studi terkait.

b) **Pengumpulan dan Pengolahan Data:**

Dalam penelitian, penting untuk menentukan kebutuhan data berdasarkan pertanyaan penelitian dan memilih sumber data yang tepat. Mengumpulkan dan memproses data secara akurat sangat penting untuk analisis lebih lanjut. Memilih sumber data yang tepat mendukung validitas dan reliabilitas penelitian serta membantu dalam pengambilan keputusan yang informatif

c) **Analisis Data:**

Analisis data adalah langkah penting dalam penelitian yang bertujuan untuk mendapatkan wawasan yang bermakna dari data yang dikumpulkan. Tahapan analisis data meliputi pengumpulan, pembersihan, pengkodean, dan eksplorasi data. Metode analisis data dapat dibagi menjadi analisis statistik deskriptif dan inferensial sebagaimana yang telah di sampaikan sebelumnya.

d) **Interpretasi Hasil**

Interpretasikan temuan statistik dalam konteks pertanyaan penelitian, memberikan makna dan konteks pada hasil analisis data. Misalnya, jika sebuah studi menemukan bahwa pengobatan baru menurunkan tekanan darah rata-rata sebesar 10 mmHg dengan nilai $p = 0.01$, interpretasinya adalah pengobatan tersebut efektif secara signifikan dalam menurunkan tekanan darah dibandingkan dengan plasebo. Nilai p yang kecil menunjukkan bahwa hasil ini tidak mungkin terjadi secara kebetulan. Namun, penting juga untuk mengakui keterbatasan studi seperti ukuran sampel yang kecil, yang dapat mempengaruhi generalisasi temuan ini. Dalam mengkomunikasikan hasil, penting untuk menggunakan bahasa yang jelas dan visualisasi data yang mudah dipahami, serta mengaitkan temuan dengan implikasi praktis dan kebijakan yang relevan.

e) **Penarikan Kesimpulan**

Kesimpulan adalah bagian terakhir dalam penelitian di sertai pemberian rekomendasi berdasarkan analisis dengan menyarankan arah penelitian masa depan yang diperlukan. Pendekatan terstruktur ini memastikan analisis

data yang sistematis dan menyeluruh, menghasilkan kesimpulan yang andal dan valid. Dalam validasi hipotetis, Statistik membantu dalam pengujian hipotesis, menentukan apakah ada cukup bukti untuk mendukung atau menolak hipotesis penelitian. Sedangkan Keandalan Temuan menggunakan statistik, peneliti dapat mengukur keandalan dan signifikansi temuan mereka, memastikan bahwa hasil tidak terjadi secara kebetulan. Dengan mengikuti langkah-langkah ini, peneliti dapat memastikan bahwa penelitian didasarkan pada data yang kuat dan analisis yang tepat.

BAB II

TEKNIK PENYAJIAN DATA STATISTIK

Penggunaan teknik penyajian data statistik menjadi semakin penting dalam berbagai bidang seperti bisnis, ilmu sosial, kedokteran, dan lainnya. Hal ini disebabkan oleh kebutuhan untuk menganalisis dan memahami informasi yang terkandung dalam data. Teknik-teknik ini membantu dalam mengorganisir data, mengidentifikasi pola, dan mengambil keputusan yang lebih baik berdasarkan pemahaman yang lebih dalam terhadap fenomena yang diamati.

Dalam konteks ini bertujuan untuk memberikan pemahaman yang komprehensif tentang teknik-teknik penyajian data statistik. Materi yang disajikan mencakup dasar-dasar statistik seperti pengertian variabel, tipe data, dan ukuran-ukuran pemusatan serta penyebaran data. Selain itu, buku ini juga akan membahas teknik-teknik penyajian data secara praktis, seperti pembuatan tabel frekuensi, diagram, dan analisis statistik lanjutan.

Adapun ruang lingkup mencakup pengenalan konsep-konsep dasar hingga teknik-teknik lanjutan. Materi disajikan dengan bahasa yang sederhana dan

dilengkapi dengan contoh kasus yang relevan untuk membantu pembaca memahami penerapan konsep-konsep tersebut dalam situasi nyata. Diharapkan setelah membaca buku ini, pembaca akan memiliki pemahaman yang kuat tentang teknik-teknik penyajian data statistik dan mampu mengaplikasikannya dalam analisis data yang mereka kerjakan.

2.1. Tabel Frekuensi

Tabel frekuensi adalah salah satu teknik penyajian data statistik yang digunakan untuk mengorganisir data kategori atau data diskrit. Tujuan utama dari tabel frekuensi adalah untuk menunjukkan jumlah frekuensi atau kemunculan setiap kategori dalam suatu variabel. Hal ini membantu dalam visualisasi distribusi data dan mempermudah analisis statistik lebih lanjut.

Pada dasarnya, tabel frekuensi terdiri dari dua kolom utama: kolom pertama berisi kategori atau nilai-nilai yang diamati dalam data, sedangkan kolom kedua berisi frekuensi atau jumlah kemunculan setiap kategori tersebut. Selain itu, tabel frekuensi juga dapat dilengkapi dengan kolom persentase untuk menunjukkan proporsi atau persentase dari masing-masing kategori terhadap total data.

Berikut langkah-langkah untuk membuat tabel

frekuensi:

1. Identifikasi Kategori: Tentukan kategori atau nilai-nilai yang ingin dianalisis dalam data. Misalnya, jika memiliki data mengenai jenis-jenis buah yang dijual di sebuah toko, kategori-kategori tersebut mungkin adalah apel, pisang, jeruk, dan mangga.
2. Hitung Frekuensi: Lakukan penghitungan untuk setiap kategori guna mengetahui berapa kali masing-masing kategori muncul dalam data. Misalnya, jika dalam data penjualan buah terdapat 50 transaksi dengan apel, 30 transaksi dengan pisang, 20 transaksi dengan jeruk, dan 40 transaksi dengan mangga, maka ini adalah frekuensi masing-masing kategori.
3. Buat Tabel: Susun tabel dengan kolom kategori pada bagian pertama dan kolom frekuensi pada bagian kedua, juga dapat menambahkan kolom persentase jika diinginkan.
4. Isi Tabel: Masukkan nilai-nilai kategori dan frekuensi yang telah dihitung ke dalam tabel.
5. Interpretasi: Analisis tabel frekuensi untuk memahami pola distribusi data. Perhatikan kategori yang memiliki frekuensi tinggi atau

rendah, serta lihat apakah ada pola atau tren yang dapat dilihat dari distribusi data.

Tabel frekuensi sangat berguna dalam berbagai bidang seperti survei, riset pasar, dan analisis data secara umum. Dengan menggunakan tabel frekuensi, dapat dengan mudah mengorganisir dan memahami distribusi data yang dimiliki.

2.2. Diagram Batang

Diagram batang adalah salah satu jenis diagram yang digunakan untuk menggambarkan data dalam bentuk bar atau batang yang memiliki panjang yang berbeda-beda. Diagram ini sering digunakan untuk membandingkan data antara beberapa kategori yang berbeda. Berikut adalah penjelasan lengkap mengenai diagram batang beserta daftar pustaka yang dapat kamu gunakan untuk mendalami topik ini lebih lanjut.

1. Definisi Diagram batang adalah representasi visual dari data menggunakan batang atau bar yang mewakili nilai-nilai yang diukur. Panjang bar menggambarkan nilai-nilai tersebut, dan bar bisa berada dalam orientasi vertikal (baris) atau horizontal (kolom), tergantung pada preferensi atau jenis data yang ingin ditampilkan.

2. Komponen Utama

- Bar/Batang: Merepresentasikan nilai atau kategori yang diukur. Panjang bar menunjukkan nilai yang terkait.
- Sumbu-X (Horizontal): Menunjukkan kategori atau variabel independen.
- Sumbu-Y (Vertikal): Menunjukkan nilai atau jumlah dari variabel dependen.

3. Keuntungan

- Mudah dipahami dan digunakan untuk membandingkan data secara visual.
- Cocok untuk menunjukkan perubahan nilai dari waktu ke waktu atau perbandingan antara beberapa kategori.
- Dapat digunakan dengan berbagai jenis data, termasuk data kualitatif dan kuantitatif.

4. Jenis-Jenis Diagram Batang

- Diagram Batang Vertikal: Baris atau batang tegak mewakili kategori atau nilai.
- Diagram Batang Horizontal: Baris atau batang mendatar mewakili kategori atau nilai.

5. Contoh Penggunaan

- Perbandingan penjualan bulanan per produk dalam sebuah toko.
- Perbandingan suhu rata-rata bulanan dalam setahun.
- Perbandingan populasi antar negara dalam sebuah wilayah.

2.3. Diagram Lingkaran

Diagram lingkaran adalah representasi visual dari data yang disusun dalam bentuk lingkaran yang dibagi menjadi beberapa bagian atau sektor. Setiap sektor mewakili proporsi atau persentase yang berbeda dari total keseluruhan. Diagram ini berguna untuk memperlihatkan perbandingan relatif antara beberapa kategori data.

Secara umum, diagram lingkaran terdiri dari beberapa komponen utama:

1. Judul Diagram: Ini adalah bagian atas diagram yang memberikan judul atau topik dari data yang direpresentasikan.
2. Lingkaran: Merupakan bentuk utama dari diagram ini yang melambangkan keseluruhan data.
3. Sektor: Bagian-bagian lingkaran yang mewakili proporsi atau persentase dari keseluruhan data.

Setiap sektor biasanya diberi label atau keterangan yang menunjukkan kategori atau nilai yang direpresentasikan.

4. Label atau Keterangan: Informasi tambahan yang diberikan untuk menjelaskan setiap sektor atau kategori dalam diagram.
5. Legenda: Digunakan untuk menjelaskan makna dari warna atau pola yang digunakan dalam diagram jika ada lebih dari satu variabel yang direpresentasikan.
6. Persentase: Nilai persentase atau proporsi setiap sektor sering ditampilkan di dalam atau di sekitar sektor tersebut.
7. Garis Tengah atau Pusat: Garis atau titik tengah lingkaran yang menjadi pusat dari semua sektor dan sering kali berfungsi sebagai titik fokus diagram.
8. Diagram lingkaran sangat berguna dalam menyajikan data kualitatif, seperti perbandingan proporsi penjualan berbagai produk dalam suatu perusahaan, distribusi persentase populasi berdasarkan kelompok usia, atau alokasi anggaran berbagai departemen dalam sebuah organisasi.

2.4. Diagram Garis

Diagram garis adalah jenis diagram yang menampilkan hubungan antara dua variabel dengan menggunakan garis yang menghubungkan titik-titik data yang terpisah. Biasanya, sumbu horizontal (x) digunakan untuk variabel independen, sementara sumbu vertikal (y) digunakan untuk variabel dependen. Diagram garis digunakan untuk menunjukkan tren atau pola perubahan dari waktu ke waktu atau dari satu variabel terhadap variabel lainnya.

Berikut adalah beberapa komponen utama dalam diagram garis:

1. Sumbu X (Horizontal): Ini adalah sumbu yang mewakili variabel independen atau waktu. Biasanya, titik-titik data untuk variabel ini ditempatkan di sepanjang sumbu ini.
2. Sumbu Y (Vertikal): Sumbu ini mewakili variabel dependen yang ingin ditampilkan. Data untuk variabel ini diplot di sepanjang sumbu ini.
3. Garis Data: Garis ini menghubungkan titik-titik data yang mewakili nilai-nilai variabel dependen pada titik waktu atau titik-titik lainnya.
4. Judul Diagram: Memberikan informasi tentang apa yang direpresentasikan dalam diagram.

5. Label Sumbu: Label-label yang ditempatkan di sepanjang sumbu x dan sumbu y untuk menunjukkan skala dan unit dari masing-masing variabel.
6. Keterangan: Jika ada lebih dari satu garis data dalam diagram, keterangan dapat digunakan untuk menjelaskan makna dari setiap garis.
7. Legenda: Digunakan untuk menjelaskan makna dari warna atau pola yang digunakan dalam diagram jika ada lebih dari satu variabel yang direpresentasikan.
8. Diagram garis sering digunakan dalam analisis data, terutama untuk menunjukkan tren, pola, atau hubungan antara variabel yang berubah seiring waktu. Contoh penggunaannya meliputi grafik pergerakan harga saham dari waktu ke waktu, penjualan produk dalam periode waktu tertentu, atau perubahan suhu harian selama beberapa bulan.

2.5. Histogram

Sebuah histogram adalah representasi grafis dari distribusi frekuensi data. Ini biasanya digunakan untuk menampilkan distribusi data numerik yang terkumpul dalam interval atau kelas tertentu. Histogram

membantu dalam memvisualisasikan bagaimana data tersebar dan menunjukkan kecenderungan, pola, atau anomali dalam data tersebut.

Berikut adalah langkah-langkah untuk membuat histogram:

1. Mengelompokkan Data: Pertama, data harus dikelompokkan ke dalam interval atau kelas. Misalnya, jika Anda memiliki data tentang tinggi badan orang, Anda dapat mengelompokkannya dalam interval seperti 150-160 cm, 161-170 cm, dan seterusnya.
2. Menghitung Frekuensi: Hitung berapa banyak data yang jatuh ke dalam setiap interval atau kelas. Ini adalah langkah penting untuk membangun histogram.
3. Membuat Grafik: Sumbu horizontal histogram akan mewakili interval atau kelas, sedangkan sumbu vertikal akan mewakili frekuensi atau jumlah data dalam interval tersebut. Untuk setiap interval, gambar bar atau balok dengan tinggi sesuai dengan frekuensi data di interval tersebut.
4. Menambahkan Label dan Judul: Tambahkan label sumbu dan judul untuk menjelaskan isi histogram dengan jelas.

5. Histogram dapat memberikan informasi yang berharga tentang distribusi data. Ini dapat membantu mengidentifikasi apakah data cenderung normal (disebar secara merata), memiliki skewness (kecondongan), atau memiliki kumpulan data yang terpusat pada nilai tertentu.

2.6. Poligon Frekuensi

Poligon frekuensi adalah jenis diagram garis yang digunakan untuk menggambarkan distribusi frekuensi data numerik. Ini serupa dengan histogram, tetapi bedanya adalah bahwa dalam poligon frekuensi, titik pusat atas setiap interval atau kelas digunakan untuk membuat garis yang menghubungkan titik-titik ini untuk membentuk poligon. Poligon frekuensi membantu dalam memvisualisasikan pola distribusi data dengan lebih halus daripada histogram, karena garis yang menghubungkan titik-titik pusat atas interval memberikan representasi yang lebih kontinu.

Berikut adalah langkah-langkah untuk membuat poligon frekuensi:

1. Mengelompokkan Data: Seperti pada histogram, data harus dikelompokkan ke dalam interval atau kelas.

2. Menghitung Frekuensi: Hitung frekuensi atau jumlah data yang jatuh ke dalam setiap interval atau kelas.
3. Menentukan Titik Pusat: Untuk setiap interval, tentukan titik pusat atas. Ini dapat dihitung dengan mengambil rata-rata dari batas atas dan batas bawah interval.
4. Membuat Grafik: Gunakan titik pusat atas dari setiap interval untuk menggambar titik pada sumbu horizontal (sumbu x), dan gunakan frekuensi atau jumlah data sebagai nilai pada sumbu vertikal (sumbu y). Kemudian, hubungkan titik-titik ini dengan garis lurus untuk membentuk poligon frekuensi.
5. Menambahkan Label dan Judul: Tambahkan label sumbu dan judul untuk menjelaskan isi poligon frekuensi dengan jelas.
6. Poligon frekuensi sering digunakan dalam analisis statistik untuk menampilkan distribusi data secara visual. Ini memberikan cara yang lebih halus untuk melihat pola-pola dalam data daripada histogram, terutama ketika data terkumpul dalam interval yang lebar.

BAB III

TEKNIK STATISTIKA UNTUK MENETAPKAN UKURAN PEMUSATAN DATA, PENYEBARAN DATA, DAN LOKASI

3.1. Ukuran Pemusatan Data

Ukuran pemusatan data adalah statistik yang digunakan untuk mengukur "pusat" dari sebuah kumpulan data. Ini memberikan gambaran tentang titik tengah atau nilai yang mewakili data secara keseluruhan. Beberapa ukuran pemusatan data yang umum digunakan termasuk mean (rata-rata), median (nilai tengah), dan modus (nilai yang paling sering muncul). Berikut adalah penjelasan lengkap untuk masing-masing ukuran tersebut:

1. Mean (Rata-Rata) Rata-rata adalah jumlah semua nilai dalam kumpulan data dibagi dengan jumlah nilai tersebut. Rumus matematisnya adalah:

$$\text{Mean} = \frac{\text{Jumlah semua nilai}}{\text{Jumlah nilai}}$$

Rata-rata sangat sensitif terhadap nilai ekstrem (outlier), karena nilai-nilai yang jauh dari rata-rata dapat memengaruhi hasil akhir secara signifikan.

2. **Median** Median adalah nilai tengah dari kumpulan data yang telah diurutkan. Untuk menemukan median perlu mengurutkan data dari yang terkecil hingga terbesar, lalu ambil nilai tengahnya. Jika jumlah data ganjil, median adalah nilai di tengah. Jika jumlah data genap, median adalah rata-rata dari dua nilai tengah.
3. **Modus** Modus adalah nilai yang paling sering muncul dalam kumpulan data. Kumpulan data dapat memiliki satu modus (unimodal), dua modus (bimodal), atau lebih dari dua modus (multimodal). Modus digunakan terutama dalam data kategorikal, seperti jenis-jenis yang berbeda.

Selain tiga ukuran tersebut, terdapat juga ukuran pemusatan data lain yang kurang umum, seperti geometrik mean, harmonic mean, dan weighted mean.

3.2. Ukuran Penyebaran Data

Ukuran penyebaran data adalah statistik yang

digunakan untuk mengukur sejauh mana data tersebar di sekitar nilai pusatnya. Ini memberikan informasi tentang variasi dan distribusi data dalam suatu sampel atau populasi. Berikut adalah beberapa ukuran penyebaran data yang umum digunakan:

1. Rentang (Range): Rentang adalah perbedaan antara nilai maksimum dan minimum dalam data. Formula untuk menghitung rentang adalah:

$$\text{Rentang} = \text{Nilai maksimum} - \text{Nilai minimum}$$

2. Jangkauan Antar Kuartil (Interquartile Range, IQR): IQR adalah perbedaan antara kuartil atas (Q3) dan kuartil bawah (Q1) dari data. Ini memberikan gambaran tentang seberapa tersebarnya data dalam "bagian tengah" distribusi. Formula IQR adalah:

$$IQR = Q3 - Q1$$

3. Variansi (Variance): Variansi mengukur seberapa jauh data tersebar dari nilai rata-rata. Rumus variansi S^2 untuk sampel adalah:

$$s^2 = \frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n-1}$$

Di mana x_i adalah nilai data ke- i , \bar{x} adalah rata-rata data, dan n adalah jumlah data.

4. Deviasi Standar (Standard Deviation): Deviasi standar adalah akar kuadrat dari variansi. Ini memberikan gambaran tentang seberapa bervariasi data dalam satu set. Rumus deviasi standar S adalah:

$$s = \sqrt{\frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n-1}}$$

5. Koefisien Variasi (Coefficient of Variation, CV): CV adalah ukuran relatif dari deviasi standar terhadap rata-rata. Ini berguna untuk membandingkan variabilitas antara dua set data yang memiliki rata-rata yang berbeda-beda. Formula CV adalah:

$$CV = \frac{s}{\bar{x}} \times 100\%$$

Di mana S adalah deviasi standar dan x adalah rata-rata data.

6. Jangkauan Intersempat (Interquartile Range, IQR): IQR adalah ukuran penyebaran data yang dihitung dari perbedaan antara kuartil atas dan kuartil bawah. Formula untuk menghitung IQR adalah:

$$IQR = Q3 - Q1$$

Di mana $Q3$ adalah nilai kuartil atas dan $Q1$ adalah nilai kuartil bawah.

3.3. Ukuran Lokasi

Ukuran lokasi dalam statistik adalah metode untuk menentukan titik sentral atau "nilai tengah" dari suatu kumpulan data. Ini membantu dalam memberikan gambaran tentang pusat distribusi data. Berikut adalah beberapa ukuran lokasi yang umum digunakan:

1. Rata-Rata (Mean): Rata-rata adalah jumlah dari semua nilai dalam kumpulan data dibagi dengan jumlah nilai tersebut. Ini adalah ukuran lokasi yang paling umum digunakan dan seringkali merepresentasikan "nilai tengah" dari data. Rumus rata-rata untuk suatu sampel adalah:

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

Di mana x adalah rata-rata, x_1 adalah nilai data ke- i , dan n adalah jumlah data.

2. Median: Median adalah nilai tengah dalam kumpulan data ketika data diurutkan dari terkecil hingga terbesar. Jika jumlah data ganjil, median adalah nilai di tengah-tengah setelah diurutkan. Jika jumlah data genap, median

adalah rata-rata dari dua nilai tengah setelah diurutkan.

3. Modus: Modus adalah nilai yang paling sering muncul dalam kumpulan data. Suatu kumpulan data bisa memiliki lebih dari satu modus jika ada beberapa nilai dengan frekuensi yang sama tertinggi.
4. Kuartil (Quartiles): Kuartil membagi data menjadi empat bagian yang sama besar. Kuartil utama adalah:
 - Kuartil Pertama (Q1): Nilai yang membagi 25% data terendah.
 - Kuartil Kedua atau Median (Q2): Nilai yang membagi data menjadi dua bagian yang sama besar.
 - Kuartil Ketiga (Q3): Nilai yang membagi 75% data terendah.
5. Persentil (Percentiles): Persentil adalah nilai yang membagi data menjadi seratus bagian yang sama besar. Persentil ke-50 sama dengan median, sedangkan persentil ke-25 adalah kuartil pertama, dan persentil ke-75 adalah kuartil ketiga.
6. Rata-rata Terpotong (Trimmed Mean): Rata-rata terpotong adalah rata-rata dari sebagian

data yang dibuang pada titik-titik ekstrem, misalnya, 5% data terkecil dan 5% data terbesar.

7. Midrange: Midrange adalah rata-rata dari nilai maksimum dan minimum dalam kumpulan data. Rumusnya adalah:

$$\text{Midrange} = \frac{\text{Nilai maksimum} + \text{Nilai minimum}}{2}$$

3.4. Teknik Statistika Lanjutan

Teknik statistika lanjutan melibatkan penggunaan konsep dan metode statistika yang lebih kompleks untuk menganalisis data yang lebih mendalam dan rumit. Berikut adalah beberapa teknik statistika lanjutan yang umum digunakan:

1. Analisis Regresi: Analisis regresi digunakan untuk memahami hubungan antara satu atau lebih variabel independen (penyebab) dengan variabel dependen (hasil). Regresi linier sederhana adalah kasus di mana hanya ada satu variabel independen, sementara regresi linier berganda melibatkan lebih dari satu variabel independen. Teknik regresi lainnya termasuk regresi logistik untuk variabel dependen biner dan regresi non-linier untuk hubungan yang tidak linier.

2. Analisis Varians (ANOVA): ANOVA digunakan untuk membandingkan rata-rata antara tiga atau lebih kelompok data. ANOVA dapat digunakan untuk analisis satu faktor (satu variabel independen), dua faktor (dua variabel independen), atau lebih. Ini membantu dalam menentukan apakah perbedaan antara rata-rata kelompok-kelompok tersebut signifikan secara statistik.
3. Analisis Komponen Utama (Principal Component Analysis, PCA): PCA adalah teknik reduksi dimensi yang digunakan untuk mengurangi kompleksitas data dengan mengidentifikasi pola-pola dalam data dan mengekstraksi komponen-komponen utama yang menjelaskan variasi data sebanyak mungkin.
4. Analisis Kluster (Cluster Analysis): Analisis kluster digunakan untuk mengelompokkan data ke dalam kelompok-kelompok yang serupa berdasarkan kemiripan antara data tersebut. Teknik ini membantu dalam mengidentifikasi pola-pola atau segmen-segmen dalam data yang mungkin tidak terlihat pada pandangan pertama.

5. Analisis Seri Waktu (Time Series Analysis): Analisis seri waktu digunakan untuk memahami pola-pola, tren, dan pola musiman dalam data yang diukur sepanjang waktu. Ini melibatkan penggunaan model-model seperti moving average, autoregressive integrated moving average (ARIMA), dan model-model lainnya untuk memprediksi perilaku masa depan dari data seri waktu.
6. Analisis Regresi Logistik (Logistic Regression Analysis): Regresi logistik digunakan ketika variabel dependen adalah biner atau kategorikal. Ini membantu dalam memprediksi probabilitas kejadian suatu peristiwa berdasarkan variabel-variabel independen.
7. Analisis Faktor (Factor Analysis): Analisis faktor adalah teknik yang digunakan untuk mengidentifikasi pola-pola atau faktor-faktor yang mendasari korelasi antara berbagai variabel. Ini membantu dalam menyederhanakan interpretasi data dengan mengelompokkan variabel-variabel yang saling terkait menjadi faktor-faktor yang lebih besar.
8. Analisis Regresi Nonparametrik: Analisis regresi nonparametrik adalah pendekatan yang

digunakan ketika asumsi-asumsi distribusi tertentu tidak dipenuhi dalam data. Ini melibatkan metode seperti regresi spline, regresi loess, dan regresi kernel.

3.5. Penerapan Teknik Statistika dalam Penelitian

Penerapan teknik statistika dalam penelitian adalah proses penggunaan metode statistika untuk mengumpulkan, menganalisis, dan menafsirkan data dalam rangka menjawab pertanyaan penelitian atau menguji hipotesis. Berikut adalah langkah-langkah umum dalam penerapan teknik statistika dalam penelitian:

1. Perumusan Pertanyaan Penelitian atau Hipotesis: Langkah pertama adalah merumuskan pertanyaan penelitian yang jelas atau hipotesis yang akan diuji. Misalnya, apakah terdapat hubungan antara variabel X dan variabel Y? Apakah perlakuan A lebih efektif daripada perlakuan B?
2. Perencanaan Desain Penelitian: Selanjutnya, perencanaan desain penelitian dilakukan, termasuk pemilihan sampel, metode pengumpulan data, dan perancangan eksperimen jika diperlukan. Desain penelitian

yang baik penting untuk memastikan validitas dan reliabilitas hasil.

3. Pengumpulan Data: Data dikumpulkan sesuai dengan desain penelitian yang telah direncanakan. Metode pengumpulan data dapat berupa survei, observasi, eksperimen, atau pengumpulan data sekunder dari sumber yang ada.
4. Preprocessing Data: Data yang telah dikumpulkan kemudian diproses untuk membersihkan data yang tidak valid atau tidak lengkap, melakukan transformasi data jika diperlukan, dan menyiapkan data untuk analisis statistika.
5. Analisis Data: Setelah data diproses, langkah selanjutnya adalah melakukan analisis statistika. Berikut adalah beberapa teknik statistika yang sering digunakan dalam penelitian:
 - Statistik Deskriptif: Digunakan untuk menggambarkan dan merangkum data, seperti rata-rata, median, modus, deviasi standar, dan visualisasi data.
 - Uji Hipotesis: Digunakan untuk menguji hipotesis penelitian. Contohnya adalah uji t,

uji ANOVA, uji chi-square, uji regresi, dan lain sebagainya. Tujuannya adalah untuk menentukan apakah perbedaan atau hubungan antara variabel adalah hasil dari kebetulan atau ada hubungan yang signifikan secara statistik.

- Analisis Regresi: Digunakan untuk memahami hubungan antara variabel dependen dan variabel independen, serta untuk memprediksi nilai variabel dependen berdasarkan variabel independen.
- Analisis Multivariat: Digunakan untuk memeriksa hubungan antara beberapa variabel pada saat yang sama, seperti analisis faktor, analisis kluster, dan analisis komponen utama.
- Analisis Seri Waktu: Digunakan untuk menganalisis pola-pola dan tren dalam data yang diukur sepanjang waktu.
- Analisis Survival: Digunakan untuk menganalisis waktu hingga terjadinya suatu peristiwa tertentu.
- Interpretasi Hasil: Setelah analisis statistika dilakukan, hasilnya diinterpretasikan untuk menjawab pertanyaan penelitian atau

menguji hipotesis. Hasil yang signifikan secara statistik menunjukkan adanya hubungan atau perbedaan yang tidak dapat dijelaskan oleh kebetulan.

6. Penyusunan Laporan: Akhirnya, hasil penelitian dan analisis statistika disusun dalam bentuk laporan penelitian yang mencakup ringkasan temuan, interpretasi hasil, kesimpulan, dan rekomendasi jika diperlukan.

BAB IV

TEKNIK STATISTIK UNTUK MENGHITUNG VARIABILITAS DATA

4.1. Pengertian Variabilitas Data

Variabilitas data merujuk pada tingkat variasi atau perbedaan antara nilai-nilai dalam sebuah kumpulan data. Secara sederhana, variabilitas mencerminkan seberapa jauh data tersebar dari nilai tengah atau pusatnya. Dalam analisis statistik, memahami variabilitas data penting karena memberikan informasi tentang seberapa stabil atau bervariasi data yang diamati. Berikut adalah penjelasan lebih lanjut tentang pengertian variabilitas data:

- 1. Definisi Variabilitas**

Variabilitas data mengacu pada perbedaan antara nilai-nilai dalam kumpulan data. Ini mencakup variasi atau dispersi dari nilai-nilai tersebut. Semakin besar variabilitasnya, semakin besar pula variasi antara nilai-nilai individual dalam data.

- 2. Pentingnya Memahami Variabilitas**

Memahami variabilitas data penting karena memberikan wawasan tentang karakteristik dan sifat data yang diamati. Variabilitas dapat mengindikasikan seberapa stabil atau fluktuatif suatu fenomena atau variabel, yang dapat memiliki implikasi besar dalam pengambilan keputusan.

3. Jenis-jenis Variabilitas

Variabilitas data dapat muncul dalam berbagai bentuk. Ini dapat mencakup variasi dalam rentang nilai, sebaran data di sekitar nilai tengah, ketidakteraturan dalam pola data, atau asimetri dalam distribusi data. Memahami jenis-jenis variabilitas ini membantu dalam pemilihan teknik analisis yang tepat.

Misalnya, data yang memiliki variabilitas tinggi mungkin memerlukan pendekatan analisis yang lebih cermat atau penggunaan teknik statistik yang lebih lanjut untuk menginterpretasikan pola atau tren yang tersembunyi di dalamnya. Di sisi lain, data dengan variabilitas rendah mungkin menunjukkan stabilitas yang lebih besar dalam fenomena yang diamati. Dengan memahami variabilitas data, analis dapat membuat kesimpulan yang lebih tepat dan memperoleh wawasan

yang lebih dalam tentang sifat dan karakteristik data yang diamati. Hal ini memungkinkan pengambilan keputusan yang lebih baik dalam berbagai bidang, mulai dari ilmu pengetahuan dan teknik hingga bisnis dan keuangan.

4.2. Penggunaan Teknik Statistik dalam Konteks Praktis

Penjelasan tentang penggunaan teknik statistik dalam konteks praktis mencakup berbagai aplikasi di berbagai bidang, seperti keuangan, ilmu sosial, ilmu alam, dan lainnya. Berikut adalah beberapa contoh penggunaannya:

a. Keuangan

Teknik statistik digunakan untuk mengukur variabilitas data ekonomi, seperti fluktuasi harga saham, nilai tukar mata uang, atau harga komoditas. Misalnya, standar deviasi digunakan untuk mengukur tingkat risiko investasi dalam portofolio saham.

b. Ilmu Sosial

Dalam penelitian sosial, teknik statistik digunakan untuk menganalisis data survei atau data perilaku sosial. Contohnya, median digunakan untuk menggambarkan pendapatan

rumah tangga di suatu wilayah, sementara analisis regresi digunakan untuk memahami hubungan antara variabel sosial seperti pendidikan dan pendapatan.

c. Ilmu Alam

Dalam ilmu alam, teknik statistik digunakan untuk menganalisis data dari eksperimen atau pengamatan alam. Misalnya, dalam biologi, analisis variansi (ANOVA) digunakan untuk membandingkan rata-rata antara beberapa kelompok eksperimental, sedangkan dalam meteorologi, distribusi normal digunakan untuk memodelkan dan memprediksi pola cuaca.

d. Teknik Analisis Data

Teknik statistik seperti regresi linier, regresi logistik, dan analisis faktor digunakan untuk memahami hubungan antara variabel dalam konteks yang lebih kompleks. Misalnya, regresi linier digunakan untuk memprediksi harga rumah berdasarkan faktor-faktor seperti luas tanah, lokasi, dan jumlah kamar tidur.

e. Pengambilan Keputusan

Statistik deskriptif seperti mean, median, dan modus digunakan untuk memberikan gambaran yang jelas tentang data, sementara teknik grafis

seperti histogram dan box plot membantu dalam visualisasi dan interpretasi data secara efektif. Hal ini sangat berguna dalam pengambilan keputusan strategis di berbagai bidang, dari bisnis hingga pemerintahan.

Dengan menggunakan teknik statistik yang tepat, para peneliti, analis, dan pengambil keputusan dapat mengungkap pola-pola yang tersembunyi dalam data, mengidentifikasi tren, memahami variabilitas, dan membuat keputusan yang didasarkan pada bukti empiris yang kuat.

4.3. Distribusi Data

Distribusi data mengacu pada pola atau cara di mana nilai-nilai dalam sebuah set data terdistribusi atau tersebar. Ini mencakup informasi tentang frekuensi kemunculan nilai-nilai tertentu serta pola atau karakteristik yang mungkin terdapat dalam data tersebut. Pemahaman tentang distribusi data penting karena dapat memberikan wawasan tentang sifat-sifat statistik dan memungkinkan analisis yang lebih tepat. Berikut adalah beberapa konsep yang terkait dengan distribusi data:

1. Distribusi Normal:

- a. Distribusi normal adalah distribusi data yang simetris di sekitar nilai rata-rata, dengan sebagian besar nilai terkumpul di sekitar rata-rata dan ekor distribusi menyebar secara simetris di kedua sisi.
 - b. Distribusi normal sering digunakan dalam statistik karena banyak fenomena alami dan proses pengukuran yang mengikuti pola ini.
 - c. Kurva distribusi normal memiliki bentuk lonceng dan dapat digambarkan dengan parameter rata-rata (mean) dan standar deviasi.
2. Distribusi Tidak Normal:
- a. Distribusi tidak normal mengacu pada pola distribusi data yang tidak mengikuti pola distribusi normal.
 - b. Contoh distribusi tidak normal termasuk distribusi asimetris seperti distribusi skedastis (positif atau negatif), distribusi bimodal (dua puncak), atau distribusi diskrit yang tidak memiliki pola kontinu.
3. Skewness dan Kurtosis:
- a. Skewness adalah ukuran statistik yang mengukur sejauh mana kurva distribusi data condong ke satu arah. Skewness positif

menunjukkan bahwa ekor distribusi lebih panjang di sisi kanan kurva (ekor positif), sementara skewness negatif menunjukkan kebalikannya.

- b. Kurtosis adalah ukuran statistik yang mengukur seberapa tajam atau datar puncak distribusi data. Kurtosis tinggi menunjukkan distribusi dengan puncak yang lebih tajam (kurva lebih tinggi dan lebih lancip), sedangkan kurtosis rendah menunjukkan distribusi dengan puncak yang lebih datar.

Memahami distribusi data membantu dalam pemilihan teknik analisis yang sesuai dan memahami karakteristik dan sifat data yang diamati. Hal ini juga penting dalam membuat asumsi yang tepat tentang model statistik yang akan digunakan untuk menganalisis data.

4.4. Grafik dan Diagram

Grafik dan diagram adalah alat visual yang digunakan untuk merepresentasikan data secara grafis. Maka membantu mengorganisir, meringkas, dan menyajikan informasi dalam bentuk yang mudah

dipahami. Berikut adalah beberapa penjelasan tentang grafik dan diagram:

1. Diagram Batang (Bar Charts):
 - a. Diagram batang adalah grafik yang menggunakan batang vertikal atau horizontal untuk mewakili frekuensi atau proporsi dari kategori atau variabel.
 - b. Maka sering digunakan untuk membandingkan nilai-nilai kategori yang berbeda atau menunjukkan perubahan dalam satu variabel dari waktu ke waktu.
2. Histogram:
 - a. Histogram adalah grafik yang digunakan untuk mewakili distribusi frekuensi dari data numerik yang berkelanjutan.
 - b. Maka terdiri dari serangkaian balok (bins) yang mewakili rentang nilai data, dengan tinggi balok menunjukkan jumlah observasi dalam setiap bin.
3. Diagram Garis (Line Charts):
 - a. Diagram garis menggunakan garis untuk menghubungkan titik data pada sumbu-x dan sumbu-y.

- b. Maka sering digunakan untuk menunjukkan tren atau perubahan dalam data sepanjang periode waktu tertentu.
- 4. Diagram Pie (Pie Charts):
 - a. Diagram pie menggunakan lingkaran untuk mewakili proporsi atau persentase dari keseluruhan.
 - b. Maka berguna untuk menunjukkan kontribusi relatif dari bagian-bagian dalam satu keseluruhan.
- 5. Diagram Kotak (Box-and-Whisker Plot):
 - a. Diagram kotak adalah grafik yang digunakan untuk menunjukkan distribusi dari data numerik melalui representasi visual dari kuartil.
 - b. Maka memberikan gambaran tentang sebaran data serta menunjukkan keberadaan nilai-nilai ekstrem atau pencilan.
- 6. Scatter Plot:
 - a. Scatter plot menggunakan titik-titik untuk mewakili pasangan data pada dua variabel yang berbeda.
 - b. Maka membantu dalam menunjukkan hubungan atau korelasi antara dua variabel.

Grafik dan diagram adalah alat yang kuat untuk menyajikan data secara visual dan memungkinkan pemahaman yang lebih cepat dan mendalam tentang pola dan tren dalam data. Pemilihan grafik atau diagram yang tepat tergantung pada jenis data yang akan disajikan dan tujuan dari analisis yang dilakukan.

4.5. Koefisien Variasi (Coefficient of Variation)

Koefisien variasi (Coefficient of Variation atau disingkat CV) adalah ukuran statistik yang digunakan untuk mengukur seberapa besar variasi relatif dari suatu set data dalam hubungannya dengan nilai rata-ratanya. Ini adalah metode standar yang digunakan untuk membandingkan variabilitas antara dua atau lebih set data yang memiliki skala atau satuan pengukuran yang berbeda. Berikut adalah penjelasan lebih rinci tentang koefisien variasi:

1. Formula Koefisien Variasi:

Koefisien variasi dinyatakan sebagai persentase dan dihitung dengan rumus: $CV = \frac{\text{Standar Deviasi}}{\text{Rata-rata}} \times 100\%$

Di mana:

- Standar Deviasi adalah ukuran penyebaran data.
- Rata-rata adalah nilai tengah dari data.

2. Interpretasi:
 - a. Koefisien variasi menyediakan informasi tentang seberapa besar variasi dalam set data relatif terhadap rata-ratanya.
 - b. Semakin tinggi nilai koefisien variasi, semakin tinggi tingkat variabilitas relatif dalam data.
 - c. Sebaliknya, semakin rendah nilai koefisien variasi, semakin rendah variabilitas relatif dalam data.
3. Keuntungan Penggunaan:
 - a. Koefisien variasi memungkinkan perbandingan langsung antara variabilitas dari dua atau lebih set data yang memiliki unit atau skala yang berbeda.
 - b. Ini adalah alat yang berguna dalam situasi di mana data memiliki satuan yang berbeda dan memungkinkan perbandingan yang adil antara kelompok-kelompok dengan karakteristik yang berbeda.
4. Keterbatasan Penggunaan:
 - a. Koefisien variasi tidak dapat dihitung jika rata-rata data adalah nol atau sangat mendekati nol, karena pembaginya adalah rata-rata data.

- b. Jika rata-rata data sangat kecil, koefisien variasi mungkin menjadi tidak stabil dan tidak dapat diinterpretasikan dengan akurat.

Koefisien variasi adalah alat yang berguna untuk membandingkan variabilitas relatif antara set data yang berbeda dalam konteks analisis statistik. Ini membantu dalam memahami sebaran data dan memberikan wawasan tentang karakteristik relatif dari dua atau lebih kelompok data yang berbeda.

4.6. Analisis Lanjutan Mengenai Variabilitas

Analisis lanjutan mengenai variabilitas melibatkan penerapan teknik statistik yang lebih kompleks untuk memahami variasi data dengan lebih mendalam. Berikut adalah beberapa konsep dan teknik yang sering digunakan dalam analisis lanjutan mengenai variabilitas:

1. Analisis Regresi

Analisis regresi digunakan untuk memahami hubungan antara satu atau lebih variabel independen (penyebab) dengan variabel dependen (hasil). Ini membantu mengidentifikasi seberapa kuat atau lemah

hubungan antara variabel tersebut, serta mengukur variabilitas data yang tidak dapat dijelaskan oleh model regresi.

2. Analisis Variansi (ANOVA)

ANOVA digunakan untuk membandingkan rata-rata antara dua atau lebih kelompok data. Tujuan utamanya adalah untuk menentukan apakah terdapat perbedaan signifikan antara kelompok-kelompok ini. ANOVA memungkinkan kita untuk memahami seberapa besar variabilitas di antara kelompok-kelompok tersebut dan seberapa besar variabilitas di dalam kelompok-kelompok itu sendiri.

3. Analisis Seri Waktu

Analisis seri waktu melibatkan pengamatan data dalam interval waktu yang berurutan. Teknik ini digunakan untuk mengidentifikasi pola, tren, dan siklus dalam data waktu serta untuk mengukur variabilitas dan fluktuasi dalam rentang waktu tertentu.

4. Analisis Multivariat

Analisis multivariat melibatkan pengamatan dan analisis simultan dari lebih dari satu variabel. Ini membantu untuk memahami interaksi kompleks antara variabel-variabel

tersebut dan pengaruhnya terhadap variabilitas data secara keseluruhan.

5. Analisis Faktor

Analisis faktor digunakan untuk mengidentifikasi pola atau struktur yang mendasari hubungan antara sejumlah besar variabel. Tujuannya adalah untuk mengelompokkan variabel-variabel ini ke dalam faktor-faktor yang lebih kecil dan lebih terukur, sehingga membantu mengurangi kompleksitas analisis dan memahami variasi yang kompleks.

6. Metode Bayesian

Metode Bayesian memungkinkan integrasi dari informasi sebelumnya atau pengetahuan yang ada (prior) dengan data observasi baru untuk membuat estimasi yang lebih akurat tentang variabilitas data. Pendekatan ini menghasilkan distribusi probabilitas posterior yang memberikan informasi tentang seberapa yakin kita terhadap estimasi tersebut.

Analisis lanjutan mengenai variabilitas tidak hanya membantu dalam memahami kompleksitas data, tetapi juga memungkinkan pengambilan keputusan yang lebih tepat dan informasi yang lebih dalam tentang fenomena

yang diamati. Dengan menggunakan teknik-teknik ini, analis dapat menggali lebih dalam untuk menemukan pola-pola yang signifikan dan memahami sumber-sumber variabilitas yang mungkin tersembunyi dalam data.

BAB V

KONSEP DISTRIBUSI NORMAL DALAM TERAPAN PENELITIAN KUANTITATIF

5.1. Mengapa Distribusi Normal Penting dalam Penelitian Kuantitatif?

Dalam penelitian secara kuantitatif, konsep distribusi normal memiliki peran yang sangat penting dan menjadi dasar bagi berbagai analisis statistik. Distribusi normal, yang juga dikenal sebagai distribusi Gaussian atau distribusi kurva lonceng, merupakan salah satu distribusi probabilitas yang paling banyak digunakan dalam statistika. Pentingnya distribusi normal dalam konteks penelitian kuantitatif dapat dijelaskan melalui beberapa alasan yang mendasar.

1. **Representasi Realitas.** Dalam banyak kasus, data yang diamati dalam penelitian kuantitatif cenderung memiliki pola distribusi yang mendekati distribusi normal. Fenomena ini dikenal sebagai "phenomenon of normality". Oleh karena itu, distribusi normal memberikan representasi yang baik terhadap distribusi data di alam nyata, memungkinkan peneliti untuk

membuat asumsi yang kuat tentang sifat data yang mereka amati.

2. **Dasar untuk Pengujian Statistik.** Distribusi normal menyediakan dasar matematis untuk berbagai teknik pengujian statistik yang umum digunakan dalam penelitian kuantitatif, seperti uji hipotesis, analisis regresi, dan analisis varians. Dalam pengujian hipotesis, misalnya, distribusi normal digunakan sebagai dasar untuk menghitung nilai-nilai kritis dan menguji signifikansi perbedaan antara sampel dan populasi.
3. **Memungkinkan Penggunaan Statistik Inferensial.** Statistik inferensial adalah salah satu alat utama dalam penelitian kuantitatif yang digunakan untuk membuat generalisasi dari sampel ke populasi yang lebih luas. Distribusi normal memainkan peran kunci dalam statistik inferensial dengan memungkinkan penghitungan interval kepercayaan, estimasi parameter, dan pengujian hipotesis.
2. **Asumsi dalam Analisis Statistik.** Dalam banyak analisis statistik, seperti regresi linear dan analisis varians, salah satu asumsi penting

yang harus dipenuhi adalah bahwa residu atau kesalahan mengikuti distribusi normal. Melanggar asumsi ini dapat mengarah pada kesalahan interpretasi dan kesimpulan yang tidak dapat diandalkan. Oleh karena itu, pemahaman tentang distribusi normal sangat penting dalam memvalidasi hasil analisis statistik.

3. **Interpretasi yang Mudah Dipahami.**

Distribusi normal memiliki sifat-sifat yang terdokumentasi dengan baik dan mudah dipahami. Misalnya, sebagian besar nilai dalam distribusi normal terkonsentrasi di sekitar nilai rata-rata, sementara nilai-nilai ekstrim jarang terjadi. Hal ini memudahkan peneliti untuk menginterpretasikan hasil analisis statistik dengan cara yang jelas dan bermakna.

Dengan demikian, pemahaman yang baik tentang konsep distribusi normal sangat penting bagi peneliti kuantitatif dalam memahami, menganalisis, dan menafsirkan data dengan benar. Distribusi normal tidak hanya memberikan kerangka kerja matematis yang kuat untuk analisis statistik, tetapi juga memungkinkan peneliti untuk membuat kesimpulan

yang valid tentang fenomena yang mereka teliti.

5.2. Dasar-Dasar Distribusi Normal

Distribusi normal, juga dikenal sebagai distribusi Gauss adalah salah satu konsep fundamental dalam statistika yang memiliki aplikasi luas dalam penelitian kuantitatif. Pemahaman yang mendalam tentang dasar-dasar distribusi normal merupakan langkah awal yang penting bagi peneliti untuk menjelajahi lebih lanjut tentang analisis statistik yang menggunakan distribusi ini sebagai dasar. Bagian ini akan membahas konsep dasar dari distribusi normal serta propertinya yang memengaruhi analisis statistik.

5.3. Konsep Dasar Distribusi Normal.

Distribusi normal adalah distribusi probabilitas yang memiliki bentuk bell-shaped symmetrical. Dalam distribusi normal, titik tengahnya adalah nilai rata-rata (mean) dan kurva distribusinya simetris terhadap titik ini. Karakteristik penting lainnya dari distribusi normal adalah bahwa sebagian besar nilai berada di sekitar nilai rata-rata, dengan nilai-nilai yang lebih jauh dari rata-rata menjadi semakin jarang terjadi. Distribusi normal dapat didefinisikan dengan dua parameter utama: mean (μ) dan standar deviasi (σ). Distribusi

normal merupakan distribusi dengan bentuk berupa kurva normal. Adapun Distribusi normal mempunyai sifat-sifat yang penting adalah

1. grafiknya selalu berada diatas sumbu datar x
2. bentuknya simetris terhadap $x = \mu$
3. mempunyai satu modus
4. Grafiknya berasimtotkan sumbu datar di x, baik dikiri maupun kanan
5. luas daerah grafik selalu sama dengan satu

Untuk setiap pasang μ dan σ sifat-sifat seperti tertulis di atas selalu terpenuhi, hanya bentuk kurvanya saja yang berlainan. Jika σ semakin besar maka kurvanya semakin rendah (kurva platikurtis) sedangkan apabila σ semakin kecil kurvanya semakin tinggi (kurva leptikurtis). Model persamaan distribusi normal umum adalah

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

dengan batas nilai x, antara $-\infty < x < \infty$ dimana:

π merupakan nilai konstan yang besarnya 3,1416

e merupakan bilangan konstan yang besarnya 2,7183

μ merupakan rata-rata hitung untuk distribusi

σ merupakan simpangan baku untuk distribusi

Dari persamaan distribusi normal umum ini dapat diturunkan menjadi distribusi normal standar. Persamaan Distribusi Normal Standar adalah

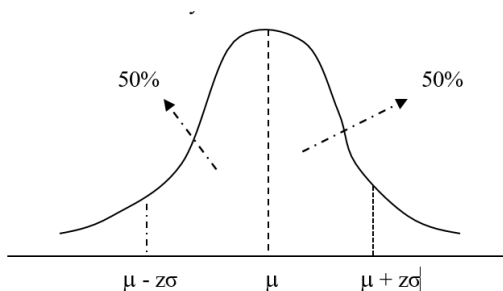
$$z = \frac{x - \mu}{\sigma}$$

dimana:

Z merupakan harga Z Score. Harga Z Score dapat dilihat pada tabel z distribusi normal

x merupakan selisih dari harga pada posisi X dikurangi dengan μ rata-rata σ merupakan standart deviasi

Pada distribusi normal standar harga x merupakan sumbu simetri yang membagi daerah kurva normal menjadi 2 yang sama besar. Pada bagian kanan sumbu simetri mempunyai harga positif yang besarnya sama dengan 50% sedangkan pada bagian kiri sumbu simetri mempunyai harga negatif yang besarnya sama dengan sebelah kanan sumbu simetri yaitu 50%. Lihat Gambar dibawah.



Contoh. Pada Wilayah Kantor Wilayah DIY yang terediri dari 4 kabupaten dan 1 kota mempunyai juru ukur sebanyak 100 orang. Produktifitas rata-rata pertahun adalah 80 bidang, dan mempunyai standart deviasi sebesar 42. Tentukan berapa orang juru ukur yang dapat mengukur 110 bidang pertahun.

Jawab:

Jumlah Juru Ukur yang mempunyai produktifitas 110 bidang, berarti harga $X = 110$, $\mu = 80$,

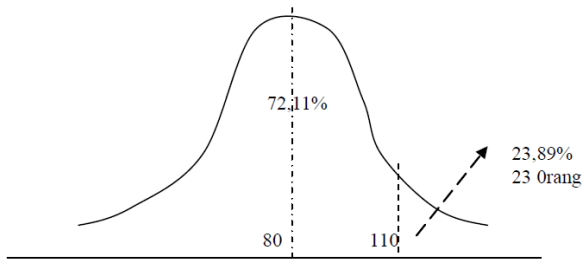
$$\sigma = 42, \quad N = 100$$

$$Z = \frac{x - \mu}{\sigma}$$

$$z = \frac{110 - 80}{42} = 0.714286 \text{ atau } Z = 0.71$$

Hasil pembacaan dari tabel distribusi normal harga untuk harga $Z = 0,714286$, $Z = 26,11\%$, atau luas daerah sebelumnya = (50% + 26,11 %) = 76,11 % Jadi jumlah juru ukur yang mempunyai produktifitas diatas 110 sebanyak:

100% - 76,11% = 23,89% atau sekitar 23 orang juru ukur. Lihat Gambar dibawah



5.4. Properti Distribusi Normal.

Properti dasar dari distribusi normal mencakup sejumlah konsep kunci yang memengaruhi analisis statistik:

- **Mean, Median, dan Mode:** Dalam distribusi normal, mean, median, dan mode semuanya memiliki nilai yang sama karena distribusinya simetris.
- **Simetri dan Kurtosis:** Distribusi normal memiliki kurva yang simetris dan kurtosis yang sesuai. Nilai kurtosis distribusi normal adalah 0, menunjukkan bahwa distribusi tersebut tidak terlalu "spiked" atau "flat".
- **Standar Deviasi dan Varians:** Standar deviasi adalah ukuran dispersi yang penting dalam distribusi normal, menentukan seberapa jauh data tersebar dari rata-rata. Varians, yang merupakan kuadrat dari standar deviasi, juga memberikan gambaran tentang sebaran data.

Pemahaman tentang dasar-dasar distribusi normal memberikan landasan yang kuat bagi peneliti untuk menggali lebih dalam tentang berbagai teknik statistik yang menggunakan distribusi ini sebagai dasar. Dengan memahami sifat-sifat distribusi normal, peneliti dapat membuat asumsi yang tepat, menginterpretasikan hasil analisis statistik dengan benar, dan membuat kesimpulan yang valid dari penelitian kuantitatif.

5.5. Penggunaan Distribusi Normal dalam Penelitian Kuantitatif

Dalam penelitian kuantitatif, distribusi normal memainkan peran yang sangat penting dalam berbagai aspek analisis statistik. Bagian ini akan membahas berbagai cara di mana distribusi normal digunakan dalam konteks penelitian kuantitatif, termasuk penggunaan Z-Score, pengujian hipotesis, dan penghitungan interval kepercayaan.

1. **Pengenalan Z-Score.** Z-Score adalah metode yang digunakan untuk mengukur seberapa jauh suatu nilai individu berjarak dari mean dalam satuan standar deviasi dalam distribusi normal. Penggunaan Z-Score penting dalam penelitian kuantitatif karena memungkinkan peneliti untuk membandingkan dan menstandarisasi

nilai-nilai yang berasal dari distribusi normal yang berbeda.

2. **Penggunaan Distribusi Normal untuk Pengujian Hipotesis.** Pengujian hipotesis adalah teknik statistik yang digunakan untuk menguji asumsi atau pernyataan tentang populasi berdasarkan sampel data. Distribusi normal sering digunakan sebagai dasar untuk pengujian hipotesis karena memungkinkan penentuan nilai-nilai kritis yang diperlukan untuk membuat keputusan statistik yang tepat.
3. **Interval Kepercayaan.** Interval kepercayaan adalah rentang nilai yang mungkin berisi parameter populasi dengan tingkat kepercayaan tertentu. Distribusi normal digunakan untuk menghitung interval kepercayaan..

Pemahaman tentang penggunaan distribusi normal dalam penelitian kuantitatif memungkinkan peneliti untuk menggunakan teknik statistik dengan benar dan mengambil kesimpulan yang valid dari data. Dengan memanfaatkan distribusi normal secara efektif, peneliti dapat menghasilkan hasil penelitian yang kredibel dan dapat dipercaya.

5.6. Aplikasi Distribusi Normal dalam Analisis Data

Dalam penelitian kuantitatif, distribusi normal memiliki beragam aplikasi dalam analisis data yang meliputi regresi linear, pengujian normalitas, dan analisis varians (ANOVA). Bagian ini akan membahas bagaimana distribusi normal digunakan dalam konteks ini.

1. Regresi Linear dan Distribusi Normal.

Regresi linear adalah teknik statistik yang digunakan untuk memodelkan hubungan antara satu atau lebih variabel independen dan variabel dependen. Distribusi normal sering diasumsikan dalam analisis regresi linear karena memungkinkan untuk membuat estimasi yang akurat dan menghasilkan interval kepercayaan yang dapat diandalkan.

2. Pengujian Normalitas.

Pengujian normalitas adalah proses untuk menentukan apakah data mengikuti distribusi normal. Distribusi normal sering diuji dalam penelitian kuantitatif untuk memastikan bahwa asumsi-asumsi yang diperlukan untuk analisis statistik telah terpenuhi. Metode-metode seperti uji Kolmogorov-Smirnov, uji Shapiro-Wilk, dan uji

Lilliefors digunakan untuk menguji normalitas data.

- 3. Analisis Varians (ANOVA) dan Distribusi Normal.** Analisis Varians (ANOVA) adalah teknik statistik yang digunakan untuk membandingkan rata-rata antara tiga atau lebih kelompok. Distribusi normal sering dianggap sebagai asumsi yang penting dalam ANOVA, terutama jika sampel-sampel dari kelompok-kelompok tersebut cukup besar. ANOVA memerlukan distribusi normal untuk memastikan bahwa hasilnya valid dan dapat diinterpretasikan dengan benar.

Pemahaman tentang aplikasi distribusi normal dalam analisis data memungkinkan peneliti untuk menggunakan teknik statistik dengan tepat dan menghasilkan hasil penelitian yang dapat dipercaya. Dengan memastikan bahwa asumsi-asumsi yang mendasari analisis data telah terpenuhi, peneliti dapat memastikan bahwa kesimpulan yang mereka ambil dari analisis tersebut relevan dan dapat diandalkan.

5.7. Studi Kasus dan Contoh Penerapan Distribusi Normal

Studi kasus dan contoh penerapan distribusi normal dalam penelitian kuantitatif memberikan pemahaman yang lebih konkret tentang bagaimana konsep ini diterapkan dalam konteks nyata. Bagian ini akan menggambarkan studi kasus yang menerapkan distribusi normal dalam analisis data.

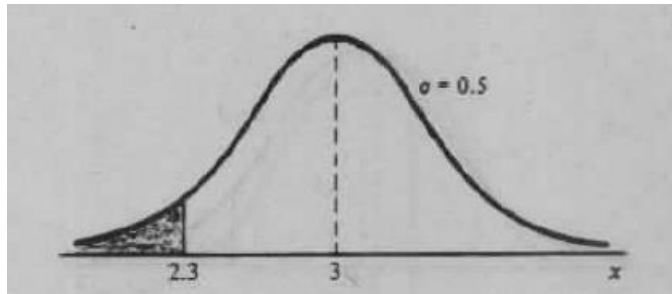
1. **1. Penelitian Kasus: Penggunaan Distribusi Normal dalam Penelitian Psikologi.** Dalam studi kasus penelitian psikologi, dapat dilihat bagaimana distribusi normal digunakan dalam penelitian psikologi untuk menganalisis data dari sebuah eksperimen. Seorang psikolog ingin menguji apakah terdapat perbedaan signifikan dalam tingkat kecemasan antara kelompok yang diberi terapi kognitif dan kelompok kontrol. Data yang diperoleh dari kedua kelompok tersebut kemudian dianalisis menggunakan uji-z.
2. **Penggunaan Distribusi Normal dalam Kasus Penelitian Sains.** Dalam contoh studi ini, kita akan melihat bagaimana distribusi normal digunakan dalam penelitian sains untuk menganalisis data dari sebuah produksi aki.

Seorang insinyur ingin menguji umur aki. Data yang diperoleh kemudian dianalisis menggunakan distribusi z, yang bergantung pada asumsi distribusi normal dari data. Melalui analisis distribusi normal, peneliti dapat memastikan bahwa asumsi-asumsi yang diperlukan untuk pengujian statistik telah terpenuhi.

Berikut contoh suatu jenis aki yang diproduksi oleh perusahaan X mencapai umur rata-rata 3,0 tahun dengan simpangan baku 0,5 tahun. Bila umur aki menyebar normal, hitunglah peluang sebuah aki tertentu akan mencapai umur kurang dari 2,3 tahun.

Jawab.

Pertama-tama buatlah diagram seperti gambar dibawah yang menunjukkan sebaran aki yang diinginkan. Untuk menghitung $P(x < 2,3)$ kita harus mendapatkan luas daerah dibawah kurva normal disebelah kiri 2,3. Ini dapat dicapai dengan menentukan luas daerah sebelah kiri dari nilai z yang dihitung.



$$z = \frac{2,3 - 3}{0,5} = -1,4$$

Dan kemudian menggunakan z score, kita peroleh
 $P(x < 2,3) = P(z < -1,4) = 0,0808$ (lihat nilai z dari tabel distribusi normal)

Tabel Z Distribusi Normal

x	0	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
-3.5	0.0002	0.0002	0.0002	0.0002	0.0002	0.0002	0.0002	0.0002	0.0002	0.0002
-3.4	0.0003	0.0003	0.0003	0.0003	0.0003	0.0003	0.0003	0.0003	0.0003	0.0003
-3.3	0.0005	0.0005	0.0005	0.0004	0.0004	0.0004	0.0004	0.0004	0.0004	0.0003
-3.2	0.0007	0.0007	0.0006	0.0006	0.0006	0.0006	0.0006	0.0005	0.0005	0.0005
-3.1	0.0010	0.0009	0.0009	0.0009	0.0008	0.0008	0.0008	0.0008	0.0007	0.0007
-3.0	0.0013	0.0013	0.0013	0.0012	0.0012	0.0011	0.0011	0.0011	0.0010	0.0010
-2.9	0.0019	0.0018	0.0018	0.0017	0.0016	0.0016	0.0015	0.0015	0.0014	0.0014
-2.8	0.0026	0.0025	0.0024	0.0023	0.0023	0.0022	0.0021	0.0021	0.0020	0.0019
-2.7	0.0035	0.0034	0.0033	0.0032	0.0031	0.0030	0.0029	0.0028	0.0027	0.0026
-2.6	0.0047	0.0045	0.0044	0.0043	0.0041	0.0040	0.0039	0.0038	0.0037	0.0036
-2.5	0.0062	0.0060	0.0059	0.0057	0.0055	0.0054	0.0052	0.0051	0.0049	0.0048
-2.4	0.0082	0.0080	0.0078	0.0075	0.0073	0.0071	0.0069	0.0068	0.0066	0.0064
-2.3	0.0107	0.0104	0.0102	0.0099	0.0096	0.0094	0.0091	0.0089	0.0087	0.0084
-2.2	0.0139	0.0136	0.0132	0.0129	0.0125	0.0122	0.0119	0.0116	0.0113	0.0110
-2.1	0.0179	0.0174	0.0170	0.0166	0.0162	0.0158	0.0154	0.0150	0.0146	0.0143
-2.0	0.0228	0.0222	0.0217	0.0212	0.0207	0.0202	0.0197	0.0192	0.0188	0.0183
-1.9	0.0287	0.0281	0.0274	0.0268	0.0262	0.0256	0.0250	0.0244	0.0239	0.0233
-1.8	0.0359	0.0351	0.0344	0.0336	0.0329	0.0322	0.0314	0.0307	0.0301	0.0294
-1.7	0.0446	0.0436	0.0427	0.0418	0.0409	0.0401	0.0392	0.0384	0.0375	0.0367
-1.6	0.0548	0.0537	0.0526	0.0516	0.0505	0.0495	0.0485	0.0475	0.0465	0.0455
-1.5	0.0668	0.0655	0.0643	0.0630	0.0618	0.0606	0.0594	0.0582	0.0571	0.0559
-1.4	0.0808	0.0793	0.0778	0.0764	0.0749	0.0735	0.0721	0.0708	0.0694	0.0681
-1.3	0.0968	0.0951	0.0934	0.0918	0.0901	0.0885	0.0869	0.0853	0.0838	0.0823
-1.2	0.1151	0.1131	0.1112	0.1093	0.1075	0.1056	0.1038	0.1020	0.1003	0.0985
-1.1	0.1357	0.1335	0.1314	0.1292	0.1271	0.1251	0.1230	0.1210	0.1190	0.1170
-1.0	0.1587	0.1562	0.1539	0.1515	0.1492	0.1469	0.1446	0.1423	0.1401	0.1379
-0.9	0.1841	0.1814	0.1788	0.1762	0.1736	0.1711	0.1685	0.1660	0.1635	0.1611
-0.8	0.2119	0.2090	0.2061	0.2033	0.2005	0.1977	0.1949	0.1922	0.1894	0.1867
-0.7	0.2420	0.2389	0.2358	0.2327	0.2296	0.2266	0.2236	0.2206	0.2177	0.2148
-0.6	0.2743	0.2709	0.2676	0.2643	0.2611	0.2578	0.2546	0.2514	0.2483	0.2451
-0.5	0.3085	0.3050	0.3015	0.2981	0.2946	0.2912	0.2877	0.2843	0.2810	0.2776
-0.4	0.3449	0.3409	0.3372	0.3336	0.3300	0.3264	0.3228	0.3192	0.3156	0.3121
-0.3	0.3821	0.3783	0.3745	0.3707	0.3669	0.3632	0.3594	0.3557	0.3520	0.3483
-0.2	0.4207	0.4168	0.4129	0.4090	0.4052	0.4013	0.3974	0.3936	0.3897	0.3859
-0.1	0.4602	0.4562	0.4522	0.4483	0.4443	0.4404	0.4364	0.4325	0.4286	0.4247
-0.0	0.5000	0.4960	0.4920	0.4880	0.4840	0.4801	0.4761	0.4721	0.4681	0.4641
0.0	0.5000	0.5040	0.5080	0.5120	0.5160	0.5199	0.5239	0.5279	0.5319	0.5359
0.1	0.5398	0.5438	0.5478	0.5517	0.5557	0.5596	0.5636	0.5675	0.5714	0.5753
0.2	0.5793	0.5832	0.5871	0.5910	0.5948	0.5987	0.6026	0.6064	0.6103	0.6141
0.3	0.6179	0.6217	0.6255	0.6293	0.6331	0.6368	0.6406	0.6443	0.6480	0.6517
0.4	0.6554	0.6591	0.6628	0.6664	0.6700	0.6736	0.6772	0.6808	0.6844	0.6879
0.5	0.6915	0.6950	0.6985	0.7019	0.7054	0.7088	0.7123	0.7157	0.7190	0.7224
0.6	0.7257	0.7291	0.7324	0.7357	0.7389	0.7422	0.7454	0.7486	0.7517	0.7549
0.7	0.7580	0.7611	0.7642	0.7673	0.7704	0.7734	0.7764	0.7794	0.7823	0.7852
0.8	0.7881	0.7910	0.7939	0.7967	0.7995	0.8023	0.8051	0.8078	0.8106	0.8133
0.9	0.8159	0.8186	0.8212	0.8238	0.8264	0.8289	0.8315	0.8340	0.8365	0.8389
1.0	0.8413	0.8438	0.8461	0.8485	0.8508	0.8531	0.8554	0.8577	0.8599	0.8621
1.1	0.8643	0.8665	0.8686	0.8708	0.8729	0.8749	0.8770	0.8790	0.8810	0.8830
1.2	0.8849	0.8869	0.8888	0.8907	0.8925	0.8944	0.8962	0.8980	0.8997	0.9015
1.3	0.9032	0.9049	0.9066	0.9082	0.9099	0.9115	0.9131	0.9147	0.9162	0.9177
1.4	0.9192	0.9207	0.9222	0.9236	0.9251	0.9265	0.9279	0.9292	0.9306	0.9319
1.5	0.9332	0.9345	0.9357	0.9370	0.9382	0.9394	0.9406	0.9418	0.9429	0.9441
1.6	0.9452	0.9463	0.9474	0.9484	0.9495	0.9505	0.9515	0.9525	0.9535	0.9545
1.7	0.9554	0.9564	0.9573	0.9582	0.9591	0.9599	0.9608	0.9616	0.9625	0.9633
1.8	0.9641	0.9649	0.9656	0.9664	0.9671	0.9678	0.9686	0.9693	0.9699	0.9706
1.9	0.9713	0.9719	0.9726	0.9732	0.9738	0.9744	0.9750	0.9756	0.9761	0.9767
2.0	0.9772	0.9778	0.9783	0.9788	0.9793	0.9798	0.9803	0.9808	0.9812	0.9817
2.1	0.9821	0.9826	0.9830	0.9834	0.9838	0.9842	0.9846	0.9850	0.9854	0.9857
2.2	0.9861	0.9864	0.9868	0.9871	0.9875	0.9878	0.9881	0.9884	0.9887	0.9890
2.3	0.9893	0.9896	0.9898	0.9901	0.9904	0.9906	0.9909	0.9911	0.9913	0.9916
2.4	0.9918	0.9920	0.9922	0.9925	0.9927	0.9929	0.9931	0.9932	0.9933	0.9936
2.5	0.9938	0.9940	0.9941	0.9943	0.9945	0.9946	0.9948	0.9949	0.9951	0.9952
2.6	0.9953	0.9955	0.9956	0.9957	0.9959	0.9960	0.9961	0.9962	0.9963	0.9964
2.7	0.9965	0.9966	0.9967	0.9968	0.9969	0.9970	0.9971	0.9972	0.9973	0.9974
2.8	0.9974	0.9975	0.9976	0.9977	0.9977	0.9978	0.9979	0.9979	0.9980	0.9981
2.9	0.9981	0.9982	0.9982	0.9983	0.9984	0.9984	0.9985	0.9985	0.9986	0.9986
3.0	0.9987	0.9987	0.9987	0.9988	0.9988	0.9989	0.9989	0.9989	0.9990	0.9990
3.1	0.9990	0.9991	0.9991	0.9991	0.9992	0.9992	0.9992	0.9992	0.9993	0.9993
3.2	0.9993	0.9993	0.9994	0.9994	0.9994	0.9994	0.9994	0.9995	0.9995	0.9995
3.3	0.9995	0.9995	0.9995	0.9996	0.9996	0.9996	0.9996	0.9996	0.9996	0.9997
3.4	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9998
3.5	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998

Sumber: <https://lmsspada.kemdikbud.go.id>

BAB VI

TEKNIK STATISTIK UNTUK MENGUJI HIPOTESIS MENGENAI UJI BEDA LEBIH DARI DUA KELOMPOK

6.1. Pendahuluan

Dalam penelitian ilmiah, seringkali muncul kebutuhan untuk membandingkan lebih dari dua kelompok untuk menentukan apakah terdapat perbedaan signifikan di antara mereka. Teknik statistik yang digunakan untuk menguji hipotesis mengenai perbedaan lebih dari dua kelompok memainkan peran penting dalam analisis data, membantu peneliti menarik kesimpulan yang valid dan dapat diandalkan.

Uji beda lebih dari dua kelompok melibatkan perbandingan antara rata-rata atau median dari tiga kelompok atau lebih untuk melihat apakah ada perbedaan yang signifikan secara statistik. Metode-metode ini tidak hanya digunakan dalam ilmu-ilmu alam dan teknik, tetapi juga dalam ilmu sosial, kesehatan, pendidikan, dan ekonomi.

6.2. Analisis Varian (ANOVA)

Analisis Varian (ANOVA) adalah sebuah teknik statistik yang digunakan untuk menguji perbedaan rata-rata antara tiga atau lebih kelompok atau perlakuan. ANOVA menguji hipotesis nol bahwa tidak ada perbedaan signifikan antara rata-rata kelompok-kelompok tersebut terhadap hipotesis alternatif bahwa setidaknya satu pasangan rata-rata kelompok memiliki perbedaan yang signifikan.

1. Pengertian ANOVA

ANOVA membandingkan variasi antara kelompok dengan variasi di dalam kelompok untuk menentukan apakah terdapat perbedaan yang signifikan antara kelompok-kelompok tersebut. Dalam konteks ini, variasi merujuk pada variasi antara nilai-nilai individu dalam kelompok dibandingkan dengan rata-rata kelompok.

2. Asumsi-Asumsi ANOVA

Beberapa asumsi yang perlu dipertimbangkan sebelum menggunakan ANOVA meliputi:

- a. Normalitas: Data dalam setiap kelompok harus terdistribusi normal.

- b. Homogenitas Varians: Varians (ragam) dari data dalam setiap kelompok harus sekitar sama.
- c. Independensi: Observasi dalam setiap kelompok harus bersifat independen satu sama lain.

Jika asumsi-asumsi ini tidak terpenuhi, maka mungkin diperlukan pendekatan alternatif atau transformasi data sebelum menggunakan ANOVA.

3. Jenis-Jenis ANOVA

- a. ANOVA Satu Arah (One-Way ANOVA): Digunakan untuk menguji perbedaan rata-rata antara tiga atau lebih kelompok yang independen terhadap satu variabel bebas (faktor).
- b. ANOVA Dua Arah (Two-Way ANOVA): Menggabungkan dua faktor atau variabel bebas untuk menguji efek masing-masing faktor secara individu dan efek interaksi di antara keduanya.
- c. ANOVA Tiga Arah (Three-Way ANOVA) dan lebih: Berlaku untuk desain eksperimen yang melibatkan tiga atau lebih faktor atau variabel bebas yang mempengaruhi respons.

4. Contoh Penggunaan ANOVA

Misalnya, dalam sebuah studi pengaruh diet terhadap berat badan, ANOVA bisa digunakan untuk membandingkan rata-rata berat badan setelah diberikan tiga jenis diet yang berbeda. Hipotesis nolnya adalah tidak ada perbedaan signifikan antara rata-rata berat badan di tiga kelompok diet, sementara hipotesis alternatifnya adalah setidaknya satu dari mereka memiliki rata-rata yang berbeda secara signifikan.

ANOVA memberikan keuntungan dalam menguji lebih dari dua kelompok tanpa meningkatkan risiko kesalahan tipe I secara signifikan, seperti yang terjadi jika kita menggunakan uji t secara berulang.

ANOVA dapat dilakukan baik dalam konteks parametrik (ketika data memenuhi asumsi normalitas dan homogenitas varians) maupun non-parametrik (ketika asumsi ini tidak terpenuhi), meskipun pendekatan non-parametrik memiliki kekurangan dalam hal daya statistiknya.

6.3. One-Way ANOVA

One-Way ANOVA (Analysis of Variance Satu Arah) adalah sebuah teknik statistik yang digunakan untuk menguji perbedaan rata-rata antara tiga atau lebih kelompok yang berbeda terhadap satu variabel bebas (faktor). Faktor ini bisa berupa perlakuan atau kondisi yang diberikan kepada kelompok-kelompok tersebut.

1. Pengertian One-Way ANOVA:

One-Way ANOVA digunakan untuk menguji apakah terdapat perbedaan signifikan antara rata-rata kelompok-kelompok yang berbeda terhadap satu variabel dependen (response variable). Teknik ini bekerja dengan membandingkan variasi antara kelompok (variansi antar kelompok) dengan variasi di dalam kelompok (variansi dalam kelompok).

2. Langkah-Langkah dalam One-Way ANOVA:

a. Penetapan Hipotesis:

- Hipotesis nol (H_0): Tidak ada perbedaan signifikan antara rata-rata kelompok.
- Hipotesis alternatif (H_a): Setidaknya satu kelompok memiliki rata-rata yang berbeda secara signifikan.

b. Perhitungan Statistik Uji:

- F-ratio (F-test): Menghitung rasio variasi antar kelompok terhadap variasi dalam kelompok.
 - Nilai F yang dihitung kemudian dibandingkan dengan nilai kritis dari distribusi F untuk menentukan signifikansi statistik.
- c. Interpretasi Hasil:

Jika nilai p (signifikansi) yang dihitung kurang dari tingkat signifikansi yang ditetapkan (biasanya $\alpha = 0.05$), maka kita menolak hipotesis nol dan menyimpulkan bahwa terdapat perbedaan yang signifikan antara rata-rata kelompok-kelompok.

Jika nilai p lebih besar dari α , maka tidak cukup bukti untuk menolak hipotesis nol.

3. Asumsi dalam One-Way ANOVA

Untuk mendapatkan hasil yang valid, One-Way ANOVA mengasumsikan bahwa:

- a. Data dalam setiap kelompok terdistribusi normal.
- b. Homogenitas (keseragaman) varians, artinya varians dari data dalam setiap kelompok sekitar sama.

c. Observasi atau data dalam setiap kelompok bersifat independen satu sama lain.

Jika salah satu dari asumsi ini tidak terpenuhi, maka hasil dari One-Way ANOVA mungkin tidak dapat diandalkan. Dalam kasus ketika asumsi normalitas atau homogenitas varians tidak terpenuhi, pendekatan alternatif seperti transformasi data atau uji non-parametrik dapat dipertimbangkan.

4. Contoh Penggunaan One-Way ANOVA

Misalnya, dalam studi tentang pengaruh tiga jenis pupuk berbeda terhadap pertumbuhan tanaman, One-Way ANOVA dapat digunakan untuk mengevaluasi apakah ada perbedaan yang signifikan dalam tinggi tanaman di antara ketiga jenis pupuk tersebut. Dengan kata lain, kita dapat menentukan apakah satu jenis pupuk memiliki efek yang signifikan terhadap pertumbuhan tanaman dibandingkan dengan yang lainnya.

One-Way ANOVA adalah alat yang kuat dan sering digunakan dalam berbagai bidang ilmu pengetahuan, termasuk sains, ekonomi, sosial, dan lainnya, untuk

menguji perbedaan yang signifikan antara tiga atau lebih kelompok dalam satu faktor perlakuan.

6.4. Two-Way ANOVA

Two-Way ANOVA adalah sebuah teknik statistik yang digunakan untuk menguji pengaruh dari dua faktor atau variabel bebas (independen) terhadap satu variabel dependen (response variable). Teknik ini memungkinkan untuk mengevaluasi efek masing-masing faktor secara individual serta interaksi antara kedua faktor tersebut terhadap variabel dependen.

1. Pengertian Two-Way ANOVA:

Two-Way ANOVA melibatkan analisis variasi antara dua faktor atau variabel bebas terhadap variasi dalam variabel dependen. Dua faktor ini dapat berupa faktor yang terpisah dan tidak terkait satu sama lain, atau dapat juga berinteraksi dalam mempengaruhi variabel dependen.

2. Faktor-faktor dalam Two-Way ANOVA

- a. Faktor Pertama (Factor A): Faktor ini adalah salah satu dari dua variabel bebas yang digunakan dalam Two-Way ANOVA.

- b. Faktor Kedua (Factor B): Faktor kedua adalah variabel bebas kedua yang juga digunakan dalam analisis.
 - c. Interaksi (Interaction): Interaksi antara Factor A dan Factor B adalah efek gabungan dari kedua faktor yang mempengaruhi variabel dependen. Interaksi ini memungkinkan untuk mengevaluasi apakah efek satu faktor bergantung pada tingkat faktor lainnya.
3. Langkah-Langkah dalam Two-Way ANOVA
- a. Penetapan Hipotesis:
 - Hipotesis nol (H_0): Tidak ada interaksi antara Factor A dan Factor B, serta tidak ada efek signifikan dari Factor A maupun Factor B terhadap variabel dependen.
 - Hipotesis alternatif (H_a): Terdapat interaksi antara Factor A dan Factor B, dan/atau setidaknya satu dari kedua faktor tersebut memiliki efek yang signifikan terhadap variabel dependen.
 - b. Perhitungan Statistik Uji:
 - F-ratio (F-test): Seperti pada One-Way ANOVA, F-ratio dihitung untuk masing-

masing faktor dan untuk interaksi antara kedua faktor.

- Nilai F yang dihitung kemudian dibandingkan dengan nilai kritis dari distribusi F untuk menentukan signifikansi statistik.

c. Interpretasi Hasil:

Hasil Two-Way ANOVA memberikan informasi tentang apakah terdapat efek signifikan dari masing-masing faktor secara individual, serta apakah terdapat interaksi antara kedua faktor tersebut yang signifikan terhadap variabel dependen.

4. Asumsi dalam Two-Way ANOVA

Two-Way ANOVA memiliki asumsi yang mirip dengan One-Way ANOVA, yaitu:

- a. Data dalam setiap sel (kombinasi dari faktor A dan faktor B) harus terdistribusi normal.
- b. Homogenitas (keseragaman) varians di seluruh kelompok data.
- c. Observasi dalam setiap sel harus bersifat independen satu sama lain.

Seperti halnya dengan One-Way ANOVA, jika asumsi-asumsi ini tidak terpenuhi, maka hasil

dari Two-Way ANOVA mungkin tidak dapat diandalkan.

5. Contoh Penggunaan Two-Way ANOVA

Misalnya, dalam studi tentang pengaruh dua faktor: jenis pupuk (Factor A) dan intensitas cahaya (Factor B), terhadap pertumbuhan tanaman (variabel dependen), Two-Way ANOVA dapat digunakan untuk mengevaluasi efek masing-masing faktor secara individu (jenis pupuk dan intensitas cahaya) serta interaksi antara kedua faktor tersebut terhadap tinggi tanaman. Dengan demikian, kita dapat mengetahui apakah jenis pupuk, intensitas cahaya, atau kombinasi keduanya secara bersama-sama mempengaruhi hasil pertumbuhan tanaman.

Two-Way ANOVA merupakan alat statistik yang kuat dan sering digunakan dalam penelitian eksperimen untuk memahami pengaruh dari dua faktor yang berbeda terhadap satu variabel dependen.

6.5. ANOVA dengan Repeated Measures

ANOVA dengan Repeated Measures, juga dikenal sebagai ANOVA berulang, adalah teknik statistik yang

digunakan untuk menguji efek perlakuan atau kondisi yang diukur berulang kali pada subjek yang sama atau unit percobaan yang sama. Teknik ini sering digunakan dalam desain penelitian di mana pengukuran berulang terhadap variabel dependen dilakukan pada setiap subjek dalam berbagai kondisi atau waktu yang berbeda.

1. Pengertian ANOVA dengan Repeated Measures:
ANOVA dengan Repeated Measures digunakan ketika kita ingin mengevaluasi efek perlakuan atau kondisi yang diberikan terhadap subjek atau unit percobaan yang sama, yang diukur beberapa kali pada berbagai waktu atau kondisi. Teknik ini mengatasi masalah variabilitas antar subjek yang mungkin terjadi dalam desain eksperimen.
2. Langkah-Langkah dalam ANOVA dengan Repeated Measures:
 - a. Penetapan Hipotesis:
 - Hipotesis nol (H_0): Tidak ada perbedaan signifikan dalam rata-rata pengukuran berulang terhadap variabel dependen antara kondisi atau waktu yang berbeda.

- Hipotesis alternatif (H_a): Setidaknya satu kondisi atau waktu memiliki efek yang signifikan terhadap variabel dependen.
- b. Desain Eksperimen:
- Subjek atau unit percobaan diukur dalam setiap kondisi atau waktu yang berbeda.
 - Pengukuran yang berulang dilakukan untuk masing-masing subjek.
- c. Perhitungan Statistik Uji:
- F-ratio (F-test): Seperti dalam ANOVA lainnya, F-ratio dihitung untuk membandingkan variasi antar kondisi atau waktu dengan variasi dalam kondisi atau waktu.
 - Nilai F yang dihitung kemudian dibandingkan dengan nilai kritis dari distribusi F untuk menentukan signifikansi statistik.
3. Interpretasi Hasil:
- Jika nilai p (signifikansi) yang dihitung kurang dari tingkat signifikansi yang ditetapkan (biasanya $\alpha = 0.05$), maka kita menolak hipotesis nol dan menyimpulkan bahwa terdapat perbedaan yang signifikan dalam rata-rata pengukuran berulang terhadap variabel

dependen antara kondisi atau waktu yang berbeda.

Jika nilai p lebih besar dari α , maka tidak cukup bukti untuk menolak hipotesis nol.

4. Asumsi dalam ANOVA dengan Repeated Measures:

Beberapa asumsi yang perlu dipertimbangkan dalam ANOVA dengan Repeated Measures meliputi:

- a. Sphericity: Asumsi bahwa kovarians antar pengukuran berulang adalah sama untuk semua pasangan kondisi atau waktu.
- b. Normalitas: Data dalam setiap kondisi atau waktu harus terdistribusi normal.
- c. Homogenitas Varians: Varians dari data dalam setiap kondisi atau waktu sekitar sama.

Jika asumsi sphericity tidak terpenuhi, mungkin perlu dilakukan koreksi atau transformasi data sebelum melakukan analisis.

5. Contoh Penggunaan ANOVA dengan Repeated Measures:

Misalnya, dalam studi psikologis tentang efek terapi kognitif pada memori, ANOVA dengan Repeated Measures dapat digunakan untuk

membandingkan tingkat memori subjek sebelum terapi (baseline), setelah satu minggu terapi, dan setelah dua minggu terapi. Dengan menggunakan teknik ini, peneliti dapat menentukan apakah terapi kognitif memiliki efek yang signifikan terhadap perbaikan memori dalam waktu yang berbeda.

ANOVA dengan Repeated Measures memungkinkan untuk memeriksa perubahan dalam variabel dependen dari waktu ke waktu atau kondisi ke kondisi pada subjek yang sama, yang memberikan keuntungan dalam mengurangi variasi antar subjek dan meningkatkan kekuatan statistik dari analisis.

BAB VII

TEKNIK STATISTIK UNTUK MENGUJI HIPOTESIS MENGENAI HUBUNGAN DUA VARIABEL

7.1. Hubungan Antar Variabel

7.1.1. Definisi dan Jenis Hubungan

Hubungan antar variabel merujuk pada cara dua atau lebih variabel berinteraksi satu sama lain. Hubungan ini dapat diklasifikasikan menjadi beberapa jenis berdasarkan bagaimana perubahan dalam satu variabel mempengaruhi perubahan dalam variabel lain.

a. Korelasi

Korelasi adalah ukuran statistik yang menggambarkan sejauh mana dua variabel berubah bersama. Korelasi tidak menyiratkan sebab-akibat, tetapi menunjukkan adanya hubungan antara variabel-variabel tersebut. Korelasi dapat positif, negatif, atau nol.

- Korelasi Positif: Ketika satu variabel meningkat, variabel lain juga meningkat. Contoh: tinggi badan dan berat badan.
- Korelasi Negatif: Ketika satu variabel meningkat, variabel lain menurun. Contoh: kecepatan mobil dan waktu tempuh.
- Korelasi Nol: Tidak ada hubungan antara kedua variabel. Contoh: nomor telepon seseorang dan tinggi badan mereka.

b. Kausalitas

Kausalitas menyiratkan bahwa perubahan dalam satu variabel menyebabkan perubahan dalam variabel lain. Untuk menetapkan kausalitas, diperlukan eksperimen yang terkontrol atau metode analisis yang sangat kuat.

Hubungan Sebab-Akibat: Misalnya, merokok menyebabkan kanker paru-paru.

7.1.2. Korelasi vs. Kausalitas

Korelasi dan kausalitas adalah konsep yang berbeda meskipun sering membingungkan:

- Korelasi hanya menunjukkan bahwa dua variabel berkaitan secara statistik.

- Kausalitas menunjukkan bahwa satu variabel secara langsung mempengaruhi variabel lainnya.

Contoh klasik untuk menjelaskan perbedaan ini adalah hubungan antara es krim dan tenggelam: keduanya mungkin berkorelasi (lebih banyak orang makan es krim dan berenang di musim panas), tetapi makan es krim tidak menyebabkan tenggelam.

7.2. Korelasi

7.2.1. Pengertian Korelasi

Korelasi adalah teknik statistik yang digunakan untuk mengukur kekuatan dan arah hubungan linier antara dua variabel. Korelasi membantu peneliti untuk memahami sejauh mana dua variabel berubah bersama. Meskipun korelasi tidak menyiratkan kausalitas, ini merupakan langkah awal yang penting dalam analisis hubungan antar variabel.

7.2.2. Hipotesis Korelasi

Dalam pengujian hipotesis menggunakan korelasi, hipotesis nol (H_0) dan hipotesis alternatif

(H1) dirumuskan sebagai berikut:

- Hipotesis Nol (H_0): Tidak ada hubungan linier yang signifikan antara dua variabel ($r = 0$).
- Hipotesis Alternatif (H_1): Ada hubungan linier yang signifikan antara dua variabel ($r \neq 0$).

7.2.3. Jenis-Jenis Korelasi

a. Korelasi Pearson (Pearson's r)

Digunakan untuk data kontinu yang berdistribusi normal.

Mengukur kekuatan dan arah hubungan linier antara dua variabel.

Nilai r berkisar antara -1 hingga 1:

- $r = 1$: Korelasi positif sempurna.
- $r = -1$: Korelasi negatif sempurna.
- $r = 0$: Tidak ada korelasi linier.

b. Korelasi Spearman (Spearman's rho)

Digunakan untuk data ordinal atau data yang tidak berdistribusi normal.

Mengukur kekuatan dan arah hubungan monotonik antara dua variabel.

Berdasarkan peringkat data, bukan nilai mentahnya.

- c. Korelasi Kendall (Kendall's tau)
Alternatif untuk korelasi Spearman, cocok untuk data kecil atau data dengan banyak nilai yang sama.
Mengukur kekuatan dan arah hubungan ordinal antara dua variabel.

7.2.4. Keterbatasan Korelasi

- a. Korelasi Tidak Menunjukkan Kausalitas: Meskipun dua variabel mungkin berkorelasi, itu tidak berarti bahwa satu variabel menyebabkan perubahan pada variabel lain.
- b. Pengaruh Outliers: Nilai yang jauh berbeda dari yang lain dapat mempengaruhi nilai korelasi secara signifikan.
- c. Hanya Mengukur Hubungan Linier: Korelasi Pearson hanya mengukur hubungan linier dan mungkin tidak mendeteksi hubungan non-linier antara variabel.

Korelasi adalah teknik statistik yang esensial dalam menguji hipotesis mengenai hubungan dua variabel. Dengan memahami kekuatan, arah, dan signifikansi hubungan antar variabel, peneliti dapat mendapatkan wawasan berharga tentang data

mereka. Namun, penting untuk diingat bahwa korelasi tidak menyiratkan kausalitas dan harus digunakan dengan bijak dalam interpretasi hasil penelitian.

7.3. Regresi Linier

7.3.1. Pengertian Regresi Linier

Regresi linier adalah teknik statistik yang digunakan untuk memodelkan dan menganalisis hubungan antara dua variabel. Ini melibatkan satu variabel dependen (respons) dan satu variabel independen (prediktor). Tujuannya adalah untuk menemukan persamaan garis lurus yang paling sesuai (fit) dengan data yang menggambarkan hubungan antara kedua variabel.

7.3.2. Hipotesis dalam Regresi Linier

Dalam regresi linier, hipotesis nol (H_0) dan hipotesis alternatif (H_1) umumnya dirumuskan sebagai berikut:

- Hipotesis Nol (H_0): Tidak ada hubungan linier yang signifikan antara variabel independen dan variabel dependen ($\beta_1 = 0$).

- Hipotesis Alternatif (H1): Ada hubungan linier yang signifikan antara variabel independen dan variabel dependen ($\beta_1 \neq 0$).

7.3.3. Langkah-Langkah Analisis Regresi Linier

- a. Plot Data: Buat scatter plot dari data untuk melihat apakah hubungan linier antara dua variabel masuk akal.
- b. Estimasi Parameter: Hitung nilai β_0 dan β_1 menggunakan metode kuadrat terkecil.
- c. Persamaan Garis Regresi: Tentukan persamaan garis regresi dari parameter yang diestimasi.
- d. Uji Signifikansi: Uji apakah parameter β_1 signifikan secara statistik menggunakan uji t.
- e. Analisis R^2 (Koefisien Determinasi): Hitung R^2 untuk melihat seberapa baik model regresi menjelaskan variabilitas dalam data.

7.3.4. Keterbatasan Regresi Linier

- a. Asumsi Linearitas: Model mengasumsikan bahwa hubungan antara variabel dependen dan independen adalah linier.

- b. Outliers: Outliers dapat mempengaruhi hasil model secara signifikan.
- c. Multikolinearitas: Dalam regresi berganda, variabel independen seharusnya tidak berkorelasi tinggi satu sama lain.
- d. Homoskedastisitas: Varians error harus konstan untuk semua nilai X.

Regresi linier adalah teknik statistik yang kuat untuk menguji hipotesis mengenai hubungan linier antara dua variabel. Dengan memahami bagaimana mengestimasi, menguji, dan menginterpretasi parameter regresi, peneliti dapat membuat prediksi dan memperoleh wawasan berharga tentang data mereka. Namun, penting untuk selalu mempertimbangkan asumsi model dan keterbatasan yang mungkin mempengaruhi validitas hasil.

7.4. Uji Chi-Square untuk Hubungan Dua Variabel Kategorikal

1. Pengertian Uji Chi-Square

Uji Chi-Square (χ^2) adalah teknik statistik yang digunakan untuk menguji hubungan atau asosiasi antara dua variabel kategorikal. Uji ini

menentukan apakah distribusi pengamatan di satu variabel berbeda secara signifikan dengan distribusi yang diharapkan jika kedua variabel itu independen.

2. Jenis Uji Chi-Square

a. Uji Chi-Square untuk Independensi:

Digunakan untuk menentukan apakah ada hubungan yang signifikan antara dua variabel kategorikal dalam satu populasi.

b. Uji Chi-Square untuk Kecocokan:

Digunakan untuk menguji apakah distribusi sampel sesuai dengan distribusi teoritis yang diharapkan.

3. Uji Chi-Square untuk Independensi

a. Hipotesis

- Hipotesis Nol (H_0): Tidak ada hubungan (independensi) antara kedua variabel kategorikal.
- Hipotesis Alternatif (H_1): Ada hubungan (dependensi) antara kedua variabel kategorikal.

b. Langkah-Langkah Uji Chi-Square untuk Independensi

1) Membuat Tabel Kontingensi:

Tabel yang menunjukkan frekuensi pengamatan untuk kombinasi kategori dari dua variabel.

2) Menghitung Frekuensi Harapan:

Frekuensi harapan (E) dihitung berdasarkan asumsi bahwa tidak ada hubungan antara variabel. Rumus untuk frekuensi harapan adalah:

$$E = \frac{(total\ baris) \times (total\ kolom)}{total\ keseluruhan}$$

Menghitung Nilai Chi-Square:

Rumus untuk menghitung nilai χ^2 adalah:

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

di mana O adalah frekuensi pengamatan, dan E adalah frekuensi harapan.

3) Menentukan Derajat Kebebasan (df):

df untuk uji independensi adalah

$$df = (jumlah\ baris - 1) \times (jumlah\ kolom - 1)$$

Menentukan Nilai P dan Keputusan:

Bandingkan nilai χ^2 hitung dengan nilai kritis dari tabel distribusi χ^2

berdasarkan df dan tingkat signifikansi (α , biasanya 0.05).

Atau, gunakan p-value untuk menentukan apakah H_0 ditolak atau tidak. Jika $p\text{-value} < \alpha$, maka H_0 ditolak.

4. Keterbatasan Uji Chi-Square

- Frekuensi Harapan Rendah: Uji Chi-Square kurang dapat diandalkan jika frekuensi harapan dalam satu atau lebih sel kurang dari 5.
- Data Kategorikal Saja: Uji ini hanya berlaku untuk variabel kategorikal.
- Sensitif terhadap Ukuran Sampel: Hasil uji bisa dipengaruhi oleh ukuran sampel yang sangat besar atau sangat kecil.

Uji Chi-Square adalah alat yang kuat untuk menguji hubungan antara dua variabel kategorikal. Dengan memahami cara kerja uji ini dan interpretasi hasilnya, peneliti dapat menyimpulkan apakah ada asosiasi yang signifikan antara variabel-variabel tersebut dalam populasi yang diteliti.

BAB VIII

ANALISIS REGRESI SEDERHANA

8.1. Pendahuluan

Analisis regresi adalah analisis statistik yang mempelajari hubungan antara dua atau lebih variabel kuantitatif sehingga satu variabel dapat diramalkan (*predicted*) dari variabel lainnya. Hubungan antara dua variabel dapat dibedakan menjadi dua, yaitu *hubungan fungsional* dan *hubungan statistik*. Hubungan fungsional antara dua variabel dapat dinyatakan secara matematis; jika X variabel bebas (*independent variable*) dan Y variabel tak bebas (*dependent variable*), hubungan fungsional ditulis dalam bentuk

8.2. Analisis Regresi Sederhana

Analisis regresi sederhana merupakan analisis statistik yang mempelajari hubungan dua variabel dengan satu variabel bebas (*independent*). Tujuan dari analisis ini adalah untuk memmbuat prediksi dari nilai suatu variabel bebas (variabel *independent*) melalui variabel terikat (dependen).

Dalam melakukan prediksi digunakan persamaan garis yang dapat diperoleh dengan metode pendugaan parameter. Metode pendugaan parameter yang sering digunakan adalah metode kuadrat terkecil (*least squar*). Metode kadrat terkecil adalah teknik untuk membaut garis regresi dengan cara mengurangi sebanyak mungkin jumlah kuadrat jarak antara nilai Y yang teramati dan Y yang diperdiksi.

Model regresi sederhana dengan satu variabel bebas X adalah sebagai berikut:

$$Y = \beta_0 + \beta_1 X_1 + \varepsilon$$

dimana

Y : variabel tak bebas (*dependent*)

β_0 : intersep

X : variabel bebas (*independent*)

Analisis regresi pada umumnya digunakan untuk keperluan uji hipotesis yang berkaitan dengan prediksi dan eksplanasi. Contoh pemakaian analisis regresi untuk keperluan prediksi adalah penggunaan skor tes SNMPTN untuk memprediksikan/meramalkan keberhasilan belajar calon mahasiswa di perguruan tinggi. Sedangkan pemakaian analisis regresi untuk keperluan eksplanasi adalah untuk mengungkapkan faktor-faktor yang menentukan suatu fenomena.

Misalnya, untuk mengungkapkan faktor-faktor yang mempengaruhi prestasi belajar seseorang. Permasalahan yang pertama tentang prediksi dan permasalahan yang kedua tentang eksplanasi, keduanya bisa diselesaikan dengan teknik analisis regresi. Pada dasarnya teknik ini berusaha menganalisis peranan setiap ubahan (prediktor) dalam menentukan besarnya varians dari suatu ubahan yang ada pada fenomena yang diteliti (ubahan kriterium).

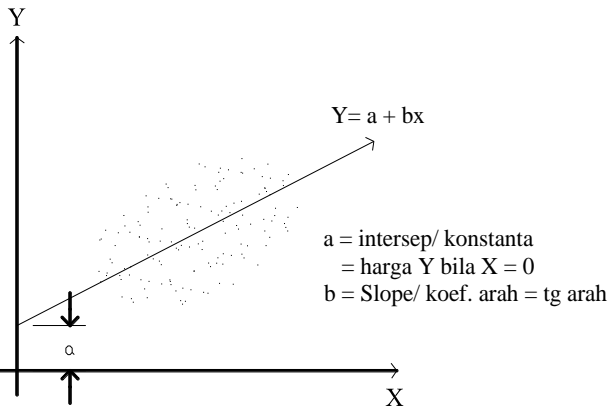
Seperti pada teknik korelasi, teknik analisis regresi menggunakan asumsi adanya hubungan yang linier atau berupa garis lurus antara variabel prediktor dengan variabel kriterium. Teknik analisis regresi yang menggunakan asumsi hubungan linier disebut regresi linier. Adapun bentuk umum persamaan garis regresi Y atas X, dimana X adalah ubahan prediktor adalah:

$Y' = a + b X$ → regresi sederhana dengan sebuah ubahan X

$$Y' = a + b_1.X_1 + b_2.X_2 ++b_n.X_n$$

Pada prinsipnya teknik analisis regresi ini adalah menentukan suatu garis lurus (garis regresi/garis prediksi) yang paling tepat diantara penyebaran titik-titik plotting (scatter-plot). Garis regresi yang tepat ini harus memenuhi syarat sebagai berikut:

1. Berupa garis lurus atau memenuhi persamaan linier
2. Jumlah kuadrat penyimpangan skor pengamatan dan skor yang diramalkan adalah minimum $\gg \Sigma(Y - \hat{Y})^2 = \text{minimum}$ (metode least square).



Berdasarkan konsep kuadrat terkecil (least square) inilah, maka koefisien a dan b dapat dihitung.

Rumus:
$$a = \frac{(\Sigma Y)(\Sigma X^2) - (\Sigma X)(\Sigma XY)}{n \cdot \Sigma X^2 - (\Sigma X)^2}$$

$$b = \frac{n \cdot \Sigma XY - (\Sigma X)(\Sigma Y)}{n \cdot \Sigma X^2 - (\Sigma X)^2} \rightarrow b = \frac{\Sigma xy}{\Sigma x^2}$$

Contoh:

Akan diteliti mengenai peranan motivasi (X) terhadap prestasi belajar (Y) pada sekelompok siswa yang terdiri dari 30 orang. Dari tiap-tiap siswa yang diteliti tersebut dicatat tentang dua karakteristik, yaitu

motivasi (X) dan prestasi belajarnya (Y). Adapun datanya disajikan dalam tabel berikut.

Resp.	X	Y	X²	Y²	XY
1	34	32	1156	1024	1088
2	38	36	1444	1296	1368
3	34	31	1156	961	1054
4	40	38	1600	1444	1520
5	30	29			
6	40	35	dst	dst	dst
7	40	33			
8	34	30			
9	35	32			
10	39	36			
11	33	31			
12	32	31			
13	42	36			
14	40	37			
15	42	35			
16	42	38			
17	41	37			
18	32	30			
19	34	30			
20	36	30			
21	37	33			

22	36	32			
23	37	34			
24	39	35			
25	40	36			
26	33	32			
27	34	32			
28	36	34			
29	37	32			
30	38	34			
n= 30	1105	1001	41.029	33.599	37.094

Berdasarkan data-data diatas, maka koefisien a dan b dapat dihitung sebagai berikut:

$$a = \frac{(\Sigma Y)(\Sigma X^2) - (\Sigma X)(\Sigma XY)}{n \cdot \Sigma X^2 - (\Sigma X)^2}$$

$$a = \frac{(1001)(41.029) - (1.105)(37.094)}{30 \cdot (41.029) - (1.105)^2} = \mathbf{8,24}$$

$$b = \frac{n \cdot \Sigma XY - (\Sigma X)(\Sigma Y)}{n \cdot \Sigma X^2 - (\Sigma X)^2} \text{ atau } b = \frac{\Sigma xy}{\Sigma x^2}$$

$$b = \frac{30 \cdot (37.094) - (1.105)(1.001)}{30 \cdot (41.029) - (1.105)^2} = \mathbf{0,68}$$

Dengan demikian, persamaan garis regresinya adalah:

$$\hat{Y} = 8,24 + 0,68x$$

Dalam hal ini, Y diganti \hat{Y} (Y topi) \Rightarrow untuk membedakan antara Y ramalan (yang bentuk dasarnya: $Y = \alpha + \beta.X$) dengan Y hasil pengamatan ($\hat{Y} = a + bx$).

8.3. Uji Kelinearan dan Keberartian Regresi

Sebelum digunakan untuk membuat kesimpulan, maka regresi yang diperoleh harus diuji dulu yang berkaitan dengan: (1) kelinearan bentuk regresi (uji linieritas hubungan) ; dan (2) keberartian regresi (koefisien regresi). Pemeriksaan kelinearan bentuk regresi dilakukan melalui pengujian hipotesis bahwa: regresi linier melawan hipotesis tandingan. bahwa: regresi non linier.

Sementara itu, uji keberartian regresi diperiksa melalui pengujian hipotesis bahwa koefisien-koefisien regresi atau koefisien arah b samadengan nol (tidak berarti) melawan hipotesis tandingan bahwa koefisien arah regresi atau b tidak sama dengan nol.

$H_0: \text{Pop. } b = 0 \rightarrow H_0: \beta = 0 \rightarrow$ tidak berarti

$H_0: \text{Pop. } b \neq 0 \rightarrow H_0: \beta \neq 0 \rightarrow$ berarti

Untuk uji kelinearan diperlukan adanya beberapa pengulangan atau kelompok data x . Tiap kelompok terdiri atas beberapa data x yang berharga sama,

sementara harga-harga Y pasangannya tetap seperti pasangan data semula.

Dengan pengelompokan data x yang berharga sama tersebut, akan diperoleh bahwa pada kelompok pertama, ada n_1 buah data x yang masing-masing harganya sama yaitu x_1 berpasangan dengan Y yang umumnya harganya berbeda-beda. Pada kelompok dua, ada n_2 buah data x yang masing-masing harganya sama yaitu x_2 berpasangan dengan Y yang harganya berbeda-beda dan begitu seterusnya. Seluruhnya ada k buah kelompok yang dalam tiap kelompoknya harga-harga x nya sama besar. Berdasarkan data di muka, maka pengelompokkan harga-harga x yang sama tersebut dapat dilihat pada table berikut:

X	Kelompok	Ni	Y
30	1	1	29
32	2	2	31
32			30
33	3	2	31
33			32
34	4	5	32
34			31
34			30
34			30
34			32

35	5	1	32
36	6	3	30
36			32
36			34
37	7	3	33
37			34
37			32
38	8	2	36
38			34
39	9	2	36
39			35
40	10	5	38
40			35
40			33
40			37
40			36
41	11	1	37
42	12	3	36
42			35
42			38

Berdasarkan data pengelompokan harga-harga yang telah disajikan pada table diatas, maka uji kelinieran dan keberartian regresi dapat dilakukan terlebih dahulu menghitung jumlah kuadrat (jk) untuk

berbagai sumber variasi. Sumber variasi tersebut adalah total atau jk (T), regresi (a), regresi (bla), sisa atau resida, tuna cocok dan galat atau kekeliruan (error). Jk untuk berbagai sumber variasi tersebut berturut-turut diberi symbol dengan jk (T), jk (a), jk (bla), jk (s) atau jk res, jk (TC) dan jk (G) atau jk (E), dan dapat dihitung dengan rumus-rumus sebagai berikut:

$$Jk (T) = \Sigma Y^2$$

$$Jk (a) = \frac{(\Sigma Y)^2}{n}$$

Dalam persamaan umum dipersatukan manjadi Jk total

$$\Rightarrow Jk_{tot} = \Sigma Y^2 = \Sigma Y^2 - \frac{(\Sigma Y)^2}{n}$$

$$Jk (bla) = b \cdot \left\{ \Sigma xy - \frac{(\Sigma X)(\Sigma Y)}{n} \right\} \Rightarrow Jk_{reg}$$

= b. \Sigma xy biasa disebut jk regresi

$$Jk (s) = Jk_{res} = Jk (T) - Jk (a) - Jk (bla) \Rightarrow Jk_{res} = Jk_{tot} - J$$

$$Jk (G) = JK (E) = \Sigma_{xi} \left\{ \Sigma Y^2 - \frac{(\Sigma Y)^2}{ni} \right\}$$

$$Jk (TC) = Jk (s) - Jk (G) = Jk_{res} = Jk G$$

Setiap sumber variasi tersebut memiliki derajat kebebasan atau dk, yang besarnya berturut-turut n untuk total, 1 untuk regresi (a), 1 untuk regresi (bla), (n-2) untukn sisa/ residu, (k-2) untuk tuna cocok dan

110 | Statistik Dasar

$(n-k)$ untuk galat/ error. Selanjutnya dapat ditentukan harga rerata jumlah kuadrat atau Rjk untuk masing-masing sumber variasi yakni dengan cara membagi jk oleh dk-nya masing-masing.

Sementara itu, Rjk (bIa) sering pula dilambangkan dengan s^2_{reg} (Varians regresi), dan Rjk (s) atau Rjk (res) sering dilambangkan oleh s^2_{sis} atau s^2_{res} . Demikian pula Rjk (TC) juga dapat dilambangkan dengan s^2_{TC} dan Rjk (G) atau Rjk (E) sering dilambangkan dengan s^2_G atau s^2_E . Semua besaran yang diperoleh selanjutnya disusun dalam daftar analisis varians (ANOVA) untuk regresi linier sederhana (Tabel Rangkuman Regresi).

Besaran-besaran yang ada dalam daftar anava tersebut, khususnya pada kolom Rjk, digunakan untuk menguji hipotesis:

1. Koefisien arah regresi tidak berarti melawan hipotesis tandingan bahwa koefisien arah regresi berarti $H_0: \beta_i = 0$ VS $H_a: \beta_i \neq 0$.
2. Bentuk regresi linier, melawan hipotesis tandingan bahwa bentuk regresi non linier.
Kedua hipotesis diatas diuji dengan menggunakan statistic F yang dibentuk oleh perbandingan dua Rjk.

Untuk menguji hipotesis (1) dipakai statistic $F = s^2_{reg} / s^2_{res}$ dan selanjutnya digunakan table distribusi F dengan dk pembilang 1 dan dk penyebut (n-2).

Kriteria:

Tolak hipotesis bahwa koefisien arah regresi tidak berarti ($H_0: \beta = 0$) jika $F_h = s^2_{reg} / s^2_{res} \geq F_{(1-\alpha)(1;n-2)}$.

Untuk menguji hipotesis (2) dipakai statistik $F = s^2_{TC} / s^2_E$ yang selanjutnya digunakan tabel distribusi F dengan dk pembilang (k-2) dan dk penyebut (n-k).

Kriteria:

Tolak hipotesis bahwa bentuk regresi linier, jika $F = s^2_{TC} / s^2_E \geq F_{(1-\alpha)(k-2; n-k)}$.

Dalam hal lainnya, hipotesis-hipotesis yang disebutkan diatas diterima.

Tabel Anava Regresi linier sederhana

Sumber Variasi	dk	Jk	Rjk	F
Total	n	ΣY^2	ΣY^2	-
Regresi (a)	1	Jk (a)	Jk (a)	

Regresi (b1a)	1	Jk (b1a)	$S_{reg}^2 = Jk (b1a)$	$\frac{S_{reg}^2}{S_{res}^2}$
Residu / sisa	n-2	Jk res	$S_{res}^2 = \frac{Jk res}{n - 2}$	-
Tuna cocok	k-2	Jk (TC)	$S_{TC}^2 = \frac{Jk (TC)}{k - 2}$	$\frac{S_{TC}^2}{S_E^2}$
Galat /Error	n-k	Jk (E)	$S_E^2 = \frac{Jk (E)}{n - k}$	

Berdasarkan rumus-rumus diatas, maka berdasarkan data yang telah disajikan dimuka selanjutnya akan diuji apakah persamaan garis regresi $\hat{Y} = 8,24 + 0,68x$ tersebut memiliki bentuk linier dan koefisien arah regresi yang berarti atau tidak.

$$Jk_{total} = Jk (T) = \Sigma Y^2 = 33,599$$

$$Jk (a) = \frac{(\Sigma Y)^2}{n} = \frac{(1.001)^2}{30} = 33.400,03$$

$$Jk (b1a) = b. \left\{ \Sigma xy - \frac{(\Sigma X)(\Sigma Y)}{n} \right\}$$

$$= 0,68 \left\{ 37.094 - \frac{(1.105)(1.001)}{30} \right\} = 151,75$$

$$Jk_{res} = Jk (s) = Jk (T) - Jk (a) - Jk (b1a)$$

$$Jk_{res} = 33.599 - 33.400,03 - 151,75 \Rightarrow Jk_{res} = 47,22$$

$$Jk (G) = Jk (E) = \sum_{xi} \left\{ \Sigma Y^2 - \frac{(\Sigma Y)^2}{ni} \right\}$$

Berdasarkan pengelompokan harga x yang sama, ternyata ada 12 kelompok (k=12) dengan n1=1, n2=2, n3=2,.....,n12=3. Maka Jk (G) atau Jk (E) dapat dihitung sebagai berikut:

$$\begin{aligned} Jk (G) = Jk (E) &= \left\{ 29^2 - \frac{29^2}{1} \right\} + \left\{ 31^2 + 30^2 - \frac{(31+30)^2}{2} \right\} + \\ &\left\{ 31^2 + 32^2 - \frac{(31+30)^2}{2} \right\} + \dots \dots \\ &= \left\{ 36^2 + 35^2 + 38^2 - \frac{(36+35+38)^2}{3} \right\} \end{aligned}$$

$$Jk (G) = Jk (E) = 37,67$$

$$Jk (TC) = Jk_{res} - Jk (E) = 47,22 - 37,67 \Rightarrow Jk (TC) = 9,55$$

\(\Rightarrow\) Selanjutnya harga-harga tersebut diatas dimasukkan dalam tabel anava sebagai berikut:

Anava untuk regresi linier $\hat{Y} = 8,24 + 0,68x$

Sumber Variasi	dk	Jk	Rjk	F
Total	30	33,599		
Regresi (a)	1	33,400,03	-	-
Regresi (bIa)	1	151,75	151,75	89,79
Residu / sisa	28	47,22	1,69	
Tuna cocok	10	9,55	0,96	0,45
Galat /Error	18	37,67	2,09	

Misalkan ditentukan $\alpha = 0,05$

\Rightarrow untuk menguji hipotesis (1), maka harga F tabel

$$F_{(1-\alpha);(1,28)} = 4,20$$

Ternyata bahwa : F tabel = $F_{(1-\alpha);(1,28)} < F_{hitung}$

Sehingga: $H_0: \rho = 0 \Rightarrow$ ditolak

Kesimpulan: Koefisien arah regresi berarti ($F_h > F_{tab}$).

\Rightarrow Untuk menguji hipotesis (2), maka harga F tabel

$$F_{(1-\alpha);(10,18)} = 2,43$$

Ternyata F hitung $< F_{(1-\alpha);(10,18)} \Rightarrow H_0$: diterima

Kesimpulan: Kita terima bahwa bentuk regresi adalah linier.

8.4. Ketelitian Prediksi

Kesalahan prediksi dapat dinyatakan dengan simpangan baku dari Y untuk harga x tertentu. Jadi, regresi $\hat{Y} = a + bx$ yang diperoleh berdasarkan data pengamatan adalah merupakan taksiran terhadap regresi $Y = \alpha + \beta x$ dalam populasi, dimana n buah pasangan data yang dipakai untuk menentukan $\hat{Y} = a + bx$ telah diambil. Dalam menaksir Y sebenarnya (dalam populasi) berdasarkan regresi $\hat{Y} = a + bx$ akan kita temui suatu penyimpangan (Error). Perbedaan ini merupakan galas (error) yang apabila dihitung Jk-nya kemudian dibagi dengan dk yaitu (n-2), maka akan diperoleh varians galas (error) taksiran Y untuk x

tertentu, yang dilambangkan dengan $S^2y.x$, dan dapat dihitung dengan rumus sebagai berikut:

$$S^2y.x = \frac{\sum (Y - \hat{Y})^2}{n-2} \Rightarrow \text{Jk residu}$$

Varians $S^2y.x$ ini merupakan taksiran untuk varians $\sigma^2y.x$ dalam populasi dengan model regresi $Y = \alpha + \beta x$.

Asumsi:

1. Besarnya varians $\sigma^2y.x$ (Varians Y) untuk tiap kelompok harga x tertentu adalah sama besar.
 \Rightarrow Variansnya homogen (disebut: homoscedasticity)
2. Distribusi Y untuk tiap kelompok harga x tertentu adalah normal, dengan rata-rata $(\alpha + \beta x)$.
 \Rightarrow Pengujian dua asumsi ini dapat dilakukan dengan menggunakan scatter plot (scatter diagram).
 \Rightarrow Selanjutnya dengan adanya taksiran simpangan baku Y ini, berdasarkan regresi $\hat{Y} = a + bx$, akan dapat dihitung pula varians untuk konstanta a yaitu S_a^2 dan varians untuk koefisien arah regresi b, yaitu S_b . Adapun rumusnya sebagai berikut:

$$S_a^2 = S_{y.x}^2 \left\{ \frac{1}{n} + \frac{\hat{X}^2}{\Sigma x^2 - \frac{(\Sigma X)^2}{n}} \right\}$$

\hat{X} adalah rata-rata X untuk semua pengamatan

$$S_b^2 = \frac{S_{y.x}^2}{\Sigma x^2 - \frac{(\Sigma X)^2}{n}}$$

BAB IX

UJI STATISTIK NON PARAMETRIK UJI BEDA DUA KELOMPOK

9.1. Uji Wilcoxon Rank-Sum (Uji Mann-Whitney U)

Uji Wilcoxon Rank-Sum, yang sering juga dikenal sebagai Uji Mann-Whitney U, adalah salah satu uji statistik non-parametrik yang digunakan untuk membandingkan median dari dua kelompok data yang independen. Uji ini sangat berguna ketika data tidak terdistribusi normal atau tidak memenuhi asumsi lain dari uji parametrik.

1. Tujuan Penggunaan:

Membandingkan distribusi dua kelompok data yang independen.

Cocok digunakan untuk data dalam skala ordinal atau interval ketika asumsi normalitas tidak terpenuhi.

2. Asumsi:

- Data dalam bentuk ordinal atau interval.
- Independensi antar kelompok (kelompok pertama tidak dipengaruhi oleh kelompok kedua).

- Variabel dependen mengikuti distribusi yang sama di antara kelompok.
3. Langkah-langkah Uji:
- a. Penetapan Hipotesis:
 - Hipotesis nol (H_0): Tidak ada perbedaan signifikan antara kedua kelompok dalam median.
 - Hipotesis alternatif (H_1): Ada perbedaan signifikan antara kedua kelompok dalam median.
 - b. Peringkat Data:

Gabungkan data dari kedua kelompok dan peringkatkan dari yang terkecil hingga yang terbesar.
 - c. Perhitungan Uji Statistik

Hitung nilai U Mann-Whitney U berdasarkan perbedaan peringkat antara dua kelompok. Nilai U ini mencerminkan seberapa sering peringkat dari satu kelompok lebih besar dari peringkat dari kelompok lain.
 - d. Penentuan Signifikansi

Bandingkan nilai U yang dihitung dengan nilai kritis dari tabel distribusi U Mann-Whitney U untuk menentukan apakah

perbedaan antara kelompok-kelompok itu signifikan secara statistik.

4. Interpretasi Hasil

Jika nilai U yang dihitung lebih kecil dari nilai kritis yang ditentukan, maka hipotesis nol ditolak, dan terdapat bukti yang cukup untuk menyimpulkan bahwa ada perbedaan signifikan antara kedua kelompok dalam median.

Jika tidak, maka tidak cukup bukti untuk menolak hipotesis nol.

5. Contoh Penggunaan Uji Wilcoxon Rank-Sum

Misalnya, Anda ingin membandingkan hasil tes kognitif antara dua kelompok siswa yang berbeda. Data hasil tes tersebut tidak terdistribusi normal. Dengan menggunakan Uji Wilcoxon Rank-Sum, Anda dapat menentukan apakah terdapat perbedaan median yang signifikan antara kedua kelompok tersebut.

9.2. Uji Wilcoxon Signed-Rank

Uji Wilcoxon Signed-Rank adalah salah satu uji statistik non-parametrik yang digunakan untuk membandingkan dua kondisi terkait dari satu kelompok data. Uji ini cocok digunakan ketika data tidak

terdistribusi normal atau memiliki asimetri yang signifikan.

1. Tujuan Penggunaan

Membandingkan dua kondisi terkait dari satu kelompok data, misalnya sebelum dan sesudah perlakuan.

Cocok digunakan ketika data bersifat ordinal atau interval dan tidak memenuhi asumsi normalitas.

2. Asumsi

- Data dalam bentuk ordinal atau interval.
- Kedua kondisi terkait (misalnya, pengukuran sebelum dan sesudah) diambil dari populasi yang sama.
- Perbedaan antar pengamatan dalam pasangan tidak terdistribusi secara normal.

3. Langkah-langkah Uji

a. Penetapan Hipotesis

- Hipotesis nol (H_0): Tidak ada perbedaan signifikan antara kedua kondisi.
- Hipotesis alternatif (H_1): Terdapat perbedaan signifikan antara kedua kondisi.

b. Peringkat Data

- Hitung selisih antara pasangan data (misalnya, perbedaan antara skor sebelum dan sesudah).
- Peringkatkan nilai absolut dari selisih ini dari yang terkecil hingga yang terbesar.

c. Perhitungan Uji Statistik

Hitung nilai Wilcoxon Signed-Rank W berdasarkan peringkat dari selisih yang diurutkan.

Nilai W dapat digunakan untuk menguji hipotesis nol melalui distribusi Wilcoxon Signed-Rank.

d. Penentuan Signifikansi

Bandungkan nilai W yang dihitung dengan nilai kritis dari tabel distribusi Wilcoxon Signed-Rank untuk menentukan apakah perbedaan antara kondisi tersebut signifikan secara statistik.

4. Interpretasi Hasil

Jika nilai W yang dihitung lebih kecil dari nilai kritis yang ditentukan, maka hipotesis nol ditolak, dan terdapat bukti yang cukup untuk menyimpulkan bahwa ada perbedaan signifikan antara kedua kondisi.

Jika tidak, maka tidak cukup bukti untuk menolak hipotesis nol.

5. Contoh Penggunaan Uji Wilcoxon Signed-Rank
Misalnya, Anda ingin menguji efektivitas suatu intervensi terhadap tingkat kecemasan pada individu. Dengan menggunakan Uji Wilcoxon Signed-Rank, Anda dapat menentukan apakah terdapat perbedaan yang signifikan antara tingkat kecemasan sebelum dan sesudah intervensi.

9.3. Uji Kolmogorov-Smirnov

Uji Kolmogorov-Smirnov adalah salah satu uji statistik non-parametrik yang digunakan untuk menguji kesamaan antara distribusi kumulatif dari dua kelompok data atau untuk menguji apakah sampel data berasal dari populasi yang terdistribusi sesuai dengan distribusi tertentu.

1. Tujuan Penggunaan:
 - Menguji kesamaan antara dua distribusi data.
 - Menguji apakah suatu sampel data terdistribusi sesuai dengan distribusi teoritis tertentu (misalnya, distribusi normal).

2. Asumsi:
 - Data dalam bentuk interval atau rasio.
 - Data tidak perlu berdistribusi normal.
3. Langkah-langkah Uji:
 - a. Penetapan Hipotesis:
 - Hipotesis nol (H_0): Distribusi kedua kelompok data adalah sama.
 - Hipotesis alternatif (H_1): Distribusi kedua kelompok data adalah berbeda.
 - b. Perhitungan Statistik Uji:
 - Hitung nilai uji Kolmogorov-Smirnov D , yang merupakan nilai maksimum dari selisih absolut antara distribusi kumulatif empiris dari dua kelompok data.
 - Nilai D dihitung berdasarkan perbedaan antara distribusi kumulatif empiris dari dua kelompok data.
 - c. Penentuan Signifikansi:

Bandingkan nilai D yang dihitung dengan nilai kritis dari tabel distribusi Kolmogorov-Smirnov untuk menentukan apakah perbedaan antara distribusi kedua kelompok data itu signifikan secara statistik.

4. Interpretasi Hasil:

Jika nilai D yang dihitung lebih besar dari nilai kritis yang ditentukan, maka hipotesis nol ditolak, dan terdapat bukti yang cukup untuk menyimpulkan bahwa distribusi kedua kelompok data berbeda secara signifikan.

Jika tidak, maka tidak cukup bukti untuk menolak hipotesis nol.

5. Contoh Penggunaan Uji Kolmogorov-Smirnov:

Misalnya, Anda memiliki dua kelompok data yang masing-masing mewakili waktu reaksi terhadap dua jenis stimulasi yang berbeda. Dengan menggunakan Uji Kolmogorov-Smirnov, Anda dapat menentukan apakah distribusi waktu reaksi dari kedua kelompok itu berbeda secara signifikan.

9.4. Uji Chi-Square untuk Data Berpasangan

Uji Chi-Square untuk data berpasangan, atau yang sering juga disebut sebagai Uji Chi-Square McNemar, adalah salah satu metode uji statistik yang digunakan untuk menguji perbedaan signifikan antara dua variabel kategorikal yang diukur pada waktu yang berbeda pada satu kelompok subjek atau objek yang sama.

1. Tujuan Penggunaan

Menguji perbedaan antara dua variabel kategorikal yang diukur pada dua waktu atau kondisi yang berbeda, tetapi pada kelompok subjek yang sama.

Cocok digunakan untuk data yang tidak terdistribusi secara normal dan berjenis data nominal atau ordinal.

2. Asumsi

- Data dalam bentuk kategorikal (misalnya, yes/no, sukses/gagal, hadir/tidak hadir).
- Variabel-variabel yang diukur pada waktu yang berbeda memiliki hubungan atau terkait satu sama lain.

3. Langkah-langkah Uji

a. Penetapan Hipotesis

- Hipotesis nol (H_0): Tidak ada perbedaan yang signifikan antara proporsi kategori pada waktu pertama dan waktu kedua.
- Hipotesis alternatif (H_1): Terdapat perbedaan yang signifikan antara proporsi kategori pada waktu pertama dan waktu kedua.

b. Perhitungan Tabel Kontingensi

Bentuk tabel kontingensi 2x2 yang memperlihatkan frekuensi dari kombinasi

perbedaan antara dua variabel kategori pada waktu pertama dan waktu kedua.

c. Perhitungan Statistik Uji

Hitung nilai Chi-Square McNemar berdasarkan selisih antara frekuensi yang diamati dan frekuensi yang diharapkan jika hipotesis nol benar.

Rumus Chi-Square McNemar:

$$X^2 = \frac{(|b - c| - 1)^2}{b + c}$$

Keterangan:

- b: frekuensi di sel yang berbeda antara kategori positif → negative.
- c: frekuensi di sel yang berbeda antara kategori negatif → positif.

d. Penentuan Signifikansi

Bandingkan nilai Chi-Square yang dihitung dengan nilai kritis dari tabel distribusi Chi-Square untuk menentukan apakah perbedaan antara waktu pertama dan waktu kedua signifikan secara statistik.

4. Interpretasi Hasil

Jika nilai Chi-Square yang dihitung lebih besar dari nilai kritis yang ditentukan, maka hipotesis nol ditolak, dan terdapat bukti yang cukup

untuk menyimpulkan bahwa terdapat perbedaan yang signifikan antara variabel kategori pada waktu pertama dan waktu kedua. Jika tidak, maka tidak cukup bukti untuk menolak hipotesis nol.

5. Contoh Penggunaan Uji Chi-Square untuk Data Berpasangan

Misalnya, Anda ingin menguji apakah terdapat perubahan signifikan dalam preferensi warna antara dua waktu pengukuran pada sekelompok mahasiswa. Dengan menggunakan Uji Chi-Square untuk data berpasangan, Anda dapat menentukan apakah terdapat perbedaan yang signifikan dalam distribusi preferensi warna pada dua waktu pengukuran tersebut.

BAB X

UJI STATISTIK NON PARAMETRIK UJI BEDA LEBIH DARI DUA KELOMPOK

10.1. Uji Kruskal-Wallis H

1. Pengertian Uji Kruskal-Wallis H

Uji Kruskal-Wallis H adalah uji statistik non-parametrik yang digunakan untuk menentukan apakah terdapat perbedaan yang signifikan antara tiga atau lebih kelompok independen. Uji ini adalah versi non-parametrik dari ANOVA satu arah dan digunakan ketika asumsi normalitas tidak dapat dipenuhi.

2. Prinsip Dasar

Uji Kruskal-Wallis H bekerja dengan membandingkan peringkat data daripada nilai mentahnya. Data dari semua kelompok digabungkan dan kemudian diurutkan. Perbedaan peringkat antara kelompok kemudian dianalisis untuk melihat apakah ada perbedaan yang signifikan.

3. Asumsi-Asumsi

- a. Data terdiri dari sampel independen.

- b. Data diukur pada skala ordinal atau lebih tinggi.
- c. Distribusi kelompok memiliki bentuk yang serupa.

4. Langkah-Langkah Analisis

a. Penggabungan dan Peringkat Data

- Gabungkan semua data dari kelompok yang berbeda menjadi satu set data.
- Urutkan data dari nilai terkecil hingga terbesar dan berikan peringkat pada data ini.

b. Menghitung Statistik Uji

- Hitung total peringkat untuk setiap kelompok.
- Hitung statistik H menggunakan rumus:

$$H = \frac{12}{N(N + 1)} \sum \frac{R_i^2}{n_i} - 3(N + 1)$$

Keterangan:

- N: total jumlah observasi.
 - n_i : jumlah observasi dalam kelompok i.
 - R_i : jumlah peringkat dalam kelompok i.
- c. Menentukan Nilai Kritis:
- Nilai H yang dihitung dibandingkan dengan nilai kritis dari distribusi chi-square dengan

derajat kebebasan (df) sebesar jumlah kelompok - 1.

d. Menarik Kesimpulan:

Jika nilai H yang dihitung lebih besar daripada nilai kritis chi-square, hipotesis nol (tidak ada perbedaan) ditolak.

10.2. Uji Friedman

1. Pengertian Uji

Friedman Uji Friedman adalah uji statistik non-parametrik yang digunakan untuk mendeteksi perbedaan signifikan antara kelompok-kelompok yang berpasangan. Uji ini sering digunakan dalam desain eksperimen yang melibatkan pengukuran berulang atau data yang berpasangan. Uji Friedman merupakan alternatif non-parametrik dari ANOVA dua arah dengan pengukuran berulang.

2. Prinsip Dasar

Uji Friedman bekerja dengan membandingkan peringkat data daripada nilai mentahnya. Data dari setiap kelompok diperingkatkan, dan kemudian perbedaan peringkat antara kelompok dianalisis untuk melihat apakah ada perbedaan yang signifikan.

3. Asumsi-Asumsi

- Data terdiri dari sampel berpasangan atau pengukuran berulang.
- Data diukur pada skala ordinal atau lebih tinggi.
- Distribusi bentuk data harus serupa antar kelompok.

4. Langkah-Langkah Analisis

a. Penggabungan dan Peringkat Data:

- Data diorganisir dalam tabel di mana setiap baris mewakili satu subjek dan setiap kolom mewakili satu kondisi atau waktu.
- Peringkat diberikan untuk setiap baris (subjek) secara independen, dengan peringkat 1 untuk nilai terkecil dalam setiap baris.

b. Menghitung Statistik Uji:

Hitung total peringkat (R_j) untuk setiap kolom (kondisi).

Hitung statistik Q menggunakan rumus:

$$Q = \frac{12}{nk(k+1)} \sum R_j^2 - 3n(k+1)$$

Keterangan:

- n: jumlah subjek.

- k: jumlah kondisi atau waktu.
 - R_j adalah jumlah peringkat untuk kondisi j.
- c. Menentukan Nilai Kritis:
Nilai Q yang dihitung dibandingkan dengan nilai kritis dari distribusi chi-square dengan derajat kebebasan (df) sebesar jumlah kondisi - 1.
- d. Menarik Kesimpulan:
Jika nilai Q yang dihitung lebih besar daripada nilai kritis chi-square, hipotesis nol (bahwa tidak ada perbedaan) ditolak.

10.3. Uji Cochran

a. Pengertian Uji Cochran

Uji Cochran adalah uji statistik non-parametrik yang digunakan untuk menentukan apakah terdapat perbedaan signifikan antara proporsi atau persentase dari beberapa kelompok yang berbeda. Uji ini umumnya digunakan untuk data kategori atau binomial yang diperoleh dari kelompok-kelompok yang berbeda.

b. Prinsip Dasar Uji Cochran

Uji Cochran membandingkan proporsi atau persentase dari beberapa kelompok untuk

melihat apakah terdapat perbedaan yang signifikan di antara mereka. Uji ini cocok untuk menguji hipotesis bahwa proporsi atau persentase dari beberapa kelompok itu sama atau tidak sama.

c. Asumsi-Asumsi

- Data yang digunakan adalah data kategori atau data binomial (dua kategori).
- Kelompok-kelompok adalah independen satu sama lain.
- Setiap kelompok memiliki jumlah sampel yang sama atau hampir sama.

d. Langkah-Langkah Analisis

a. Pengumpulan Data

Kumpulkan data dari masing-masing kelompok untuk variabel kategori atau binomial.

b. Menentukan Proporsi

Hitung proporsi atau persentase dari setiap kelompok untuk variabel kategori atau binomial yang diamati.

c. Uji Statistik

Hitung nilai statistik Cochran Q menggunakan rumus:

$$Q = \frac{(k - 1) \cdot (n - 1)}{n} \left(\frac{\sum (X_{ij} - X_i)^2}{k} \right)$$

Keterangan:

- k: jumlah kelompok.
- n: jumlah observasi (sampel) dalam setiap kelompok.
- X_{ij} : frekuensi atau jumlah kejadian kategori tertentu dalam kelompok i dan kategori j.
- X_i : total frekuensi atau jumlah kejadian dalam kelompok i.

d. Menentukan Nilai Kritis

Bandingkan nilai Q yang dihitung dengan nilai kritis dari distribusi chi-square dengan derajat kebebasan $k-1$ untuk tingkat signifikansi yang dipilih.

e. Interpretasi Hasil

Jika nilai Q yang dihitung lebih besar dari nilai kritis chi-square, maka kita dapat menolak hipotesis nol dan menyimpulkan bahwa terdapat perbedaan signifikan antara setidaknya dua kelompok.

Jika tidak, kita gagal menolak hipotesis nol, yang berarti tidak ada cukup bukti untuk

menyatakan adanya perbedaan signifikan antara kelompok-kelompok tersebut.

10.4. Uji Jonckheere-Terpstra

1. Pengertian Uji Jonckheere-Terpstra

Uji Jonckheere-Terpstra adalah uji statistik non-parametrik yang digunakan untuk menguji apakah terdapat tren monotonik (naik atau turun) di antara beberapa kelompok yang tidak terhubung. Uji ini sering digunakan dalam studi medis, sosial, dan perilaku untuk menguji perbedaan yang berurutan di antara kelompok-kelompok tersebut.

2. Prinsip Dasar

Uji Jonckheere-Terpstra membandingkan peringkat dari masing-masing kelompok data untuk melihat apakah terdapat tren yang signifikan antara mereka. Tren bisa berupa peningkatan atau penurunan, dan uji ini sensitif terhadap perbedaan-perbedaan ini tanpa mengharuskan data mengikuti distribusi tertentu.

3. Asumsi-Asumsi

- Data diukur pada skala ordinal atau lebih tinggi.

- Kelompok-kelompok adalah independen satu sama lain.
- Terdapat tren monotonik yang diharapkan di antara kelompok-kelompok.

4. Langkah-Langkah Analisis

a. Pengumpulan Data

Kumpulkan data dari masing-masing kelompok untuk variabel yang diamati.

b. Menghitung Peringkat

Peringkatkan data dari masing-masing kelompok secara terpisah.

c. Uji Statistik

Hitung nilai statistik Jonckheere-Terpstra (JT) menggunakan perbedaan peringkat antar kelompok.

Uji hipotesis menggunakan nilai JT yang dihitung dengan membandingkannya dengan nilai kritis dari distribusi JT untuk derajat kebebasan yang sesuai.

d. Menarik Kesimpulan

Jika nilai JT yang dihitung lebih besar dari nilai kritis pada tingkat signifikansi yang ditentukan (misalnya 0.05), maka kita dapat menolak hipotesis nol dan menyimpulkan bahwa terdapat tren

monotik yang signifikan di antara kelompok-kelompok.

Jika tidak, kita tidak memiliki cukup bukti untuk menolak hipotesis nol.

10.5. Uji Median

1. Pengertian Uji Median

Uji Median adalah metode statistik non-parametrik yang digunakan untuk menentukan apakah median dua kelompok data berbeda secara signifikan. Uji ini sering digunakan ketika data tidak mengikuti distribusi normal atau ketika hanya ingin mengetahui perbedaan median antara dua kelompok.

2. Prinsip Dasar Uji Median

Uji Median membandingkan median dari dua kelompok data. Median adalah nilai tengah dari data yang diurutkan. Jika terdapat perbedaan yang signifikan antara median kedua kelompok, maka uji ini dapat mendeteksinya.

3. Asumsi-Asumsi

- Data diukur pada skala ordinal atau lebih tinggi.
- Tidak ada asumsi distribusi tertentu seperti pada uji parametrik.

- Data dari dua kelompok adalah independen satu sama lain.
4. Langkah-Langkah Analisis
 - a. Pengumpulan Data

Kumpulkan data dari dua kelompok yang ingin dibandingkan.
 - b. Menghitung Median

Hitung median dari masing-masing kelompok data.
 - c. Uji Statistik

Gunakan uji statistik non-parametrik yang sesuai untuk membandingkan median dua kelompok, seperti Uji U Mann-Whitney (Uji Wilcoxon Rank Sum) atau Uji Median.
 - d. Menarik Kesimpulan

Interpretasikan hasil uji untuk menentukan apakah terdapat perbedaan median yang signifikan antara dua kelompok.

DAFTAR PUSTAKA

- Agresti, A. (2002). *Categorical Data Analysis*. 2nd Edition. Wiley-Interscience.
- Agresti, A. (2015). *Foundations of Linear and Generalized Linear Models*. John Wiley & Sons.
- Agresti, A., & Finlay, B. (2009). *Statistical Methods for the Social Sciences*. Pearson.
- Agresti, A., & Finlay, B. (2009). *Statistical Methods for the Social Sciences* (4th ed.). Prentice Hall.
- Agresti, Alan, and Barbara Finlay. "Statistical Methods for the Social Sciences." 4th ed., Pearson, 2008.
- Al Dianty, M., Yahaya, AS., Ahmad,F. (2014). *Probability Distribution of Engineering Properties of Soil at Telecommunication Sites in Indonesia*, International Journal of Scientific Research in Knowledge. 143-150.
- Bertin, Jacques. "Semiology of Graphics: Diagrams, Networks, Maps." Esri Press, 2010.
- Bulmer, M. G. (1979). *Principles of Statistics*. Dover Publications.
- Chambers, John M. "Graphical Methods for Data Analysis." Wadsworth Pub Co, 1983.
- Cleveland, W. S. (1993). "Visualizing Data." Hobart Press.

- Conover, W.J. (1999). Practical Nonparametric Statistics. 3rd Edition. John Wiley & Sons.
- Devore, J. L. (2015). Probability and Statistics for Engineering and the Sciences. Cengage Learning.
- Everitt, B. S., & Dunn, G. (2001). Applied Multivariate Data Analysis. London: Arnold Publishers.
- Everitt, B.S., & Skrondal, A. (2010). The Cambridge Dictionary of Statistics. 4th Edition. Cambridge University Press.
- Few, S. (2009). "Now You See It: Simple Visualization Techniques for Quantitative Analysis." Analytics Press.
- Field, A. (2013). Discovering Statistics Using IBM SPSS Statistics. Sage Publications.
- Field, A. (2013). Discovering Statistics Using IBM SPSS Statistics. Sage Publications.
- Field, A. (2013). Discovering Statistics Using IBM SPSS Statistics. 4th Edition. SAGE Publications.
- Freedman, David, et al. "Statistics." 4th ed., Norton & Company, 2007.
- Freund, J. E., & Perles, B. M. (2010). Modern Elementary Statistics, 12th ed. Upper Saddle River, NJ: Prentice Hall.
- Gibbons, J.D., & Chakraborti, S. (2011). Nonparametric Statistical Inference. 5th Edition. Chapman & Hall/CRC.

- Gravetter, Frederick J., and Larry B. Wallnau. "Statistics for the Behavioral Sciences." 10th ed., Cengage Learning, 2017.
- Hair Jr, J. F., Black, W. C., Babin, B. J., & Anderson, R. E. (2014). *Multivariate Data Analysis* (7th ed.). Pearson Education Limited.
- Hair, J. F., Black, W. C., Babin, B. J., & Anderson, R. E. (2018). *Multivariate data analysis* (8th ed.). Cengage Learning.
- Hair, J. F., Black, W. C., Babin, B. J., & Anderson, R. E. (2019). *Multivariate Data Analysis*. Pearson.
- Hasan MK., Mohamed, H., Sahari, Sahran S., Hanawi SA. (2002). *Statistik untuk Teknologi Maklumat dan Industri. McGraw- Hill (Malaysia) Sdn. Bhd.*
- Hastie, T., Tibshirani, R., & Friedman, J. (2009). *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*. Springer Science & Business Media.
- Helsel, Dennis R., and Richard M. Hirsch. "Statistical Methods in Water Resources." Elsevier, 2002.
- Hinkle, D.E., Wiersma, W., & Jurs, S.G. (2003). *Applied Statistics for the Behavioral Sciences*. 5th Edition. Houghton Mifflin.
- Hollander, M., Wolfe, D.A., & Chicken, E. (2013). *Nonparametric Statistical Methods*. 3rd Edition. John Wiley & Sons.
- Howell, D. C. (2013). *Statistical Methods for Psychology* (8th ed.). Cengage Learning.

Howell, D.C. (2010). *Statistical Methods for Psychology*. 7th Edition. Wadsworth Cengage Learning.

<https://www.semanticscholar.org/paper/Probability-Distribution-of-Engineering-Properties-Dianty-Yahaya/e4dfd45404e4414aaef7d925e6158c44086f95cb>

Introduction to the Practice of Statistics" oleh David S. Moore, George P. McCabe, dan Bruce A. Craig.

Johnson, R. A., & Wichern, D. W. (2007). *Applied Multivariate Statistical Analysis*. Pearson Prentice Hall.

Kanellopoulos, Ilias, dan Nicholas S. Floudas. "The Importance of Diagrams in Scientific Papers and Presentations." *AIChE Journal* 54, no. 2 (2008): 302-320.

Keller, G. (2019). *Statistics for management and economics* (11th ed.). Cengage Learning.

Kenney, J. F., & Keeping, E. S. (1951). *Mathematics of Statistics, Pt. 1*, 3rd ed. Princeton, NJ: Van Nostrand.

Kirk, R.E. (2013). *Experimental Design: Procedures for the Behavioral Sciences*. 4th Edition. Sage Publications.

Kosslyn, Stephen M. "Graph Design for the Eye and Mind." Oxford University Press, 2006.

Lane, D. (2017). *Online Statistics Education: An Interactive Multimedia Course of Study*. Retrieved from <http://onlinestatbook.com/>.

- Manly, B.F.J. (2007). *Randomization, Bootstrap and Monte Carlo Methods in Biology*. 3rd Edition. Chapman & Hall/CRC.
- McClave, J. T., Benson, P. G., & Sincich, T. (2019). *Statistics for business and economics* (13th ed.). Pearson.
- Monmonier, Mark. "How to Lie with Maps." University of Chicago Press, 1996.
- Montgomery, D. C., Peck, E. A., & Vining, G. G. (2012). *Introduction to Linear Regression Analysis*. John Wiley & Sons.
- Montgomery, D. C., Peck, E. A., & Vining, G. G. (2021). *Introduction to Linear Regression Analysis* (6th ed.). Wiley.
- Montgomery, D.C. (2017). *Design and Analysis of Experiments*. 9th Edition. John Wiley & Sons.
- Montgomery, Douglas C., et al. "Introduction to Linear Regression Analysis." 5th ed., Wiley, 2012.
- Moore, D. S., & McCabe, G. P. (2018). *Introduction to the Practice of Statistics*. W. H. Freeman.
- Moore, D. S., McCabe, G. P., & Craig, B. A. (2016). *Introduction to the Practice of Statistics*. W. H. Freeman.
- Ott, R. L., Longnecker, M., & Witt, E. (2016). *An Introduction to Statistical Methods and Data Analysis*. Cengage Learning.

- Ott, R. Lyman, and Michael Longnecker. "An Introduction to Statistical Methods and Data Analysis." 7th ed., Cengage Learning, 2015.
- Priyatno, Duwi, SPSS untuk Analisis Korelasi, Regresi, dan Multivariate, Yogyakarta: Gaya Media, 2009.
- Brigham, E. F., & Daves, P. R. (2014).
- Rafi'i S. (1983). Metode Statistika Analisis (untuk penarikan kesimpulan). *Jakarta: Penerbit Binacipta Anggota IKAPI*
- Rencher, A. C., & Christensen, W. F. (2012). *Methods of Multivariate Analysis*, 3rd ed. New York, NY: John Wiley & Sons.
- Rencher, A. C., & Schaalje, G. B. (2012). *Linear models in statistics* (2nd ed.). Wiley.
- Rosner, B. (2015). *Fundamentals of Biostatistics*. Cengage Learning.
- Rumsey, Deborah J. "Statistics For Dummies." Wiley, 2020.
- Siegel, S., & Castellan, N.J. (1988). *Nonparametric Statistics for the Behavioral Sciences*. 2nd Edition. McGraw-Hill.
- Statistics for Business and Economics" oleh Paul Newbold, William L. Carlson, dan Betty Thorne.
- Sudarmanto, R. Gunawan, Analisis Regresi Linier Ganda dengan SPSS, edisi pertama, Yogyakarta: Graha Ilmu, 2005.

- Tabachnick, B. G., & Fidell, L. S. (2019). Using Multivariate Statistics (7th ed.). Pearson.
- Tabachnick, B. G., & Fidell, L. S. (2019). Using Multivariate Statistics (7th ed.). Pearson.
- Tabachnick, B.G., & Fidell, L.S. (2013). Using Multivariate Statistics. 6th Edition. Pearson Education.
- Triola, M. F. (2018). Elementary Statistics (13th ed.). Pearson.
- Tufte, E. R. (2001). "The Visual Display of Quantitative Information." Graphics Press.
- Utts, J. M., & Heckard, R. F. (2019). Mind on statistics (6th ed.). Cengage Learning.
- [Walpole](#), R. E. (1990). Pengantar Statistika (Edisi 3). Penerbit PT Gramedia Pustaka Utama.
- Wickham, H. (2016). "ggplot2: Elegant Graphics for Data Analysis." Springer.
- Wooldridge, J. M. (2019). Introductory Econometrics: A Modern Approach (7th ed.). Cengage Learning.
- Zar, J.H. (2010). Biostatistical Analysis. 5th Edition. Pearson.

STATISTIK DASAR

Statistik adalah bahasa yang digunakan untuk menggambarkan dan memahami fenomena kompleks di sekitar kita. Dari ilmu pengetahuan hingga keputusan bisnis, statistik memainkan peran penting dalam membantu kita mengambil keputusan yang lebih baik dan lebih informasional.

Buku ini dirancang untuk memberikan landasan yang kuat dalam konsep dasar statistik, dari pengukuran pemusatan data hingga pengenalan analisis inferensial. Dengan penjelasan yang jelas dan contoh yang relevan, pembaca akan dibimbing melalui berbagai teknik dan aplikasi statistik yang penting.



Penerbit Yayasan
Cendikia Mulia Mandiri



ISBN 978-623-8576-80-7



9 786238 576807