STATISTIK DASAR

Syaiful Mansyur, M.Sc

"Statistik adalah ilmu yang mempelajari ketidakpastian" - Ronald Fisher

UP45 PRESS

Penerbit UP45 PRESS

STATISTIK DASAR

Syaiful Mansyur, M.Sc

PENERBIT: UP45 PRESS – Universitas proklamasi 45

BUKU

STATISTIK DASAR

Penulis	:	Syaiful Mansyur, M.Sc
ISBN	:	
Editor	:	UP45 PRESS
Sampul & Tata Letak	:	UP45 PRESS
Penerbit	:	UP45 PRESS – UNIVERSITAS PROKLAMASI 45

Alamat:
Jl. Proklamasi No.1 Babarsari, Yogyakarta 55281
Tlp. (0274) 485517
Email: up45jogjapress@gmail.com
Cetak Pertama : 2025

KATA PENGANTAR

Puji syukur kami panjatkan ke hadirat Tuhan Yang Maha Esa atas selesainya penyusunan buku **"Statistik Dasar"** ini. Buku ini disusun sebagai referensi bagi pembaca, khususnya mahasiswa yang ingin memahami konsep dasar statistik secara sistematis dan aplikatif.

Statistik merupakan disiplin ilmu yang sangat penting dalam berbagai bidang, termasuk sains, teknik, ekonomi, dan manajemen. Dengan memahami statistik, kita dapat menganalisis data secara efektif, mengambil keputusan berdasarkan informasi yang akurat, serta melakukan prediksi yang lebih tepat. Oleh karena itu, buku ini dirancang untuk memberikan pemahaman yang komprehensif tentang prinsip-prinsip dasar statistik, mulai dari konsep dasar dengan metode analisis yang sederhana.

Buku ini terdiri dari tujuh bab yang mencakup topik-topik utama dalam statistik dasar:

1. **Bab I: Pengantar Statistik** – Membahas sejarah, definisi, tujuan, dan jenis-jenis statistik, termasuk perbedaan antara statistik deskriptif dan inferensial.
2. **Bab II: Statistik Deskriptif** – Menjelaskan ukuran pemusatan (mean, median, modus), ukuran variabilitas (rentang dan deviasi standar), serta penyajian data dalam bentuk grafik dan diagram.
3. **Bab III: Statistik Inferensial** – Mengupas konsep estimasi parameter, uji hipotesis, confidence interval, serta statistik parametrik dan non-parametrik.
4. **Bab IV: Probabilitas** – Menguraikan konsep dasar probabilitas, probabilitas bersyarat, Teorema Bayes, serta konsep permutasi dan kombinasi.
5. **Bab V: Distribusi Peluang** – Membahas distribusi probabilitas diskrit (binomial dan Poisson) serta distribusi probabilitas kontinu (normal, t-Student, dan F).
6. **Bab VI: Regresi dan Korelasi** – Mengajarkan cara menganalisis hubungan antara variabel melalui regresi linier dan korelasi.

7. **Bab VII: Analisis Varians (ANOVA)** – Membahas metode untuk membandingkan rata-rata lebih dari dua kelompok serta penerapan uji lanjut (Post Hoc Test).

Kami berharap buku ini dapat membantu pembaca dalam memahami dan mengaplikasikan konsep-konsep statistik secara lebih mendalam. Kami juga menyadari bahwa tidak ada karya yang sempurna, sehingga kami sangat terbuka terhadap kritik dan saran yang membangun untuk perbaikan buku ini di masa mendatang.

Akhir kata, semoga buku ini dapat memberikan manfaat bagi semua yang membacanya dan menjadi referensi yang berguna dalam studi serta penerapan statistik di berbagai bidang.

Syaiful Mansyur

Yogyakarta, 2025

DAFTAR ISI

BAB I

Tujuan Pembelajaran

Setelah mempelajari bab ini, pembaca diharapkan mampu:

1. Memahami konsep dasar statistik dan perkembangannya dalam berbagai bidang.
2. Menjelaskan definisi statistik serta perbedaan antara populasi dan sampel.
3. Menganalisis berbagai teknik pengambilan sampel untuk penelitian.
4. Menjelaskan tujuan penggunaan statistik dalam pengolahan data.
5. Membedakan antara statistik deskriptif dan statistik inferensial serta penerapannya.

Pokok Bahasan

1. **Pengantar Statistik**
2. **Definisi Statistik**
3. **Tujuan Statistik**
4. **Jenis-Jenis Statistik**

1. PENGANTAR STATISTIK

Dalam setiap upaya untuk memahami dunia di sekitar kita, pengetahuan menjadi fondasi yang tak tergantikan. Di belantara hutan Afrika, seekor kera hidup dalam ketidakpastian yang terus-menerus. Ia tidak memiliki kemampuan untuk memprediksi kapan pemangsa akan menyerang. Dalam situasi seperti ini, kera tersebut harus mengandalkan dua hal: insting dan analisis situasi. Instingnya, yang telah terasah melalui pengalaman dan evolusi, membantunya merasakan bahaya yang mungkin mengintai. Namun, insting saja tidak cukup. Kera tersebut juga harus mampu menganalisis lingkungan sekitarnya—mendengarkan suara-suara, mengamati gerakan, dan merasakan perubahan dalam atmosfer hutan. Dengan cara ini, ia dapat meningkatkan peluangnya untuk bertahan hidup.

Gambar 1. Cheetah yang sedang mengawasi monyet diatas pohon

Manusia, meskipun memiliki kemampuan kognitif yang lebih kompleks, sering kali menghadapi situasi serupa. Dalam banyak kasus, kita juga harus mengandalkan insting ketika situasi tidak memungkinkan untuk dianalisis secara mendalam. Misalnya, dalam situasi darurat, keputusan harus diambil dengan cepat, dan sering kali kita tidak memiliki cukup waktu untuk melakukan analisis yang menyeluruh. Dalam momen-momen seperti itu, insting yang sering kali merupakan hasil dari pengalaman dan pembelajaran sebelumnya menjadi alat yang berharga.

Namun, seiring dengan perkembangan zaman, manusia telah mengembangkan metode yang lebih sistematis untuk memahami dan memprediksi berbagai kondisi. Salah satu metode tersebut adalah analisis statistik probabilitas. Dengan menggunakan statistik, kita dapat mengukur seberapa besar peluang terjadinya suatu kondisi. Misalnya, dalam dunia bisnis, perusahaan menggunakan analisis statistik untuk memprediksi tren pasar, memahami perilaku konsumen, dan mengidentifikasi risiko. Dengan cara ini, mereka dapat membuat keputusan yang lebih terinformasi dan mengurangi ketidakpastian.

Ilmu statistik tidak hanya penting dalam konteks bisnis, tetapi juga dalam berbagai aspek kehidupan sehari-hari. Dalam bidang kesehatan, statistik digunakan untuk menganalisis data epidemiologi, membantu kita memahami penyebaran penyakit dan efektivitas pengobatan. Dalam pendidikan, analisis statistik membantu dalam mengevaluasi metode pengajaran dan hasil belajar siswa. Bahkan dalam kehidupan pribadi, kita sering kali menggunakan statistik secara tidak sadar—misalnya, saat kita mempertimbangkan peluang keberhasilan suatu keputusan berdasarkan pengalaman masa lalu.

Oleh karena itu, penting bagi setiap orang untuk memahami dasar-dasar ilmu statistik. Pengetahuan ini tidak hanya memberikan kita alat untuk menganalisis situasi dengan lebih baik, tetapi juga membantu kita membuat keputusan yang lebih bijaksana. Dalam dunia yang semakin kompleks dan penuh ketidakpastian, kemampuan untuk menganalisis data dan memahami probabilitas menjadi keterampilan yang sangat berharga.

Dengan menggabungkan insting dan analisis statistik, kita dapat meningkatkan kemampuan kita untuk menghadapi tantangan dan membuat keputusan yang lebih baik. Seperti kera di hutan yang mengandalkan kedua kemampuan ini untuk bertahan hidup, kita pun dapat belajar untuk memanfaatkan pengetahuan dan insting kita dalam menjalani kehidupan sehari-hari. Dalam perjalanan ini, kita tidak hanya menjadi lebih cerdas, tetapi juga lebih siap untuk menghadapi segala kemungkinan yang ada di depan kita.

Sejarah Statistik

Sejarah statistik dapat ditelusuri kembali ke zaman kuno, ketika manusia mulai mengumpulkan data untuk tujuan administratif dan ekonomi. Berikut adalah beberapa tonggak penting dalam perkembangan statistik:

1. **Zaman Kuno**: Pengumpulan data pertama kali dilakukan oleh peradaban kuno seperti Mesir dan Babilonia, yang mencatat informasi tentang populasi, hasil pertanian, dan pajak. Data ini digunakan untuk tujuan pemerintahan dan perencanaan ekonomi.

2. **Abad Pertengahan**: Pada abad pertengahan, statistik mulai berkembang di Eropa, terutama dalam konteks pengumpulan data untuk tujuan pajak dan administrasi. Namun, pada masa ini, metode yang digunakan masih sangat sederhana dan tidak sistematis.

3. **Abad ke-17**: Statistik modern mulai muncul pada abad ke-17 dengan kontribusi dari ilmuwan seperti John Graunt, yang dikenal sebagai "bapak statistik." Ia menerbitkan karya berjudul "Natural and Political Observations Made upon the Bills of Mortality," yang

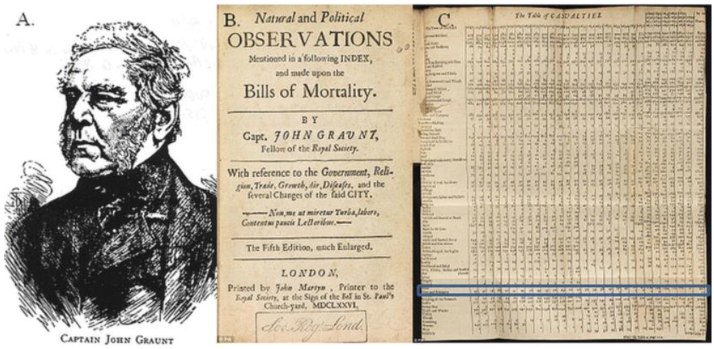
menganalisis data kematian di London dan memperkenalkan konsep dasar probabilitas.

4. **Abad ke-18**: Pada abad ini, statistik mulai digunakan dalam konteks ilmiah dan sosial. Pierre-Simon Laplace dan Carl Friedrich Gauss mengembangkan teori probabilitas dan metode statistik yang lebih formal. Konsep distribusi normal juga diperkenalkan, yang menjadi dasar bagi banyak analisis statistik modern.

5. **Abad ke-19**: Statistik semakin berkembang dengan munculnya metode inferensial dan regresi. Sir Francis Galton dan Karl Pearson berkontribusi pada pengembangan analisis regresi dan korelasi, yang memungkinkan peneliti untuk memahami hubungan antara variabel.

6. **Abad ke-20**: Statistik menjadi semakin penting dalam penelitian ilmiah dan pengambilan keputusan. Ronald A. Fisher memperkenalkan metode analisis varians (ANOVA) dan desain eksperimen, yang menjadi standar dalam penelitian ilmiah. Selain itu, penggunaan komputer mulai mengubah cara statistik diterapkan, memungkinkan analisis data yang lebih kompleks dan cepat.

7. **Era Modern**: Saat ini, statistik telah menjadi bagian integral dari hampir semua disiplin ilmu. Dengan kemajuan teknologi dan ketersediaan data besar (big data), statistik terus berkembang, dengan metode baru dan alat analisis yang muncul untuk menangani tantangan data yang semakin kompleks.



Gambar 2. Buku karya John Graunt tentang observasi
(https://www.milestone-books.de/)

2. DEFINISI STATISTIK

Statistik adalah cabang ilmu Matematika yang berkaitan dengan pengumpulan, analisis, interpretasi, presentasi, dan pengorganisasian data. Definisi statistik mencakup berbagai konsep dan prinsip yang digunakan untuk menyajikan informasi dengan cara yang sistematis. Dari definisi tersebut, secara konseptual dapat definisikan satu persatu:

1) Pengumpulan Data

Pengumpulan data statistik merupakan tahap awal dalam proses statistik yang mencakup pengambilan informasi atau pengukuran dari satu atau lebih populasi atau sampel untuk keperluan analisis. Proses pengumpulan data ini dilakukan dengan metode tertentu untuk memastikan data yang dikumpulkan akurat, representatif, dan dapat diandalkan. Untuk mengumpulkan data ada tahap yang perlu dilakukan

a. Membuat Perencanaan:
   - ✓ Menentukan tujuan pengumpulan data dan pertanyaan penelitian yang ingin dijawab.
   - ✓ Mendefinisikan populasi atau sampel yang akan diteliti.

- ✓ Memilih metode pengumpulan data yang sesuai dengan tujuan penelitian.

b. Menentukan sumber data, baik itu sumber primer (data yang dikumpulkan secara langsung) atau sumber sekunder (data yang sudah ada sebelumnya).

c. Metode Pengumpulan Data:

- ✓ **Survei**: Penggunaan kuesioner atau wawancara untuk mengumpulkan data dari responden.
- ✓ **Pengukuran**: Pengumpulan data melalui pengukuran langsung (misalnya, pengukuran fisik) atau pengukuran tidak langsung (misalnya, menggunakan instrumen).
- ✓ **Observasi**: Pengumpulan data dengan mengamati kejadian atau perilaku tanpa interaksi langsung dengan subjek.
- ✓ **Eksperimen**: Pengumpulan data melalui manipulasi variabel bebas untuk mengamati dampaknya pada variabel terikat.

d. Desain Sampel

Jika menggunakan sampel, perlu merancang cara yang representatif untuk memilih elemen sampel dari populasi. Kemudian dapat menggunakan teknik pengambilan sampel dengan metode acak, stratifikasi, atau cluster, tergantung pada kebutuhan penelitian.

e. Instrumen Pengumpulan Data

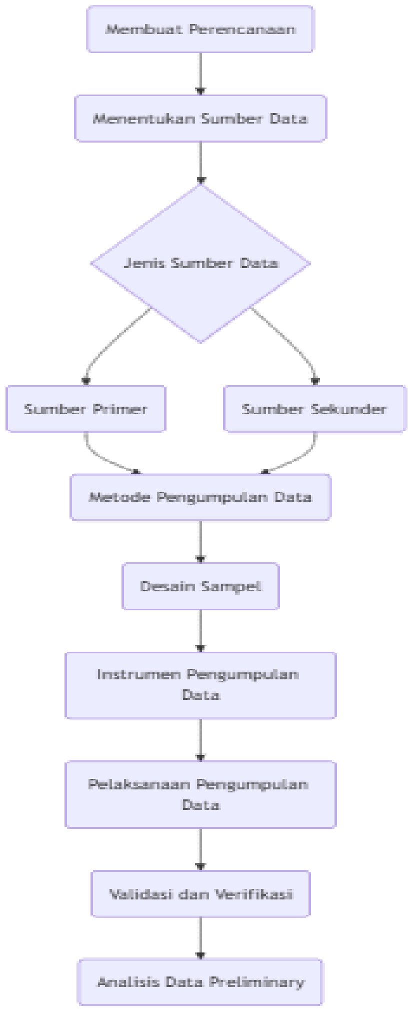
Menentukan alat atau instrumen yang akan digunakan untuk mengumpulkan data, seperti kuesioner, alat ukur, atau perangkat lunak khusus.

f. Pelaksanaan Pengumpulan Data
   - ✓ Mengumpulkan data sesuai dengan desain dan metode yang telah ditentukan.
   - ✓ Memastikan konsistensi dan keakuratan selama proses pengumpulan data.

g. Validasi dan Verifikasi
   - ✓ Memastikan bahwa data yang dikumpulkan valid dan dapat dipercaya.
   - ✓ Melakukan pengujian dan verifikasi data untuk mengidentifikasi potensi kesalahan atau bias.

h. Analisis Data Preliminary

   Melakukan analisis awal terhadap data yang dikumpulkan untuk mendapatkan pemahaman awal tentang pola atau tren.



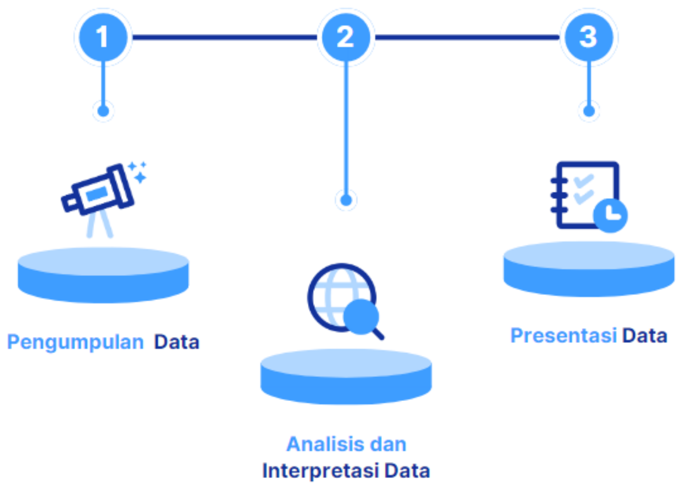
Gambar 3. Diagram alir pengumpulan data

2) Analisis dan Interpretasi

Analisis dan interpretasi data bekerja bersama untuk memberikan gambaran lengkap dari informasi yang ditemukan dalam data. Analisis memungkinkan kita menggali struktur dan pola data, sementara interpretasi memberikan konteks dan makna pada hasil tersebut. Keduanya sangat penting untuk mendukung proses pengambilan keputusan dan pengembangan pemahaman yang mendalam terhadap fenomena yang diamati

3) Presentasi Data

Presentasi yang baik membantu orang untuk lebih mudah memahami informasi yang disajikan. Statistik menyediakan cara efektif untuk menyajikan data, baik melalui tabel ataupun grafik.



Gambar 4. Definisi Statistik

Populasi dan Sampel

Populasi: Populasi dalam statistik merujuk kepada seluruh sekumpulan individu, objek, atau peristiwa yang memiliki karakteristik tertentu dan menjadi objek dari suatu penelitian. Populasi dapat berupa manusia, hewan, objek, atau konsep.

- **Populasi Manusia:** Semua penduduk di suatu negara atau wilayah geografis merupakan contoh populasi manusia. Misalnya, populasi Indonesia adalah semua penduduk yang tinggal di wilayah Indonesia.

- **Populasi Mahasiswa:** Semua mahasiswa yang terdaftar di sebuah perguruan tinggi atau universitas adalah contoh populasi mahasiswa. Contohnya, populasi mahasiswa Universitas XYZ adalah semua mahasiswa yang terdaftar di universitas tersebut.
- **Populasi Produk:** Semua produk yang diproduksi oleh suatu perusahaan adalah contoh populasi produk. Misalnya, populasi produk dari perusahaan sepatu A adalah semua jenis sepatu yang diproduksi oleh perusahaan tersebut.
- **Populasi Hewan:** Semua spesies hewan di suatu habitat atau kawasan geografis adalah contoh populasi hewan. Contohnya, populasi burung di hutan Amazon adalah semua jenis burung yang ada di hutan tersebut.

Gambar 5. Populasi

Sampel: Sampel adalah bagian yang diambil dari populasi untuk diobservasi atau diukur dalam suatu penelitian. Sampel dipilih dengan tujuan untuk mewakili populasi secara keseluruhan, sehingga hasil dari analisis sampel dapat digunakan untuk membuat kesimpulan tentang populasi secara umum.

- **Sampel Mahasiswa:** Sebuah penelitian bertujuan untuk mengukur tingkat kepuasan mahasiswa terhadap layanan kampus. Dalam hal ini, 100 mahasiswa yang dipilih secara acak dari keseluruhan populasi mahasiswa universitas tersebut akan menjadi contoh sampel.

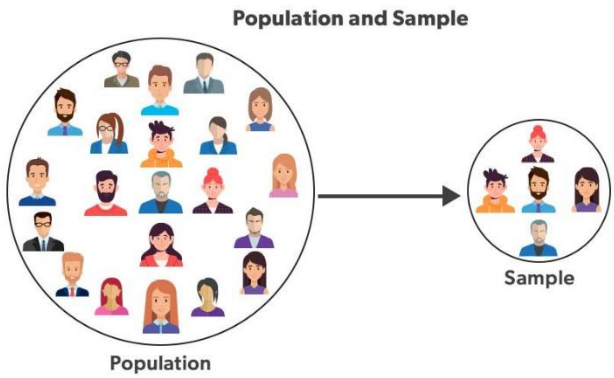
- **Sampel Produk:** Sebuah perusahaan ingin menilai kualitas produk barunya. Mereka memilih 50 produk secara acak dari seluruh produk yang diproduksi untuk diuji dan dievaluasi. Produk-produk yang dipilih ini merupakan contoh sampel.

- **Sampel Penduduk:** Sebuah lembaga survei ingin mengumpulkan data tentang preferensi politik penduduk sebuah negara. Mereka mengambil sampel acak dari seluruh penduduk yang terdaftar sebagai pemilih dalam daftar pemilih. Sampel ini merupakan contoh sampel dari populasi penduduk.

Tujuan Penggunaan Sampel:

- **Representasi**: Sampel yang diambil harus mewakili karakteristik utama dari populasi agar hasil analisis sampel dapat diperluas ke populasi secara keseluruhan.
- **Efisiensi**: Mengambil sampel daripada seluruh populasi memungkinkan untuk menghemat waktu, biaya, dan sumber daya lainnya yang diperlukan dalam penelitian.
- **Keterjangkauan**: Terkadang sulit atau tidak mungkin untuk mengumpulkan data dari seluruh populasi, sehingga pengambilan sampel menjadi solusi yang lebih realistis.

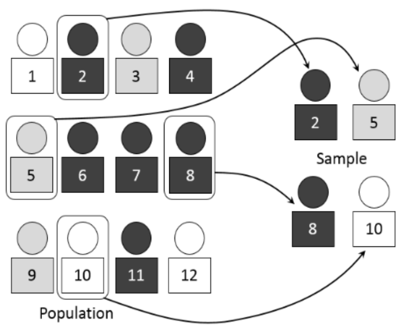
- **Ketepatan**: Dalam banyak kasus, analisis sampel dapat memberikan hasil yang cukup akurat dan dapat diandalkan tentang populasi dengan risiko yang diterima.



Gambar 6. Populasi Dan Sampel

Teknik Pengambilan Sampel

Teknik Pengambilan Sampel adalah metode atau cara untuk memilih sebagian individu atau elemen dari suatu populasi sehingga sampel yang diperoleh dapat mewakili keseluruhan populasi. Tujuannya adalah untuk mengumpulkan data yang valid dan dapat diandalkan tanpa harus melakukan penelitian pada seluruh populasi.



Gambar 7. Teknik Pengambil Sampel

1. *Probability Sampling*

Probability sampling adalah teknik pengambilan sampel yang memberikan peluang yang sama bagi setiap unsur (anggota) populasi untuk dipilih menjadi anggota sampel. Teknik ini merupakan teknik yang memungkinkan peneliti atau evaluator untuk membuat generalisasi dari karakteristik sampel menjadi karakteristik populasi.

a. *Simple Random Sampling*

Pengambilan sampel acak sederhana (simple random sampling) adalah suatu metode statistik yang menyatakan bahwa setiap orang dalam suatu populasi mempunyai peluang yang sama untuk dipilih menjadi sampel. Sampel mewakili sebagian kecil dan lebih mudah dikelola dari orang-orang yang dapat dipelajari dan dianalisis. Ini adalah teknik mendasar untuk mengumpulkan data dan membuat kesimpulan tentang suatu populasi. Cara melakuan sampel acak sederhana dapat dilihat pada langkah-langkah berikut:

Langkah 1: Buat Daftar

Untuk memulai pengambilan sampel acak sederhana, pertama-tama, buatlah daftar lengkap seluruh 500 karyawan dalam organisasi. Daftar tersebut harus mencantumkan nama setiap karyawan untuk menjamin bahwa setiap orang dipertimbangkan. Daftar yang tepat dan menyeluruh sangat penting untuk memastikan pengambilan sampel mencerminkan seluruh populasi secara akurat.

Daftar lengkap karyawan (ilustrasi)

No	Nama Karyawan
1	Andi Prasetyo
2	Budi Santoso
3	Citra Dewi
4	Dedi Supriyadi
5	Eko Setiawan

6	Fani Nurul
7	Gita Rahmawati
8	Hendra Wijaya
9	Indah Permatasari
10	Joko Santoso
...	...
500	Zainal Abidin

Langkah 2: Tetapkan Nomor Urut

Setelah membuat daftar karyawan, selanjutnya yang dilakukan adalah memberi nomor pada setiap karyawan secara berurutan. Ini adalah kerangka pengambilan sampel Anda (daftar tempat Anda mengambil sampel). Penomoran ini membantu mengatur daftar, membuat identifikasi setiap orang dalam kelompok menjadi lebih mudah. Setiap pegawai harus mempunyai nomornya masing-masing, mulai dari 1 sampai dengan n (1-500), yang merupakan jumlah seluruh pegawai dalam organisasi.

Langkah 3: Pilih Ukuran Sampel

Memilih ukuran sampel yang tepat penting dalam pengambilan sampel acak sederhana. Dalam situasi ini, kami telah memilih sampel sebanyak 100 karyawan dari total populasi 500 orang. Penting untuk memilih ukuran sampel yang cukup besar untuk mendapatkan hasil yang dapat diandalkan namun tetap praktis untuk dianalisis.

Langkah 4: Gunakan Penghasil Angka Acak

Untuk memilih sampel dari grup, gunakan generator nomor acak. Pertama, cari jumlah total orang (Langkah 2) dan putuskan berapa banyak yang kita inginkan dalam sampel kita (Langkah 3). Kemudian, gunakan tabel atau generator angka acak untuk membuat 100 angka acak berbeda antara 1 dan 500. Angka-angka ini cocok dengan urutan yang diberikan kepada setiap karyawan, yang membantu Anda memilih siapa yang akan dijadikan sampel.

5, 12, 23, 34, 45, 56, 67, 78, 89, 90,
101, 112, 123, 134, 145, 156, 167, 178, 189, 190,
201, 212, 223, 234, 245, 256, 267, 278, 289, 290,
301, 312, 323, 334, 345, 356, 367, 378, 389, 390,
401, 412, 423, 434, 445, 456, 467, 478, 489, 490,
501, 512, 523, 534, 545, 556, 567, 578, 589, 590,
601, 612, 623, 634, 645, 656, 667, 678, 689, 690,
701, 712, 723, 734, 745, 756, 767, 778, 789, 790,
801, 812, 823, 834, 845, 856, 867, 878, 889, 890,
901, 912, 923, 934, 945, 956, 967, 978, 989, 990

Catatan: Angka-angka di atas hanya contoh dan harus berada dalam rentang 1 hingga 500.

Hasil Sampel

Setelah mencocokkan angka acak dengan nomor urut karyawan, kita mendapatkan 100 karyawan yang terpilih untuk survei. Misalnya, jika angka acak yang dihasilkan adalah 5, 12, 23, dan seterusnya, maka karyawan yang terpilih adalah:

No	Nama Karyawan
5	Eko Setiawan
12	Fani Nurul
23	Gita Rahmawati
...	...
100	Zainal Abidin

Metode ini memastikan bahwa setiap karyawan mempunyai kesempatan yang sama dalam memilih, menjaga keadilan dan ketidakberpihakan dalam pemilihan sampel.

b. *Systematic Sampling*

Systematic Sampling atau pengambilan sampel secara sistematis adalah salah satu teknik pengambilan sampel probabilitas di mana elemen-elemen dari populasi dipilih pada interval tertentu. Dalam metode ini, pemilihan elemen sampel dilakukan dengan menentukan titik awal secara acak, kemudian mengambil elemen berikutnya berdasarkan interval tetap yang disebut dengan *sampling interval*. Teknik ini sering digunakan karena mudah diterapkan dan efisien, terutama jika populasi memiliki struktur yang teratur.

Misalkan sebuah perusahaan memiliki daftar 500 karyawan, dan peneliti ingin mengambil sampel sebanyak 50 orang. Langkah-langkahnya sebagai berikut:

Langkah 1: Identifikasi Populasi Target

Tentukan populasi yang akan diteliti. Misalnya, populasi adalah seluruh karyawan dalam perusahaan.

Langkah 2: Tentukan ukuran sampel (n)

Tentukan jumlah elemen yang ingin diambil sebagai sampel dari populasi. Misalnya, dari populasi sebanyak 500 orang, ingin diambil sampel sebanyak 50 orang.

Langkah 3: Hitung Sampling Interval (k)

Sampling interval dihitung dengan membagi jumlah elemen dalam populasi (N) dengan ukuran sampel (n):

$$k = \frac{N}{n}$$

Misalnya, jika populasi 500 dan sampel yang diinginkan 50, maka intervalnya k=10.

Langkah 4: Pilih Titik Awal secara Acak

Pilih elemen pertama dari populasi secara acak antara 1 hingga k. Sebagai contoh, jika k=10, titik awal bisa dipilih secara acak, misalnya angka 7

Elemen yang dipilih: ke-7, ke-17, ke-27, ke-37, dan seterusnya hingga ke-497.

Dengan demikian, sampel terdiri dari 50 karyawan yang terpilih berdasarkan interval sistematis.

c. *Stratified Sampling*

Stratified Sampling adalah metode pengambilan sampel dalam statistik di mana populasi dibagi menjadi beberapa kelompok yang homogen, yang disebut **strata**, berdasarkan karakteristik tertentu. Setelah populasi dibagi menjadi strata, sampel diambil dari setiap strata secara acak atau proporsional terhadap ukuran stratum. Teknik ini bertujuan untuk memastikan bahwa semua kelompok dalam populasi diwakili dalam sampel.

Misalkan, sebuah sekolah memiliki 1.000 siswa yang terdiri dari:

- **Kelas X**: 400 siswa,
- **Kelas XI**: 300 siswa,
- **Kelas XII**: 300 siswa.

Langkah 1: Identifikasi Populasi Target

Peneliti ingin mengambil sampel sebanyak 100 siswa.

Langkah 2: Tentukan Variabel untuk Stratifikasi

Peneliti membagi populasi berdasarkan **kelompok kelas** (X, XI, XII) sebagai variabel stratifikasi

Langkah 3: Bagi Populasi ke dalam Strata

Kelas X: 400 siswa,

Kelas XI: 300 siswa,

Kelas XII: 300 siswa.

Langkah 4: Tentukan Ukuran Sampel

Peneliti memutuskan untuk mengambil **100 orang** sebagai sampel.

Langkah 5: Hitung Proporsi Sampel dari Setiap Stratum

Gunakan rumus proporsi untuk menghitung jumlah sampel dari masing-masing strata:

$$Sampel\ Stratum\ = \frac{Jumlah\ Elemen\ Stratum}{Jumlah\ elemen\ Populasi} \times n$$

Kelas X = $\left(\frac{400}{1000}\right) \times 100\ =\ 40$

Kelas XI = $\left(\frac{300}{1000}\right) \times 100\ =\ 30$

Kelas XII = $\left(\frac{300}{1000}\right) \times 100\ =\ 30$

Langkah 6: Hitung Proporsi Sampel dari Setiap Stratum

Setelah menentukan jumlah sampel dari setiap strata, peneliti memilih elemen sampel secara acak. Misalnya:

- Kelas X: Pilih secara acak 40 orang dari Kelas X.
- Kelas XI: Pilih secara acak 30 orang dari Kelas XI
- Kelas XII: Pilih secara acak 30 orang dari Kelas XII

Hasil Akhir

Total sampel = **40 (Kelas X) + 30 (Kelas XI) + 30 (Kelas XII) = 100 orang**.

Setiap kelas terwakili sesuai proporsi dalam populasi.

d. *Cluster Sampling*

Cluster Sampling adalah teknik pengambilan sampel di mana populasi dibagi menjadi kelompok-kelompok yang disebut kluster. Setiap kluster biasanya terdiri dari individu-individu yang memiliki karakteristik serupa atau berada dalam lokasi geografis yang sama. Dalam metode ini, beberapa kluster dipilih secara acak, dan semua anggota dalam kluster yang terpilih menjadi bagian dari sampel. Teknik ini sering digunakan ketika populasi tersebar secara luas, sehingga lebih praktis untuk mengambil sampel dari kelompok daripada individu secara acak.

Sebuah peneliti ingin mengetahui tingkat kepuasan siswa terhadap fasilitas di sekolah-sekolah di kota X. Terdapat 10 sekolah di kota tersebut, dan peneliti memutuskan untuk menggunakan teknik cluster sampling.

Langkah 1: Identifikasi Populasi

Populasi adalah semua siswa di 10 sekolah di kota X.

Langkah 2: Pembagian ke dalam Kluster

Setiap sekolah dianggap sebagai kluster. Jadi, kita memiliki 10 kluster (Sekolah A, Sekolah B, Sekolah C, dst.).

Langkah 3: Pilih kluster secara acak

Peneliti menggunakan generator angka acak dan memilih 3 sekolah secara acak dari 10 sekolah. Misalnya, sekolah yang terpilih adalah:

Sekolah A

Sekolah C

Sekolah E

Langkah 4: Ambil Sampel dari Kluster yang Dipilih

Peneliti kemudian mengambil semua siswa dari Sekolah A, Sekolah C, dan Sekolah E sebagai sampel. Misalnya, jika Sekolah A memiliki 200 siswa, Sekolah C memiliki 150 siswa, dan Sekolah E memiliki 180 siswa, maka total sampel adalah 530 siswa.

Analisis Data

Peneliti mengumpulkan data tentang kepuasan siswa terhadap fasilitas di ketiga sekolah tersebut dan menganalisis hasilnya untuk menarik kesimpulan tentang tingkat kepuasan siswa di seluruh kota X.

e. Multistage Sampling

Multistage Sampling adalah teknik pengambilan sampel yang menggabungkan beberapa metode pengambilan sampel. Dalam teknik ini, populasi dibagi menjadi beberapa kelompok atau kluster, dan kemudian sampel diambil secara bertahap melalui beberapa tahap. Metode ini sering digunakan ketika populasi sangat besar dan tersebar luas, sehingga lebih praktis untuk mengambil sampel dari kelompok-kelompok yang lebih kecil daripada dari individu secara langsung.

Sebuah peneliti ingin mengetahui tingkat pendidikan masyarakat di suatu negara yang memiliki 34 provinsi. Peneliti memutuskan untuk menggunakan teknik multistage sampling.

Langkah 1: Identifikasi Populasi

Populasi adalah semua warga negara di 34 provinsi.

Langkah 2: Pembagian ke dalam Kluster

Setiap provinsi dianggap sebagai kluster. Jadi, kita memiliki 34 kluster.

Langkah 3: Pilih kluster secara acak

Peneliti menggunakan generator angka acak dan memilih 5 provinsi secara acak dari 34 provinsi. Misalnya, provinsi yang terpilih adalah:

Provinsi A

Provinsi B

Provinsi C

Provinsi D

Provinsi E

Langkah 4: Pembagian Kluster yang Dipilih

Setelah memilih provinsi, peneliti membagi setiap provinsi menjadi beberapa kabupaten atau kota.

Langkah 5: Pilih Subkelompok Secara Acak

Setelah memilih provinsi, peneliti membagi setiap provinsi menjadi beberapa kabupaten atau kota.

Dari setiap provinsi yang terpilih, peneliti memilih 2 kabupaten secara acak. Misalnya:

Provinsi A: Kabupaten 1 dan Kabupaten 2

Provinsi B: Kabupaten 3 dan Kabupaten 4

Provinsi C: Kabupaten 5 dan Kabupaten 6

Provinsi D: Kabupaten 7 dan Kabupaten 8

Provinsi E: Kabupaten 9 dan Kabupaten 10

Langkah 6: Ambil Sampel dari Subkelompok yang Dipilih

Dari setiap kabupaten yang terpilih, peneliti mengambil sampel acak dari 100 rumah tangga. Dengan demikian, jika ada 10 kabupaten yang terpilih, total sampel adalah 1.000 rumah tangga.

Analisis Data

Peneliti mengumpulkan data tentang tingkat pendidikan dari rumah tangga yang terpilih dan menganalisis hasilnya untuk menarik kesimpulan tentang tingkat pendidikan masyarakat di seluruh provinsi.

2. *Non-Probability Sampling*

Tidak semua elemen populasi memiliki peluang yang sama untuk dipilih. Teknik ini biasanya digunakan jika representasi sempurna dari populasi tidak diperlukan.

a. *Convenience Sampling*

Convenience Sampling adalah teknik pengambilan sampel di mana peneliti memilih individu yang paling mudah diakses atau tersedia untuk dijadikan sampel. Metode ini sering digunakan dalam penelitian awal atau eksploratori ketika peneliti tidak memiliki sumber daya atau waktu yang cukup untuk melakukan pengambilan sampel yang lebih sistematis. Meskipun metode ini mudah dan cepat, *convenience sampling* memiliki kelemahan dalam hal representativitas, karena sampel yang diambil mungkin tidak mencerminkan keseluruhan populasi *Purposive Sampling (Judgmental Sampling)*

Seorang peneliti ingin mengetahui kepuasan mahasiswa terhadap layanan perpustakaan di universitasnya. Karena keterbatasan waktu dan sumber daya, peneliti memutuskan untuk menggunakan *convenience sampling*.

Langkah 1: Identifikasi Populasi

Populasi adalah semua mahasiswa di universitas tersebut

Langkah 2: Tentukan kriteria pemilihan

Peneliti memutuskan untuk memilih mahasiswa yang berada di perpustakaan pada hari tertentu.

Langkah 3: Pilih sampel

Peneliti pergi ke perpustakaan dan memilih mahasiswa yang sedang berada di sana untuk dijadikan sampel. Misalnya, peneliti memilih 30 mahasiswa yang sedang membaca atau belajar di perpustakaan.

Langkah 4: Kumpulkan data

Peneliti memberikan kuesioner kepada 30 mahasiswa tersebut untuk mengukur kepuasan mereka terhadap layanan perpustakaan.

Analisis Data

Setelah mengumpulkan data, peneliti menganalisis hasilnya. Namun, peneliti harus menyadari bahwa hasil survei ini mungkin tidak mencerminkan kepuasan seluruh mahasiswa di universitas, karena hanya melibatkan mahasiswa yang kebetulan berada di perpustakaan pada saat itu.

b. *Quota Sampling*

Quota Sampling adalah teknik pengambilan sampel non-probabilitas di mana peneliti menetapkan kuota untuk setiap subkelompok dalam populasi. Dalam metode ini, peneliti memastikan bahwa setiap subkelompok terwakili dalam sampel sesuai dengan proporsi yang telah ditentukan. Meskipun lebih cepat dan lebih murah dibandingkan dengan metode probabilitas, quota sampling dapat menghasilkan bias karena pemilihan responden tidak dilakukan secara acak.

Sebuah perusahaan ingin melakukan survei kepuasan pelanggan terhadap produk baru mereka. Mereka ingin memastikan bahwa sampel mencakup pelanggan dari berbagai usia dan jenis kelamin.

Langkah 1: Identifikasi Populasi

Pelanggan yang membeli produk baru dalam 6 bulan terakhir.

Langkah 2: Identifikasi subkelompok

Bagi populasi menjadi subkelompok berdasarkan usia (18-25 tahun, 26-35 tahun, 36-45 tahun, dan 46 tahun ke atas) dan jenis kelamin (laki-laki dan perempuan).

Langkah 3: Tentukan Kuota

Peneliti memutuskan untuk mengambil 20 responden dari setiap subkelompok, sehingga total kuota adalah 80 responden.

Langkah 4: Pengambilan sampel

Peneliti memilih 20 laki-laki dan 20 perempuan dari setiap kelompok usia hingga kuota terpenuhi.

Langkah 5: Kumpulkan data

Setelah kuota terpenuhi, peneliti mengumpulkan data mengenai kepuasan pelanggan dari 80 responden tersebut.

c. *Snowball Sampling*

Snowball Sampling adalah teknik pengambilan sampel non-probabilitas yang sering digunakan dalam penelitian sosial, terutama ketika populasi yang diteliti sulit diakses atau tidak terdefinisi dengan jelas. Dalam metode ini, peneliti memulai dengan beberapa individu yang dikenal (responden awal) dan meminta mereka untuk merekomendasikan atau menghubungkan peneliti dengan individu lain yang juga memenuhi kriteria penelitian. Proses ini berlanjut hingga jumlah sampel yang diinginkan tercapai.

Ciri-ciri *Snowball Sampling:*

- ✓ Non-Probabilitas: Pemilihan responden tidak dilakukan secara acak, sehingga hasilnya mungkin tidak dapat digeneralisasi ke seluruh populasi.
- ✓ Rekomendasi: Responden awal merekomendasikan individu lain, yang dapat membantu peneliti menemukan lebih banyak responden.
- ✓ Bermanfaat untuk Populasi Tersembunyi: Sangat berguna untuk penelitian yang melibatkan kelompok-kelompok yang sulit dijangkau, seperti pengguna narkoba, pelacur, atau individu dengan kondisi kesehatan tertentu.
- ✓ Proses Berkelanjutan: Proses pengambilan sampel dapat berlanjut hingga peneliti merasa cukup dengan jumlah data yang diperoleh.

Seorang peneliti ingin melakukan studi tentang pengalaman hidup para penyintas kanker. Karena populasi ini sulit diakses, peneliti memutuskan untuk menggunakan metode snowball sampling.

Langkah 1: Identifikasi populasi

Penyintas kanker yang telah menjalani perawatan dalam 5 tahun terakhir.

Langkah 2: Mulai dengan responden awal

Peneliti mulai dengan menghubungi 3 penyintas kanker yang dikenal.

Langkah 3: Minta rekomendasi

Peneliti meminta ketiga penyintas tersebut untuk merekomendasikan penyintas kanker lainnya.

Langkah 4: Kumpulkan data

Peneliti mengumpulkan data dari individu yang direkomendasikan

Langkah 5: Ulangi proses

Proses ini dilanjutkan hingga peneliti mendapatkan 30 responden.

3. TUJUAN STATISTIK

Bagaimana cara mengamati, sehingga tujuan statistik tercapai?

Statistik hadir untuk menggambarkan bagaimana perilaku, gambaran atau deskripsi dari sebuah fenomena yang diamati. Sebagai contoh sederhana, bagaimana ciri-ciri akan terjadinya hujan di suatu kawasan atau daerah. Statistik adalah alat yang sangat penting dalam memahami dan menggambarkan fenomena yang terjadi di sekitar kita. Dengan menggunakan metode statistik, kita dapat mengumpulkan, menganalisis, dan menginterpretasikan data untuk mendapatkan wawasan yang lebih dalam tentang berbagai aspek kehidupan. Namun, untuk mencapai tujuan statistik yang diinginkan, langkah pertama yang krusial adalah melakukan pengamatan yang tepat.

Pengamatan adalah proses sistematis dalam mengumpulkan informasi tentang objek atau fenomena yang sedang diteliti. Dalam konteks statistik, pengamatan ini dapat dilakukan melalui berbagai cara, seperti survei, eksperimen, atau pengumpulan data sekunder. Misalnya, jika kita ingin memahami ciri-ciri yang berhubungan dengan terjadinya hujan di suatu daerah, kita perlu melakukan pengamatan yang cermat terhadap berbagai faktor yang dapat mempengaruhi cuaca.

Langkah pertama dalam pengamatan adalah menentukan variabel yang relevan. Dalam contoh hujan, variabel yang mungkin perlu diamati meliputi suhu udara, kelembapan, tekanan atmosfer, dan pola angin. Dengan mengidentifikasi variabel-variabel ini, kita dapat merancang metode pengumpulan data yang lebih efektif. Misalnya, kita dapat menggunakan alat pengukur cuaca untuk mendapatkan data suhu dan kelembapan secara akurat.

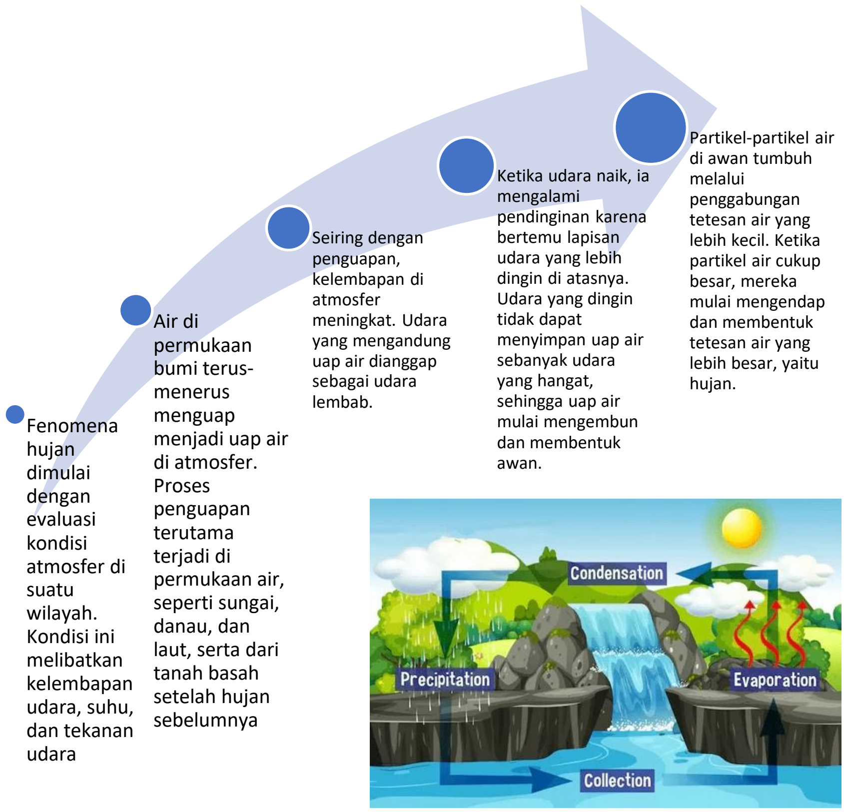
Setelah variabel ditentukan, langkah selanjutnya adalah mengumpulkan data. Pengumpulan data harus dilakukan secara sistematis dan konsisten untuk memastikan keakuratan dan keandalan informasi yang diperoleh. Dalam kasus hujan, kita bisa melakukan pengamatan harian selama periode tertentu untuk mencatat kondisi cuaca dan mencatat kapan hujan terjadi. Data yang terkumpul kemudian dapat dianalisis untuk mencari pola atau hubungan antara variabel yang diamati.

Analisis data adalah tahap penting dalam mencapai tujuan statistik. Dengan menggunakan teknik analisis yang tepat, seperti analisis regresi atau analisis korelasi, kita dapat mengidentifikasi hubungan antara variabel dan membuat prediksi tentang kemungkinan terjadinya hujan di masa depan. Misalnya, jika analisis menunjukkan bahwa peningkatan kelembapan berhubungan erat dengan kemungkinan hujan, kita dapat menggunakan informasi ini untuk memberikan peringatan dini kepada masyarakat.

Selain itu, penting untuk mempertimbangkan faktor eksternal yang dapat mempengaruhi hasil pengamatan. Misalnya, perubahan iklim atau aktivitas manusia dapat memengaruhi pola cuaca di suatu daerah. Oleh karena itu, pengamatan harus dilakukan dengan mempertimbangkan konteks yang lebih luas agar hasil analisis dapat diinterpretasikan dengan benar.

Akhirnya, komunikasi hasil pengamatan dan analisis kepada pemangku kepentingan adalah langkah yang tidak kalah penting. Hasil statistik yang diperoleh harus disajikan dengan cara yang jelas dan mudah dipahami, sehingga dapat digunakan untuk pengambilan keputusan yang lebih baik. Dalam konteks cuaca, informasi tentang kemungkinan hujan dapat membantu petani merencanakan waktu tanam atau membantu masyarakat mempersiapkan diri menghadapi cuaca buruk.

Dengan demikian, pengamatan yang sistematis dan terencana adalah kunci untuk mencapai tujuan statistik. Melalui pengamatan yang cermat, pengumpulan data yang akurat, analisis yang tepat, dan komunikasi yang efektif, kita dapat memahami fenomena yang kompleks dan membuat keputusan yang lebih baik berdasarkan data yang ada.



Gambar 8. Proses terjadinya hujan dalam pengamatan

4. JENIS-JENIS STATISTIK

Statistik dapat dibagi menjadi dua jenis utama yaitu statistik deskriptif dan statistik inferensial. Mari kita lihat perbedaan antara keduanya.

Statistik Deskriptif

Statistik deskriptif adalah cabang ilmu statistik yang berkaitan dengan pengumpulan, penyajian, pengorganisasian. Statistik deskriptif sangat berguna dalam memberikan gambaran awal tentang data dan membantu peneliti atau analis untuk memahami pola, tendensi, dan karakteristik dasar dari dataset tanpa melakukan inferensi terhadap populasi secara keseluruhan. Beberapa teknik statistik deskriptif meliputi:

1) Ukuran Pusat:
   - ✓ Mean (Rata-rata): Nilai tengah dari data.
   - ✓ Median: Nilai tengah saat data diurutkan.
   - ✓ Modus: Nilai yang muncul paling sering.
2) Ukuran Sebaran:
   - ✓ Rentang: Selisih antara nilai maksimum dan minimum.
   - ✓ Varians dan Deviasi Standar: Mengukur sebaran data.
   - ✓ Kuartil dan Persentil: Membagi data menjadi interval.
3) Grafik dan Diagram:
   - ✓ Histogram: Grafik batang yang menunjukkan distribusi frekuensi.
   - ✓ Diagram Batang dan Lingkaran: Menyajikan data secara visual.
   - ✓ Diagram Pencar *(Scatter Plot)*: Menunjukkan hubungan antara dua variabel.

Statistik Inferensial

Statistik inferensial adalah cabang statistik yang bertujuan untuk membuat inferensi atau generalisasi tentang populasi berdasarkan analisis sampel yang diambil dari populasi tersebut. Dalam statistik inferensial, informasi dari sampel digunakan untuk membuat kesimpulan atau memperkirakan sifat-sifat populasi secara lebih luas. Metode ini juga digunakan untuk menentukan apakah perbedaan antara kelompok-kelompok dalam sampel tersebut adalah signifikan atau hanya terjadi secara kebetulan.

Beberapa konsep kunci dalam statistik inferensial melibatkan:

1) Parameter dan Statistik:
   - ✓ Parameter: Sifat-sifat numerik yang menggambarkan populasi secara keseluruhan.
   - ✓ Statistik: Sifat-sifat numerik yang dihitung dari sampel yang digunakan untuk mengestimasi parameter.
2) Estimasi:
   - ✓ Proses menghitung nilai estimasi (pendekatan) dari parameter populasi berdasarkan informasi yang diperoleh dari sampel.
3) Uji Hipotesis:
   - ✓ Proses pengujian klaim atau hipotesis tentang karakteristik populasi berdasarkan data sampel. Tujuannya adalah membuat keputusan apakah klaim tersebut dapat diterima atau ditolak.
4) Interval Kepercayaan:
   - ✓ Menyajikan rentang nilai yang mungkin mengandung parameter populasi dengan tingkat kepercayaan tertentu. Ini memberikan estimasi kisaran yang dapat diandalkan.
5) Analisis Regresi dan Korelasi:
   - ✓ Menganalisis hubungan antara variabel dan memodelkan pola yang mungkin ada dalam populasi berdasarkan data sampel.

Statistik digunakan sebagai alat pengukuran juga sebagai berperan sangat penting dalam pengambilan keputusan yang dapat digunakan dalam kasus kehidupan sehari-hari. Misalnya dalam pengelolaan keuangan, belanja dan konsumsi, pemantauan kesehatan, partisipasi dalam politik, pengelolaan energi, dan lain-lain.

Tugas

1. Apa yang dimaksud dengan statistik dan apa tujuan utamanya?
2. Apa perbedaan antara statistik deskriptif dan statistik inferensial?
3. Jelaskan pengertian populasi dan sampel dalam statistik. Bagaimana hubungan keduanya?
4. Sebutkan tiga ukuran pemusatan data utama dalam statistik!
5. Apa tujuan dari ukuran variabilitas dalam statistik?
6. Apa perbedaan antara parameter dan statistik?
7. Apa yang dimaksud dengan variabel dalam statistik dan apa jenis-jenis utamanya?
8. Apa tujuan utama dari penyajian data dalam statistik?
9. Apa yang dimaksud dengan pengujian hipotesis dan mengapa hal ini penting?
10. Apa yang dimaksud dengan interval kepercayaan dan apa yang diinformasikannya kepada kita?

BAB II

Tujuan Pembelajaran

Setelah mempelajari bab ini, pembaca diharapkan mampu:

1. Memahami konsep statistik deskriptif dalam menganalisis data.
2. Menghitung dan menginterpretasikan ukuran pemusatan data seperti mean, median, dan modus.
3. Menganalisis ukuran variabilitas data, termasuk rentang (range) dan deviasi standar.
4. Menyajikan data secara visual menggunakan grafik dan diagram, seperti histogram, diagram batang, dan pie chart.

Pokok Bahasan

1. **Pengukuran Pusat**
2. **Pengukuran Variabilitas**
3. **Grafik dan Diagram**

STATISTIK DESKRIPTIF

1. PENGUKURAN PUSAT: Mean, Median, Modus

Mean

Mean, atau rata-rata, adalah salah satu ukuran pusat yang paling umum digunakan dalam statistik deskriptif. Mean dihitung dengan menjumlahkan semua nilai dalam suatu set data dan kemudian membaginya dengan jumlah total nilai tersebut. Rata-rata adalah representasi "titik tengah" dari suatu distribusi data.

Rumus matematis untuk mean ($\bar{x}$) dari suatu set data adalah:

$$\bar{x} = \frac{\sum xi}{N}$$

Dimana

$\bar{\mathrm{x}}$: mean atau rata − rata

Σ : Simbol sigma yang menujukkan penjumlahan

xi: setiap nilai dalam data

N : Jumlah total nilai dalam data

Contoh: Misalkan kita memiliki data nilai ujian matematika siswa

Data: 75,80,85,90,95

$$\bar{x} = 75 + 80 + 85 + 90 + 955 = \frac{425}{5} = 85$$

Jadi, rata-rata atau mean dari data tersebut adalah 80. Rata-rata memberikan gambaran tentang nilai tengah atau pusat dari suatu distribusi data.

Median

Median adalah ukuran pusat lainnya yang digunakan dalam statistik deskriptif. Median adalah nilai tengah dari suatu set data yang diurutkan dari yang terkecil hingga yang terbesar (atau sebaliknya). Untuk menemukan median, data harus diurutkan terlebih dahulu, dan nilai yang berada di tengah-tengahnya menjadi median.

Jika jumlah data (n) ganjil, median adalah nilai yang berada tepat di tengah setelah diurutkan. Jika jumlah data genap, median adalah rata-rata dari dua nilai yang berada di tengah setelah diurutkan.

Misalkan kita memiliki data nilai ujian matematika siswa: 70,75,80,85,90

Karena jumlah data ganjil (5 data), median adalah nilai yang berada di tengah, yaitu 80.

Jika kita memiliki data nilai ujian yang lebih banyak: 70,75,80,85,90,95

Karena jumlah data genap (6 data), median adalah rata-rata dari dua nilai tengah, yaitu $\frac{80+85}{2} = 82,5$

Median memiliki keunggulan dibandingkan dengan mean (rata-rata) ketika data mengandung pencilan atau outlier, karena median tidak dipengaruhi oleh nilai ekstrem. Median memberikan gambaran tentang nilai tengah yang lebih stabil dalam distribusi data.

Modus

Modus adalah ukuran kecenderungan sentral dalam statistik deskriptif yang mewakili nilai atau kategori yang paling sering muncul dalam suatu set data. Modus dapat diterapkan pada data kualitatif (kategori) dan data kuantitatif (nilai numerik).

Untuk data kualitatif, modus adalah nilai atau kategori yang memiliki frekuensi tertinggi, artinya muncul paling sering dalam data. Untuk data kuantitatif, modus adalah nilai numerik yang memiliki frekuensi tertinggi.

Data Kualitatif:

Misalkan kita memiliki data jenis buah yang dibeli di sebuah toko buah:

Data: Apel, Jeruk, Apel, Pisang, Apel, Jeruk, Pisang

Modusnya adalah "Apel" karena Apel muncul paling sering dalam data tersebut.

Data Kuantitatif

Misalkan kita memiliki data nilai ujian matematika siswa: 70,75,80,85,90,80,75,80

Dalam contoh ini, modus adalah "80," karena nilai 80 muncul lebih sering daripada nilai lainnya.

Modus berguna dalam memberikan gambaran tentang nilai yang paling umum atau karakteristik yang paling mendominasi dalam suatu set data. Meskipun modus adalah ukuran kecenderungan sentral yang sederhana, itu dapat memberikan informasi yang bermanfaat terutama ketika kita tertarik pada nilai yang sering muncul.

2. PENGUKURAN VARIABILITAS: Rentang dan Deviasi Standar

Rentang *(Range)*

Rentang *(range)* adalah ukuran penyebaran atau sebaran dalam statistik yang mengukur perbedaan antara nilai maksimum dan nilai minimum dalam suatu set data. Rentang memberikan gambaran tentang sejauh mana data tersebar atau berkisar. Rumus matematis untuk menghitung rentang adalah:

$$Rentang = Nilai\ Maksimum - Nilai\ Minimum$$

Dengan kata lain, rentang adalah selisih antara nilai tertinggi (maksimum) dan nilai terendah (minimum) dalam suatu set data.

Contoh:

Misalkan kita memiliki data tinggi siswa dalam sentimeter

Data: 160,165,170,175,180

$$Rentang = 180 - 160 = 20$$

Jadi, rentang tinggi siswa adalah 20 sentimeter.

Rentang memberikan informasi tentang sebaran keseluruhan data, tetapi perhatikan bahwa rentang dapat dipengaruhi oleh nilai ekstrem *(outliers)* dalam set data. Oleh karena itu, sementara rentang memberikan gambaran umum, statistik deskriptif lainnya seperti deviasi standar lebih sensitif terhadap perbedaan dalam sebaran data.

Deviasi Standar (Standard Deviation)

Deviasi standar *(standard deviation)* adalah ukuran sebaran atau dispersi yang mengukur sejauh mana nilai-nilai dalam suatu set data menyebar dari nilai rata-rata *(mean)* dari set data tersebut. Deviasi standar memberikan gambaran tentang seberapa homogen atau heterogen distribusi data.

$$Deviasi\ Standar\ (\sigma) = \sqrt{\frac{\sum(xi - \bar{x})^2}{N}}$$

Dimana:

σ = Deviasi standar

xi = setiap nilai dalam data

$\bar{x}$ = mean atau rata-rata

N = jumlah total nilai dalam data

Perlu untuk diketahui, dalam rumus deviasi standar ada dua versi. Pertama, hanya dengan menggunakan N, ini artinya bahwa total nilai data tersebut adalah untuk seluruh populasi.

Kedua, (N-1), artinya jumlah total nilainya adalah data sampel, ada yang menyebutnya sebagai "penyesuaian bessel".

Contoh:

Misalnya kita mempunyai nilai ujian statistik untuk 10 siswa, untuk dicari berapa deviasi standar pada seluruh nilai itu.

No	**Nilai Ujian**
1	80
2	85
3	90
4	98
5	88
6	82
7	90
8	85
9	81

10	89

Langkah pertama, yaitu mencari nilai *mean* dari nilai-nilai tersebut

$$\bar{x} = \frac{80 + 85 + 90 + 98 + 88 + 82 + 90 + 85 + 81 + 89}{10} = 86{,}8$$

Langkah kedua, hitung selisih antara setiap nilai dan *mean*, kemudian kuadratkan selisihnya:

No	Nilai Ujian	(Nilai Ujian-*mean*)^2
1	80	46,2
2	85	3,2
3	90	10,2
4	98	125,4
5	88	1,4
6	82	23,0
7	90	10,2
8	85	3,2
9	81	33,6
10	89	4,8
mean	86,8	26,2

Ketiga, hitung rata-rata dari nilai-nilai yang telah di kuadratkan:

$$\frac{\sum(xi-\bar{x})^2}{N}$$

$$=\frac{46{,}2+3{,}2+10{,}2+125{,}4+1{,}4+23+10{,}2+3{,}2+33{,}6+4{,}8}{10}$$

$$=26{,}2$$

Terakhir, ambil akar kuadrat dari rata-rata nilai-nilai yang telah di kuadratkan:

$$\sqrt{26{,}2}=5{,}1$$

Jadi, deviasi standar dari data tersebut adalah sekitar 5,1. Semakin besar deviasi standar, semakin besar sebaran atau dispersi data. Deviasi standar memberikan informasi tentang sejauh mana nilai-nilai dalam data tersebar dari nilai tengah (mean).

Jika nilai deviasi standar jauh lebih besar dibandingkan nilai *mean*, maka nilai *mean* merupakan representasi yang buruk dari keseluruhan data. Sedangkan jika nilai deviasi standar sangat kecil dibandingkan nilai *mean*, maka nilai *mean* dapat digunakan sebagai representasi dari keseluruhan data.

Sedangkan varians adalah nilai-nilai yang telah dikuadratkan tadi yaitu 26,2. Varians memberikan gambaran tentang sebaran atau variasi data dari *mean*. Namun, karena varians memiliki satuan pengukuran yang berbeda dan sulit diinterpretasikan, deviasi standar (akar kuadrat dari varian) lebih umum digunakan.

3. Grafik dan Diagram: Histogram, Diagram Batang dan Pie Chart

Grafik dan diagram adalah visualisasi data yang digunakan untuk menyajikan informasi secara grafis. Mereka membantu kita dalam memahami pola, tren, dan distribusi data dengan cara yang lebih intuitif dan mudah dipahami daripada data dalam bentuk tabel atau daftar angka.

Histogram

Histogram adalah jenis grafik batang yang digunakan untuk menyajikan distribusi frekuensi dari suatu set data kontinu. Histogram membagi rentang nilai data ke dalam interval atau kelas dan menunjukkan seberapa sering nilai-nilai berada dalam setiap interval tersebut.

Sebagai contoh, kita memiliki nilai ujian sebagai berikut:

No	Nilai Ujian
1	80
2	85
3	90
4	98
5	88
6	82
7	90
8	85
9	81
10	89

Langkah-langkah dalam membuat histogram sebagai berikut:

1) Tentukan Rentang Nilai:
   Identifikasi rentang nilai data dan pilih interval atau kelas.
2) Buat Tabel Frekuensi:

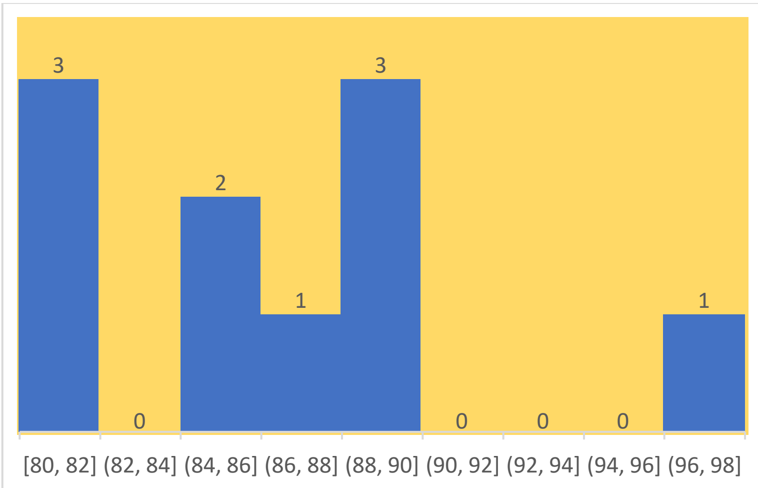
Hitung frekuensi atau jumlah kemunculan nilai-nilai dalam setiap interval.

3) Gambar Batang:

Gambarkan batang untuk setiap interval pada sumbu horizontal, dan tinggi batang sesuai dengan frekuensinya pada sumbu vertikal.

4) Berikan Label:

Sertakan label pada sumbu-sumbu dan berikan judul pada histogram untuk memberikan konteks.



Gambar 9. Histogram Grafik Nilai Ujian

Histogram diatas menujukkan sebaran nilai yang bervariasi menurut nilai-nilai apa saja yang muncul. Sebagai contoh rentang nilai 80-82 memiliki 3 nilai didalamnya. Yaitu 80, 81, dan 82.

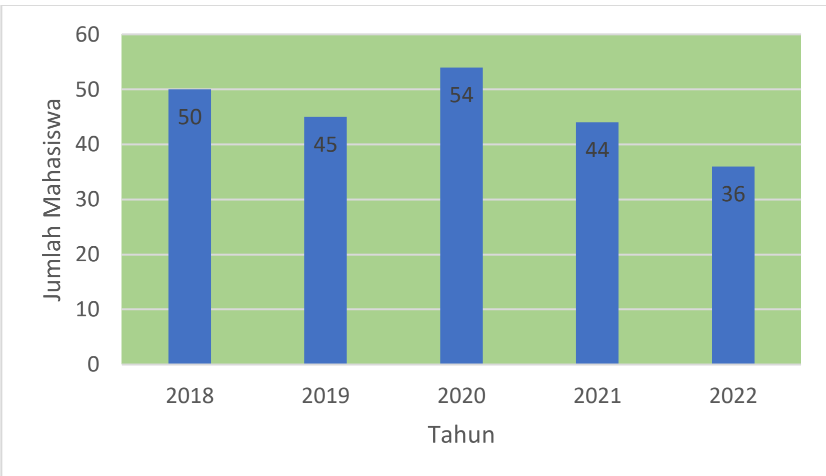
Histogram memberikan gambaran visual tentang distribusi data, memungkinkan kita melihat pola, bentuk kurva, dan pusat kerumunan data. Histogram sering digunakan dalam statistik untuk menganalisis distribusi frekuensi dari data kontinu.

Diagram Batang

Diagram batang adalah jenis visualisasi data yang menggunakan batang vertikal atau horizontal untuk mewakili frekuensi atau kuantitas dari kategori atau nilai-nilai tertentu. Diagram batang membantu membandingkan kuantitas antar kategori atau menunjukkan distribusi data dengan cara yang mudah dipahami.

Sebagai contoh, bagaimana membuat diagram batang untuk jumlah mahasiswa:

No	Tahun	Jumlah
1	2018	50
2	2019	45
3	2020	54
4	2021	44
5	2022	36



Gambar 10. Diagram Batang Mahasiswa Teknik Industri 2018-2022

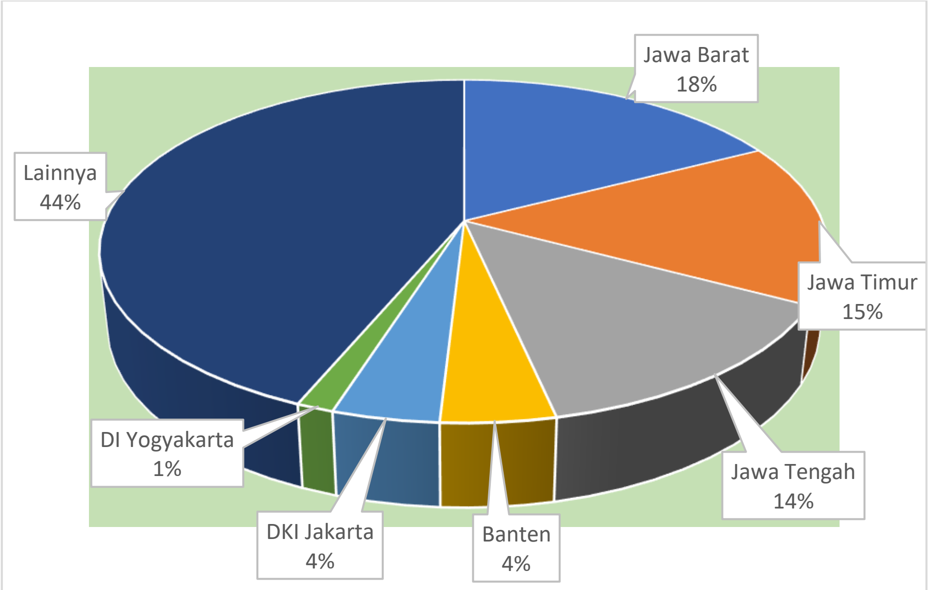
Diagram batang memberikan representasi visual yang jelas dan sederhana dari distribusi atau perbandingan data dalam bentuk batang-batang.

Pie Chart

Pie chart atau diagram lingkaran adalah jenis grafik yang menggambarkan data dalam bentuk lingkaran, dibagi menjadi beberapa bagian yang mewakili proporsi relatif dari keseluruhan. Setiap bagian dari lingkaran disebut sebagai "sektor" atau "slice," dan besar sudut setiap sektor sesuai dengan proporsi data yang diwakilinya.

Misalkan kita memiliki data pemilih tetap pemilu 2024 di pulau Jawa, kemudian data tersebut kita membuat pie chart yang didalamnya memuat informasi persentase daftar pemilih di Pulau Jawa.

No	Provinsi	Jumlah Pemilih	Persentase
1	Jawa Barat	35.710.000	17%
2	Jawa Timur	31.400.000	15%
3	Jawa Tengah	28.290.000	14%
4	Banten	8.840.000	4%
5	DKI Jakarta	8.250.000	4%
6	DI Yogyakarta	2.880.000	1%
7	Lainnya	89.437.222	44%
8	Total DPT Seluruh Indonesia	204.807.222	100%



Gambar 11. Persentase Daftar Pemilih Tetap Pemilu Pulau Jawa 2024

Kuartil dan Persentil: Membagi data menjadi interval dan menyajikan posisi nilai tertentu.

Kuartil dan persentil adalah konsep dalam statistik deskriptif yang digunakan untuk membagi data ke dalam bagian-bagian tertentu. Kuartil membagi data menjadi empat bagian, sedangkan persentil membaginya menjadi seratus bagian.

Contoh:

Kita memiliki data tinggi siswa dalam sentimeter: 160, 165, 170, 175, 180, 185, 190, 195, 200

Kuartil pada data tunggal:

Kuartil Pertama (Q1):

$$Q1 = \frac{1}{4}(N + 1)$$

$$Q1 = \frac{1}{4}(10) = 2,5$$

$$Q1 = \frac{165 + 170}{2} = 167,5$$

Kuartil Kedua (Q2):

$$Q2 = \frac{N + 1}{2}$$

$$Q2 = \frac{10}{2} = 5$$

$$Q2 = 180$$

Kuartil Ketiga (Q3):

$$Q3 = \frac{3}{4}(N + 1)$$

$$Q3 = \frac{3}{4}(10) = 7,5$$

$$Q3 = \frac{190 + 195}{2} = 192,5$$

Persentil pada data tunggal:

Persentil 20 (P20)

Persentil 20 adalah nilai di mana 20% data berada di bawahnya. Untuk menghitung persentil 20:

$$P20 = \frac{20}{100} * 9 = 1{,}8$$

$$P20 = 165$$

Persentil 55 (P55)

Persentil 55 adalah nilai di mana 55% data berada di bawahnya. Untuk menghitung persentil 55:

$$P55 = \frac{55}{100} * 9 = 4{,}95$$

$$P55 = 180$$

Tugas

1. Pengukuran Pusat (Mean) Data nilai ujian matematika dari 10 siswa adalah sebagai berikut: 75, 80, 85, 90, 75, 80, 85, 90, 95, 75. Hitunglah rata-rata (mean) dari nilai ujian tersebut.
2. Pengukuran Pusat (Median) Urutkan data nilai ujian matematika berikut dan tentukan mediannya: 75, 80, 85, 90, 75, 80, 85, 90, 95, 75.
3. Pengukuran Pusat (Modus) Tentukan modus dari data nilai ujian matematika berikut: 75, 80, 85, 90, 75, 80, 85, 90, 95, 75.
4. Pengukuran Variabilitas (Rentang) Hitunglah rentang dari data nilai ujian matematika berikut: 75, 80, 85, 90, 75, 80, 85, 90, 95, 75
5. Pengukuran Variabilitas (Varians) Hitunglah varians dari data nilai ujian matematika berikut: 75, 80, 85, 90, 75, 80, 85, 90, 95, 75
6. Pengukuran Variabilitas (Standar Deviasi) Hitunglah standar deviasi dari data nilai ujian matematika berikut: 75, 80, 85, 90, 75, 80, 85, 90, 95, 75.
7. Grafik (Histogram) Buatlah histogram untuk data nilai ujian matematika berikut: 75, 80, 85, 90, 75, 80, 85, 90, 95, 75.
8. Grafik (Diagram Lingkaran) Data pekerjaan penduduk sebuah desa adalah sebagai berikut: Petani: 30%, Pedagang: 25%, PNS: 15%, Wiraswasta: 30% Buatlah diagram lingkaran untuk data tersebut.
9. Grafik (Diagram Batang) Data penjualan bulanan sebuah toko buku adalah sebagai berikut: Januari: 120, Februari: 150, Maret: 140, April: 180, Mei: 200 Buatlah diagram batang untuk data tersebut.
10. Interpretasi Data Berdasarkan data nilai ujian matematika (75, 80, 85, 90, 75, 80, 85, 90, 95, 75), interpretasikan hasil perhitungan statistik berikut: Mean: 83 Median: 82,5 Standar Deviasi: 7,15

BAB III

Tujuan Pembelajaran

Setelah mempelajari bab ini, pembaca diharapkan mampu:

1. Memahami konsep dasar statistik inferensial dan perbedaannya dengan statistik deskriptif.
2. Menjelaskan dan menerapkan teknik estimasi parameter dalam analisis data.
3. Melakukan uji hipotesis untuk menguji kebenaran suatu pernyataan berdasarkan data sampel.
4. Menghitung dan menginterpretasikan interval kepercayaan (confidence interval).
5. Membedakan serta memahami penerapan statistik parametrik dan non-parametrik.

Pokok Bahasan

1. Estimasi Parameter
2. Uji Hipotesis
3. Confidence Interval (Interval Kepercayaan)
4. Statistik Parametrik dan Non-Parametrik

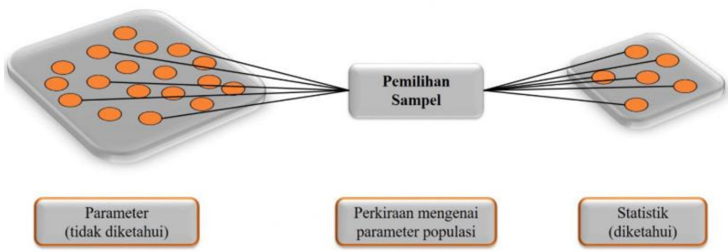
STATISTIK INFERENSIAL

Inferensi statistik adalah proses mengambil kesimpulan atau keputusan mengenai populasi berdasarkan sampel yang diambil dari populasi tersebut. Tujuan utama dari inferensi statistik adalah untuk membuat estimasi dan menguji hipotesis tentang parameter populasi menggunakan data sampel.

Ada dua jenis utama inferensi statistik:

Estimasi Parameter

Mengestimasi parameter populasi adalah salah satu aspek fundamental dalam statistik inferens. Ketika peneliti ingin memahami karakteristik suatu populasi, mereka sering kali tidak dapat mengumpulkan data dari seluruh populasi tersebut. Sebagai gantinya, mereka mengambil sampel yang representatif dari populasi untuk melakukan analisis. Dengan menggunakan data sampel ini, peneliti dapat memperkirakan nilai parameter populasi, seperti rata-rata, proporsi, atau varians.



Gambar 12. Estimasi Parameter

Rata-rata Populasi

Salah satu parameter yang paling umum diestimasi adalah rata-rata populasi (μ). Misalnya, jika seorang peneliti ingin mengetahui rata-rata tinggi badan mahasiswa di sebuah universitas, ia mungkin tidak dapat mengukur tinggi badan semua mahasiswa. Sebagai gantinya, ia akan mengambil sampel acak dari mahasiswa dan menghitung rata-rata tinggi badan dalam sampel tersebut ($\bar{X}$). Dengan menggunakan rata-rata sampel ini, peneliti dapat memperkirakan rata-rata tinggi badan populasi dengan interval kepercayaan tertentu. Interval kepercayaan memberikan rentang nilai di mana parameter populasi kemungkinan besar berada, berdasarkan data sampel yang diperoleh.

Proporsi Populasi

Selain rata-rata, peneliti juga sering kali tertarik untuk mengestimasi proporsi populasi (p). Misalnya, jika peneliti ingin mengetahui proporsi mahasiswa yang menggunakan transportasi umum untuk pergi ke kampus, ia dapat melakukan survei terhadap sejumlah mahasiswa dan mencatat berapa banyak dari mereka yang menggunakan transportasi umum. Dari data ini, peneliti dapat menghitung proporsi dalam sampel ($\hat{p}$) dan menggunakannya untuk memperkirakan proporsi populasi dengan interval kepercayaan.

Varians Populasi

Varians populasi (σ^2) adalah parameter lain yang penting untuk diestimasi, terutama dalam konteks analisis variabilitas data. Varians memberikan informasi tentang seberapa jauh data menyebar dari rata-ratanya. Dengan menghitung varians sampel (s^2) dari data yang diperoleh, peneliti dapat memperkirakan varians populasi dan memahami tingkat ketidakpastian dalam data yang dianalisis.

Uji Hipotesis

Uji hipotesis adalah metode statistik yang digunakan untuk menguji pernyataan atau klaim tentang parameter populasi. Proses ini melibatkan dua hipotesis yang saling bertentangan: hipotesis nol (H_0) dan hipotesis alternatif (H_1). Hipotesis nol biasanya menyatakan bahwa tidak ada efek atau perbedaan yang signifikan, sedangkan hipotesis alternatif menyatakan bahwa ada efek atau perbedaan yang signifikan.

Langkah-langkah Uji Hipotesis

1. **Menentukan Hipotesis**: Peneliti harus merumuskan hipotesis nol dan hipotesis alternatif. Misalnya, jika peneliti ingin menguji apakah rata-rata waktu belajar mahasiswa berbeda dari 15 jam, maka H_0: $\mu = 15$ dan H_1: $\mu \neq 15$.

2. **Menentukan Tingkat Signifikansi**: Tingkat signifikansi (α) adalah probabilitas untuk menolak hipotesis nol ketika sebenarnya benar. Umumnya, nilai α yang digunakan adalah 0,05 atau 0,01. Ini menunjukkan bahwa peneliti bersedia menerima risiko 5% atau 1% untuk membuat kesalahan tipe I (menolak H_0 yang benar).

3. **Mengumpulkan Data dan Menghitung Statistik Uji**: Peneliti mengumpulkan data dari sampel dan menghitung statistik uji, seperti nilai t atau z, tergantung pada jenis uji yang dilakukan. Statistik uji ini digunakan untuk menentukan seberapa jauh data sampel mendukung hipotesis alternatif.

4. **Menentukan Nilai Kritis atau Nilai p**: Peneliti kemudian membandingkan statistik uji dengan nilai kritis dari distribusi yang sesuai (misalnya, distribusi t atau normal) untuk menentukan

apakah hasilnya signifikan. Alternatifnya, peneliti dapat menghitung nilai p, yang menunjukkan probabilitas untuk mendapatkan hasil yang sama atau lebih ekstrem jika hipotesis nol benar.

5. **Mengambil Keputusan**: Berdasarkan perbandingan antara statistik uji dan nilai kritis, atau nilai p dan α, peneliti akan memutuskan untuk menolak atau gagal menolak hipotesis nol. Jika hasilnya signifikan, peneliti dapat menyimpulkan bahwa ada cukup bukti untuk mendukung hipotesis alternatif.

Kesalahan Tipe I dan Tipe II

Dalam uji hipotesis, kita dapat membuat dua jenis kesalahan:

- **Kesalahan Tipe I (α)**: Terjadi ketika kita menolak hipotesis nol yang sebenarnya benar *(false positive)*. Contoh: Menganggap obat baru efektif padahal sebenarnya tidak.
- **Kesalahan Tipe II (β)**: Terjadi ketika kita gagal menolak hipotesis nol yang sebenarnya salah *(false negative)*. Contoh: Menganggap obat baru tidak efektif padahal sebenarnya efektif.

Hubungan antara α dan β:

- **Tingkat Signifikansi (α)**: Menentukan seberapa besar kemungkinan kita melakukan kesalahan tipe I.
- **Kekuatan Uji (1 - β)**: Probabilitas untuk mendeteksi perbedaan yang sebenarnya ada, dan berhubungan dengan ukuran sampel. Semakin besar ukuran sampel, semakin kecil kemungkinan kesalahan tipe II.

Contoh:

Misalkan kita menguji apakah obat baru efektif, dengan hipotesis:

- H_0: Obat tidak efektif.
- H_1: Obat efektif.

Jika kita menetapkan α = 0,05, artinya ada 5% peluang untuk membuat kesalahan tipe I (menyatakan obat efektif padahal tidak). Jika obat memang efektif tetapi kita gagal mendeteksinya, itu akan menjadi kesalahan tipe II (β).

Mengurangi Kesalahan:

- Mengurangi α (misalnya dengan menetapkan α lebih rendah, seperti 0,01) akan mengurangi kesalahan tipe I, tetapi dapat meningkatkan kesalahan tipe II.
- Meningkatkan ukuran sampel dapat meningkatkan kekuatan uji dan mengurangi kesalahan tipe II.

Confidence Interval (Interval Kepercayaan)

Confidence Interval (CI) adalah rentang nilai yang digunakan untuk memperkirakan parameter populasi, dengan tingkat keyakinan tertentu. CI memberikan gambaran seberapa pasti kita bisa memperkirakan suatu parameter berdasarkan data sampel yang tersedia.

Rumusan Umum CI:

$$CI = \bar{x} \pm Z_{\alpha/2} \times \frac{\sigma}{\sqrt{n}}$$

Dimana:

$\bar{x}$ = rata-rata sampel

$Z_{\alpha/2}$ = nilai z yang sesuai dengan tingkat kepercayaan (misalnya 1,96 untuk 95% CI)

σ = deviasi standar populasi (jika diketahui) atau estimasi deviasi standar sampel

n = ukuran sampel

Misalnya, jika kita memiliki interval kepercayaan 95% untuk rata-rata populasi, kita bisa mengatakan bahwa 95% dari sampel-sampel yang diambil dari populasi yang sama akan menghasilkan interval yang mengandung nilai rata-rata populasi yang sebenarnya.

Contoh:

Misalkan, kita ingin memperkirakan rata-rata tinggi badan mahasiswa di suatu universitas dengan 95% tingkat kepercayaan. Jika hasil dari sampel adalah rata-rata tinggi badan 165 cm dengan deviasi standar 10 cm dan ukuran sampel 100, maka CI-nya adalah:

$$CI = 165 \pm 1{,}96 \times \frac{10}{\sqrt{100}} = 165 \pm 1{,}96 \times 1$$

Kemudian,

$$165 + 1{,}96 \times 1 = 166{,}96$$

$$165 - 1{,}96 \times 1 = 163{,}04$$

Sehingga, CI = (163,04 cm; 166,96 cm). Artinya, kita 95% yakin bahwa rata-rata tinggi badan mahasiswa universitas tersebut berada di antara 163,04 cm dan 166,96 cm.

Statistik Parameterik dan Non-Parameterik

Dalam analisis statistik, pemilihan metode yang tepat sangat penting untuk mendapatkan hasil yang valid dan dapat diandalkan. Dua kategori utama dari metode statistik adalah statistik parameterik dan statistik non-parametrik. Masing-masing memiliki karakteristik, asumsi, dan aplikasi yang berbeda.

Statistik Parameterik

Statistik parameterik adalah metode analisis yang mengasumsikan bahwa data mengikuti distribusi tertentu, biasanya distribusi normal. Metode ini digunakan untuk mengestimasi parameter populasi, seperti rata-rata dan varians, berdasarkan data sampel. Statistik parameterik sering kali lebih kuat dan efisien dibandingkan dengan metode non-parametrik, terutama ketika asumsi yang mendasarinya terpenuhi.

Karakteristik Statistik Parameterik

1. **Asumsi Distribusi**

   Statistik parameterik mengasumsikan bahwa data mengikuti distribusi tertentu, seperti distribusi normal. Ini berarti bahwa data harus memiliki bentuk yang simetris dan tidak memiliki outlier yang signifikan.

2. **Penggunaan Parameter**

Metode ini berfokus pada estimasi parameter populasi, seperti rata-rata (μ) dan varians (σ^2). Contoh statistik parameterik termasuk uji t, ANOVA (Analisis Varians), dan regresi linier.

3. **Kekuatan dan Efisiensi**

   Statistik parameterik cenderung lebih kuat dan efisien dalam hal penggunaan data. Ketika asumsi terpenuhi, metode ini dapat memberikan hasil yang lebih akurat dan memiliki daya deteksi yang lebih tinggi untuk perbedaan yang signifikan.

Contoh Penggunaan Statistik Parameterik

- **Uji t**: Digunakan untuk membandingkan rata-rata dua kelompok. Misalnya, seorang peneliti ingin mengetahui apakah ada perbedaan signifikan dalam rata-rata tinggi badan antara dua kelompok mahasiswa.
- **ANOVA**: Digunakan untuk membandingkan rata-rata lebih dari dua kelompok. Misalnya, seorang peneliti ingin mengetahui apakah ada perbedaan rata-rata nilai ujian antara tiga kelas yang berbeda.
- **Regresi Linier**: Digunakan untuk menganalisis hubungan antara dua variabel. Misalnya, seorang peneliti ingin mengetahui pengaruh jumlah jam belajar terhadap nilai ujian.

Statistik Non-Parameterik

Statistik non-parametrik adalah metode analisis yang tidak mengasumsikan distribusi tertentu untuk data. Metode ini lebih fleksibel dan dapat digunakan untuk data yang tidak memenuhi asumsi distribusi normal. Statistik non-parametrik sering kali digunakan ketika data bersifat ordinal atau ketika ukuran sampel kecil.

Karakteristik Statistik Non-Parameterik

1. **Tanpa Asumsi Distribusi**

   Statistik non-parametrik tidak memerlukan asumsi tentang distribusi data. Ini membuat metode ini lebih cocok untuk data yang tidak terdistribusi normal atau memiliki outlier yang signifikan.

2. **Penggunaan Rank**

   Banyak metode non-parametrik menggunakan peringkat (rank) data daripada nilai absolut. Ini membantu mengurangi pengaruh outlier dan memberikan hasil yang lebih robust.

3. **Keterbatasan Daya Deteksi**

   Meskipun statistik non-parametrik lebih fleksibel, mereka sering kali memiliki daya deteksi yang lebih rendah dibandingkan dengan statistik parameterik ketika asumsi parameterik terpenuhi. Ini berarti bahwa mereka mungkin tidak se-efisien dalam mendeteksi perbedaan yang signifikan.

Contoh Penggunaan Statistik Non-Parameterik

- **Uji Mann-Whitney U**: Digunakan untuk membandingkan dua kelompok independen ketika data tidak terdistribusi normal. Misalnya, seorang peneliti ingin membandingkan peringkat kepuasan pelanggan antara dua produk yang berbeda.

- **Uji Wilcoxon Signed-Rank**: Digunakan untuk membandingkan dua kelompok terkait ketika data tidak terdistribusi normal. Misalnya, seorang peneliti ingin mengetahui apakah ada perbedaan dalam skor sebelum dan sesudah intervensi pada kelompok yang sama.
- **Kruskal-Wallis H Test**: Digunakan untuk membandingkan lebih dari dua kelompok independen ketika data tidak terdistribusi normal. Misalnya, seorang peneliti ingin membandingkan peringkat kepuasan pelanggan antara tiga merek yang berbeda.

Aspek	**Statistik Parameterik**	**Statistik Non-Parameterik**
Asumsi Distribusi	Mengasumsikan distribusi normal	Tidak mengasumsikan distribusi tertentu
Penggunaan Data	Data interval atau rasio	Data ordinal atau nominal
Kekuatan	Lebih kuat dan efisien jika asumsi terpenuhi	Lebih fleksibel, tetapi kurang kuat
Contoh Metode	Uji t, ANOVA, regresi linier	Uji Mann-Whitney U, Kruskal-Wallis H

Tugas

1. Seorang peneliti mengumpulkan data nilai ujian dari 30 siswa. Dari sampel tersebut diperoleh rata-rata 75 dan standar deviasi 10. Hitunglah estimasi parameter untuk rata-rata populasi dengan tingkat kepercayaan 95%.
2. Sebuah perusahaan mengklaim bahwa rata-rata waktu pengiriman paket mereka adalah 30 menit. Dari sampel 25 pengiriman, diperoleh rata-rata 32.5 menit dengan standar deviasi 4 menit. Uji klaim tersebut dengan $\alpha = 0.05$
3. Sebuah survei terhadap 100 rumah tangga menunjukkan bahwa rata-rata pengeluaran bulanan adalah Rp 5.000.000 dengan standar deviasi Rp 800.000. Hitunglah interval kepercayaan 90% untuk rata-rata pengeluaran bulanan populasi.

BAB IV

Tujuan Pembelajaran

Setelah mempelajari bab ini, pembaca diharapkan mampu:

1. Memahami konsep dasar probabilitas dan penerapannya dalam analisis data.
2. Menghitung probabilitas sederhana, probabilitas gabungan, dan probabilitas bersyarat.
3. Menganalisis probabilitas menggunakan Teorema Bayes.
4. Memahami konsep permutasi dan kombinasi serta penerapannya dalam statistik dan probabilitas.

Pokok Bahasan

1. Pengenalan Probabilitas
2. Penerapan Probabilitas Sederhana
3. Penerapan Probabilitas Gabungan
4. Probabilitas Bersyarat (Conditional Probability)
5. Teorema Bayes

PERMUTASI DAN KOMBINASI

1. Permutasi
2. Kombinasi

PROBABILITAS

1. Pengenalan Probabilitas

Probabilitas adalah peluang atau kemungkinan terjadinya salah satu keadaan dari seluruh kemungkinan yang ada. Teori Probabilitas adalah cabang dari matematika yang mengkaji peluang dan kejadian acak. Probabilitas mencakup konsep, teknik, dan prinsip yang digunakan untuk mengukur, menganalisis, dan memodelkan kejadian acak. Teori probabilitas memberikan landasan matematis bagi analisis statistik, teori keputusan, dan berbagai bidang ilmu pengetahuan lainnya yang melibatkan ketidakpastian dan peluang. Probabilitas dapat digambarkan dengan nilai yang berkisar antara 0 (peristiwa tidak mungkin terjadi) hingga 1 (peristiwa pasti terjadi). Jika probabilitas suatu peristiwa adalah 0,5; dimana hal itu berarti peristiwa tersebut memiliki peluang 50% untuk terjadi.

Ruangan Sampel *(Sample Space)* dalam teori probabilitas adalah himpunan semua hasil yang mungkin dari suatu percobaan acak. Setiap elemen dalam ruang sampel mewakili suatu hasil atau kejadian yang dapat terjadi dalam konteks percobaan tersebut. Ruang sampel biasanya dilambangkan dengan huruf S. Contohnya, jika kita melempar sebuah koin, ruang sampelnya adalah {H, T} (untuk kepala dan ekor).

Peristiwa *(Event)* adalah sub-himpunan dari ruang sampel yang terdiri dari satu atau lebih hasil. Sebagai contoh, "munculnya kepala" adalah peristiwa dalam melempar koin.

2. Penerapan Probabilitas Sederhana

Jika simbol E melambangkan jumlah hasil yang dapat terjadi, sedangkan simbol N mencerminkan jumlah seluruh hasil yang dapat terjadi dari percobaan/kejadian yang berlangsung, maka notasi probabilitasnya adalah:

$$P(E) = \frac{E}{N}, dimana: 0 \leq P(E) \leq 1$$

Peristiwa terjadinya bukan E: $P(\bar{E}) = 1 - P(E)$

Percobaan Tunggal

Percobaan tunggal dalam teori probabilitas merujuk pada situasi di mana hanya satu percobaan dilakukan untuk mengamati hasil atau kejadian tertentu. Hasil dari percobaan tersebut dapat termasuk dalam ruang sampel yang terdefinisi sebelumnya.

Contoh 1:

Sebuah koin yang memiliki dua sisi atau ruang sampelnya “kepala” dan “ekor”, dilemparkan sekali, kemudian pertanyaannya adalah:

a. Berapa probabilitas keluarnya “kepala”?
b. Berapa probabilitas keluarnya bukan “kepala”?

Penyelesainnya:

a. $P(kepala) = \frac{1}{2}$
b. $P(\overline{kepala}) = 1 - P(kepala)$

$$P(\overline{kepala}) = 1 - \frac{1}{2} = \frac{1}{2}$$

Percobaan Majemuk (probabilitas bersyarat)

"Mutually exclusive" adalah istilah yang digunakan dalam teori probabilitas untuk menyatakan bahwa dua atau lebih peristiwa tidak dapat

terjadi bersamaan dalam satu percobaan. Artinya, jika satu peristiwa terjadi, maka peristiwa lainnya tidak bisa terjadi pada saat yang sama. Peristiwa yang mutually exclusive juga disebut sebagai peristiwa saling eksklusif.

Jika A dan B adalah dua peristiwa yang saling eksklusif, maka $P(A \cup B) = P(A) + P(B)$.

Seperti persoalan pada contoh 1, dengan pertanyaan:

Berapa probabilitas keluarnya "kepala" atau "ekor"?

Penyelesaiannya:

$P(kepala) + P(ekor) = \frac{1}{2} + \frac{1}{2} = 1$

Contoh 2

Sebuah dadu dilemparkan keatas, probabilitas munculnya permukaan pertama diatas sama dengan munculnya permukaan empat diatas.

$$P(1) = 1/6$$

$$P(4) = 1/6$$

$$P(1\ atau\ 4) = P(1) + P(4) = \frac{2}{6} = \frac{1}{3}$$

Perlu diingat, operator "+" dapat juga bermakna "atau", yang berarti peristiwa yang terjadi salah satu saja

Percobaan majemuk juga dapat dikatakan percobaan "independen". Dua percobaan atau lebih dikatakan independen apabila hasil percobaan yang satu tidak berpengaruh terhadap hasil percobaan yang lain.

Jika A dan B adalah dua peristiwa yang saling independen, maka $P(A \cap B) = P(A) \times P(B)$.

$$Probabilitas\ terjadi\ bersama - sama\ =\ P\ (A\ dan\ B) =\ PA \times PB$$

Contoh 3

Sebuah dadu dilemparkan keatas bersama-sama dengan sebuah koin. Berapakah probabilitas untuk mendapatkan permukaan "kepala" pada koin dan permukaan nomor 1 pada dadu yang tampak diatas

Penyelesaiannya:

$$P(kepala) = \frac{1}{2}$$

$$P(1) = \frac{1}{6}$$

$$P(kepala\ dan\ 1)\ =\ P(kepala) \times P(1) = \frac{1}{2} \times \frac{1}{6} = \frac{1}{12}$$

3. Penerapan Probabilitas Gabungan

Gabungan Kejadian (Union of Events)

Gabungan kejadian (union of events) dalam teori probabilitas merujuk pada peristiwa yang setidaknya satu di antaranya terjadi. Gabungan dua peristiwa A dan B, dilambangkan dengan $A \cup B$, terjadi jika setidaknya satu dari A atau B terjadi, atau keduanya bersamaan.

Rumus umum untuk menghitung probabilitas gabungan dua peristiwa $A \cup B$ adalah:

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

Di mana:

- $P(A)$ adalah probabilitas peristiwa A terjadi.
- $P(B)$ adalah probabilitas peristiwa B terjadi.
- $P(A \cap B)$ adalah probabilitas peristiwa A dan B terjadi bersamaan.

Penambahan $P(A)+P(B)$ memberikan probabilitas terjadi setidaknya satu dari A atau B. Namun, jika A dan B tidak mutually exclusive (bisa terjadi bersamaan), maka kita perlu mengurangkan $P(A \cap B)$ agar kita tidak menggandakan probabilitas peristiwa yang tumpang tindih.

Contoh 4: Misalkan kita memiliki dua peristiwa:

- A: Munculnya angka genap pada lemparan sebuah dadu $P(A)=\frac{3}{6}=\frac{1}{2}$
- B: Munculnya angka 4 pada lemparan sebuah dadu $P(B)=\frac{1}{6}$

Probabilitas gabungan $A \cup B$ (setidaknya satu dari A atau B terjadi) adalah:

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

$$P(A \cup B) = \frac{1}{2} + \frac{1}{6} - \frac{1}{12} = \frac{7}{12} = 0{,}583$$

Jadi, probabilitas gabungan $A \cup B$ adalah $\frac{7}{12}$. Ini berarti setidaknya satu dari A atau B terjadi pada lemparan dadu.

4. Probabilitas Bersyarat (Conditional)

Probabilitas bersyarat (conditional probability) adalah probabilitas bahwa suatu peristiwa terjadi, diberikan bahwa peristiwa lain telah terjadi. Notasi probabilitas bersyarat $P(A|B)$ menunjukkan probabilitas peristiwa A terjadi, setelah peristiwa B telah terjadi.

Rumus umum untuk menghitung probabilitas bersyarat adalah:

$$P(A \mid B) = \frac{P(A \cap B)}{P(B)} \rightarrow P(A \cap B) = P(B) \times P(A|B)$$

Di mana:

- $P(A|B)$ probabilitas peristiwa A terjadi, setelah peristiwa B telah terjadi
- $P(A \cap B)$ adalah probabilitas irisan dari A dan B, yaitu probabilitas bahwa keduanya terjadi bersamaan.
- $P(B)$ adalah probabilitas bahwa peristiwa B terjadi.

Jika A dan B independen (tidak saling bergantung), maka $P(A|B)=P(A)$, karena terjadinya B tidak memberikan informasi tambahan tentang terjadinya A.

Contoh 5:

1) Seorang mahasiswa memiliki peluang lulus statistik adalah 70%. Jika setelah lulus ujian statistik, dia melanjutkan mengikuti ujian matematika dan peluang lulusnya 80%. Berapa peluang mahasiswa tersebut untuk lulus ujian statistik dan ujian matematika?

Penyelesaiaan:

P(A) = Peluang lulus statistik = 70%

P(B|A) = Peluang lulus matematika, setelah lulus statistik = 80%

$$P(A \cap B) = P(A) \times P(B|A)$$
$$= 70\% \times 80\% = 56\% = 0{,}56$$

Jadi peluang mahasiswa tersebut untuk lulus pada kedua jenis ujian tersebut adalah 56% atau 0,56

2) Dua buah dadu dilempar keatas secara bersama, tentukan peluang muncul jumlah mata dadu lebih besar dari 9, dengan syarat dadu pertama muncul 5

Penyelesaian:

$P(\,A \mid B\,) = \frac{P(A \cap B)}{P(B)}$

$P(A \cap B) =$
$dua\ mata\ dadu\ dilempar\ bersamaan\ kemudian\ muncul\ angka\ 9\ dengan\ dad$

$P(A \cap B) = \{(5{,}5); (5{,}6)\} = \frac{2}{36}$

$P(B) = peluang\ dadu\ pertama\ angka\ 5$

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

Sehingga,

$$P(\,A \mid B\,) = \frac{\frac{2}{36}}{\frac{6}{36}} = \frac{2}{36} \times \frac{36}{6} = \frac{72}{216} = \frac{1}{3}$$

Atau dapat di selesaikan menggunakan cara peluang sederhana, yaitu:

Ruang sampel: muncul mata dadu pertama 5, sehingga

: { (5,1) ; (5,2) ; (5,3) ; (5,4) ; (5,5) ; (5,6) } → n(Sampel) = 6

Kejadian A adalah kejadian mata dadu yang berjumlah lebih dari 9 kedalam ruang sampel:

{ (5,5) ; (5,6) } → n(A) = 2

$$Peluang \ = \frac{n(A)}{n(Sampel)} = \frac{2}{6} = \frac{1}{3}$$

3) Dalam sebuah kotak terdapat 6 bola biru dan 4 bola hitam. Jika sebuah bola diambil dalam kotak itu berturut-turut sebanyak dua kali <u>tanpa pengembalian</u>. Tentukan peluang yang terambil kedua-duanya bola biru!

Penyelesaiannya:

P(A) : peluang terambilnya bola biru pada pengambilan pertama = $\frac{6}{10}$

P(B|A) : peluang terambilnya bola biru kedua dengan syarat pengambilan pertama bola biru = $\frac{5}{9}$

$P(A \cap B) = P(A) \times P(B|A)$

$P(A \cap B) = \frac{6}{10} \times \frac{5}{9} = \frac{30}{90} = \frac{1}{3}$

4) Dalam sebuah kotak terdapat 6 bola biru dan 4 bola hitam. Jika sebuah bola diambil dalam kotak itu berturut-turut sebanyak dua kali <u>tanpa pengembalian</u>. Tentukan peluang yang terambil bola biru dan hitam!

Penyelesaiannya:

P(A) : peluang terambilnya bola biru pada pengambilan pertama = $\frac{6}{10}$

P(B|A) : peluang terambilnya bola hitam dengan syarat pengambilan pertama bola biru = $\frac{4}{9}$

$P(A \cap B) = P(A) \times P(B|A)$

$P(A \cap B) = \frac{6}{10} \times \frac{4}{9} = \frac{24}{90} = \frac{4}{15}$

5) Didalam sebuah mall, terdapat 12 ibu-ibu dan 4 orang remaja yang sedang berbelanja. Kemudian akan dipilih 3 orang secara acak untuk mendapatkan hadiah. Aturannya, setiap orang hanya akan mendapatkan 1 hadiah. Berapa peluang kejadian jika ketiga hadiah tersebut akan dimenangkan ibu-ibu semua

Penyelesaiaan:

P(A) : peluang tepilihnya ibu-ibu yang pertama = $\frac{12}{16}$

P(B|A) : peluang terpilihnya ibu-ibu yang kedua dengan syarat pertama ibu-ibu = $\frac{11}{15}$

P(C|B,A) : peluang terpilihnya ibu-ibu yang ketiga dengan syarat pertama dan kedua ibu-ibu = $\frac{10}{14}$

$P(A \cap B) = P(A) \times P(B|A) \times P(C|B, A)$

$P(A \cap B) = \frac{12}{16} \times \frac{11}{15} \times \frac{10}{14} = \frac{11}{28}$

5. Teorema Bayes

Teorema Bayes adalah konsep fundamental dalam teori probabilitas yang memberikan cara untuk memperbarui probabilitas suatu peristiwa dengan mempertimbangkan informasi baru. Teorema ini dinamai dari ilmuwan Inggris abad ke-18, Thomas Bayes, yang pertama kali merumuskannya.

Teorema Bayes menyatakan bahwa probabilitas kondisional suatu peristiwa A setelah suatu peristiwa B terjadi dapat dihitung menggunakan rumus berikut:

$$P(A \mid B) = \frac{\big(P(B \mid A) \cdot P(A)\big)}{P(B)}$$

PERMUTASI DAN KOMBINASI

1. PERMUTASI

Dalam matematika dan kombinatorik, permutasi adalah pengaturan atau susunan dari suatu himpunan objek yang berbeda. Permutasi menunjukkan berapa banyak cara objek-objek tersebut dapat diatur. Ada dua jenis permutasi: permutasi dengan pengulangan dan permutasi tanpa pengulangan.

Permutasi Tanpa Pengulangan

Permutasi tanpa pengulangan adalah susunan objek-objek yang berbeda di mana setiap objek hanya muncul sekali. Jumlah permutasi tanpa pengulangan dari *n* objek yang diambil *r* pada suatu waktu dapat dihitung dengan rumus:

$$P(n, r) = \frac{n!}{(n - r)!}$$

di mana:

- $P(n,r)$: Jumlah permutasi dari n objek yang diambil r pada suatu waktu.
- $n!$: Faktorial dari n, yaitu $n\times(n-1)\times(n-2)\times\ldots\times2\times1$.

Permutasi dengan Pengulangan

Permutasi dengan pengulangan mempertimbangkan pengulangan objek-objek, artinya satu objek dapat muncul lebih dari sekali dalam susunan tersebut. Jumlah permutasi dengan pengulangan dari n objek yang diambil r pada suatu waktu dapat dihitung dengan rumus:

$$P_{dengan\ pengulangan}(n,r) = n^r$$

di mana:

- $P_{dengan\ pengulangan}(n,r)$: Jumlah permutasi dengan pengulangan dari n objek yang diambil r pada suatu waktu.

Contoh 6:

Jika terdapat 5 buah buku dan kita ingin menata 3 buku di atas meja, berapa banyak cara kita dapat menata buku-buku tersebut?

Penyelesaian :

$$P(5,3) = \frac{5!}{(5-3)!} = \frac{5\times4\times3\times2\times1}{(2\times1)} = 60\ cara.$$

Contoh 7:

Jika terdapat 3 warna cat (merah, biru, hijau) dan kita ingin mengecat pagar dengan 5 balok, berapa banyak cara kita dapat mengecat setiap balok dengan warna yang berbeda?

Penyelesaian :

$$P_{denganpengulangan}(3{,}5) = 3^5 = 243\ cara.$$

2. KOMBINASI

Kombinasi adalah konsep dalam matematika dan kombinatorika yang berkaitan dengan cara memilih *r* objek dari suatu himpunan objek total *n*, tanpa memperhatikan susunan atau urutan objek tersebut. Kombinasi sering digunakan untuk menghitung berapa banyak cara kita dapat memilih sekelompok objek dari suatu himpunan.

Kombinasi Tanpa Pengulangan

Kombinasi tanpa pengulangan adalah cara memilih *r* objek dari *n* objek yang berbeda, tanpa memperhatikan urutan atau susunan objek tersebut. Jumlah kombinasi tanpa pengulangan dari *n* objek yang diambil *r* pada suatu waktu diberikan oleh rumus:

$$C(n, r) = \frac{n!}{r!\,(n-r)!}$$

di mana:

- $C(n,r)$: Jumlah kombinasi dari *n* objek yang diambil *r* pada suatu waktu.
- $n!$: Faktorial dari *n*, yaitu $n\times(n-1)\times(n-2)\times\ldots\times 2\times 1$.

- $r!$: Faktorial dari r.

Kombinasi dengan Pengulangan

Kombinasi dengan pengulangan adalah cara memilih r objek dari n objek yang dapat diulang, tanpa memperhatikan urutan atau susunan objek tersebut. Jumlah kombinasi dengan pengulangan dari n objek yang diambil r pada suatu waktu diberikan oleh rumus:

$$C_{denganpengulangan(n,r)} = \frac{(n+r-1)!}{r!\,(n-1)!}$$

di mana:

- $C_{denganpengulangan(n,r)}$: Jumlah kombinasi dengan pengulangan dari n objek yang diambil r pada suatu waktu.

Contoh 8:

Jika terdapat 7 siswa dan kita ingin memilih tim berisi 3 siswa, berapa banyak cara kita dapat memilih tim tersebut?

Penyelesaian:

$$C(7,3) = \frac{7!}{3!\,(7-3)!} = \frac{7 \times 6 \times 5}{(3 \times 2 \times 1)(4 \times 3 \times 2 \times 1)}$$
$$= 35\ cara.$$

Contoh 9:

Jika terdapat 4 warna cat (merah, biru, kuning, hijau) dan kita ingin memilih tim berisi 2 warna cat untuk sebuah proyek, berapa banyak cara kita dapat memilih tim tersebut?

Penyelesaian:

$$C_{denganpengulangan(4,2)} = \frac{(4+2-1)!}{2!\,(4-1)!} = \frac{5 \times 4 \times 3 \times 2 \times 1}{(2 \times 1)(3 \times 2 \times 1)}$$

$$= 10\ cara$$

Tugas

1. Sebuah kotak berisi 8 bola merah, 5 bola hijau, dan 7 bola biru. Jika satu bola diambil secara acak, berapa probabilitas terambilnya bola hijau?
2. Dalam sebuah kelas, 60% siswa suka matematika dan 70% siswa suka fisika. Jika 40% siswa suka keduanya, berapa probabilitas seorang siswa yang dipilih secara acak suka matematika atau fisika atau keduanya?
3. Dalam sebuah survei, 70% responden memiliki smartphone, 80% memiliki laptop, dan 60% memiliki keduanya. Jika seorang responden dipilih secara acak dan diketahui memiliki smartphone, berapa probabilitas dia juga memiliki laptop?
4. Sebuah tes kesehatan memiliki tingkat akurasi 95% untuk mendeteksi penyakit tertentu. Penyakit ini memengaruhi 2% populasi. Jika seseorang dites positif, berapa probabilitas orang tersebut benar-benar memiliki penyakit?
5. Berapa banyak cara untuk mengatur 7 buku berbeda di rak?
6. Berapa banyak cara untuk mengatur huruf-huruf dalam kata "INDONESIA"?
7. Dari 15 anggota klub, berapa banyak cara untuk memilih 4 orang untuk membentuk sebuah komite?
8. Sebuah restoran menawarkan 6 jenis nasi, 8 jenis lauk, dan 4 jenis sayur. Berapa banyak cara untuk menyusun menu dengan 1 jenis nasi, 2 jenis lauk, dan 1 jenis sayur?
9. Dalam sebuah kelas dengan 25 siswa, 15 di antaranya perempuan. Jika 5 siswa dipilih secara acak untuk sebuah proyek, berapa probabilitas bahwa tepat 3 dari mereka adalah perempuan?

BAB V

Tujuan Pembelajaran

Setelah mempelajari bab ini, pembaca diharapkan mampu:

1. Memahami konsep distribusi peluang dalam statistik.
2. Membedakan antara distribusi probabilitas diskrit dan kontinu serta penerapannya.
3. Menghitung dan menganalisis distribusi binomial dan distribusi Poisson.
4. Memahami dan menerapkan distribusi normal (Gaussian), distribusi t-Student, dan distribusi F dalam analisis statistik.

Pokok Bahasan

1. Distribusi Probabilitas Diskrit
    - Distribusi Binomial
    - Distribusi Poisson
2. Distribusi Probabilitas Kontinu
    - Distribusi Normal (Gaussian)
    - Distribusi T (t-Student)
    - Distribusi F

DISTRIBUSI PELUANG

Distribusi peluang, atau distribusi probabilitas, adalah suatu himpunan dari semua kemungkinan hasil atau nilai suatu variabel acak beserta dengan probabilitas masing-masing hasil atau nilai tersebut. Dalam hal ini, probabilitas mengukur sejauh mana suatu kejadian tertentu mungkin terjadi. Distribusi probabilitas dapat bersifat diskrit, di mana nilainya terbatas pada himpunan nilai yang terpisah, atau kontinu, di mana nilainya dapat mengambil rentang nilai dalam interval tertentu. Dua karakteristik penting dari distribusi probabilitas adalah bahwa probabilitas dari setiap hasil harus selalu berada di antara 0 dan 1, dan jumlah probabilitas dari seluruh hasil yang saling lepas (mutually exclusive) dalam distribusi tersebut adalah 1.

1. Distribusi Probabilitas Diskrit (Binomial, Poisson)

Distribusi probabilitas diskrit adalah distribusi peluang yang terkait dengan variabel acak diskrit. Variabel acak diskrit adalah variabel acak yang dapat mengambil satu set nilai terpisah, biasanya berupa bilangan bulat. Setiap nilai dalam himpunan nilai variabel acak memiliki probabilitas tertentu terjadi.

A. Binomial

Distribusi binomial adalah suatu distribusi probabilitas yang muncul ketika suatu eksperimen dilakukan sejumlah n kali dan setiap percobaan bersifat independen, hanya memiliki dua hasil yang mungkin: sukses atau gagal. Distribusi ini dinamakan "binomial" karena setiap percobaan dapat dianggap sebagai uji coba Bernoulli, yang hanya memiliki dua kemungkinan hasil.

Karakteristik Distribusi Binomial:

1. **Bernoulli Trial:**
    - Setiap percobaan dianggap sebagai uji coba Bernoulli, yang memiliki dua hasil mungkin: sukses atau gagal.
2. **Bilangan Percobaan (n):**
    - Menunjukkan jumlah percobaan atau uji coba independen yang dilakukan.
3. **Probabilitas Keberhasilan (p):**
    - Menunjukkan probabilitas sukses dalam satu percobaan atau uji coba.
4. **Probabilitas Gagal (q=1−p):**
    - Menunjukkan probabilitas kegagalan dalam satu percobaan atau uji coba.

Notasi dan Rumus Distribusi Binomial:

X: Variabel acak yang mewakili jumlah keberhasilan dalam n percobaan.

x: Nilai yang diambil oleh variabel acak X

Fungsi massa peluang (PMF) untuk distribusi binomial:

$$P(X = x) = \binom{n}{x} p^x q^{n-x}$$

di mana:

- $\binom{n}{x}$ adalah kombinasi dari n pilih x atau "n atas x" yang dapat dihitung dengan $\frac{n!}{x!(n-x)!}$

- p adalah probabilitas sukses.
- q adalah probabilitas kegagalan (q=1−p).
- n adalah jumlah percobaan.
- x adalah jumlah keberhasilan yang ingin dihitung probabilitasnya.

Sifat Distribusi Binomial

1. **Bentuk Distribusi:**
   - Distribusi binomial memiliki bentuk lonceng atau simetris jika n cukup besar.
2. **Nilai Ekspektasi (Rata-rata):**
   - Rata-rata distribusi binomial adalah $\mu = n \times p$
3. **Variansi:**
   - Variansi distribusi binomial adalah $\sigma^2 = n \times p \times q$
4. **Aturan Kombinasi (n atas x) $\binom{n}{x}$:**
   - Digunakan untuk menghitung berapa cara x keberhasilan dapat terjadi dalam n percobaan.

Contoh 10:

Seorang penulis artikel berita online menulis artikel pendek setiap hari. Berdasarkan pengalamannya, rata-rata 20% artikel yang ditulisnya menjadi viral dan mendapatkan banyak pembaca, sedangkan 80% artikel lainnya tidak mendapatkan banyak perhatian. Hari ini, penulis tersebut merencanakan untuk menulis 5 artikel. Tentukan probabilitas bahwa tepat

3 artikel yang dia tulis akan menjadi viral. Berapa probabilitas bahwa tepat 3 artikel dari 5 artikel yang ditulis oleh penulis tersebut akan menjadi viral?

Penyelesaian:

Kita dapat menggunakan distribusi binomial dengan parameter n=5 (jumlah percobaan) dan p=0,2 (probabilitas sukses).

$$P(X = 3) = \binom{5}{3} 0{,}2^3 0{,}8^{5-3}$$

$$\binom{n}{x} = \frac{n!}{x!\,(n-x)!}$$

$$\binom{5}{3} = \frac{5!}{3!\,(5-3)!} = \frac{5 \times 4 \times 3 \times 2 \times 1}{(3 \times 2 \times 1)(2 \times 1)} = 10$$

$$P(X = 3) = 10 \times 0{,}008 \times 0{,}64 = 0{,}0512$$

Probabilitas bahwa tepat 3 artikel dari 5 artikel yang ditulis oleh penulis tersebut akan menjadi viral adalah 0,0512 atau 5,12%.

B. Poisson

Distribusi Poisson adalah suatu distribusi probabilitas diskrit yang menyatakan jumlah peristiwa langka dalam suatu interval waktu atau ruang tertentu. Distribusi ini dinamai dari matematikawan Prancis Simeon Denis Poisson, yang memberikan kontribusi besar dalam pengembangan teori probabilitas.

Karakteristik Distribusi Poisson:

1. **Parameter λ:**

- Distribusi Poisson memiliki satu parameter, biasanya disimbolkan sebagai λ (lambda).
- λ menyatakan tingkat kejadian rata-rata dalam suatu interval atau ruang tertentu.

2. **Ruang Sampel:**
    - Distribusi Poisson cocok digunakan ketika peristiwa langka terjadi secara acak dan saling independen di dalam suatu ruang sampel.
3. **Nilai Discret:**
    - Distribusi Poisson berlaku untuk nilai-nilai diskrit, khususnya jumlah peristiwa yang terjadi.

Fungsi Massa Peluang *(Probability Mass Function)*:

Fungsi massa peluang (PMF) dari distribusi Poisson diberikan oleh rumus:

$$P(X = k) = \frac{e^{-\lambda}.\lambda^{k}}{k!}$$

di mana:

- $P(X=k)$: Probabilitas bahwa terjadi k peristiwa.
- e: Konstanta Euler (sekitar 2,71828).
- λ: Tingkat kejadian rata-rata.
- k: Jumlah peristiwa yang terjadi.

Contoh Penggunaan Distribusi Poisson:

1. Kejadian Kecelakaan Lalu Lintas:

Menghitung jumlah kecelakaan lalu lintas dalam suatu kota dalam satu hari dengan asumsi tingkat kejadian rata-rata.

2. Kerusakan Mesin Pabrik:
   Menghitung berapa kali mesin di pabrik mengalami kerusakan dalam sebulan dengan tingkat kejadian rata-rata tertentu.
3. Jumlah Panggilan ke Layanan Darurat:
   Menghitung jumlah panggilan ke layanan darurat dalam satu jam di suatu wilayah dengan tingkat kejadian rata-rata.

Contoh 11:

Sebuah restoran cepat saji menerima rata-rata 5 pesanan pengiriman per jam. Hitung probabilitas bahwa dalam satu jam restoran tersebut akan menerima tepat 3 pesanan.

Penyelesaian:

$$P(X = k) = \frac{e^{-\lambda}.\lambda^{k}}{k!}$$

Di mana:

- $P(X=3)$: Probabilitas menerima tepat 3 pesanan.
- $\lambda=5$: Tingkat kejadian rata-rata (jumlah pesanan per jam).
- $k=3$: Jumlah pesanan yang kita cari probabilitasnya.

Substitusi nilai ke dalam rumus:

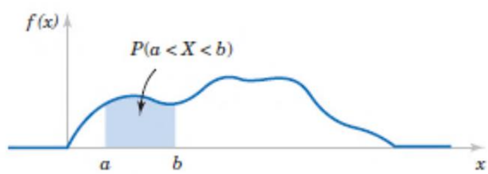
$$P(X = 3) = \frac{e^{-5}.5^{3}}{6} = 0{,}1403$$

Jadi, probabilitas bahwa dalam satu jam restoran tersebut akan menerima tepat 3 pesanan adalah sekitar 0,1403 atau sekitar 14%.

2. Distribusi Probabilitas Kontinu (Normal, T, F)

Distribusi probabilitas kontinu adalah suatu distribusi probabilitas di mana variabel acak dapat mengambil nilai dalam suatu rentang kontinu. Dalam konteks distribusi probabilitas kontinu, setiap nilai tunggal memiliki probabilitas nol karena ada takhingga banyak nilai yang mungkin di dalam rentang tertentu. Sedangkan menurut Montgomery dalam bukunya "*Applied Statistics and Probability for Engineers*" Distribusi Probabilitas Kontinu adalah daftar atau sebaran probabilitas dari setiap nilai variabel random Kontinu. Variabel random Kontinu adalah variabel random dengan interval (baik terbatas maupun tidak terbatas) dalam suatu jarak dari bilangan nyata.

Distribusi Probabilitas Kontinu dapat digambarkan dengan fungsi kepadatan probabilitas f(x) yang mempunyai nilai-nilai dalam variabel Kontinu. Seperti pada gambar dibawah ini, daerah dibawah kurva a sampai b merupakan distribusi probabilitas Kontinu yang nilainya berada pada interval dua buah angka a dan b yang termasuk dalam variabel x atau variabel Kontinu.



Gambar 13. Contoh Fungsi Kepadatan Probabilitas Variabel Random Kontinu Probabilitas daerah interval a dan b

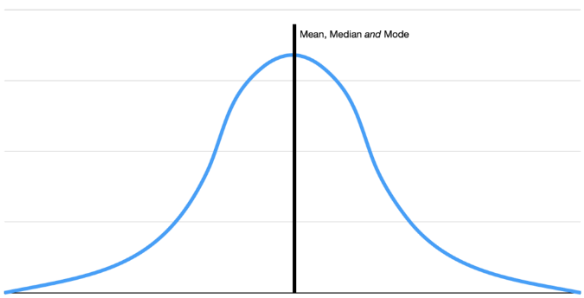
A. Distribusi Normal (Gaussian)

Distribusi Normal, juga dikenal sebagai distribusi Gaussian atau distribusi *bell-shaped*, adalah salah satu distribusi probabilitas paling penting dan umum dalam statistika. Distribusi Normal sering digunakan untuk mewakili berbagai fenomena alam dan manusia. Distribusi ini memainkan peran krusial dalam analisis statistika inferensial dan metode-metode yang bergantung pada distribusi tersebut.

Karakteristik Distribusi Normal:

1. **Bentuk *Bell-Shaped*:**

Distribusi Normal memiliki bentuk grafik lonceng atau *bell-shaped* yang simetris. Ini berarti mayoritas data cenderung berpusat di sekitar nilai rerata dan menyebar secara simetris.



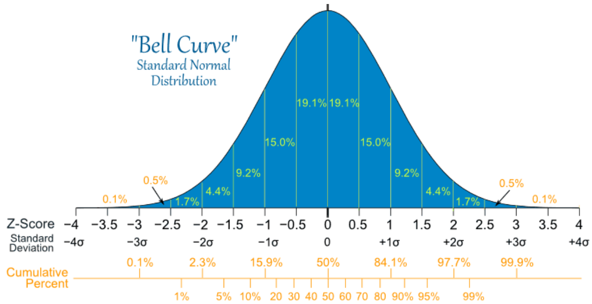
Gambar 14. Kurva Lonceng

2. **Parameter Utama:**

Distribusi Normal dapat dijelaskan oleh dua parameter utama: rerata (μ) dan deviasi standar (σ).

- Rerata (μ): Nilai tengah distribusi, menunjukkan pusat dari distribusi.

- Deviasi Standar (σ): Ukuran sebaran data di sekitar rerata, mengukur "lebar" distribusi.



Gambar 15. Kurva Lonceng dengan parameter-parameternya

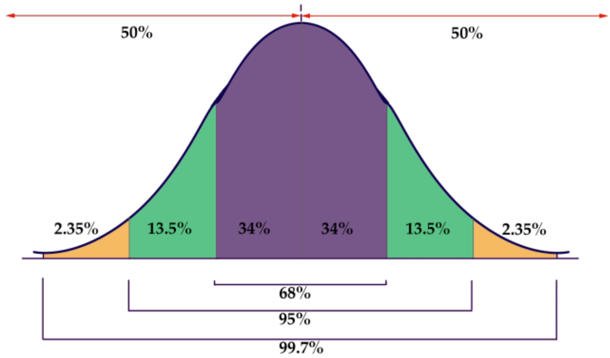
3. **Kurva Distribusi:**

Distribusi Normal dapat didefinisikan oleh fungsi kepadatan probabilitas *(Probability Density Function - PDF)*, yang menghasilkan kurva berbentuk lonceng.

Sifat Distribusi Normal:

1. **Empat Kuadran:**
    - Distribusi Normal dapat dibagi menjadi empat kuadran simetris, di mana proporsi data tertentu berada di bawah atau di atas rerata dengan sejumlah deviasi standar tertentu.
2. **Empat Hukum Empiris:**
    - Empat hukum empiris Distribusi Normal:
        - Sekitar 68% dari data berada dalam satu deviasi standar dari rerata.
        - Sekitar 95% berada dalam dua deviasi standar.

- Sekitar 99.7% berada dalam tiga deviasi standar.
- Distribusi Normal adalah asimtotik, yang berarti ekor distribusi tidak pernah menyentuh sumbu horizontal.



Gambar 16. Kurva Lonceng dengan sifat distribusi nya.

3. **Z-Score:**

Z-Score, juga dikenal sebagai skor standar, adalah ukuran statistik yang menggambarkan seberapa jauh suatu nilai dari rerata dalam satuan deviasi standar. Z-Score digunakan untuk menormalkan data dan membandingkan posisi relatif suatu nilai terhadap distribusi normal.

Z-Score dari suatu nilai x dalam distribusi dengan rerata μ dan deviasi standar σ dihitung dengan rumus:

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

di mana:

Z= Z-Score dari nilai x.

X= Nilai yang ingin dihitung Z-Score-nya.

μ= Rerata (mean) dari distribusi.

σ= Deviasi standar dari distribusi.

Interpretasi Z-Score:

- Jika Z-Score positif, itu berarti nilai x lebih tinggi dari rerata ($x>\mu$).
- Jika Z-Score negatif, itu berarti nilai x lebih rendah dari rerata ($x<\mu$).
- Semakin besar nilai absolut Z-Score, semakin jauh nilai x dari rerata dalam satuan deviasi standar.

Penggunaan Z-Score:

1. **Standarisasi Data:**
    - Z-Score digunakan untuk menstandarisasi data dengan mengubah distribusi menjadi distribusi Normal dengan rerata 0 dan deviasi standar 1.
2. **Perbandingan Data:**
    - Z-Score memungkinkan perbandingan langsung antara nilai-nilai dari distribusi yang berbeda.
3. **Analisis Outlier:**
    - Z-Score digunakan untuk mengidentifikasi outlier, yaitu nilai yang jauh dari rerata dalam distribusi.
4. **Prediksi Probabilitas:**

- Z-Score digunakan dalam memprediksi probabilitas terjadinya suatu peristiwa dalam distribusi Normal standar.

Contoh 12:

Misalkan tinggi badan manusia dewasa mengikuti Distribusi Normal dengan rerata μ=170 cm dan deviasi standar σ=10 cm. Jika seseorang memiliki tinggi 180 cm, kita dapat menghitung Z-Score-nya sebagai berikut:

Diketahui tinggi badan siswa di sebuah sekolah mengikuti distribusi normal dengan rata-rata (mean) 160 cm dan deviasi standar 10 cm. Tentukan probabilitas seorang siswa yang dipilih secara acak memiliki tinggi badan antara 155 cm dan 165 cm!

Tingkat signifikansi yang digunakan adalah α=0,05

Langkah-langkah:

1. Tentukan nilai z-score untuk kedua batas (155 cm dan 165 cm).
2. Hitung probabilitas menggunakan tabel distribusi normal standar (tabel z).

Penyelesaian:

1. **Menentukan Z-score 155cm**:

$$Z = \frac{155 - 160}{10} = -0{,}5$$

Menentukan Z-score 165cm:

$$Z = \frac{165 - 160}{10} = 0{,}5$$

2. **Mencari Probabilitas**: Untuk mencari probabilitas antara 155 cm dan 165 cm, kita akan mencari probabilitas kumulatif untuk $Z_1 = -0{,}5\ dan\ Z_2 = 0{,}5$ dari tabel distribusi normal standar (tabel z).

- Probabilitas kumulatif untuk $Z_1 = -0{,}5$ (dari tabel z) adalah sekitar **0,3085**.
- Probabilitas kumulatif untuk $Z_2 = 0{,}5$ (dari tabel z) adalah sekitar **0,6915**.

Probabilitas antara 155 cm dan 165 cm adalah selisih dari kedua probabilitas kumulatif:

$$P(155 \leq X \leq 165) = P(Z_2) - P(Z_1)$$

$$P(155 \leq X \leq 165) = 0{,}6915 - 0{,}3085 = 0{,}3830$$

Probabilitas bahwa seorang siswa yang dipilih secara acak memiliki tinggi badan antara 155 cm dan 165 cm adalah **0,3830** atau **38,30%**.

B. Distribusi T (Distribusi t-student)

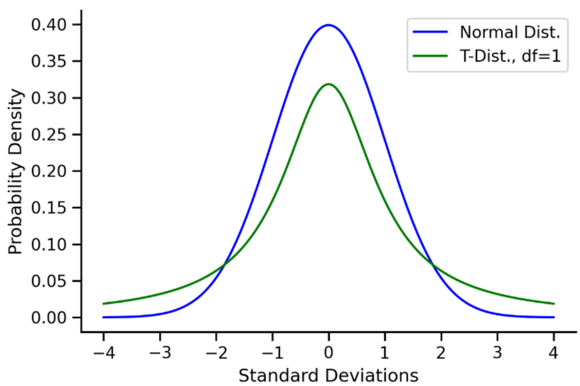
Distribusi t-Student adalah distribusi probabilitas yang digunakan untuk mengestimasi parameter populasi (seperti rata-rata) ketika ukuran sampel kecil dan varians populasi tidak diketahui. Distribusi ini diperkenalkan oleh **William Sealy Gosset** pada tahun 1908 dengan nama samaran "Student".

Gambar 17. William Sealy Gosset (1876-1937)

Distribusi t sering digunakan dalam inferensi statistik, terutama dalam uji hipotesis dan interval estimasi, ketika kita bekerja dengan sampel kecil (umumnya n < 30) atau ketika varians populasi tidak diketahui.

Sifat-sifat Distribusi T-Student

Bentuk: Distribusi t memiliki bentuk yang mirip dengan distribusi normal, tetapi lebih lebar di bagian ekor. Ini menunjukkan bahwa distribusi t lebih sensitif terhadap sampel kecil.



Gambar 18. Sifat Distribusi T-Student

Simetri: Distribusi t adalah distribusi yang simetris di sekitar angka 0, sama seperti distribusi normal.

Puncak lebih rendah dan ekor lebih lebar: Dibandingkan dengan distribusi normal, distribusi t memiliki puncak yang lebih rendah dan ekor yang lebih lebar, yang mencerminkan ketidakpastian yang lebih besar saat ukuran sampel kecil.

Derajat Kebebasan (Degrees of Freedom, df): Bentuk distribusi t tergantung pada derajat kebebasan (df). Semakin besar derajat kebebasan, distribusi t akan semakin mendekati distribusi normal standar

Derajat kebebasan dalam distribusi t adalah ukuran jumlah informasi yang tersedia untuk memperkirakan parameter populasi. Dalam konteks uji hipotesis atau estimasi rata-rata, derajat kebebasan biasanya dihitung sebagai:

$$df = n - 1$$

Dimana n adalah ukuran sampel

Kegunaan distribusi t-student

Uji t untuk sampel tunggal	Digunakan untuk menguji apakah rata-rata sampel berbeda signifikan dari nilai tertentu (misalnya, rata-rata populasi).
Uji t untuk dua sampel independen	Digunakan untuk membandingkan dua rata-rata sampel yang tidak saling bergantung.
Uji t berpasangan	Digunakan untuk membandingkan dua rata-rata dari sampel yang berhubungan atau berpasangan.
Interval Estimasi	Distribusi t digunakan untuk menghitung interval kepercayaan untuk rata-rata populasi ketika varians populasi tidak diketahui.

Uji Hipotesis dengan Distribusi t

Distribusi t digunakan dalam banyak uji hipotesis, salah satunya adalah Uji t-Student yang digunakan untuk menguji hipotesis mengenai rata-rata populasi.

Langkah-langkah Uji t

1. Tentukan hipotesis nol (H_0) dan hipotesis alternatif (H_1).
2. Pilih tingkat signifikansi (α/alpha) yang biasanya 0.05.
3. Hitung nilai t-statistic menggunakan rumus:

$$t = \frac{\bar{x} - \mu}{\frac{s}{\sqrt{n}}}$$

Dimana:

- $\bar{x}$ adalah rata-rata sampel.
- μ adalah nilai rata-rata yang diuji (nilai hipotesis nol).
- s adalah standar deviasi sampel.
- n adalah ukuran sampel.

4. Tentukan nilai kritis t dari tabel distribusi t untuk derajat kebebasan $df = n - 1$ dan tingkat signifikansi α
5. Bandingkan nilai t-statistik dengan nilai kritis t untuk membuat keputusan:

$$Jika \ \mid t \mid > \ t_{kritikal}, tolak \ H0.$$

$$Jika \ \mid t \mid \leq \ t_{kritikal}, gagal\ menolak\ H0.$$

Tabel Distribusi t

Untuk melakukan uji t atau menghitung probabilitas, kita sering merujuk ke tabel distribusi t. Tabel ini memberikan nilai kritis t untuk berbagai tingkat signifikansi dan derajat kebebasan. Sebagai contoh, untuk uji dua arah dengan tingkat signifikansi 0,05, kita mencari nilai kritis t pada kolom yang sesuai dengan derajat kebebasan yang dihitung.

Contoh: Tabel t untuk 95% kepercayaan dan df = 10

Untuk df=10 dan tingkat signifikansi 0,05; nilai kritis t adalah sekitar 2,228.

dk	α untuk Uji Satu Pihak (*one tail test*)					
	0,25	0,10	0,05	0,025	0,01	0,005
	α untuk Uji Dua Pihak (*two tail test*)					
	0,50	0,20	0,10	$\boxed{0,05}$	0,02	0,01
1	1,000	3,078	6,314	12,706	31,821	63,657
2	0,816	1,886	2,920	4,303	6,965	9,925
3	0,765	1,638	2,353	3,182	4,541	5,841
4	0,741	1,533	2,132	2,776	3,747	4,604
5	0,727	1,476	2,015	2,571	3,365	4,032
6	0,718	1,440	1,943	2,447	3,143	3,707
7	0,711	1,415	1,895	2,365	2,998	3,499
8	0,706	1,397	1,860	2,306	2,896	3,355
9	0,703	1,383	1,833	2,262	2,821	3,250
$\boxed{10}$	0,700	1,372	1,812	$\boxed{2,228}$	2,764	3,169
11	0,697	1,363	1,796	2,201	2,718	3,106

Gambar 19. Tabel T Untuk 95% Kepercayaan Dan Df = 10

Uji t-Student 1 Sampel

Contoh 13

Sebuah perusahaan ingin menguji apakah rata-rata waktu yang dihabiskan karyawan untuk menyelesaikan laporan (dalam menit) sama dengan 50 menit. Dari sampel acak yang diambil, diperoleh data waktu penyelesaian laporan berikut:

45,48,52,51,47,49,50,53,46,50

Tingkat signifikansi yang digunakan adalah $\alpha=0,05$

Langkah-langkah:

1. Tentukan hipotesis:
   - Hipotesis nol (H0): $\mu=50$ (Rata-rata waktu sama dengan 50 menit).
   - Hipotesis alternatif (H1): $\mu\neq50$ (Rata-rata waktu tidak sama dengan 50 menit).
2. Hitung statistik uji t menggunakan rumus:

$$t = \frac{\bar{x} - \mu}{\frac{s}{\sqrt{n}}}$$

3. Tentukan derajat kebebasan (df) dan cari nilai t kritis dari tabel distribusi t dengan df=n−1

Jawaban:

Data sampel: 45, 48, 52, 51, 47, 49, 50, 53, 46, 50

$$\bar{x} = \frac{45 + 48 + 52 + 51 + 47 + 49 + 50 + 53 + 46 + 50}{10} = \frac{495}{10}$$

$$= 49,5$$

Rata-rata sampel = 49,5

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

$$s = \sqrt{\frac{(45-49{,}5)+(48-49{,}5)+(52-49{,}5)\ldots+(46-49{,}5)+(50-49{,}5)}{10-1}}$$

$$= \sqrt{(6{,}94)}$$

s = 2,63

Hitung nilai t:

$$t = \frac{49{,}5 - \ 50}{\frac{2{,}63}{\sqrt{10}}} = -\frac{0{,}5}{0{,}832} = -0{,}6$$

Derajat kebebasan (df = n-1 = 10-1 = 9)

Nilai t kritis dari tabel distribusi t untuk df = 9 dan tingkat signifikansi α= 0,05. Untuk uji dua sisi, $t_{\alpha/2}$ = 2,262

Sehingga diperoleh kesimpulan:

Nilai t hitung = -0,6

Nilai t kritis = ±2,262

Karena -0,6 tidak berada di luar batas kritis ±2,262; maka **gagal menolak hipotesis nol**. Artinya, tidak ada cukup bukti untuk mengatakan bahwa rata-rata waktu yang dihabiskan berbeda dari 50 menit

Uji t-Student 2 Sampel Independen

Contoh 14

Seorang peneliti ingin mengetahui apakah ada perbedaan rata-rata tinggi badan antara pria dan wanita di sebuah universitas. Dua sampel acak diambil dari pria dan wanita. Berikut adalah data tinggi badan dalam cm:

- **Pria**: 175, 180, 169, 160, 185, 170, 177, 162, 173, 178
- **Wanita**: 160, 158, 155, 162, 159, 157, 161, 158, 154, 155

Tingkat signifikansi yang digunakan adalah α=0,05

Langkah-langkah:

1. Tentukan hipotesis:
    - Hipotesis nol (H0): $\mu_1 = \mu_2$ (tidak ada perbedaan tinggi badan antara pria dan wanita).
    - Hipotesis alternatif (H1): $\mu_1 \neq \mu_2$ (tidak ada perbedaan tinggi badan antara pria dan wanita).
2. Hitung statistik uji t menggunakan rumus:

$$t = \frac{\overline{x_1} - \overline{x_2}}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}$$

Dimana:

$\overline{x_1}$ dan $\overline{x_2}$ = rerata sampel pria dan wanita

s_1 dan s_2 = standar deviasi sampel pria dan wanita

n_1 dan n_2 = ukuran sampel pria dan wanita

3. Tentukan derajat kebebasan (df) dan cari nilai t kritis dari tabel distribusi t

Jawaban:

- **Pria**: 175, 180, 169, 160, 185, 170, 177, 162, 173, 178
- **Wanita**: 160, 158, 155, 162, 159, 157, 161, 158, 154, 155

Rata-rata dan standar deviasi sampel

Rata-rata pria:

$$\bar{x} = \frac{175 + 180 + 169 + 160 + 185 + 170 + 177 + 162 + 173 + 178}{10}$$

$$= 172{,}9$$

Rata-rata wanita:

$$\bar{x} = \frac{160 + 158 + 155 + 162 + 159 + 157 + 161 + 158 + 154 + 155}{10}$$

$$= 157{,}9$$

Standar deviasi pria:

s=7,83

Standar deviasi pria:

s=2,68

hitung nilai t:

$$t = \frac{172{,}9 - 157{,}9}{\sqrt{\frac{7{,}83^2}{10} + \frac{2{,}68^2}{10}}} = \frac{15}{2{,}47 + 0{,}84} = 4{,}51$$

Tentukan derajat kebebasan df:

$$df = \frac{\left(\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}\right)^2}{\frac{\left(\frac{s_1^2}{n_1}\right)^2}{n_1 - 1} + \frac{\left(\frac{s_2^2}{n_2}\right)^2}{n_2 - 1}}$$

Subtitusi nilai persamaan:

$$\frac{s_1^2}{n_1} = \frac{61{,}43}{10} = 6{,}143;\ \frac{s_2^2}{n_2} = \frac{7{,}21}{10} = 0{,}721$$

$$(6{,}143\ +\ 0{,}721)^2 = 45{,}07$$

$$\frac{(6{,}143)^2}{9} + \frac{(0{,}721)^2}{9} = \frac{37{,}75}{9} + \frac{0{,}52}{9} = 4{,}2$$

$$df = \frac{45{,}07}{4{,}2} = 11{,}08$$

Menentukan nilai t kritis dari table distribusi untuk df = 11 dan α= 0,05 (uji dua sisi). Nilai t kritis = ±2,109.

Sehingga diperoleh kesimpulan:

Nilai t hitung = 4,51

Nilai t kritis = ±2,109

Karena 4,51 berada di luar batas kritis ±2,109; maka **menolak hipotesis nol**. Artinya, ada perbedaan signifikan rata-rata tinggi badan antara pria dan wanita.

Tugas

1. Sebuah koin dilempar 10 kali. Berapa probabilitas mendapatkan tepat 3 kali gambar?
2. Rata-rata 3 pelanggan datang ke sebuah toko setiap jam. Berapa probabilitas bahwa tepat 5 pelanggan akan datang dalam satu jam tertentu?
3. Skor IQ dalam populasi mengikuti distribusi normal dengan mean 100 dan standar deviasi 15. Berapa persentase populasi yang memiliki IQ di atas 130?
4. Sebuah sampel acak dari 25 produk memiliki rata-rata berat 50 gram dan standar deviasi 2 gram. Hitung interval kepercayaan 95% untuk rata-rata berat populasi.

BAB VI

Tujuan Pembelajaran

Setelah mempelajari bab ini, pembaca diharapkan mampu:

1. Memahami konsep dasar regresi dan korelasi dalam analisis data.
2. Menggunakan regresi linier sederhana untuk memodelkan hubungan antara dua variabel.
3. Menghitung dan menginterpretasikan nilai korelasi untuk menentukan kekuatan hubungan antarvariabel.
4. Menganalisis hasil regresi dan korelasi untuk pengambilan keputusan dalam berbagai bidang.

Pokok Bahasan

1. Regresi
2. Korelasi

REGRESI DAN KORELASI

Regresi

Regresi adalah teknik analisis statistik yang sangat penting dalam berbagai bidang, termasuk ekonomi, ilmu sosial, kesehatan, dan ilmu alam. Dengan regresi, peneliti dapat membangun model matematis yang menggambarkan hubungan antara satu atau lebih variabel independen dan satu variabel dependen.

Regresi Linier Sederhana

Regresi linier sederhana adalah metode statistik yang digunakan untuk menganalisis hubungan antara satu variabel independen (X) dan satu variabel dependen (Y). Tujuan dari regresi linier sederhana adalah untuk menemukan garis terbaik yang dapat memprediksi nilai variabel dependen berdasarkan nilai variabel independen. Model regresi ini dapat dinyatakan dalam bentuk persamaan garis lurus:

$$Y = a + bX$$

Di mana:

Y : adalah variabel dependen (yang ingin diprediksi).

X : adalah variabel independen (yang digunakan untuk memprediksi).

a : adalah intercept (nilai Y ketika X = 0).

b : adalah koefisien regresi (perubahan rata-rata pada Y untuk setiap unit perubahan pada X).

Contoh Regresi Linier Sederhana

Misalkan kita ingin menganalisis hubungan antara jumlah jam belajar (X) dan nilai ujian (Y) siswa. Kita mengumpulkan data dari 10 siswa sebagai berikut:

Tabel 6.1 Data hubungan jam belajar dengan Nilai Ujian

Jam Belajar (X)	Nilai Ujian (Y)
1	50
2	55
3	65
4	70
5	75
6	80
7	85
8	90
9	95
10	100

Diketahui:

$n = 10$

$\bar{X} = 76{,}5$

$\bar{Y} = 5{,}5$

$\Sigma X = 55$

$\Sigma Y = 765$

$\Sigma X2 = 385$

$\Sigma XY = 4660$

Rumus Kemiringan (b)

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

$$b = \frac{10(4660) - (55)(765)}{10(385)^2 - (55)^2}$$

$$b = \frac{4525}{825} = 5{,}4848$$

Rumus Intersep (a)

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

$$a = \left(76{,}5 - (5{,}4848\ \times 5{,}5)\right) = 46{,}333$$

Persamaan Regresi

$$Y\ =\ 46{,}33 + 5{,}48X$$

Interpretasi:

Intercept ($\boldsymbol{a = 46,33}$): Jika jumlah jam belajar (X) adalah 0, maka nilai ujian (Y) diperkirakan sebesar 46,33.

Koefisien regresi ($\boldsymbol{b = 5,8}$): Setiap tambahan 1 jam belajar (X), nilai ujian (Y) diperkirakan akan meningkat sebesar 5,48.

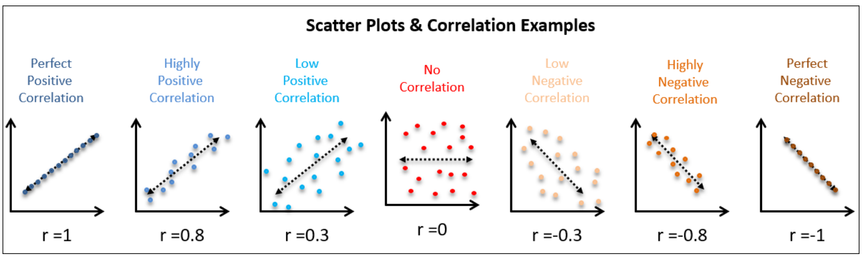
Aplikasi Regresi

Regresi digunakan dalam berbagai aplikasi, seperti:

- **Prediksi**: Misalnya, memprediksi penjualan berdasarkan pengeluaran iklan.
- **Analisis Risiko**: Dalam keuangan, regresi dapat digunakan untuk menilai risiko investasi berdasarkan faktor-faktor tertentu.
- **Evaluasi Kebijakan**: Dalam ilmu sosial, regresi dapat membantu mengevaluasi dampak kebijakan publik terhadap masyarakat.

Korelasi

Korelasi adalah suatu teknik statistik yang digunakan untuk mengukur dan menganalisis hubungan antara dua variabel. Dalam konteks ini, variabel dapat berupa data numerik yang diukur dalam skala interval atau rasio. Korelasi memberikan informasi tentang seberapa kuat hubungan antara dua variabel dan arah hubungan tersebut, apakah positif, negatif, atau tidak ada hubungan sama sekali.



Gambar 20. Scatter Plots & Correlation

Korelasi sering kali diukur menggunakan koefisien korelasi Pearson, yang memberikan nilai antara -1 hingga 1. Nilai 1 menunjukkan hubungan positif sempurna, di mana peningkatan satu variabel selalu diikuti oleh peningkatan variabel lainnya. Sebaliknya, nilai -1 menunjukkan hubungan negatif sempurna, di mana peningkatan satu variabel selalu diikuti oleh penurunan variabel lainnya. Nilai 0 menunjukkan tidak adanya hubungan antara kedua variabel.

Jenis Korelasi

1. **Korelasi Positif**: Ketika satu variabel meningkat, variabel lainnya juga meningkat. Misalnya, ada korelasi positif antara pendidikan dan pendapatan; semakin tinggi pendidikan seseorang, semakin tinggi pendapatannya.
2. **Korelasi Negatif**: Ketika satu variabel meningkat, variabel lainnya menurun. Contohnya, ada korelasi negatif antara tingkat pengangguran dan pendapatan rata-rata; ketika pengangguran meningkat, pendapatan rata-rata cenderung menurun.
3. **Korelasi Nol**: Tidak ada hubungan yang signifikan antara kedua variabel. Misalnya, tidak ada korelasi antara tinggi badan seseorang dan kemampuan mereka dalam matematika.

Koefisien Korelasi

Koefisien korelasi, seperti Pearson's r, digunakan untuk mengukur kekuatan dan arah hubungan. Nilai mendekati 1 atau -1 menunjukkan hubungan yang kuat, sedangkan nilai mendekati 0 menunjukkan hubungan yang lemah.

Aplikasi Korelasi

Korelasi sering digunakan dalam penelitian untuk:

- **Menemukan Hubungan**: Menentukan apakah ada hubungan antara dua variabel sebelum melakukan analisis lebih lanjut.

- **Pengembangan Hipotesis**: Membantu peneliti mengembangkan hipotesis yang dapat diuji lebih lanjut dengan analisis regresi.

Rumus Korelasi

Rumus untuk menghitung koefisien korelasi Pearson (r) adalah:

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

Di mana:

n = jumlah pasangan data

x = variabel pertama

y = variabel kedua

Contoh 1: Korelasi Positif

Jam Belajar (X)	Nilai Ujian (Y)
1	60
2	70
3	75
4	80
5	90

penyelesaian

tentukan $\sum x,\ \sum y,\ \sum xy,\ \sum x^2, dan \sum y^2$

Jam Belajar (X)	Nilai Ujian (Y)	*XY*	*X*²	*Y*²
1	60	60	1	3600
2	70	140	4	4900
3	75	225	9	5625
4	80	320	16	6400
5	90	450	25	8100
Total		**1195**	**55**	**28625**

Dari tabel di atas, kita dapat menghitung:

$$\sum x = 1 + 2 + 3 + 4 + 5 = 15$$

$$\sum y = 60 + 70 + 75 + 80 + 90 = 375$$

$$\sum xy = 60 + 140 + 225 + 320 + 450 = 1195$$

$$\sum x2 = 12 + 22 + 32 + 42 + 52 = 55$$

$$\sum y2 = 602 + 702 + 752 + 802 + 902 = 28625$$

Hitung korelasi (r)

$$r = \frac{5(1195) - (15)(375)}{\sqrt{[5(55) - (15)^2][5(28625) - (375)^2]}} = \frac{350}{353{,}55} = 0{,}989$$

Koefisien korelasi $r \approx 0.99$ menunjukkan adanya korelasi positif yang sangat kuat antara jam belajar dan nilai ujian. Ini berarti bahwa semakin banyak jam yang dihabiskan untuk belajar, semakin tinggi nilai ujian yang diperoleh siswa.

Tugas

1. Seorang guru ingin mengetahui hubungan antara jumlah jam belajar per hari dan nilai ujian siswa. Data yang dikumpulkan dari 7 siswa ditunjukkan dalam grafik berikut:

a) Berdasarkan grafik, bagaimana Anda mendeskripsikan hubungan antara jam belajar dan nilai ujian?
b) Hitung koefisien korelasi antara jam belajar dan nilai ujian. Interpretasikan hasilnya.
c) Tuliskan persamaan regresi linear untuk data tersebut.
d) Jika seorang siswa belajar selama 6 jam per hari, berapa nilai ujian yang diperkirakan?
e) Apa keterbatasan dari model ini dalam memprediksi nilai ujian?

2. Seorang peneliti kesehatan mengumpulkan data tinggi badan (dalam cm) dan berat badan (dalam kg) dari 8 orang dewasa:

   Tinggi (X): 160, 165, 170, 175, 180, 185, 190, 195

   Berat (Y): 55, 60, 65, 70, 75, 80, 85, 90

a) Hitung koefisien korelasi antara tinggi badan dan berat badan.

b) Interpretasikan koefisien korelasi dalam konteks kesehatan.

c) Tentukan persamaan regresi linear untuk data tersebut.
d) Estimasi berat badan seseorang dengan tinggi 178 cm.
e) Diskusikan keterbatasan penggunaan model ini untuk memprediksi berat badan.

BAB VII

Tujuan Pembelajaran

Setelah mempelajari bab ini, mahasiswa diharapkan mampu:

1. Memahami konsep dasar Analisis Varians (ANOVA) dan kegunaannya dalam statistik.
2. Menjelaskan perbedaan antara ANOVA satu arah dan ANOVA dua arah.
3. Melakukan uji ANOVA untuk membandingkan rata-rata lebih dari dua kelompok.
4. Menginterpretasikan hasil uji ANOVA dalam pengambilan keputusan statistik.

Pokok Bahasan

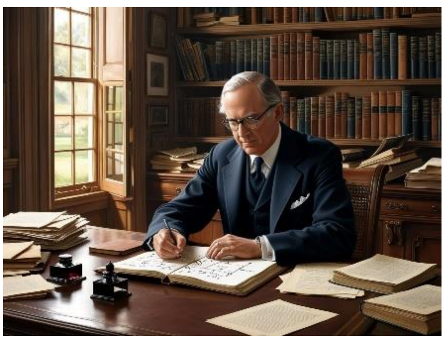
1. Konsep Dasar ANOVA
2. Jenis-Jenis ANOVA

ANALISIS VARIANS (ANOVA)

ANOVA, atau Analisis Varians, adalah metode statistik yang digunakan untuk membandingkan rata-rata dari dua atau lebih kelompok untuk menentukan apakah ada perbedaan yang signifikan di antara mereka. Metode ini sangat berguna ketika kita ingin mengetahui apakah variasi dalam data disebabkan oleh perbedaan antara kelompok atau hanya karena variasi acak dalam data.

ANOVA bekerja dengan membagi total variasi dalam data menjadi dua komponen: variasi antara kelompok dan variasi dalam kelompok. Dengan membandingkan kedua variasi ini, ANOVA dapat menentukan apakah perbedaan rata-rata antara kelompok-kelompok tersebut cukup besar untuk dianggap signifikan secara statistik.

Konsep ANOVA pertama kali diperkenalkan oleh statistikawan Inggris, Ronald A. Fisher, pada tahun 1920-an. Fisher mengembangkan metode ini sebagai bagian dari penelitian agrikultur untuk menganalisis hasil panen dari berbagai perlakuan. Ia menyadari bahwa untuk memahami pengaruh perlakuan terhadap hasil, penting untuk membandingkan variasi dalam hasil panen di antara kelompok yang berbeda.



Gambar 21. Ronald A. Fisher (1890-1962)

Fisher juga memperkenalkan istilah "analisis varians" dan mengembangkan rumus serta tabel distribusi F, yang digunakan untuk menguji hipotesis dalam ANOVA. Sejak saat itu, ANOVA telah menjadi alat penting dalam statistik dan penelitian ilmiah, digunakan dalam berbagai bidang seperti psikologi, biologi, ekonomi, dan ilmu sosial.

Jenis-jenis ANOVA

1. **ANOVA Satu Arah (One-Way ANOVA)**: Digunakan untuk membandingkan rata-rata dari tiga kelompok atau lebih berdasarkan satu variabel independen. Misalnya, membandingkan rata-rata nilai ujian siswa dari tiga kelas yang berbeda.
2. **ANOVA Dua Arah (Two-Way ANOVA)**: Digunakan untuk membandingkan rata-rata dari dua variabel independen. Misalnya, membandingkan rata-rata nilai ujian siswa berdasarkan kelas dan jenis kelamin.
3. **ANOVA Berulang (Repeated Measures ANOVA)**: Digunakan ketika pengukuran dilakukan pada subjek yang sama di beberapa waktu atau kondisi. Misalnya, mengukur tekanan darah pasien sebelum, selama, dan setelah pengobatan.

Contoh ANOVA Satu Arah

Kelas A	Kelas B	Kelas C
85	78	90
88	82	85
90	79	92

87	81	88
86	80	91

Langkah-langkah Analisis ANOVA

1. **Hipotesis**

- **Hipotesis Nol (*H_0*)**: Tidak ada perbedaan rata-rata nilai ujian antara kelas A, B, dan C.
- **Hipotesis Alternatif (*H_a*)**: Ada perbedaan rata-rata nilai ujian antara setidaknya dua kelas.

2. Hitung Rata-rata dan Varians

Rata-rata untuk Setiap Kelompok

- **Kelompok a**:

$$\bar{X}_a = \frac{85 + 88 + 90 + 87 + 86}{5} = \frac{436}{5} = 87{,}2$$

- **Kelompok b**:

$$\bar{X}_b = \frac{78 + 82 + 79 + 81 + 80}{5} = \frac{400}{5} = 80$$

- **Kelompok c**:

$$\bar{X}_b = \frac{90 + 85 + 92 + 88 + 91}{5} = \frac{446}{5} = 89{,}2$$

3. Hitung Total Rata-rata

$$\bar{\boldsymbol{X}}_{\boldsymbol{total}} = \frac{436 + 400 + 446}{15} = \frac{1282}{15} = 85{,}47$$

4. Hitung Variasi antara kelompok (SSB)

$$SSB = n \times ((\bar{X}_a - \bar{X}_{total})^2 + (\bar{X}_b - \bar{X}_{total})^2 + (\bar{X}_c - \bar{X}_{total})^2)$$

$$SSB = 5 \times ((87{,}2 - 85{,}47)^2 + (80 - 85{,}47)^2 + (89{,}2 - 85{,}47)^2) = 233{,}9335$$

5. Hitung variasi dalam kelompok (SSW)

$$SSW = \sum \left(X_{ij} - \bar{X}_i\right)^2$$

Hitung variasi untuk setiap kelompok

$$SSW_a = (85 - 87{,}2)^2 + (88 - 87{,}2)^2 + (90 - 87{,}2)^2 + (87 - 87{,}2)^2 + (86 - 87{,}2)^2$$

$$SSW_a = 4{,}84 + 0{,}64 + 7{,}84 + 0{,}04 + 1{,}14 = 14{,}8$$

$$SSW_b = (78 - 80)^2 + (82 - 80)^2 + (79 - 80)^2 + (81 - 80)^2 + (80 - 80)^2$$

$$SSW_a = 4 + 4 + 1 + 1 + 0 = 10$$

$$SSW_c = (90 - 89{,}2)^2 + (85 - 89{,}2)^2 + (92 - 89{,}2)^2 + (88 - 89{,}2)^2 + (91 - 89{,}2)^2$$

$$SSW_a = 0{,}64 + 17{,}64 + 7{,}84 + 1{,}14 + 3{,}24 = 30{,}8$$

$$Total\ SSW = SSW_a + SSW_b + SSW_c = 14{,}8 + 10 + 30{,}8 = 55{,}6$$

6. Hitung total variasi (SST)

$$SST = SSB + SSW = 233{,}9335 + 55{,}6 = 289{,}5335$$

7. Hitung derajat kebebasan (df)
   - Derajat kebebasan antara kelompok $(dfB): k - 1 = 3 - 1 = 2$
   - Derajat kebebasan antara kelompok $(dfW): N - k = 15 - 3 = 12$

Tabel Uji F

α = 0,05				**df1 =**
df2=(n-k-1)	**1**	**2**	**3**	**4**
1	161.448	199,500	215.707	224,583
2	18,513	19,000	19,164	19,247
3	10,128	9,552	9,277	9,117
4	7,709	6,944	6,591	6,388
5	6,608	5,786	5,409	5,192
6	5,987	5,143	4,757	4,534
7	5,591	4,737	4,347	4,120
8	5,318	4,459	4,066	3,838
9	5,117	4,256	3,863	3,633
10	4,965	4,103	3,708	3,478
11	4,844	3,982	3,587	3,357
12	4,747	3,885	3,490	3,259
13	4,667	3,806	3,411	3,179

8. Hitung Mean Square
   - Mean Square Between (MSB):

$$MSB = \frac{SSB}{dfB} = \frac{233{,}933}{2} = 116{,}966$$

   - Mean Square Within (MSW):

$$MSW = \frac{SSB}{dfW} = \frac{55{,}6}{2} = 4{,}633$$

9. Hitung nilai F

$$F = \frac{MSB}{MSW} = \frac{116{,}966}{4{,}633} = 25{,}3$$

10. Uji Hipotesis

Anova: Single Factor						
SUMMARY						
Groups	*Count*	*Sum*	*Average*	*Variance*		
a	5	436	87,2	3,7		
b	5	400	80	2,5		
c	5	446	89,2	7,7		
ANOVA						
Source of Variation	*SS*	*df*	*MS*	*F*	*P-value*	*F crit*
Between Groups	234,133	2	117,067	25,2662	4,99411E-05	3,88529
Within Groups	55,6	12	4,63333			
Total	289,733	14				

Nilai F tabel < F hitung (3,884 < 25,3)

P-value < alpha 0,05 (0,000049 < 0,05)

Karena nilai alpha 0,05 > dari p-value maka menolak hipotesis H_0, yang berarti menerima hipotesis H_1 yaitu ada perbedaan rata-rata nilai ujian antara setidaknya dua kelas.

Tugas

1. Seorang peneliti ingin mengetahui apakah ada perbedaan signifikan dalam nilai ujian matematika siswa berdasarkan tiga metode pengajaran yang berbeda: **ceramah**, **diskusi**, dan **praktik langsung**. Data nilai ujian dari masing-masing kelompok adalah sebagai berikut:

Metode Ceramah	Metode Diskusi	Metode Praktik Langsung
70	75	85
72	78	88
68	74	90
71	76	87
69	77	89

2. Seorang peneliti ingin mengetahui apakah **jenis pupuk** (A, B, C) dan **jenis tanah** (Tanah 1 dan Tanah 2) mempengaruhi tinggi tanaman. Peneliti menggunakan tiga jenis pupuk dan dua jenis tanah. Setiap kombinasi pupuk dan tanah diuji pada 3 tanaman, sehingga total ada 18 pengamatan. Data tinggi tanaman (dalam cm) adalah sebagai berikut

Jenis Pupuk	Tanah 1	Tanah 2
A	20	25
A	22	24
B	30	35
B	32	34
B	31	36
C	40	45
C	42	44
C	41	46

PENUTUP

Statistik merupakan alat yang sangat penting dalam berbagai bidang, membantu kita memahami data, membuat keputusan yang lebih baik, dan melakukan prediksi yang lebih akurat. Melalui buku **"Statistik Dasar"** ini, diharapkan pembaca dapat memperoleh pemahaman yang kuat mengenai konsep-konsep dasar statistik, baik dalam analisis deskriptif maupun inferensial.

Kami menyadari bahwa ilmu statistik terus berkembang, dan masih banyak hal yang dapat dipelajari lebih lanjut. Oleh karena itu, kami mengajak para pembaca untuk terus mengeksplorasi dan mengembangkan pemahaman mereka dalam bidang ini.

Semoga buku ini bermanfaat bagi pembaca, khusunya mahasiswa, dalam memahami dan menerapkan metode statistik dalam berbagai aspek kehidupan dan penelitian. Kami juga terbuka terhadap kritik dan saran yang dapat membantu meningkatkan kualitas buku ini di masa mendatang.

Terima kasih atas perhatian dan waktu yang telah diberikan untuk mempelajari buku ini. Selamat belajar dan semoga sukses dalam perjalanan akademik maupun profesional Anda!

TENTANG PENULIS

Syaiful Mansyur,

Tempat dan tanggal lahir, Kendari, 15 Oktober 1986. Menyelesaikan pendidikan sarjana (S1) pada jurusan Teknik Elektro di Universitas Islam Indonesia (2009), dan master (S2) pada jurusan Teknik Industri di Universitas Gadjah Mada (2015). Pernah bekerja sebagai dosen tetap di Akademi Komputer Global Kendari, dan sejak tahun 2015 menjadi dosen tetap di Universitas Proklamasi 45 Yogyakarta pada program studi Teknik Industri. Selain itu, dalam bidang praktisi penulis pernah menjabat sebagai *Product Design and Development* di perusahaan pengecoran logam dan permesinan PT. Mulya Jaya, Klaten. Saat ini penulis mendirikan perusahaan yang bergerak di bidang yang sama pengecoran logam dan permesinan yaitu CV. Okto Jaya Mandiri. Fokus penelitian penulis yaitu pada perancangan produk, rancang bangun mesin, pengolahan dan pengelolaan teknologi biomassa menjadi energi terbarukan.

Korespondensi: syaiful.m@up45.ac.id

Buku yang diterbitkan sebelumnya:

Black Gold; Bisnis dan Teknologi Pembuatan Produk Briket Arang dan Turunannya edisi 1 dan 2. - Penerbit Graha Ilmu (2019-2022)

STATISTIK DASAR

Statistik adalah ilmu yang tidak hanya berbicara tentang angka, tetapi juga tentang bagaimana kita memahami dunia melalui data. Dalam berbagai bidang—mulai dari bisnis, teknik, hingga sains—statistik menjadi alat penting untuk mengambil keputusan yang lebih baik dan berbasis bukti.

Buku "Statistik Dasar" ini dirancang untuk memberikan pemahaman yang jelas dan sistematis mengenai konsep-konsep statistik, mulai dari dasar hingga penerapan dalam analisis data. Beberapa topik utama yang dibahas dalam buku ini meliputi:

✅ Pengantar Statistik: Sejarah, definisi, dan tujuan statistik
✅ Statistik Deskriptif: Pengukuran pemusatan, variabilitas, serta penyajian data dalam grafik dan diagram
✅ Statistik Inferensial: Estimasi parameter, uji hipotesis, dan analisis data
✅ Probabilitas: Konsep dasar peluang, Teorema Bayes, serta permutasi dan kombinasi
✅ Distribusi Peluang: Distribusi binomial, Poisson, normal, t-Student, dan F
✅ Regresi dan Korelasi: Analisis hubungan antara variabel
✅ Analisis Varians (ANOVA): Perbandingan rata-rata lebih dari dua kelompok

Kutipan Inspiratif tentang Statistik

📌 "Statistik dapat digunakan untuk menyatakan hampir semua hal—bahkan kebalikannya."
— Karl Pearson

📌 "Tanpa data, Anda hanyalah seseorang dengan opini."
— W. Edwards Deming

📌 "Statistik adalah ilmu yang memungkinkan kita untuk menarik kesimpulan dari data yang tidak sempurna."
— George E. P. Box

Dengan buku ini, Anda tidak hanya belajar tentang angka, tetapi juga bagaimana menjadikannya alat yang kuat untuk memahami dunia. Selamat belajar dan semoga sukses! 🚀📊

UP45 PRESS

Syaiful Mansyur, M.Sc